Predicting the impact of climatic and geomorphic changes on flood hazards in Central Himalayan Rivers. An environmental modelling framework to predict potential flood hazards in the Karnali River in Nepal.

PINK, IVO,THOMAS

Predicting the impact of climatic and geomorphic changes on flood hazards in Central Himalayan Rivers

An environmental modelling framework to predict potential flood hazards in the Karnali River in Nepal

---

Ivo Thomas Pink

Thesis submitted for in fulfilment of the

requirements for the degree of Doctor of Philosophy

Department of Geography

Institute of Hazard, Risk and Resilience

Durham University

2023

Abstract

Flood hazards that originate in the Himalayas threaten the lives and livelihoods of millions of people in the downstream Ganga plain. The Himalayas are a region with high geomorphic activity that is sensitive to anthropogenic climate change. The development of effective flood management strategies, therefore, needs to predict the impact of geomorphic and climatic changes on the flood hazards which is not possible with currently existing hazard modelling frameworks.

This thesis introduces an environmental modelling framework to predict potential flood hazards for the projected climatic and related topographic conditions that can be used to adapt to future flood hazards. This framework links hydrological, statistical, morpho- and hydrodynamic models to predict: i) the climate change impact on the flood discharge into the floodplain; ii) the topographic evolution of the floodplain for the projected flood flows, and; iii) the potential spatial flood hazard characteristics in the floodplain for the projected flood discharges and new topographies. This modelling framework is developed and implemented for the Central Himalayan Karnali River in Nepal.

The predictions of the potential hazards provide evidence that flood hazards in the Karnali River will intensify with increased greenhouse gas emissions. The current 1-in-100-year flood discharge into the plain (1975 – 2014) is projected to increase by 40% (medium emissions) and by 79% (high emissions) at the end of the century (2060 – 2099) (median predictions). The return period of the current 1-in-100-year event is predicted to decrease to 11 – 16 years (medium emissions) and 2 – 7 years (high emissions). The projected increase in the flood discharge intensifies the magnitude of topographic change in the downstream floodplain.

The projected increases in the flood discharge cause deeper inundations that extend further. The current 1-in-100-years event (1975 – 2014) inundates 39% of the simulated area of the floodplain. This extent is projected to increase to 51% (medium emissions) and 60% (high emissions) at the end of the century (2060 – 2099) (median predictions). The predictions indicate that topographic change alters flood hazards on the local scale and that river and flood engineering potentially intensifies the downstream hazard.

The proposed modelling framework allows continued assessment of flood risk enabling the flood management strategies to be developed for future conditions. Recommendations are provided to improve the predictions by adjusting the modelling framework and targeted investments in surveys.

Contents

Figures

Tables

Appendices

Abbreviations

AD	Anderson-Darling
ADPC	Asian Disaster Preparedness Center
AEP	Annual Exceedance Probability
AMAX	Annual Maximum Flow
AWEI	Automated Water Extraction Index
BaRatin	Bayesian Rating Curve Framework
BaRE	Bayesian Recursive Estimation
CBS	Central Bureau of Statistics
CDF	Cumulative Density Functions
CFD	cumulative frequency distributions
CHM	Canopy Height Models
C-L	CAESAR-LISFLOOD
CMIP6	Coupled Model Intercomparison Project 6
CRN	Cosmogenic Radionuclide
CRS	Coordinate System
CRU	Climate Research Unit
DDF	Degree-Day-Factor
DEM	Digital Elevation Model
DFRS	Department of Forest Research and Survey
DHM	Department of Hydrology and Meteorology Nepal
DoD	DEM of Difference
DSM	Digital Surface Model
DTM	Digital Terrain Model
DWIP	Department of Water Induced Disaster Prevention
eGLUE	extended Generalised Likelihood Uncertainty Estimation
ET_{act}	Actual evapotranspiration
EVD	Extreme Value Distribution
EXP	Exponential distribution
FFA	Flood Frequency Analysis
FM	Flood Magnitude
FPAR	Fraction of absorbed Photosynthetically Active Radiation
GAM	General Additive Models
GCM	Global Circulation Models

GEV	Generalised Extreme Value distribution
GLO	Generalised Logistic distribution
GLOF	Glacier Lake Outburst Floods
GLUE	Generalised Likelihood Uncertainty Estimation
GPA	Generalized Pareto distribution
GSD	Grain Size Distribution
GST	Gravel-Sand-Transition
GUM	Gumbel distribution
HBV	Hydrologiska Byråns Vattenbalansavdelning
HKH	Hindu Kush Himalayas
HPC	High-Performance Computing
HRU	Hydrological Response Unit
HWSD	Harmonized World Soil Database
IDW	Inverse Distance Weighting
Kc	Crop coefficient
LAI	Leaf Area Index
LHS	Latin Hypercube Sampling
LiDAR	Light Detection And Ranging
LOF	Landslide Outburst Floods
MAAT	Mean Annual Air Temperature
MODIS	Moderate Resolution Imaging Spectroradiometer
NPC	National Planning Commission
NSE	Nash-Sutcliffe Efficiency
PBIAS	percentage BIAS
POT	Peak-Over-Threshold
PT3	Pearson Type III
Qn	The discharge predicted by the n^{th} percentile of the ensemble
R^2	Coefficient of determination
RSA	Regional Sensitivity Analysis
SD	Standard Deviation
SPHY	SPatial Processes in Hydrology
SRTM	Shuttle Radar Topography Mission
SSP	Shared Socioeconomic Pathways
SVE	Saint-Venant Equations
SWAT	Soil and Water Assessment Tool

SWE	Shallow-Water Equations
USDA	United States Department of Agriculture
WAK	Wakeby distribution
WEI	Weibull distribution
WFDEI	WATCH forcing data methodology applied to ERA-Interim dataset

Declaration

I declare that this thesis is entirely my own work unless referenced to the contrary in the text. No part of this thesis has been submitted in any previous application for a degree or qualification in this or any other university or institution.

Ivo Thomas Pink

November 2023

Statement of Copyright

Ivo Thomas Pink

November 2023

Acknowledgements

The PhD has come to an end and I owe my thanks to many people who have guided and supported me throughout my way with all its highs and lows. I cannot mention each person here but I would like to express my sincere gratitude to the following people who have been an immense help during and before this PhD.

First of all, I would like to thank my main supervisor Sim Reaney who has supported me throughout this research. Sim has guided me to develop and execute research projects, motivated me to continue during the lows and given me plenty of advice and opportunities far beyond the PhD research. I always felt Sim's support and could not have asked for a better supervisor.

I would also like my second supervisors Isabella Bovolo and Richard Hardy who have both guided me during the past four years. I thank Isabella for her constant optimism and cheerful attitude which has helped me to continue during times of frustration. I thank Rich for asking simple but fundamental questions that I have condemned many times but which were essential for my academic development.

With regard to the PhD research, I would also like to thank Callum Pearson for introducing me to the East Rapti Catchment and Elizabeth Dingle for sharing her knowledge and information of the geomorphology of Himalayan rivers. I would also like to thank Gunjan Silwal for her support in organising datasets during the pandemic and discussing high mountain hydrology.

This thesis is the product of decades of learning and I would like to thank the following persons who have guided me along the way: my high school teacher Mr Christoph Stephan for his encouragement without which I might not have aimed for an academic career; Simone Giertz who introduced me to environmental modelling in my undergraduate studies; Christiane Schmullius and Alexander Brenning for introducing me to remote sensing and geospatial data science in my postgraduate studies; and Manfred Fink, Sven Kralisch and especially Santosh Nepal who introduced me to hydrological modelling and Himalayan River systems.

I have met many nice people and made new friends who made my time in Durham memorable. A special thanks to Maria Paz for her constant support and many shared pints. I cannot write this acknowledgement without mentioning Burag, Isabella, Liping Thiago, Thuli, William, Xiaochen and Yuan. Thank you for all the moments of joy throughout the past years.

I also want to thank my long-time friends for their continuous support throughout all these years that I spent abroad. Thanks to Claus, Fred, Manuel, Phillipp and Roman. Furthermore, I would like to thank Gunjan for her support throughout the isolating times during the pandemic.

My final thanks go to my family for their unconditional love and support throughout my life. My parents have always thought about their children first before thinking about themselves. I took it for granted too often but I am very thankful for their help and guidance.

Acknowledgement of funding

This research was funded by the Charles Wilson Research Doctoral Studentship. I would like to thank the private donor for this generous research grant that enabled me to buy essential equipment and data and participate in conferences.

Additional support to cover my living expenses in the final months of this research was provided by the Institute for Hazard, Risk and Resilience and the Department of Geography of Durham University.

1 Introduction

The Himalayas provide freshwater to 10% of the global population but also cause frequent flooding in the densely populated plains downstream of the mountain range (Dilley *et al.*, 2005; Shrestha *et al.*, 2015b; Dingle, 2018; Sharma *et al.*, 2019). The steep slopes of the southern flanks of the Himalayas facilitate the fast conversion of rainfall to runoff during intense, large-scale monsoon rainfall (MacClune *et al.*, 2015; Shrestha *et al.*, 2015a; Bhandari *et al.*, 2018; DHM, 2018). The dense drainage network and the steep slopes of the rivers cause runoff concentration resulting in flood events with a high magnitude and short duration. After exiting the Himalayas, the rivers enter the flat alluvial Indo-Gangetic plain where they rework the topography which alters the flood risk for future events (Dixit, 2009; Sinha *et al.*, 2014). This thesis uses environmental modelling techniques to investigate how flood hazards may respond to the changing climate in the upstream catchment and how these large floods may change the topographies in the floodplain for the Central Himalayan Karnali River in Nepal.

The next section provides an overview of the flood hazards in Central Himalayan river systems. Section 1.2 presents the research aim and objectives, Section 1.3 introduces the stages of the research approach, and Section 1.4 provides an overview of the thesis structure.

1.1 Riverine floods in Nepal

The Indo-Gangetic plain stands out as one of the global regions most susceptible to flooding, owing to a combination of its geographical characteristics and socio-economic factors (Dilley *et al.*, 2005). Situated downstream of the southern slopes of the Himalayas, this plain encompasses territories in Pakistan, India, Nepal, and Bangladesh. Over the past five decades (1971 – 2020), these four nations collectively witnessed 508 flood events, resulting in the loss of 128,114 lives and affecting 1.27 billion people (Table 1.1). The high flood risk is underscored by the fact that these countries accommodate 23% of the global population and experience 10% of the globally reported flood events but suffer 40% of the global fatalities. Although the reported statistics encompass national data and extend beyond the boundaries of the Indo-Gangetic plain, it remains evident that this region, among other regions, exhibits the highest flood risk in terms of mortality rates and economic losses (Dilley *et al.*, 2005).

Nepal encompasses both the origin of floodwaters in the Himalayas and areas at risk of flooding within the Indo-Gangetic plain (Figure 1.1). Nepal is situated within the Central Himalayas and extends from east to west, the country's southern region transitions into the flat Terai region, which serves as the northern extension of the Indo-Gangetic plain (Upreti, 2001). This geographical configuration results in frequent and extensive riverine floods in the densely populated Terai plain during the monsoon season (DHM, 2018). These floods are of global significance due to their high lethality, averaging 150 casualties per event, and ranking as the fifth-highest fatality ratio worldwide (Table 1.1).

Table 1.1: Statistics of the casualties and affected population by flood hazards in the period 1971 – 2020 recorded in EM-DAT/CRED (2021). The population is obtained from the World Bank (2023) and describes the population for the year 2022. South Asia includes the riparian states of the Indo-Gangetic Plain, namely Bangladesh, India, Nepal and Pakistan.

	Population [Million]	**Flood Events**	**Casualties**		**Affected pop. [Million]**	
			Total	Per event	Total	Per event
Global	7,950	5,112	328,613	64.3	3,716.1	0.727
South Asia	1,858	508	128,114	252.2	1,272.2	0.250
Nepal	30.5	49	7,364	150.3	5.6	0.115

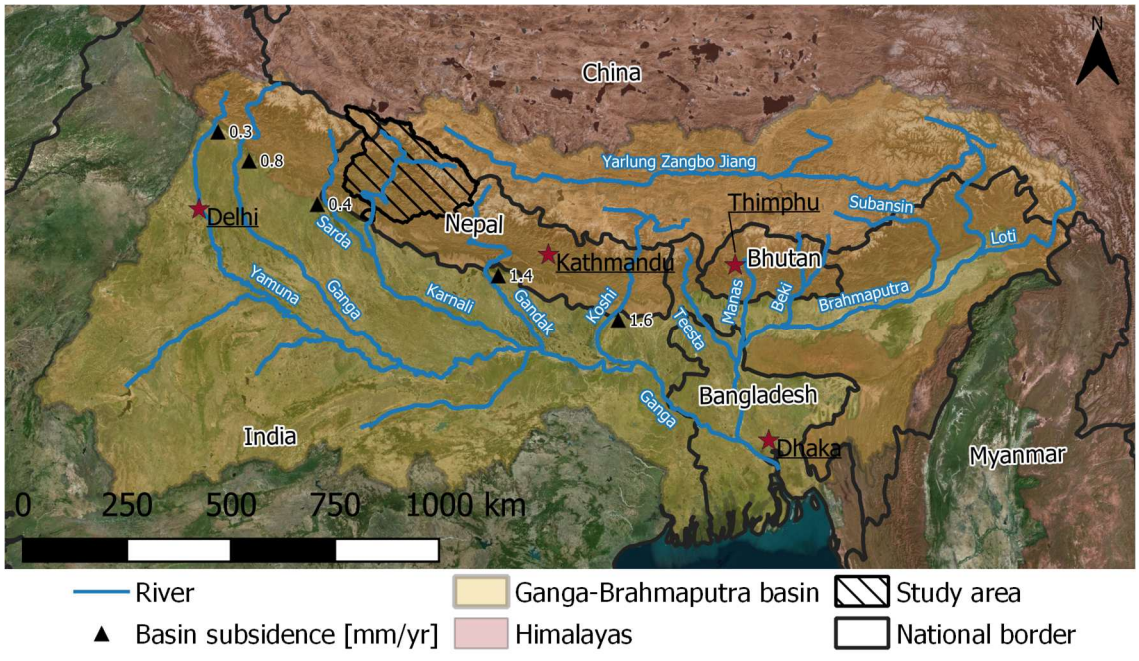


Figure 1.1: The Ganga-Brahmaputra basin (FAO, 2011) in South Asia and its main tributaries (FAO, 2022). Vast plains stretch along the Brahmaputra and Ganga rivers south of the Himalayas (ICIMOD, 2008). This plain subsides as the Indian plate submerges beneath the Eurasian plate. The average subsidence rates in the plain are obtained from Dingle (2018, Table. 3.1). The Terai is located south of the Himalayas in Nepal. The background map is a Bing Maps satellite composite.

The largest floods in the Terai region in recent years are:

i) In 2013, widespread, intense monsoon rainfall over northern India and Nepal led to vast floods killing 1109 people and affecting tens of thousands (CARITAS India, 2013; Red Cross and Red Crescent, 2013; DWIDP, 2014);
ii) In 2014, the flooding of multiple rivers across Nepal killed 202 people, displaced 50,965 people, affected 184,745 people and caused damages equivalent to 6.5% of the GDP (DWIDP, 2015; Shrestha *et al.*, 2015a);
iii) In 2017, severe floods of multiple rivers inundated 70% of the Terai region which caused 134 fatalities, destroyed or damaged 190,000 houses and affected 1.7 million people (NPC, 2017; DHM, 2018).

Flood hazards pose a severe threat to both the well-being of the population and the overall development of the country (Perera *et al.*, 2015; Okura *et al.*, 2020). The Terai region plays a pivotal role in providing housing and food for the people, with approximately half of Nepal's population residing on 13% of its land area (Khanal *et al.*, 2007; CBS, 2017). A significant portion of this populace relies on subsistence farming as their primary livelihood (Dixit *et al.*, 2007; Perera *et al.*, 2015). These livelihoods are threatened by flood hazards, which can devastate crops and livestock within a matter of hours (Dilley *et al.*, 2005; MoHA, 2009; ADPC, 2010). Floods disrupt the vital links to markets and transportation networks, increasing food prices and threatening the food security not only of local communities but the entire nation (MoHA, 2009; Okura *et al.*, 2020; Red Cross and Red Crescent, 2021). Moreover, flood hazards potentially alter pollution levels, reduce freshwater availability and can lead to epidemics in the aftermath of the event (MoHA, 2009; Hannah *et al.*, 2020; Kosow *et al.*, 2022).

The agriculture sector, which contributes approximately a quarter of the Gross Domestic Product (GDP), is of paramount importance to the Nepalese economy (MoF, 2021). When floods destroy crops and kill livestock, they not only reduce Nepal's GDP but also increase the costs associated with food imports, necessary to maintain the nation's food supply (MoHA, 2009). These economic repercussions not only hinder the ability of individuals to enhance their flood resilience but also impede the nation's capacity to mitigate flood risks on a broader scale.

The riverine floods are sourced in the Himalayan catchments and are triggered by intense monsoon rainfalls over large parts of one or multiple catchments (DHM, 2018). Rainfall rates between 100 – 500 mm within 24 hours were reported during the flood events in 2013, 2014, and 2017 (DWIDP, 2014; Shrestha *et al.*, 2015a; Bhandari *et al.*, 2018; MoHA, 2018). The environmental conditions of the Himalayas increase the flood

hazard; the shallow soils and steep slopes facilitate a rapid conversion of rainfall into runoff, which quickly drains in the dense river network creating sharp flood hydrographs with a short duration, high magnitude, and a short response time between rainfall and flood even in large river systems (Bhandari *et al.*, 2018). The floods are further intensified by land-use change, predominately deforestation, abandonment and degradation of agricultural terraces, and urbanisation which decreases the infiltration and interception of rainfall (Gardner and Gerrard, 2003; Arnáez *et al.*, 2015; Pearson, 2020).

One key driver of changes in the flood runoff is the anthropogenic climate change. These changes to the climate have the potential to alter the hydrology of Central Himalayan river systems which may intensify the flood hazards. The annual and monsoon precipitation is projected to increase alongside an increase in intensity and frequency of heavy precipitation leading to an increase in the magnitude and frequency of flood hazards (Wijngaard *et al.*, 2017; Krishnan *et al.*, 2019; Douville *et al.*, 2021; Hirabayashi *et al.*, 2021; Seneviratne *et al.*, 2021; Talchabhadel, 2021). This response is superimposed by temperature-induced changes which are projected to alter the evapotranspiration rates, the snowmelt and glacier melt and the melting seasonality but this response is complex and varies in time and space (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014; Douville *et al.*, 2021).

Our current understanding of the potential climate change impacts on flood hazards is based on global and regional modelling which does not capture the significant small-scale variations in the Himalayan climate. A range of studies have assessed the impact of climate change on the water availability on the subbasin (hundreds of square kilometres) to the regional scale (the Hindukush-Himalayas) but do not quantify the impact on the flood flows (Immerzeel *et al.*, 2013; Bharati *et al.*, 2014; Lutz *et al.*, 2014; Nepal *et al.*, 2014; Nepal, 2016; Shea and Immerzeel, 2016; Bajracharya *et al.*, 2018; Bhattarai *et al.*, 2018; Dahal *et al.*, 2020; Chandel and Ghosh, 2021; Khanal *et al.*, 2021). Several studies have quantified the changes in the flood magnitudes and frequencies for the projected climates on the regional to global scale (e.g. Hirabayashi *et al.*, 2013, 2021; Dankers *et al.*, 2014; Huang *et al.*, 2017; Krysanova *et al.*, 2017; Pechlivanidis *et al.*, 2017; Wijngaard *et al.*, 2017). The global studies use the projections of global climate models which are too coarse to capture the small-scale variation in the climate arising from the complex topography of the Himalayas (Krishnan *et al.*, 2019; Scott *et al.*, 2019). The regional studies use downscaled climate projections with a better representation of the spatial variation in the climate, but these studies are based on older generations of climate models. There is, to the best knowledge of the author, no study that uses

projections from the latest generation of climate models to project future flood magnitudes and frequencies in Himalayan catchments.

The spatial pattern of flood hazards is further influenced by the morphology of its river channels and floodplains in the Terai plain. The Terai is an alluvial plain which comprises sediments exported from the Himalayas and large megafans have formed along the mountain outlets of the major river systems (Upreti, 2001; Dingle *et al.*, 2016). The Terai is characterised by its gentle elevation gradient between the river channels and the adjoining floodplains, and in some instances, the channel's elevation surpasses that of the floodplain which facilitates large-scale inundations (Sinha, 2008; Dixit, 2009; Sinha *et al.*, 2014). The low cohesion of the Terai sediments, the high sediment delivery to the plain, and the high transport capacity of the rivers facilitate a dynamic topography shaped by bank erosion, channel aggradation and degradation and channel migration which alters flood hazards (Milliman and Syvitski, 1992; Upreti, 2001; Dixit, 2009; Dhakal, 2013; Sinha *et al.*, 2014; Dingle *et al.*, 2016, 2020a; MoHA, 2018; Scott *et al.*, 2019).

The morphodynamic evolution in the Ganga plain varies regionally and is controlled by the basin subsidence on the millennial time scale (Dingle *et al.*, 2016). High subsidence rates in the Eastern Ganga plain (e.g. the Koshi River) facilitate the channel aggradation which decreases the channel capacity and increases the frequency of channel avulsions (Figure 1.1). The subsidence decreases in the Western direction, and the rivers shift to a degrading state (channel degradation) which increases the channel capacity and decreases the avulsion frequency (e.g. the Yamuna River) (Dingle *et al.*, 2016).

The 2008 Koshi flood illustrates the necessity to account for the morphodynamic evolution in flood management. This flood event was caused by an embankment breach which affected more than 3.5 million people in Nepal and India (Sinha, 2008; Dixit, 2009). The construction of embankments in the 1950s constrained the deposition of sediments to the channel. Consequently, the topographic gradient between the channel and floodplain increased which caused the avulsion in 2008. This avulsion shifted the river course by 120 km (Sinha *et al.*, 2014).

While the 2008 Koshi flood event was a turning point that increased the awareness that geomorphological processes alter the flood risk, the impact of the morphological evolution on flood hazards remains not well understood (Sinha, 2008; Dixit, 2009). So far, only a few studies have investigated the controls of the morphological changes in the context of flood hazards. Sinha *et al.* (2014) investigated the geomorphological mechanisms that drove the 2008 avulsion of the Koshi River. Dingle *et al.* (2016) found

that the morphological evolution of rivers in the Terai is controlled by the subsidence of the Ganga Plain. Dingle *et al.* (2020a) studied the effect of topographic changes on flood hazards in the Karnali River by comparing flood inundation maps predicted for a set of artificially altered Digital Elevation Models (DEM), and Thapa *et al.* (2023) investigated how sediment fluxes during flood events alter the inundation patterns in a small sub-catchment of the Narayani River.

We currently lack the tools to predict the potential impact of the morphological evolution on future flood hazards in Himalayan river systems. Previous studies have established that morphological changes alter flood hazards and improved our understanding of the controls of the morphological evolution, but no study has investigated how the evolution may alter the future flood hazards on a decadal time scale. Furthermore, climatic changes potentially alter the runoff rates and sediment delivery which are the primary controls of the morphodynamic evolution in these river systems (Leopold and Wolman, 1957; Ferguson, 1987; Kleinhans, 2010; Kleinhans *et al.*, 2013). It remains unclear how this alteration may affect flood hazards because we currently lack the tools to investigate the relationship between climate, catchment hydrology, geomorphology and flood inundation.

A tool to predict the potential evolution of flood hazards in Himalayan river systems needs to consider climatic and morphological changes and, thus, needs to link processes in the upstream catchment and the downstream floodplain. Such a tool would improve our understanding of the spatial and temporal variability of flood hazards and the drivers of change. It would also expand our understanding of the climate change impact on flood hazards in mountainous environments because it links climatic and morphological changes which are currently investigated in isolation. Furthermore, it would enable the development of flood risk management plans for potential future flood hazards and may improve the effectiveness of flood mitigation and adaptation strategies which are commonly implemented on decadal time scales. This tool could be transferred or adjusted to other sediment-rich mountainous river systems and may be impactful beyond the Himalayas.

1.2 Research aim and objectives

This research aims to predict the evolution of flood hazards until the end of the 21st century for the Karnali River, in the Central Himalayan region of Nepal. A geospatial modelling framework is developed that combines environmental and statistical models applied on different spatial scales.

To meet this aim the research has the following objectives:

- O1: To establish a hydrological model that replicates the hydrological system of the mountainous catchment and reproduces the observed discharge at the mountain outlet.
- O2: To predict the flood discharge at the mountain outlet for the projected climates from an ensemble of climate models until the end of the century using the hydrological model established in O1.
- O3: To quantify the flood frequencies and magnitudes at the mountain outlet for the flood discharge projected in O2.
- O4: To predict the morphological evolution of the Karnali fan until the end of the century for the projected flood discharge (O2) from morphodynamic modelling.
- O5: To map the spatial flood hazard characteristics (inundation extent and depth) for the projected flood magnitudes (O3) and topographies (O4).

This case study contributes to an improved understanding of future flood hazards in Central Himalayan river systems and the sensitivity of these hazards to climatic and morphological changes. Furthermore, it improves our knowledge of the feedback between flood hazards and fan evolution in the Terai. The framework is designed as a modelling cascade where the output of one model feeds into another model. The propagation of uncertainty through the modelling cascade is inherent in the framework design and this study contributes to a better understanding of the uncertainty sources which can be used for targeted investments to improve the flood hazard projections in future studies.

The developed framework is applicable beyond the Karnali River and provides researchers, planners and decision-makers with a set of tools to design flood risk management strategies for projected hazards. However, while it is designed as a comprehensive framework, it does not incorporate all drivers of flood hazard changes (e.g. land use change, river engineering, urbanisation) and might need adjustments in different geographic settings.

1.3 Research approach: An overview of the geospatial modelling framework

The framework to predict the evolution of flood hazards on decadal time scales is developed and evaluated for the Karnali River. The catchment (45,600 km^2) drains the the southern flanks of the Central Himalayas and is characterised by heterogeneous climatic and hydrological conditions making it susceptible to both, temperature and precipitation-related changes (Lutz *et al.*, 2014). Downstream of the mountain outlet, the Karnali River traverses an alluvial fan and bifurcates into two branches. The distribution of water between these branches is controlled by the topography at the bifurcation node and affects the downstream flood hazards (Kleinhans *et al.*, 2013; Dingle *et al.*, 2020a). It is this sensitivity to climate change in the catchment and morphological changes in the fan that makes the Karnali River an ideal river system for the application of the modelling framework. This framework comprises four stages which build on each other and an overview of each of these stages is presented in Figure 1.2.

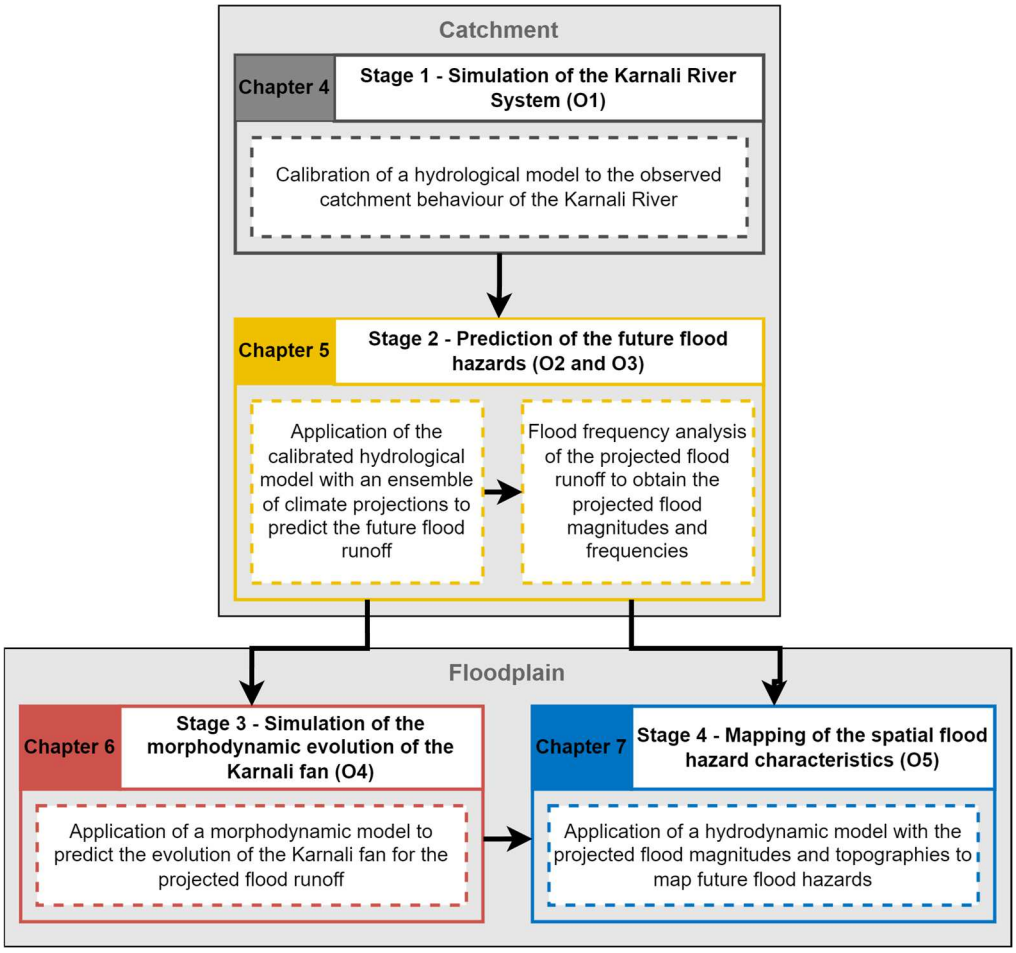


Figure 1.2: Conceptual overview of the research design to predict the potential flood hazards in the Central Himalayan Karnali River.

1.3.1 Stage 1: Predicting the catchment hydrology of the Karnali River

The flood discharge released into the fan is the sum of the hydrological processes in the upstream catchment. The prediction of future flood discharge, thus, requires an understanding of the catchment's hydrology to predict the impact of changing climatic conditions on the hydrological regime. Stage 1 establishes a hydrological model that represents the hydrological system of the Karnali River to meet the first objective. This initial stage provides the foundation of the modelling framework which the later stages build upon.

The complex hydrology of the Karnali River poses a challenge for hydrological modelling arising from the spatial and temporal heterogeneity of the climate. A model needs to simulate the hydrological processes of subtropical to alpine climates including the rainfall-runoff conversion, baseflow, and snow and glacier melt processes. In this stage, the hydrological model is calibrated and validated for the catchment until the mountain outlet / fan inlet. For the calibration, the model is established for the past climatic conditions and parameterised to optimise the agreement of the simulated and observed catchment behaviour. An ensemble of the best-performing model parameter sets is selected for the simulations in the following Stage 2 of the research.

1.3.2 Stage 2: Predicting potential flood hazards on the catchment scale

Stage 2 combines hydrological and statistical modelling with the projections of climate models to predict the flood runoff that is released into the floodplain for the projected climates (O2). For this analysis, the established hydrological model is applied with the projections of the latest generation of climate models for different climatic pathways (emission scenarios) which provides the flood runoff for the current (baseline scenario) and projected climatic conditions. An ensemble modelling approach is applied to quantify the uncertainty in the hydrological simulations by applying a hydrological ensemble with the climate projections from an ensemble of climate models.

A flood frequency analysis (FFA) is conducted to determine the projected flood magnitudes and frequencies and meet O3. The FFA offers a method to standardise flood hazards which enables a meaningful comparison of different flood records. In this study, the FFA is used to quantify the changes of the past and projected flood hazards and to

compare the projections of different climate models and climate scenarios. The uncertainty of the flood frequency analysis is quantified using a stochastic modelling approach.

1.3.3 Stage 3: Predicting the morphodynamic evolution of the alluvial fan

The previous Stages 1 and 2 are applied on the catchment scale to predict the discharge of water that is released into the downstream fan. Stages 3 and 4 are applied on the local scale of the fan to predict the pathways of the floodwater.

Stage 3 predicts the changes in the pathways of the floodwater for the hydrological conditions projected in Stage 2. A morphodynamic model is applied with the flood discharge of the projected climates (O2) to simulate the topographic changes of the channel and the floodplain to meet O4. The effect of these topographic changes on the pathways of flood water is twofold:

i) processes of lateral erosion and deposition potentially alter the direction of flood water (channel migration) and;
ii) the erosion from and deposition on the riverbed (channel aggradation and degradation) may alter the channel capacity and change the allocation of the water between the channel and floodplain or change the downstream distribution to downstream of flood water (i.e. at bifurcations).

1.3.4 Stage 4: Predicting potential flood hazards on the floodplain scale

Stage 4 is the final stage in the modelling framework that combines the projected flood magnitudes from Stage 2 (O3) with the projected topographies from Stage 3 (O4) in a hydrodynamic model to map the spatial flood hazard characteristics (i.e. inundation extent and inundation depths) and meet O5. This mapping identifies areas at risk and how these areas are changing for the projected conditions. It, thus, identifies the areas that are sensitive to changing environmental conditions which can be used for the targeted development of flood risk management strategies that account for the future hazards.

This stage synthesises the previous stages and contributes to a better understanding of the temporal and spatial variability of flood hazards, the driving mechanisms behind this variability, and the uncertainty sources and propagation within the modelling framework which may identify fields for targeted investments to reduce the uncertainty and improve the predictions in future studies.

1.4 Thesis structure

This thesis follows the structure of the geospatial modelling framework outlined in the previous section and contains eight chapters that combine to predict the potential flood hazards in the Central Himalayan Karnali River.

Chapter 2 introduces the Karnali River system for which the modelling framework is developed and evaluated. This description includes the mountainous catchment in Nepal and China which is the study area of Stages 1 and Stage 2, and the Karnali fan which is the study area of Stage 3 and Stage 4.

Chapter 3 reviews the literature of the current research state of the methods applied within the modelling framework including hydrological, morphodynamic, hydrodynamic, and statistical modelling, and frameworks for the uncertainty estimation in environmental modelling.

Chapter 4 presents the hydrological modelling of the Karnali River catchment to reproduce the hydrological system of this river to address the first objective. In this chapter, an ensemble of behavioural parameter sets is identified which is used in Stage 5 to quantify the climate change impact on flood flows.

Chapter 5 presents the climate change impact assessment on the flood discharge (O2), the flood magnitudes and the flood frequencies (O3). The hydrological modelling of this chapter builds upon the results of the previous chapter and concludes the catchments modelling part of the modelling framework.

Chapter 6 presents the prediction of the morphological evolution of the Karnali fan to determine the pathways of the floodwater for the projected climates. For this, a morphodynamic model is applied with the projected flood discharge (O2) to predict the evolution of the fan topography and meet O4.

Chapter 7 maps the spatial flood hazard characteristics in the Karnali fan to meet O5. For this, a hydrodynamic model is applied with the projected flood magnitudes of selected flood frequencies (1-in-10-years, 1-in-50-years, 1-in-100-years) of O3 and the projected topographies of O4 to map the spatial variation of the flood hazards for the projected climatic and topographic conditions.

Chapter 8 summarises the modelling framework to predict the potential flood hazards in the Karnali River system. It presents the main conclusions and provides recommendations to reduce the uncertainty and improve the predictions in further studies.

The Karnali catchment and fan

2.1 Introduction

The Karnali River is one of the three main river systems draining the Central Himalayas in Nepal. The study area is divided into the upstream catchment which generates the flood water and the downstream fan in the Terai (Figure 2.1). This fan is home to approx. 250,000 people and experienced frequent floodings which killed hundreds of people, affected tens of thousands of residents and damaged private property, infrastructure and agricultural goods over the last decade (MacClune *et al.*, 2015; Shrestha *et al.*, 2015a; Bhandari *et al.*, 2018; Practical Action, 2018).

The projected climate change in the region has the potential to impact the hydrology, and thus the flood hazards, in two ways:

- The warmer temperatures in the alpine parts alter the melting rates and seasonality of snow and glaciers which could, both, increase or decrease the flood discharge (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014; Dahal *et al.*, 2020; Chandel and Ghosh, 2021; Douville *et al.*, 2021; Khanal *et al.*, 2021);
- ii) the projected increase in the frequency and intensity of heavy rainfall events increases the flood frequencies and magnitudes (Wijngaard *et al.*, 2017; Chhetri *et al.*, 2021; Hirabayashi *et al.*, 2021; Talchabhadel, 2021).

The riverine floods occur in the Himalayan foreland where the Karnali River flows across an alluvial fan which connects the sediment sources in the Himalayas with the sediment sinks in the Ganga plain (Harvey, 2018; Quick *et al.*, 2019; Dingle *et al.*, 2020a). The topographic evolution of this fan defines the path of the flood water and, thus, alters the flood hazards (Dingle *et al.*, 2020a). This evolution is, however, shaped by flood events during which sediment pulses are delivered and the topography is reworked, and hence this evolution is affected by climate change induced changes in the flood discharge (Kleinhans *et al.*, 2013).

The sensitivity of the catchment hydrology to climate change, the high frequency and magnitude of flood hazards, and the dynamic fan topography make the Karnali River a suitable case study to develop and evaluate a modelling framework to predict the long-term evolution of flood hazards.

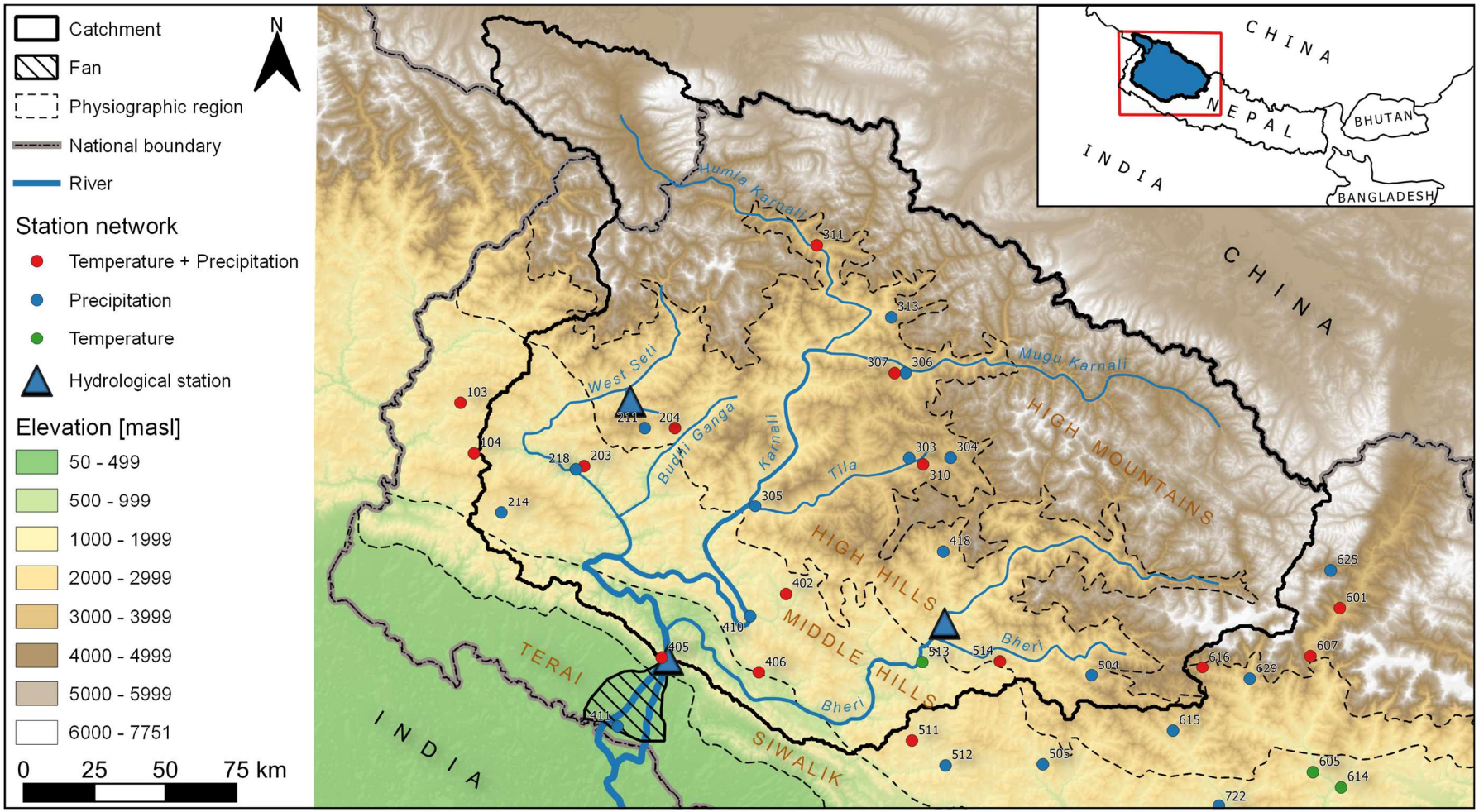


Figure 2.1: Overview of the study area comprising the Karnali catchment and the Karnali fan. The topography is derived from the SRTM 90m V4.1 (Jarvis et al.*, 2008). The station network shows the location of the hydrological and meteorological stations for which data is used in this study. The numbers on the station network relate to the gauge's identification. These stations are maintained by the Department of Hydrology and Meteorology (DHM), Nepal.*

The next section describes the location of the study area in its regional context. The environmental characteristics of the fan and the catchment are then described in Sections 2.3 and 2.4, respectively. The climatic conditions, the hydrological system, and the sediment cascade are introduced in Sections 2.5 – 2.7.

2.2 Location of the study area

The Karnali River originate in the Central Himalayas in the border region of Nepal and China and traverses the Ganga plain in Nepal and India (where it is called the Ghaghara River) before it flows into the Ganga River at Chapra in India. The catchment is located in Nepal and China and comprises the Himalayan parts of the Karnali catchment until the mountain gauge and the fan attaches south of this gauge (Figure 2.1).

The catchment covers 45,600 km^2 in the Mid-Western and Far-Western provinces of Nepal. Parts of the Humla Karnali subbasin in the North-West (3,100 km^2) are located in China. The Karnali River drains the southern flanks of the Himalayas which are demarked by the Indo-Gangetic plain in the south and the Tibetan Plateau in the north. The Karnali forms at the confluence of the Humla Karnali which drains the North-Western subbasin and the Mugu Karnali which drains the North-Eastern subbasin. The West Seti drains the Western parts of the catchment and confluences with the Karnali 40 km upstream of the outlet. The Eastern parts of the catchment are drained by the Bheri River which confluences with the Karnali 10 km upstream of the mountain gauge.

The fan investigated in this study is a 545 km^2 large subsection of the whole Karnali fan (Figure 2.2). This subsection covers the gravel reaches of the river (30 – 35 km distance to the mountain outlet) and the flood-prone fan sections in the centre and the west. The areas in the east of the river are superelevated and excluded from the study area. The Karnali River bifurcates 5-6 km downstream of the mountain gauge to a Western branch and Eastern branch which confluence back together south of the India-Nepal border (see Figure 2.1).

2.3 Fan characteristics

The fan is located in the Terai in Nepal which is the northern extension of the Indo-Gangetic plain and is demarked by the Main Frontal Thrust in the North (Upreti, 1999). The Terai is made up of Himalayan Alluvium (average thickness of 1,500 m) (Upreti, 1999). The grain sizes of this alluvium decrease with distance to the mountains and range, on the surface, from cobbles and boulders at the fan apex to fine gravel at the southern border of the study area (Quick *et al.*, 2019; Dingle *et al.*, 2020b).

The fan topography is characterised by a comparatively low elevation gradient in the southern direction from ~160 masl at the mountain outlet to ~85 – 90 masl at the downstream fan border. The highest elevations are observed on a terrace in the East that stretches along the Eastern branch. This terrace is superelevated over the channel (10 – 20 m) and this terrace has remained stable for the last four millennia (Quick, 2021). The topography decreases in the western direction and the areas west of the Western branch are located below the riverbanks and are prone to flooding (Figures 2.2 and 2.3). The topography of the river is continuously reworked as a result of the high discharge and the low cohesion of the bed and bank material (Langhorst and Pavelsky, 2023). The exception is the terrace in the east for which no changes are observed in satellite imagery over the past 40 years which indicates that the terrace prevents the lateral channel migration in the eastern direction. The most noticeable change in recent years was the alteration of the bifurcation in the year 2009 which caused a change in the drainage distribution in favour of the Western branch (Dingle *et al.*, 2020a).

The Terai is intensively used for agriculture and 80 – 90% of the fan is agricultural land dominated by irrigated agriculture (Table. 2.1). The natural vegetation of the Terai are forests dominated by Sal trees (Shorea robusta) but most forests have been cleared and only prevail in the Bardiya National Park located along the Eastern border of the fan, and in patches along the Western branch of the Karnali (Zurick, 1988; DFRS, 2014). The dominant soil type is gleysol which indicates the abundance of water in this area.

The fan is densely populated, particularly in the flood-prone areas in the Western fan and the centre between the branches. A population of around 250,000 (calculated from the population density shown in Figure 2.2) lives primarily in villages and Tikapur city. Half of the population lives in the municipalities Janaki and Tikapur which comprise most of the low-lying area West of the Western branch. Around 75,000 people live on the island between both branches. The Eastern parts of the fan are less populated because it is mostly occupied by the Bardiya National Park without permanent residencies.

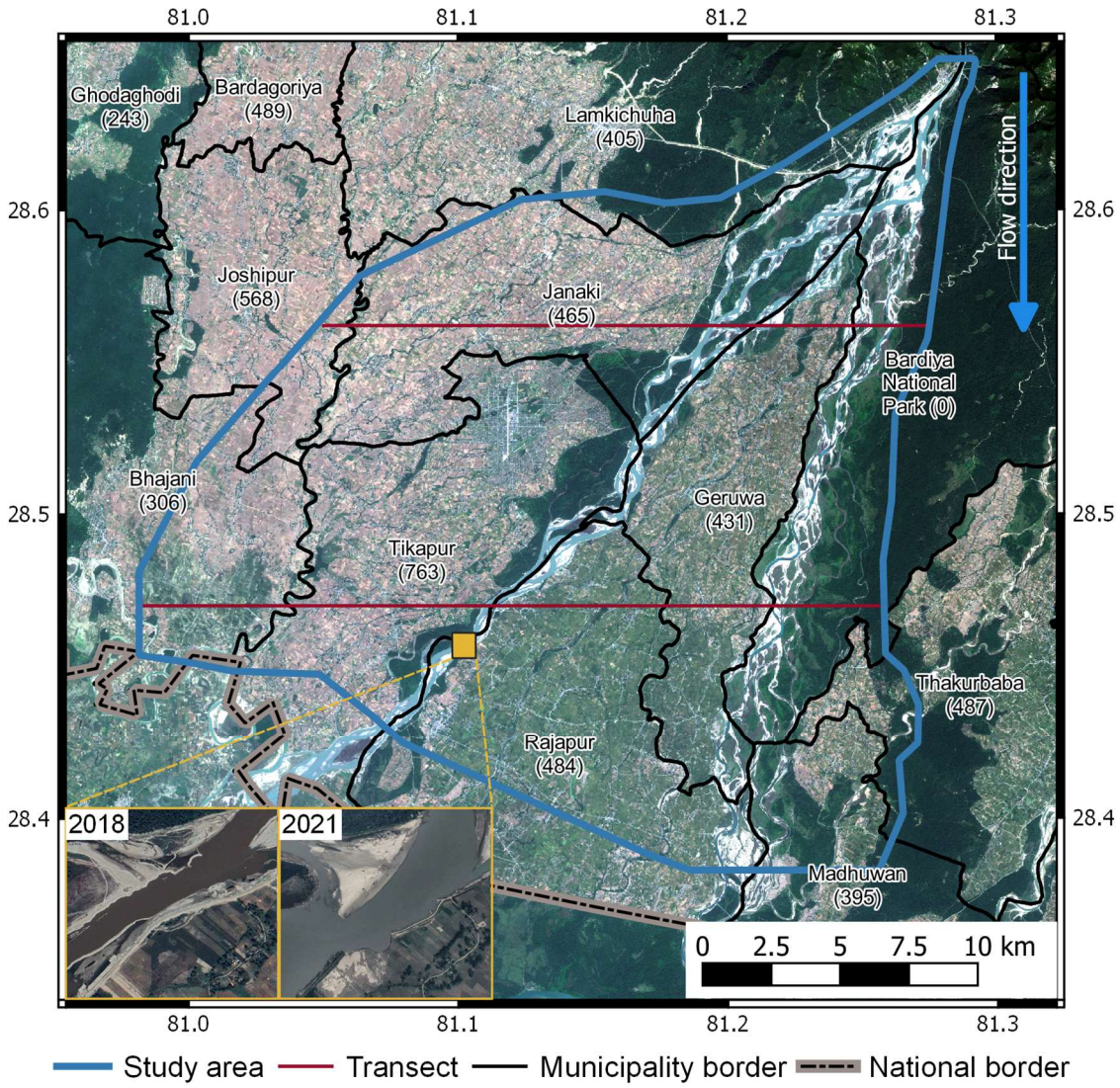


Figure 2.2: Map of the Karnali fan. The basemap is a composite of two Sentinel-2 images taken in October 2020. The insets in the yellow boxes are satellite images obtained from Google Earth (04/2018 and 11/2021). The transects show the location of the elevation profiles in Figure 2.3. The population density (people per km^2 in 2021) is shown in the brackets below the municipality borders (CBS, 2022).

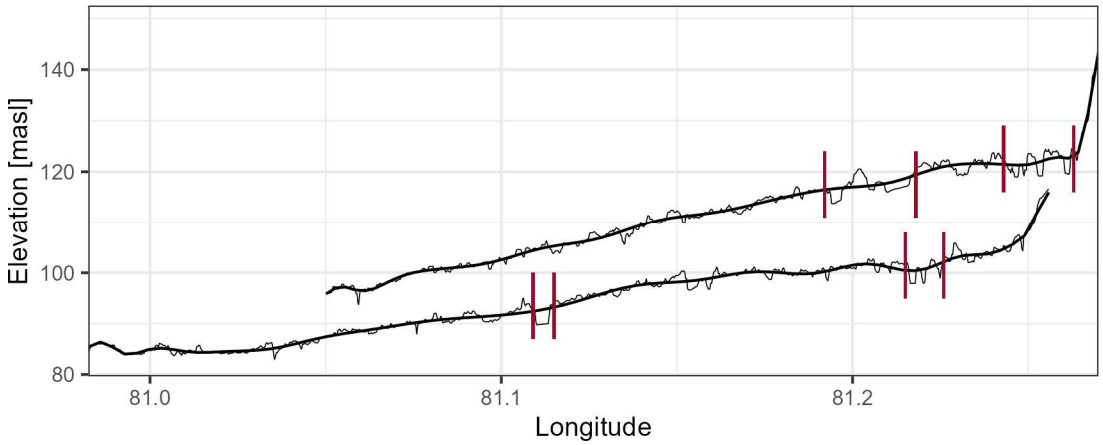


Figure 2.3: Elevation profiles of the Karnali fan at the transects shown in Figure 2.2 The topography is derived from a WorldView-2 DEM which is introduced in Chapter 6. The red lines indicate the location of the river banks of the main branches of the Karnali River.

Table 2.1: Summary statistics of the environmental conditions. The data sources are Jarvis et al. *(2008) for the elevation and slope, RGI Consortium (2017) for the glaciated areas, Scherler et al. (2018) for the glacier type, Arino* et al. *(2012) for the land cover, and Dijkshoorn and Huting (2009) for the soil type.*

	Fan	**Catchment**				
	Total	Total	Siwalik	Middle hills	High hills	High mountains
Area [km^2]	545	45,619	1,760	8,185	14,793	20,881
Fraction of catchment [%]	-	100	4	18	32	46
Mean elevation [masl]	162	3,317	691	1,353	2,699	4,745
Mean slope [degree]	1	26	16	24	29	26
Glaciated area [%]	-	3	-	-	-	6
- Debris-free [%]	-	71	-	-	-	71
- Debris-covered [%]	-	29	-	-	-	29
Dominant land cover (2009)	Irrigated croplands	Herbaceous vegetation	Shrubland	Rainfed croplands	Closed evergreen forest	Herbaceous vegetation
- Dominant land cover area [%]	73	21	24	30	36	45
Dominant soil type	Eutric Gleysols	Gelic Leptosols	Dystric Regosols	Eutric Cambisols	Eutric Regosols	Gelic Leptosols
- Dominant soil area [%]	70	34	85	67	55	72

Embankments have been built in recent years (mostly after the year 2011) along both banks of the Western branch and along parts of the Western bank of the Eastern branch to protect the population and agricultural fields from flooding. However, a 300 m long section of an embankment was destroyed between the years 2020 – 2021 and ~ 100 – 200 m of the riverbanks eroded in the process (see the insets in Figure 2.2). This erosion emphasises the erosive force of the Karnali and the limited flood protection provided by these embankments.

2.4 Catchment characteristics

The catchment is divided into four physiographic regions with distinct geologic, topographic, ecologic, and climatic (section 2.5) characteristics. These zones stretch in the E-W direction and separate the study area in the S-N direction from the foothills in the South to the high mountain ranges in the Northern parts of the catchment. The environmental characteristics of these four zones (Figure 2.4) are presented according to the topographic increase in the S-N direction in the following sections and are summarised in Table 2.1.

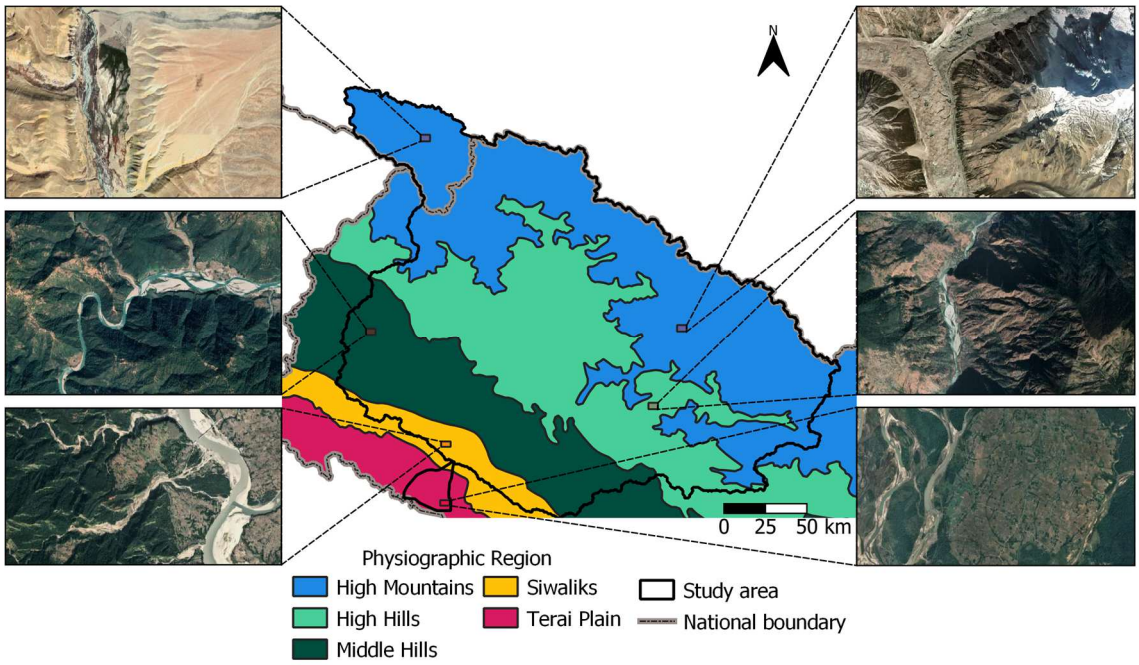


Figure 2.4: The physiographic regions of the study area (ICIMOD, 2011). The inset maps show satellite images (Google satellite accessed by QMS in QGIS) of selected areas of each physiographic division.

2.4.1 The Siwaliks

The Siwaliks are the first mountain range of the Himalayas (200 – 1,300 masl) and stretch in a narrow band (10 – 20 km width) in the E-W direction. The Siwaliks are made up of uplifted foreland deposits and comprise three units, the lower, middle, and upper Siwaliks. The lower and middle Siwaliks are made up of friable sandstone, mudstone and shale while the upper Siwaliks are made up of resistant pebble-to-bolder sized conglomerate (Zurick, 1988; Upreti, 1999; Dingle *et al.,* 2017). The Siwaliks are the geologically most active mountain range and are characterized by a steep and rugged topography with a high density of gullies and landslides, and efficient sediment transport in the rivers (Hurtrez *et al.*, 1999; Upreti, 2001; Ghimire *et al.*, 2013). However, while the sediment production is high in all three units, almost all gravel supplied to the Karnali fan originates from the upper Siwaliks because the conglomerate is resistant and does not abrade before reaching the fan (Dingle *et al.,* 2017; Quick *et al.,* 2019).

The soils are poorly developed and shallow Regosols dominate as a result of the high erosion rates. The potential vegetation in the Siwaliks are forests dominated by Sal trees but the dominant land cover type is shrublands covering one-quarter of the Siwaliks in the catchment (Zurick, 1988; Arino *et al.*, 2012). However, the different forest types (mixed, deciduous, coniferous) cover together around one-third of the mountain range (Arino *et al.*, 2012).

2.4.2 The Middle Hills

The Middle Hills are delimited to the Siwaliks in the south by the Main Boundary Thrust (Upreti, 2001). This mountain range covers 18% of the catchment in a 30 – 65 km wide belt and is characterised by the steep and rugged topography. The Middle Hills are located at elevations between 1,000 – 3,000 masl and the mountains are intersected by steep valleys which reach for the Karnali and its main tributaries as low as 300 masl. The geology is comprised of thin, interlayered sequences of low-grade metamorphic rocks (including limestone, slate, phyllite, quartzite and dolomite) (Zurick, 1988; Upreti, 1999). This area is highly susceptible to landslides due to the steep slopes and the weak geologic conditions(Upreti, 1999, 2001).

The dominant soils are Cambisols which occupy 95% of this region (Dijkshoorn and Huting, 2009). These soils are deeper, more developed soils compared to Regosols and are more suitable for agriculture. Therefore, this region is intensively used for terraced agriculture (~50% of the area) and forests have declined (~30% of the area) (Arino *et al.*, 2012). This conversion from forests to agriculture alters the hydrological processes but this impact is complex. Generally, the conversion of forests to agriculture decreases the retention of monsoon rain and increases the runoff (Gardner and Gerrard, 2003; Pearson, 2020). However, terraced agriculture increases soil infiltration and decreases the hydrological connectivity and runoff generation but this depends on the conditions of the terraces because abandoned and deteriorated terraces increase the runoff generation and soil erosion (Arnáez *et al.*, 2015; Pearson, 2020).

2.4.3 The High Hills

One-third of the catchment is located within the High Hills. This region covers a 10 – 30 km wide band parallel to the Middle Hills but it extends North into the High Mountains along the valleys of the Karnali and its tributaries. This region is characterized by steeper slopes and a more rugged terrain than the lower southern mountain ranges. The elevations range between 2,000 – 5,000 masl. but some peaks are higher and the deep valleys of the Karnali and the tributaries are located below 1,000 masl. The High Hills are a rapidly uplifting, young and fragile mountain range comprised of high-grade metamorphic rocks including gneiss, migmatites, schists, quartzites, marbles and granites (Upreti, 1999, 2001). The dominant soil types are regosols along the valleys and cambisols along the ridges (Dijkshoorn and Huting, 2009). Thick layers of bouldery colluvisols occupy the slopes (Upreti, 2001).

The valleys are primarily used for agriculture (34%) and the slopes and ridges are covered by evergreen forests (36%) (Arino *et al.*, 2012). The climate is diverse (temperate to alpine) due to the large elevation range so the vegetation is variable with oaks and rhododendrons dominating the lower areas and hemlock and fir occupying the higher slopes (Zurick, 1988).

2.4.4 The High Mountains

The High Mountains are the highest mountain range of the Himalayas separating the southern slopes from the Tibetan Plateau. The peaks exceed 5,000 masl. up to 7,751 masl at the Dhaulaghiri II located at the Western border of the catchment. This region covers almost half of the catchment (46%) and is characterised by a very rugged topography with very steep slopes. The geology comprises crystalline rocks, including gneiss, quartzite and migmatites which are stronger than those of the other mountain ranges (Upreti, 1999, 2001). However, shallow Leptosols have developed (70%) despite the intense rock weathering due to the steep slopes (Table 2.1).

The land cover is stratified with the elevation and the highest areas are covered with snow and ice (23% of this zone). All glaciers are located within this region (6% of this zone) of which 30% are covered by debris (RGI Consortium, 2017; Scherler *et al.*, 2018). Below this nival zone, herbaceous vegetation (mosses and lichens) dominate (44% of this zone) at elevations between 4,500 – 5,500 masl (Ives and Messerli, 1989). The lower-lying areas along the valleys are occupied by Rhododendron shrubs and birches (Ives and Messerli, 1989). Furthermore, barelands occupy the North-Western part of the catchment and along the northern border of the catchment (11% of this zone) (Arino *et al.*, 2012).

2.5 The climate in the study area

The climate in the study area is diverse ranging from a subtropical climate in the Terai in the south to a nival climate in the High Mountains in the North (Ives and Messerli, 1989). The precipitation is dominated by the monsoon circulation. The presented temperature and precipitation statistics are derived from observations of meteorological stations maintained by the Department of Hydrology and Meteorology Nepal (DHM) and cover the period 1985 – 2014. The observed climate is presented in Figure 2.5 and the locations of the stations are shown in Figure 2.1.

2.5.1 Temperature

The temperatures are governed by the topographic variation in the study area. The Mean Annual Air Temperature (MAAT) is 24.4 ºC at the mountain outlet (Station 405, 225 masl). The coldest month is January with an average temperature of 15.2 ºC and the warmest months are May and June with 30.3 ºC and 30.1 ºC, respectively. The climate cools with increasing elevations to warm-temperate of the ridges of the Siwalik range and to warm temperate and cold-temperate in the Middle Hills (Zurick, 1988). The MAAT is 19.6 ºC at Station 203 (1,360 masl) located in the western parts of the Middle Hills and 19.0 ºC at Station 514 (2,1000 masl) located at the border of the Middle Hills and the High Hills in the eastern part of the catchment. The temperature of the coldest month January is 11.5 ºC for both stations and the mean temperature of the warmest month June is 25.3 ºC (Station 207) and 24.0 ºC (Station 514). The climate transitions from temperate to alpine in the High Hills, and the highest station (3,050 masl) records an MAAT of 9.7 ºC. The mean monthly temperatures range from 2.2 ºC in January to 16.3 ºC in July. The climate in the High Mountains is alpine and the MAAT is below 0 ºC along the slopes and ridges of the highest mountains but no station is located within this region (Ahmad *et al.*, 2021).

Temperatures vary locally in the study area. This is indicated by the temperatures of stations 402 (1,400 masl) and 514 (2,100 masl) which are located at similar latitudes whereas station 402 is located 80 km East of station 514. The MAAT of station 402 is 18.7 ºC and, thus, 0.3 ºC colder than the recorded MAAT at station 514. This is notable because station 514 is located at a 700 m higher elevation. No distinct E-W gradient in the temperatures is observed in Nepal (Ahmad *et al.*, 2021). This difference is a local anomaly which might be driven by local variations in the dry and wet adiabatic cooling and warming patterns (Collier and Immerzeel, 2015).

The lack of meteorological records is one key challenge for hydrological assessments in the Central Himalayas (Nepal *et al.*, 2014). The station network concentrates on the accessible areas in the plains and valleys leaving the slopes and ridges vastly unobserved. No meteorological station is located within the High Mountains which occupy almost half of the catchment. One station (311) is located in the valley of the Humla Karnali close to the North-Western catchment boundary but the records have only a good coverage for the period 1997 – 2005. The annual lapse rate is -0.0054 ºC/m but this lapse rate only covers elevations until 3,050 masl (Dhami *et al.*, 2018). It is questionable if it can be extrapolated because of the precipitation patterns and, thus, the wet-adiabatic cooling change at higher elevations (Winiger *et al.*, 2005; Lutz and Immerzeel, 2016).

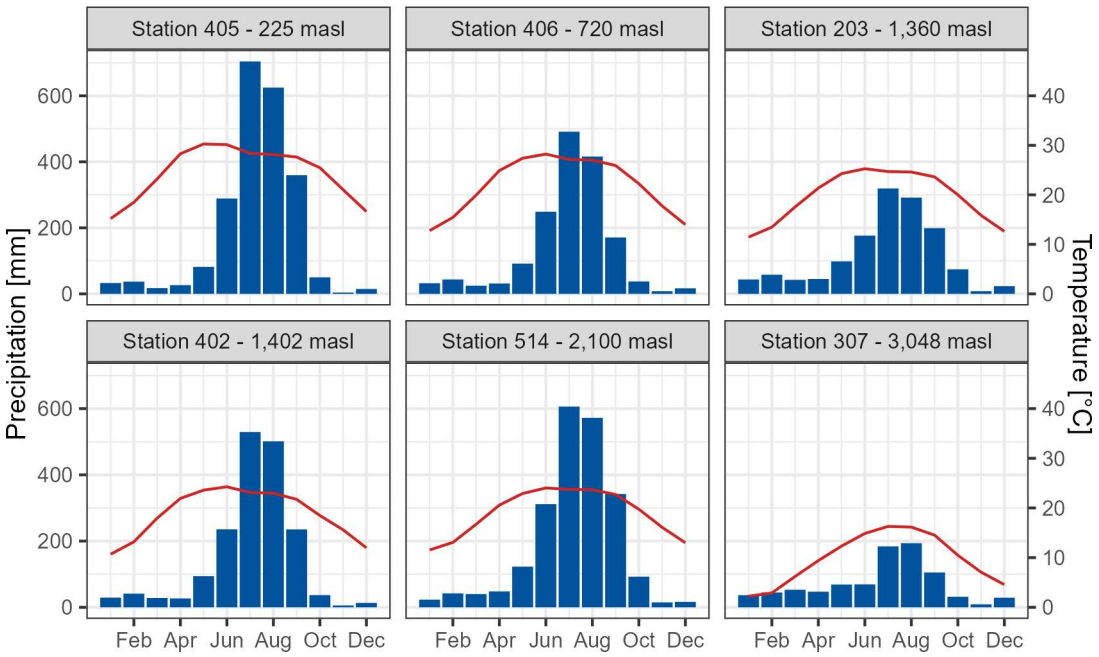


Figure 2.5: Climographs at six stations in the catchment for the period 1985 - 2014. The data is obtained from the Department of Hydrology and Meteorology Nepal. The locations of the stations are shown in Figure 2.1.

2.5.2 Precipitation

The precipitation is dominated by the monsoon circulation which brings 80% of the annual precipitation during the summer monsoon (Jun – Sep) (Bookhagen and Burbank, 2010; Krishnan *et al.*, 2019). The remaining months are comparatively dry but the Westerly circulation provides some precipitation during the winter (Dec – Feb) (Hannah *et al.*, 2005; Bookhagen and Burbank, 2010).

The large-scale precipitation patterns follow the track of the summer monsoon; Central Nepal receives the highest precipitation due to the prolonged monsoon season while this season is shorter and precipitation rates are lower in Western Nepal (Hannah *et al.*, 2005). Furthermore, the precipitation patterns are characterized by a gradient in the south-north direction with high precipitation along the southern slopes of the Central Himalayas and low precipitation in the Tibetan Plateau which is located in the North in the rain shadow of the High Mountains. The Siwalik range and the High Hills receive the highest precipitation due to their effect as an orographic barrier while the precipitation is lower in the Middle Hills where the vapour concentration is reduced as it is located in the shadow of the Siwalik range and not high enough to be an efficient orographic barrier (Bookhagen and Burbank, 2010).

This orographic effect is captured by the DHM station network (Figure 2.4); the annual precipitation is highest at the mountain outlet (Station 405: 2,200 mm) and the transition between the Middle Hills and the High Hills (Station 514: 2,200 mm). It is lower at the transition between the Siwalik range and the Middle Hills (Station 406: 1,600 mm) and in the Middle Hills (Station 203: 1,300 mm). The lowest precipitation of 900 mm is observed at the most Northern station 307. The months with the highest precipitation are July and August for all stations.

The precipitation in the northern areas at high elevations (> 4,000 – 5,000 masl) is uncertain and the information is contradictory. Several studies suggest that these areas are comparatively dry and receive annual precipitation of 250 – 1,000 mm (Upreti, 2001; ICIMOD, 1996 in Gautam and Acharya, 2012; Ahmad *et al.*, 2021). These studies are based on meteorological observations or remote sensing products. However, other studies suggest that the high mountains must receive substantially higher precipitation to sustain the observed glacier mass balances (Winiger *et al.*, 2005; Immerzeel *et al.*, 2015; Lutz and Immerzeel, 2016).

The complex topography of the Himalayas causes small-scale variations in the orographic precipitation patterns which are not captured by the coarse station network and the coarse resolution of remote sensing products or climate models (Collier and Immerzeel, 2015). Furthermore, these stations are located within the drier valleys in the rain-shadows of mountain ridges but are used for the bias correction of satellite estimates (Bookhagen and Burbank, 2010; Huffman *et al.*, 2019). Hence, the precipitation data is uncertain on the local scale and in the high mountains (see previous paragraph).

The monsoon causes frequent heavy rainfall events which cause flash floods in the sub-catchments and if occurring over larger areas, riverine floods in the Ganga plain. During such events, rainfalls of up to 500 mm within 24 hours have been observed (MacClune *et al.*, 2015). The largest rainfall of 465 mm per day was observed at station 214 in the Middle Hills (Table 2.2). Rainfall rates of ≥ 200 mm per day were recorded for one or more stations on 56 days in the period 1985 – 2014. However, it is worth noting that daily precipitation does not equal the rainfall within 24 hours of maximum intensity and hence this is merely an indicator of the heavy rainfall events which might underestimate the frequency and the intensity of these events.

Table 2.2: Highest daily rainfall observed during the period 1985 - 2014. The records are obtained from the Department of Hydrology and Meteorology. The location of these stations is shown in Figure 2.1.

Date	**Station**	**Rainfall [mm/d]**
18/08/2009	214	465.4
19/09/2008	411	355
06/09/2009	214	340.2
26/08/2006	407	308.5
12/02/2002	504	306.1

2.6 The hydrological system

The hydrological regime of the Karnali River is dominated by the monsoon seasonality with high flows during the monsoon season (Jun – Sep) and low flows during the Non-monsoon season (Oct – May) (Figure 2.6). Almost three-quarters (73%) of the annual discharge drains during the monsoon season. The monsoon hydrograph is characterised by a pronounced peak in August which is attributed to the comparatively short monsoon season in Western Nepal (Hannah *et al.*, 2005). The discharge further recedes during the post-monsoon season (Oct – Nov) which contributes 12% to the annual budget. The remaining six months of the winter (Dec – Feb) and pre-monsoon season (Mar – May) contribute 15% to the annual budget.

The hydrology of the Karnali is complex due to the diverse environmental and climatic conditions with seasonal variation in the runoff contributions. The monsoon season runoff is primarily composed of rainfall-runoff and to a lesser extent glacier melt runoff (Lutz *et al.*, 2014; Nepal *et al.*, 2014). A fraction of the monsoon rain percolates to the groundwater and drains as delayed baseflow. This baseflow maintains the flow during the winter and without it, the Karnali River would not carry water at times (Lutz *et al.*, 2014). Snowmelt becomes an important source during the pre-monsoon season as the groundwater storage depletes (Lutz *et al.*, 2014; Nepal *et al.*, 2014; Dhami *et al.*, 2018).

The loss of water from evapotranspiration Is less significant for the hydrology of mountainous sub-catchments for which it is generally below 10% (Bookhagen and Burbank, 2010). This loss is moderately higher in Central Himalayan River systems due to the warm temperatures in the lower mountain ranges of the Siwaliks and Middle Hills (Nepal, 2012). It is estimated that 14% of the annual precipitation evapotranspirates in the Karnali catchment mostly during the monsoon season (Bookhagen and Burbank, 2010). The potential evapotranspiration rates are highest during the pre-monsoon season due to the lower cloud cover (Nepal, 2012). However, evaporation observations are sparse and inaccurate and do not depict the small-scale variability in the mountainous catchments (Hannah *et al.*, 2005).

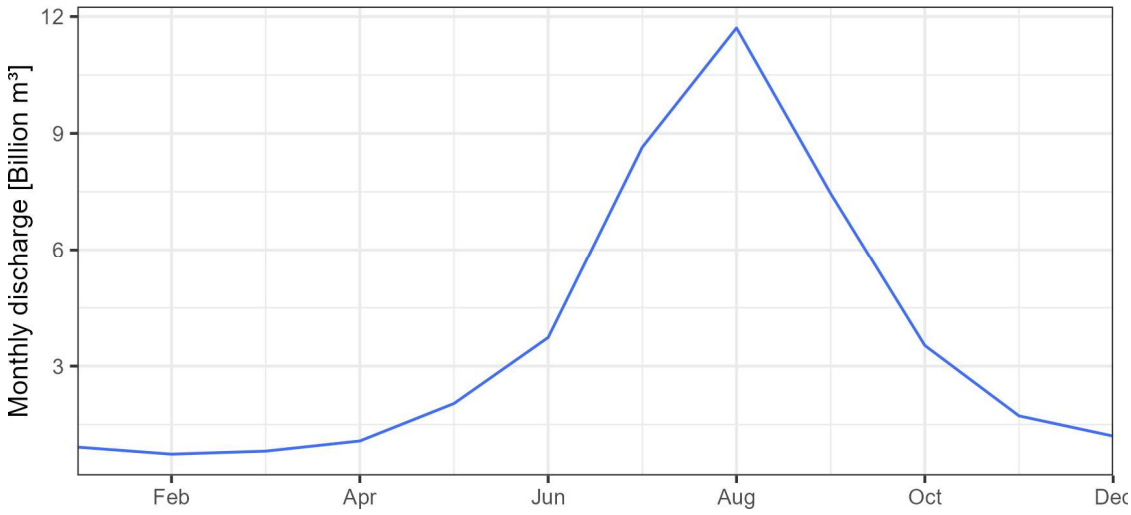


Figure 2.6: Mean monthly discharge (1985 – 2014) of the Karnali River at the mountain gauge. The gauging station is maintained by the Department of Hydrology and Meteorology, Nepal.

The intense monsoon rainfalls and the steep slopes in the catchment cause frequent floodings in the Terai region. The most recent flood events occurred in 2013 and 2014 which were with 17,400 m^3/s and 17,900 m^3/s (daily mean) the largest flood events recorded since the year 1970 (Figure 2.7). The third largest flood event occurred in 1983 (16,000 m^3/s) and stands out from the flood record because it is the only major flood which lasted more than a day (Figure 2.8). In 2009, a flood event (14,600 m^3/s) altered the bifurcation after which more water drains through the Western branch (Dingle *et al.*, 2020a). The highest observed daily discharge (17,900 m^3/s) at the mountain outlet (1985 – 2014) exceeds the lowest one (170 m^3/s) by a factor of 100 which emphasises the high range of flows.

The flood-triggering mechanisms are diverse and depend on the catchment characteristics. Glacier Lake Outburst Floods (GLOF) and Landslide Outburst Floods (LOF) cause the most catastrophic floods in the upstream sub-catchments (Cook *et al.*, 2018; Chen *et al.*, 2023). Localised heavy rainfall causes flash floods in the Siwaliks and Middle Hills (Khanal *et al.*, 2007; DHM, 2018). The riverine floods in the Terai Region are triggered by intense rainfalls over a large fraction of the catchment (MacClune *et al.*, 2015; Shrestha *et al.*, 2015a; Bhandari *et al.*, 2018; DHM, 2018). The steep slopes of Central Himalayan catchments facilitate a fast rainfall-runoff conversion and efficient transport to the Terai. Hence, these floods are characterised by a short response time, high magnitude and a quick recession.

Figure 2.7: Annual Maximum Flows recorded at the mountain outlet. The gauging station is maintained by the Department of Hydrology and Meteorology, Nepal.

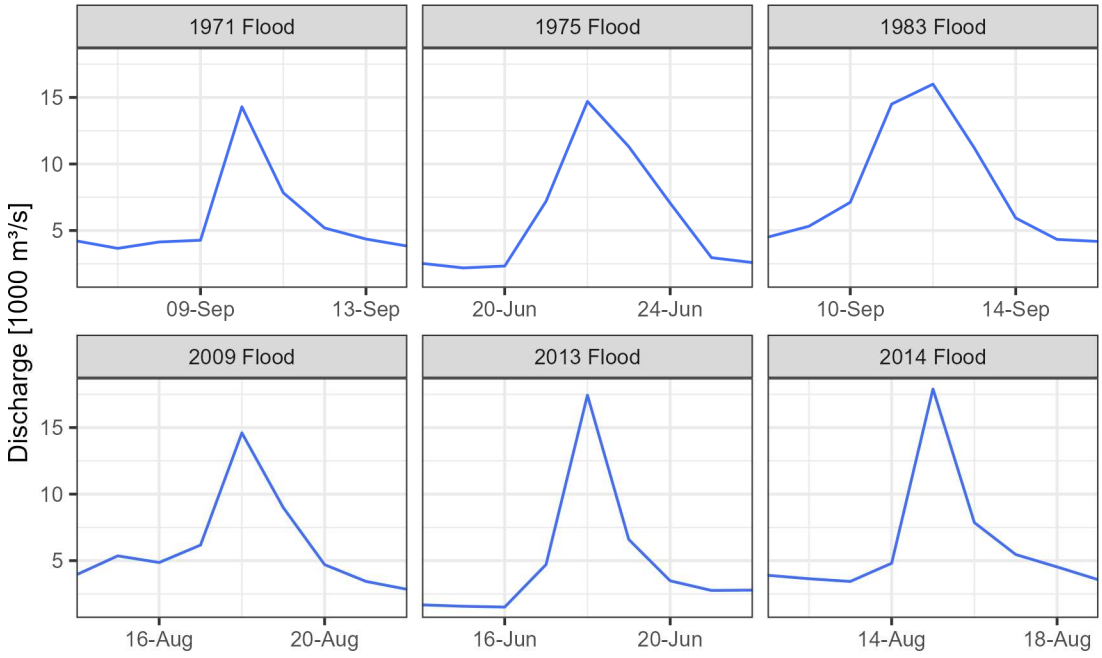


Figure 2.8: Flood hydrographs of the six largest flood events recorded at the mountain outlet in the period 1970 – 2016. The gauging station is maintained by the Department of Hydrology and Meteorology, Nepal.

This high-responsive nature of the Karnali River is illustrated by the 2014 flood event (Figure 2.8). This event was triggered by a large rain field which passed through the catchment between the 14th and 15th of August with rainfall rates of up to 500 mm within 24 hours (MacClune *et al.*, 2015; Shrestha *et al.*, 2015a). The discharge at the mountain outlet rose from 4,800 m^3/s on the 14th to 17,900 m^3/s on the 15th. The discharge receded to 7,800 m^3/s (44% of the peak runoff) on the day after the event. However, these daily averages negate the temporal variability to some extent. The stage record shows that the water level has risen from 9 m at 00:00 to 16 m at 06:00 and receded to 11 m at 23:59 (MacClune *et al.*, 2015). Hence, this river responds on the sub-daily scale and the maximum flood discharge is considerably higher than the daily average indicates.

The shape of the hydrograph and the lag time depend on the interaction between the rainfall pattern and the catchment topographic form. The lag time increases if the rain field is located further upstream and the shape of the hydrograph is affected by the timing and overlay of the subbasin hydrographs (Reaney, 2022). However, the mean runoff of the six largest floods since 1970 is 44% of the peak runoff on the day before the flood, and 57% on the day after the flood which illustrates that the Karnali is a highly responsive catchment. As a result, the floods have a high magnitude because the rainfall drains during a short period.

2.7 The sediment cascade

The sediment cascade describes the geomorphological system of a landscape with the upstream sediment source and the downstream sediment sinks whereas the processes of production, transportation and deposition vary on spatial and temporal scales (Burt and Allison, 2010; Bracken *et al.*, 2015). The young and fragile geology and the steep topography of the Himalayas in combination with the high tectonic activity and monsoon precipitation lead to high sediment production and efficient transport and hence Himalayan rivers are amongst the most sediment-laden rivers in the world (Milliman and Syvitski, 1992; Upreti, 2001; Ballard *et al.*, 2013; Scott *et al.*, 2019).

The processes governing sediment production vary spatially. Glacial erosion dominates the sediment production in the High Mountains above 5,000 masl (Gabet *et al.*, 2008). Further downstream, the most important processes are soil erosion, landslides triggered by earthquakes and heavy rainfall events, and fluvial erosion from glacial lake outburst floods and flash floods (Dingle *et al.*, 2017; Cook *et al.*, 2018; Morin *et al.*, 2018; Maharjan *et al.*, 2021; Adhikari *et al.*, 2023; Chen *et al.*, 2023). The sediment production is highest in the Siwaliks due to the weak geology of this mountain range. Landslides produce most sediments but sheet and gully erosion and to a lesser degree bank erosion contribute to the production (Upreti, 2001; Ghimire *et al.*, 2013).

The steep slopes of the rivers and the high flow rates facilitate the efficient sediment transport. This high transport capacity is generated by GLOFs and LOFs in the upstream subbasins which are the main facilitators for delivering sediment pulses downstream (Cook *et al.*, 2018; Maharjan *et al.*, 2021; Chen *et al.*, 2023). The rainfall-runoff of the monsoon precipitation gains importance with increasing catchment size and dominates the sediment delivery to the Ganga plain (Scott *et al.*, 2019). However, the coarse sediments (gravel, pebbles and boulders) which are delivered to the rivers ≥ 100 km upstream of the mountain front are stored within the channels, abrade and are then transported to the Ganga-Brahmaputra delta as fine sediments (sand or finer) (Lupker *et al.*, 2011; Dingle *et al.*, 2017).

The Upper Siwalik conglomerate is the main source of gravel for the Karnali fan. This pebble-to-bolder-sized conglomerate is made up of resistant quartzite clasts, is loose and almost unconsolidated and is efficiently transported by the Siwalik Rivers (Hurtrez *et al.*, 1999; Upreti, 2001; Dingle *et al.,* 2017; Quick *et al.*, 2019). The conglomerate is located in proximity (approx. 10 – 40 km) to the mountain front and hence does not abrade before reaching the fan (Dingle *et al.,* 2017; Quick *et al.,* 2019). The Lower

Siwalik and Middle Siwalik sediments comprise friable Silt-, Mud- and Sandstones which are, despite the proximity, eroded to fine sediments before reaching the mountain front (Dingle *et al.,* 2017).

The gravel that reaches the fan deposits upstream of the Gravel-Sand-Transition (GST) which is located ~30 km downstream of the mountain outlet (Dingle *et al.*, 2020b). The fan slope reduces from ~0.002 m m^{-1} at the mountain outlet to 0.001 m m^{-1} at the gravel-sand-transition which causes a decrease in the sediment-transport capacity and hence the coarser sediments deposit further upstream. This change is illustrated by the decrease of the median grain size (D50) from 231 mm at the bifurcation to 0.31 mm downstream of the GST (Dingle *et al.*, 2020b).

Alluvial fans are complex landscape features which connect the upstream sediment sources with the downstream sediment sinks and change their behaviour between aggradation (sediment storage) and degradation (sediment source) (Bracken *et al.*, 2015; Harvey, 2018). This fan behaviour is controlled by the sediment supply, the river discharge, and the slope of the fan (Kleinhans *et al.*, 2013; Dingle *et al.*, 2016). The fan evolution is shaped by flood events because these flows deliver most of the sediments and have the highest potential to rework the topography (Kleinhans *et al.*, 2013). On the other hand, flood hazards are affected by the fan behaviour because degrading channels are more stable and have an increased channel capacity while aggrading channels are more prone to avulsions and flood larger areas as the channel capacity is reduced and can be superelevated over the floodplain (Dingle *et al.*, 2016). However, geomorphic processes also occur during lower flows and hence flood events are not the only driver of topographic changes (Kleinhans *et al.*, 2013).

The behaviour of Central Himalayan rivers is controlled by the basin subsidence caused by the subduction of the Indian plate under the Eurasian plate (Dingle *et al.*, 2016). The Eastern Rivers (Koshi, Gandaki) are aggrading due to high subsidence rates while the slowly subsiding Western Rivers (Yarmuna, Ganga) are degrading (the subsidence rates are presented in Figure 1.1). For the Karnali fan both, the estimated subsidence rates and the position of the GST (a proxy for the basin subsidence) approach the characteristics of the degrading Eastern rivers which indicates that the Karnali fan is degrading (Dingle *et al.*, 2016). However, the river channels are superelevated over the Western floodplain which indicates that the channels aggrade and contradicts the evidence of the fan degradation (Figure 2.3). It is, therefore, not clear whether the Karnali River is currently in a degrading or aggrading phase (Kleinhans *et al.*, 2013; Sinha *et al.*, 2014; Dingle *et al.*, 2016).

3 Review of environmental modelling approaches

The prediction of potential flood hazards requires the linkage of environmental models that are applied in the mountainous Karnali catchment and the downstream Karnali fan. The first Section 3.1 reviews hydrological modelling approaches to simulate the catchment hydrology which are applied in Stages 1 and 2 (O1 and O2). Section 3.2 then presents methods to quantify the flood magnitudes and frequencies at the mountain outlet that are discharged into the fan (Stage 2, O3). Hydrodynamic models simulate the motion of water and are used to simulate the spatial flood hazard patterns (Stage 4, O5), and are combined with sediment transport models in morphodynamic models which are used to simulate the morphological evolution of the Karnali fan (Stage 3, O4). The hydrodynamic and morphodynamic models are reviewed in the Sections 3.3 and 3.4, respectively.

3.1 An overview of hydrological modelling

Flood hazards are shaped by the hydrological processes in the upstream catchment which implies that changes in the hydrological system affect the hazard (Beven, 2012; Lane *et al.*, 2011; Lane, 2017). Drivers of such change can be of natural (e.g. solar radiation, volcanic activities) or anthropogenic (e.g. land-use change, agricultural practices, urbanisation) origin and act on different spatial and temporal scales (Holden, 2008; Douville *et al.*, 2021; Pearson *et al.*, 2022). One key driver which has been shown to alter hydrological systems across the globe is anthropogenic climate change (Todd *et al.*, 2011; Prudhomme *et al.*, 2014; Giuntoli *et al.*, 2015; Bolch *et al.*, 2019; Krishnan *et al.*, 2019; IPCC, 2023). Changes in temperature and precipitation patterns affect key hydrological processes including evapotranspiration, rainfall-runoff conversion, and snow and glacier melt. Due to these changes, flood hazards are projected to increase in many regions of the world (Hirabayashi *et al.*, 2013, 2021; Dankers *et al.*, 2014; Douville *et al.*, 2021; Seneviratne *et al.*, 2021; Tellman *et al.*, 2021; Caretta *et al.*, 2022). In mountainous regions, climate change leads to elevation-dependent warming, an increase in orographic precipitation, a decrease in snow cover, a shift in snow seasonality, a decrease of glacier mass, and permafrost thawing and as a result increase in natural hazards (Wang *et al.*, 2019; Ranasinghe *et al.*, 2021). It is, therefore, important to assess the potential climate change impact on flood hazards so that flood management can be adapted to future flood hazards.

The prediction of the climate-change impact, therefore, requires an understanding of the hydrological system in the mountainous catchment that facilitates the flood water generation. This section reviews simulation techniques that provide such information, namely hydrological models (Section 3.1.1) that simulate the rainfall-runoff conversion under consideration of the antecedent conditions (e.g. soil moisture, snowmelt and glacier melt contributions), the calibration of these models to the individual catchment characteristics (Section 3.1.2), methods to estimate the uncertainty in hydrological modelling (Section 3.1.3), and methods to estimate the uncertainty in the discharge observations which are essential datasets for the calibration of hydrological models (Section 3.1.4).

3.1.1 Hydrological modelling approaches

Hydrological simulation models represent the flows of water within the landscape and river system to predict the catchment's behaviour. Despite this mutuality, the models differ in their modelling approach, the temporal and spatial scales of application, simulated processes, predicted variables (the information gain), input data requirements, and computational resources (Pechlivanidis *et al.*, 2011; Beven, 2012). These differences have implications for the suitability of a model for specific objectives and catchments. These different approaches and their applicability for the climate change impact assessment in Central Himalayan river systems are outlined in the following sections.

3.1.1.1 The process description

Empirical models simulate the catchment response to precipitation solely from observations without considering the underlying physical system. These models simulate the response (e.g. the discharge) to physical and climatic descriptors without quantifying the processes that determine the response and are, therefore, also referred to as data-driven black-box models (Pechlivanidis *et al.*, 2011; Devi *et al.*, 2015). Examples of empirical models are the Unit Hydrograph (Sherman, 1932), statistical regression and correlation models, and, more recently, data-based mechanistic models, artificial neural networks and machine learning techniques which exploit the exponential rise of computational processing power and development of new statistical methods (Pechlivanidis *et al.*, 2011; Beven, 2012).

Empirical models have low computational costs and data requirements, are simple to implement and are applicable in ungauged catchments (Pechlivanidis *et al.*, 2011; Devi *et al.*, 2015). However, these models have low explanatory depth due to their low complexity (i.e. no physical process description) (Devi *et al.*, 2015). In other words, such models are capable of predicting the catchment's response but deliver no information about the internal state of the hydrological system that causes the response. Furthermore, empirical models are not transferable across catchments and are based on the assumption of stationarity. Hence they can only be applied to stable physical and climatic conditions (Merritt *et al.*, 2003). It is further criticised that the heterogeneity of catchment characteristics and non-linearities in the system are ignored and that the uncertainty is not depicted adequately (Wheater, 2002; Pechlivanidis *et al.*, 2011; Beven, 2012).

Physics-based models (also mechanistic models) depict the hydrological system with physical equations that quantify the hydrological processes (Pechlivanidis *et al.*, 2011). The physical process description offers several advantages. Physics-based models:

- offer insights into the hydrological state and, thus, a high explanatory depth (white-box models);
- are not bound to the assumption of stationarity and can be used to predict the impact of changes in the physical or climatic conditions;
- are transferable to other catchments because the laws of physics are independent of space and;
- are physically meaningful, measurable parameters that require (in theory) no calibration (Beven, 2012; Devi *et al.*, 2015).

On the contrary, the physical process description requires extensive information about the hydrological system and considerably higher computational resources compared to empirical models (Merritt *et al.*, 2003). The theoretical foundation of the physical equations is based on laboratory or small-scale in-situ experiments. However, in practice, such models are applied on a larger grid scale which results in the lack of theoretical justification that the equations are applicable on the larger scale and the loss of physical significance of the parameters and subsequently the necessity of parameter calibration (Merritt *et al.*, 2003; Gupta *et al.*, 2005; Beven, 2012). Examples of physics-based models include the Soil and Water Assessment Tool (SWAT) (Arnold *et al.*, 1993) and LISFLOOD (De Roo *et al.*, 2000; Van Der Knijff *et al.*, 2010).

Conceptual models describe the hydrological system as a series of internal storages that represent a conceptual view of the important hydrological components. The allocation

and transfer of water to and between the storages are simulated by transfer mechanisms that are based either on empirical or physical equations. These 'grey-box' models are located in between empirical 'black-box' models and physics-based 'white-box' models as they consider the dominating hydrological processes without including specific details of process interactions (Beven, 2012; Terink *et al.*, 2015a). Conceptual models have moderate data requirements and moderate computational costs while providing information about the internal state of the system through the quantification of water fluxes. However, models of this class can vary significantly in complexity and, thus, data requirements and the selection process needs to balance the required model complexity (e.g. research objective, catchment characteristics) and the available information (Wheater, 2002; Pechlivanidis *et al.*, 2011). Conceptual models incorporate parameters that lack physical interpretation and are, thus, not measurable and require calibration which increases the uncertainty in the predictions (Beven, 2012). Examples of conceptual models include the Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström, 1976; Lindström *et al.*, 1997), J2000 (Krause, 2001), and SPatial Processes in Hydrology (SPHY) (Terink *et al.*, 2015a).

It is worth mentioning that models often utilize a hybrid approach combining elements of different types to reduce model complexity and adjust to data scarcity or to increase model complexity to increase the explanatory depth (Pechlivanidis *et al.*, 2011; Terink *et al.*, 2015a). For example, a model could use physical equations to predict evapotranspiration rates while melt processes are predicted from empirical relationships.

3.1.1.2 The spatial discretisation

Hydrological models can further be grouped based on their spatial discretisation into lumped, distributed, and semi-distributed models. Lumped models are the models with the lowest complexity as these models represent the catchment as a homogenous unit and are driven by catchment-average parameters and input data (Beven, 2012). Therefore, these are characterised by low data requirements and computational costs. However, such models are only able to predict the catchment response without providing information about the spatial variability of hydrological processes because the spatial variability in the climatic and physical conditions is not depicted. For example, the predicted discharge at the catchment outlet cannot be decomposed into the subbasin contributions.

Distributed models divide the catchment into a large number of spatial entities (i.e. grid cells) and simulate the hydrological processes on the scale of these entities (Beven, 2012). The scale of these entities is larger than the scale of the variation of processes and hence these models are somewhat lumped conceptual models at the entity scale (Beven, 2012). However, some distributed models implement routines that consider the variation at the sub-entity scale. One example is the SPHY model which simulates glacial processes for the grid cell fraction that is covered by glaciers (Terink *et al.*, 2015a). Distributed models provide more detailed information about the hydrological state of a catchment than lumped models (e.g. the spatial variation of soil moisture within a catchment). However, this comes at the cost of higher data requirements and computational costs, and these models can easily become over-parametrized (Pechlivanidis *et al.*, 2011).

Semi-distributed models divide the catchment into non-continuous entities enabling the consideration and prediction of the heterogeneity of hydrological variables (Pechlivanidis *et al.*, 2011). The spatial resolution of these entities can vary greatly between different approaches of spatial discretisation. The HBV model divides the catchment into subbasins and further into elevation bands and vegetation zones. The J2000 and SWAT models divide the catchment into Hydrological Response Units (HRUs) which are entities with similar hydrological relevant features such as land cover, soil types, elevation, slope, aspect, and groundwater aquifers. Such models aim to combine the advantages of lumped and distributed models to varying degrees. That is the representation of the heterogeneity of important hydrological features with moderate data requirements and computational costs (Pechlivanidis *et al.*, 2011). While semi-distributed models have, in general, a lower spatial resolution compared to distributed models, it is not proven that distributed approaches are superior to semi-distributed approaches in practical applications (Beven, 2012).

3.1.1.3 The temporal discretisation

Models can further be distinguished by their temporal resolution; while event-based models simulate hydrological processes for single events, continuous models simulate hydrological processes on longer time scales (Pechlivanidis *et al.*, 2011). Event-based models are commonly used for the simulation of single flood events (USDA, 2014; Nathan and Ball, 2019). Such models convert rainfall into runoff based on a simple loss model to quantify the rapid excess runoff and simulate the flood hydrograph (Berthet *et al.*, 2009; Nathan and Ball, 2019). These loss parameters are calibrated for individual events and the transfer to other events can result in a loss of the predictive quality (Al-Qurashi *et al.*, 2008). Event-based models are furthermore sensitive to the antecedent conditions (e.g. soil moisture) to accurately predict the rainfall-runoff conversion and thus require information on the hydrological catchment conditions (Berthet *et al.*, 2009).

Continuous models simulate the hydrological processes continuously over longer scales. Generally, these models are more complex because the rainfall-runoff conversion is more sophisticated than the application of a simple loss model. Such models can include algorithms to simulate processes that gain importance on longer time scales, such as snowfall and melt, or glacier melt (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014; Nepal *et al.*, 2014). These models do not require the definition of the antecedent conditions (although they might be defined) because these are predicted by the model. Therefore, continuous models are most commonly used to simulate flood hazards for projected climates for which the antecedent conditions are naturally unknown (Nathan and Ball, 2019). However, both event-based and continuous models need to be calibrated which introduces uncertainty when transferring the parameters from past to future climatic conditions (Stephens *et al.*, 2018; Bérubé *et al.*, 2022).

3.1.1.4 The inclusion of randomness

Models can also be distinguished by the inclusion of randomness. Deterministic models predict similar results as long as the parameterisation and datasets are unchanged. Stochastic models add random variables and, therefore, every model run will produce different results even if the input data and model setup remain unchanged, which allows the estimation of uncertainty. It is worth noting that every deterministic model can be used as a stochastic model by using stochastic input data (e.g. rainfall) or using stochastically generated input parameters (e.g. Monte-Carlo simulations) (Beven and Binley, 1992).

3.1.1.5 Selection of a modelling approach

The richness of model classes indicates the wide range of hydrological models whereas each model has unique characteristics that define its suitability for specific purposes. Therefore, model selection is a crucial step in any hydrological analysis. In general, model selection should be based on the research objective, dominating hydrological processes in the catchment, data availability, know-how, and financial, temporal, and computational resources (Gupta *et al.*, 2005).

In this study, the hydrological model is applied to predict the catchment hydrology for the projected climates. This cannot be achieved with empirical models which require stable conditions and therefore, either a physics-based or conceptual model is required. Any potential model needs to include routines for the prediction of snow and glacier processes to adequately reflect the hydrological system of the Karnali River. Furthermore, a distributed or semi-distributed spatial representation is preferred for two reasons: i) the wide range of climatic and physical conditions at a small scale in Himalayan watersheds cannot be captured by a lumped model and; ii) the more sophisticated output provides information about the spatial variability in the hydrological variables that shape flood hazards. The flood events in the Terai have a short duration (1-2 days) which would favour the selection of an event-based model. However, this is prevented by the lack of hydro-climatic datasets with the sub-daily resolution, and by the lack of information that defines the boundary conditions at the start of the flood event (e.g. soil moisture, groundwater aquifer saturation) which are unknown for the projected conditions. Therefore, a continuous approach is preferred because it predicts the antecedent conditions.

3.1.2 The calibration of hydrological models

In practice, all models require parameter calibration to adjust the model to the unique catchment characteristics regardless of the process description. The calibration of a hydrological model is the process of matching the predicted behaviour with the observed behaviour of the catchment by identifying the set of parameters that best match the unique characteristics of the catchment (Gupta *et al.*, 2005; Beven, 2012). The calibration aims to establish a model in which input-state-output behaviour is consistent with the observed behaviour, produces accurate (low bias) and precise (low uncertainty range) predictions, and has a model structure that reflects the current hydrological understanding of the reality (Gupta *et al.*, 2005).

The calibration process aims to identify behavioural parameter values. However, these parameters might lack physical interpretability and, thus cannot be measured, or are applied on spatial or temporal scales which cannot be measured (Gupta *et al.*, 2005). Thus, the behavioural parameter value needs to be estimated by an indirect process which compares the simulated and observed catchment response (e.g. catchment discharge) (Gupta *et al.*, 2005). This requires performance measures that evaluate the correlation between simulated and observed behaviour (Section 3.1.2.1).

The model calibration can be conducted by a manual calibration, an automated calibration, or a combination of both. In the manual calibration process, the parameter values are adjusted one at a time and the performance is evaluated. This process is subjective, dependent on the experience of the modeller and complex due to parameter interactions and the non-linear nature of the response but can produce accurate results (Gupta *et al.*, 2005). In the automated calibration process, the adjustment of parameter values is conducted by a computer which makes the process less subjective (Beven and Binley, 1992). This requires the identification of reasonable parameter spaces which is complicated by the complex interaction of the parameters and hence it is difficult to define the parameter boundaries. The quality of this approach depends on the number of samples; if the model is run for too few parameter combinations, not all behavioural parameter sets are identified, and if the sampling number is too high the calibration process is ineffective (Gupta *et al.*, 2005).

3.1.2.1 Performance measures

Performance measures (also referred to as objective functions) evaluate the correlation between simulated and observed catchment behaviour. The observed discharge is most commonly used because it is the product of all upstream processes but additional information (e.g. snow extent, evapotranspiration, groundwater levels, glacier mass balance) can be added to increase the robustness of the calibration process (Beven, 2012; Immerzeel *et al.*, 2013; Lutz *et al.*, 2014). A range of different performance measures exist that focus on different aspects of the hydrograph, and it is recommended to use multiple measures to obtain a more robust and differentiated evaluation of the performance of a parameter value or a parameter set (Legates and McCabe, 1999; Westerberg *et al.*, 2011). However, there are no standardised guidelines for the choice of the performance measure and the classification into behavioural and non-behavioural parameter values or parameter sets (Moriasi *et al.*, 2007; Beven and Binley, 2014). Performance measures can be classified into statistical and graphical techniques and Moriasi *et al.* (2007) group the statistical techniques further into standard regression, dimensionless, and error indices.

Standard regression techniques evaluate the strength of the linear relationship between simulated and observed data. One example is the coefficient of determination (R^2) which focuses on the temporal variation between simulated and observed behaviour but is sensitive to outliers, lacks information about volume differences, and should not be used as the sole criterion because also poor models can achieve high correlations (Legates and McCabe, 1999).

Dimensionless techniques provide a relative model evaluation assessment (Moriasi *et al.*, 2007, 2015a). One example is the Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), which is the normalized Mean Squared Error (MSE) (Moriasi *et al.*, 2015a):

$$NSE = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2} \qquad \text{Eq. 3.1}$$

Where O is the observed value, P is the predicted value and i is the time. This measure is sensitive towards higher flows because the differences between simulated and observed model behaviour are squared (Gupta *et al.*, 2009).

Error indices quantify the deviations between simulated and observed behaviour in absolute or relative units. Commonly used indices are mean absolute error (MAE), mean square error (MSE) and root mean square error (RMSE) which provide the modelling error in the unit (or squared unit) of the evaluated variable (e.g. m^3/s for discharge) and

are, thus, easily interpretable (Moriasi *et al.*, 2007). The percentage BIAS (PBIAS) describes the percentual difference between simulated and observed behaviour of the evaluated period. The formulas for the computation of these metrics can be found in Moriasi *et al.* (2015).

Liu *et al.* (2009) calculate the percentage of the modelling time steps where the simulated runoff falls within the confidence interval of the observed runoff to account for the uncertainty in the discharge observations (eGLUE). Furthermore, graphical techniques can be used for performance evaluation. The most common type of visualisation is the hydrograph which shows how well models reproduce the timing and magnitude of the streamflow (Legates and McCabe, 1999; Moriasi *et al.*, 2007, 2015a).

3.1.2.2 Parameter sampling

The optimal sample size for model calibration depends on the number of calibration parameters and the value points per parameter. Hydrological models are, in general, complex models with many parameters and large ranges of parameter values, resulting in large sample sizes and high requirements for computational resources (Beven and Freer, 2001). The random sampling of the parameter space of k dimensions (number of parameters) and P parameter values requires P^k parameter sets and hence a model with 10 parameters and 10 values per parameter requires a sample size of 10 Billion parameter sets (Odoni, 2007). Assuming that the runtime of a model is 1 h it would take >1 Million years of CPU time to simulate all parameter combinations. It is, therefore, necessary to reduce the required sample size wherever possible and find a compromise between parameter sampling and the available resources (e.g. computational resources, time).

One way to reduce the sample size is the application of more sophisticated sampling techniques than random sampling generators. The Latin Hypercube Sampling (LHS) technique (McKay *et al.*, 1979) divides each parameter into non-overlapping intervals according to a probability distribution which are then paired randomly with the values of the other variables (Odoni, 2007). This approach reduces the problem of under-sampling certain areas in the parameter space and reduces the required sample size to $1/10^{th}$ compared to a random sample generator (Vose, 2000; Odoni, 2007).

3.1.2.3 Parameter sensitivity analysis

The required sample size can be reduced by reducing the number of parameter values *P* and the number of calibration parameters *k*. A common approach to reducing the number of calibration parameters is the exclusion of insensitive parameters (Nepal *et al.*, 2014). The parameter sensitivity describes the degree of influence of a parameter on the model output, whereas sensitive parameters dominate the model behaviour (Nepal, 2012; Gan *et al.*, 2014). The approaches for parameter sensitivity analysis can be grouped into two categories; firstly, the local sensitivity analysis which varies one parameter while keeping the others constant and; secondly, the global sensitivity analysis which evaluates the model response by varying all parameters simultaneously and requires a larger sample size. The local sensitivity analysis is simple and intuitive but it lacks information about parameter interactions (Gan *et al.*, 2014).

The Regional Sensitivity Analysis (RSA) (Hornberger and Spear, 1981) is a global sensitivity analysis approach that splits the parameter sets into behavioural and non-behavioural sets. The RSA requires a performance measure (e.g. Nash-Sutcliffe efficiency) for grouping the parameter sets. The normalized cumulative frequency distributions (CFD) for the groups are calculated and plotted. The difference between the CFDs of the behavioural and non-behavioural groups determines the parameter sensitivity whereas the sensitivity increases with the distance of both CFDs (Nepal, 2012; Gan *et al.*, 2014). The obtained CDF plots also provide information about the model performance in the parameter space and can be applied to identify behavioural parameter ranges and decrease the parameter values P.

3.1.3 The uncertainty in hydrological modelling

Hydrological models are simplified replications of complex systems and are developed based on incomplete knowledge of the physical processes and interactions, applied to catchments with incomplete information, and therefore subject to uncertainties that need to be quantified to judge the validity of the predictions.

3.1.3.1 Uncertainty sources

The main uncertainty sources in hydrological models (and environmental models in general) are the structural, parameter and measurement uncertainties (Gupta *et al.*, 2005; Beven and Binley, 2014).

Structural uncertainty arises from the simplification of the complex hydrological system, for example, the aggregation of processes in time and space and is difficult to quantify (Gupta *et al.*, 2005; Beven, 2012). The model structure is static (e.g. global parameters in space and time) and hence the non-linear response of hydrological processes to climate variability (daily – decadal time scale) cannot be simulated (Jehanzaib *et al.*, 2020; Bérubé *et al.*, 2022). This static structure is a particular challenge for the application in Himalayan catchments due to the large climatic and hydrological variations between the monsoon and non-monsoon seasons (Nepal *et al.*, 2014). The resulting structural errors manifest in a model's inability to reproduce every aspect of the hydrograph equally well. This means that a modeller is forced to calibrate the model to the specific aspect of interest (e.g. high flows) and accept poorer performance for less relevant aspects.

One strategy to estimate structural uncertainty is to calibrate the model using a multi-criteria approach that uses multiple performance measures to evaluate the performance of different aspects of the hydrograph (Legates and McCabe, 1999). However, this approach cannot quantify the uncertainty that is inherent in the model structure (e.g. the conceptual design of the model). This model structure uncertainty can be quantified by simulating the catchment hydrology with an ensemble of different hydrological models with different model structures (Prudhomme *et al.*, 2014; Giuntoli *et al.*, 2015, 2018, 2021; Huang *et al.*, 2017; Krysanova *et al.*, 2017; Pechlivanidis *et al.*, 2017). However, this approach is complex because the input and output (e.g. data structures) potentially vary between the models.

Parameter uncertainty describes the uncertainty due to imperfect model parameterisation. The true parameter set could be estimated from a comparison of the simulated and observed catchment behaviour. This requires detailed information about the hydrological state of the components (e.g. soil moisture, groundwater recharge, etc.) which is unknown for most catchments which is why models are calibrated for a subset of information (e.g. discharge, snow cover, etc.). However, different parameter sets can produce similar results. This phenomenon is known as model equifinality and is not limited to hydrological models but applies to environmental models (Beven and Binley, 1992, 2014; Beven and Freer, 2001; Odoni, 2007). The model equifinality, imperfect catchment knowledge, and unmeasurable parameters hamper the identification of the "true" parameter set. Therefore, all parameter sets that predict the variable of interest must be considered as equally likely representations of the catchment hydrology (Beven and Freer, 2001; Gupta *et al.*, 2005). The parameter uncertainty can be estimated by applying the model with different parameter combinations, evaluating the performance of each parameter set and keeping all parameter sets that are classified as behavioural. Examples of parameter uncertainty estimation methods are the Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992) and the Bayesian Recursive Estimation (BaRE) (Thiemann *et al.*, 2001).

Measurement uncertainty is the uncertainty that relates to the observation error in the data that is used either as input (e.g. precipitation) or to calibrate and validate the model (e.g. discharge). All measurements are subject to systematic and random errors that result from the person conducting the measurements, the instruments, and the measured system and these errors in the data propagate in the modelling process (Liu *et al.*, 2009). Measurement uncertainty can be quantified depending on the utilization of the dataset. For datasets that are used as modelling input, the uncertainty can be estimated by creating synthetic datasets that reflect the measurement error and applying the model with these synthetic datasets. However, this requires information about the distribution of the error which is unknown in most cases (Liu *et al.*, 2009). A simple solution to estimate the uncertainty without any knowledge about the nature of the error is the introduction of a scaling factor as a calibration parameter (Crawford and Linsley, 1966). For simulations of the projected hydrology, it is recommended to use probabilistic climate projections (climate projections from an ensemble of climate models) to account for the uncertainty in the climate models (Giuntoli *et al.*, 2015, 2018, 2021). The measurement uncertainty of calibration and validation datasets can be considered by evaluating the model performance from the confidence intervals of the datasets (Liu *et al.*, 2009).

The structural and parameter uncertainty can be quantified within the GLUE framework, which is introduced in the next section. However, this framework requires data about the catchment's response which includes the measurement uncertainty. The extended GLUE (eGLUE) (Liu *et al.*, 2009) accounts for the measurement uncertainty in the calibration data and is introduced in Section 3.1.3.3.

3.1.3.2 The Generalized Likelihood Uncertainty Estimation framework

One of the main challenges for the calibration of hydrological models is the lack of information about the internal state of the modelled system and the model equifinality. The GLUE framework (Beven and Binley, 1992) approaches model equifinality by assuming that it is impossible to identify the true parameter set. Instead, every parameter set is evaluated regarding the likelihood of being a behavioural (acceptable) simulator of the system. It is worth noting that likelihood is not defined in the sense of Maximum Likelihood Theory which assumes a normal distribution of errors. Instead, the likelihood in GLUE is a probabilistic measure of how well a parameter set replicates the observed behaviour (Beven and Binley, 1992). A performance measure is used to indicate this likelihood and all behavioural parameter sets are maintained and treated as equally likely simulators of the system. Maintaining *n* behavioural parameter sets results in a range of *n* simulated model outputs and this range defines the confidence interval of the model predictions, and, hence, quantifies the parameter uncertainty. The GLUE framework can be extended to estimate the structural uncertainty by combining it with a multi-criteria approach (Gupta *et al.*, 2005). For this, the performance of each parameter set is evaluated for a set of performance measures and all parameter sets that are classified as behavioural for any of the performance measures are maintained. The measurement error can be estimated to some extent in GLUE by introducing calibration parameters that scale the input measurements (Crawford and Linsley, 1966; Lutz *et al.*, 2014).

The main criticism of GLUE is the lack of a formal statistical likelihood and hence the application of GLUE involves subjective decisions (Clark *et al.*, 2011, 2012; Beven and Binley, 2014). There are no universal guidelines for the use of performance measures and the identification of behavioural parameter sets. The most common approach is the definition of performance measure thresholds, whereas no universally accepted thresholds exist (Beven and Binley, 2014). Moriasi *et al.* (2007, 2015) define thresholds based on a synthesis of published studies. However, the thresholds need adjustment to the catchment and objective, i.e. data quality and quantity, model calibration procedure, evaluation time step, and the project scope. The definition of thresholds can be avoided

when using a ranked selection approach which classifies the n parameter sets with the highest efficiencies as behavioural (Pearson, 2020). This approach has the disadvantage that classification is not based on the performance of a parameter set. Non-behavioural parameter sets can be classified as behavioural if there are not *n* better-performing parameter sets, and behavioural parameter sets can be eliminated if they are not among the *n* highest-performing parameter sets.

3.1.3.3 The extended Generalised Likelihood Uncertainty Estimation

The GLUE method estimates the uncertainty from various sources by evaluating the performance of a set of model realisations (parametrization, input data) based on the observed catchment behaviour. The discharge is an essential indicator of the catchment behaviour and is most commonly used in the hydrological modelling of gauged catchments because it contains information about all upstream processes. However, the observed discharge is usually derived from the observed stage using statistical methods (see Section 3.1.4) and contains uncertainty, especially towards the extremes of the flow range (McMillan *et al.*, 2012; Kiang *et al.*, 2018). Since the model realisations are classified based on the fit between the two (simulated and observed) flow time series, the uncertainty in the observed discharge propagates to the model.

The extended GLUE (Liu *et al.*, 2009) extends the GLUE by considering the error in the observation data. This is achieved by normalising the uncertainty in the observed discharge and calculating a normalised score at any time step *t*:

$$\text{Score (t)} = \begin{cases} (Yt - Q_t)/\,(Q_t - Q_{min,t}) & Yt < Q_t \\ (Yt - Q_t)/\,(Q_{max,t} - Q_t) & Yt \geq Q_t \end{cases} \qquad \text{Eq. 3.2}$$

Where Y_t is the simulated discharge at time step *t*, Q_t is the best estimate of the observed discharge and $Q_{min,\,t}$ and $Q_{max,\,t}$ are the lower and higher confidence bands of the observed discharge at time step *t*. If the score lies within the range between -1 and 1, the model predictions fall within the uncertainty range of the observed discharge. Therefore, -1 and 1 define the lower and upper limits of acceptability.

This score is then used to identify behavioural parameter sets. However, it is unlikely that the prediction of a model falls within the limits of acceptability at every time step due to model structural errors and model input errors. The authors suggest three ways to classify behavioural parameter sets; i) the selection of parameter sets for which the simulated output falls within the confidence interval for 95% (or any other quantile) of the time steps; ii) extending the limits of acceptability until a predefined number of model

realisations are classified behavioural and; iii) the derivation of the distribution of scores over all time steps and classifying all realisations as behavioural which 5% and 95% scores fall within the range of -1 to 1. If no model satisfies this condition, the models can be ranked based on their 5% and 95% (or any other quantile) scores and the best n models are considered as behavioural.

3.1.4 The uncertainty in discharge observations

The discharge contains information about all hydrological processes in the upstream catchment. It is an essential indicator of a catchment's behaviour and hence discharge observations are used frequently (if available) for the calibration of hydrological models. However, discharge is difficult to measure directly and continuously, so most discharge time series are estimated from continuous stage measurements using stage-discharge rating curves (Kiang *et al.*, 2018). This rating curve approach is the standard procedure for discharge records in Nepal (Nepal, 2012). Despite the statistical origin of discharge time series it is most commonly referred to as observed discharge in the literature. This terminology is misguiding because it indicates a false sense of accuracy. The observed discharge contains more uncertainty than one would expect from measurements. The uncertainty typically ranges between ±50 – 100% for low flows, ±10 – 20% for medium flows, and ±40% for high flows but the uncertainty can reach up to ±200% depending on the study site and sample size (McMillan *et al.*, 2012; Kiang *et al.*, 2018).

3.1.4.1 The concept of rating curves

Stage-discharge rating curves are statistical models that relate stage height (the water level) to the discharge. This relationship is derived from a relatively small sample of discrete, concurrent observations of discharge and stage heights (Kiang *et al.*, 2018). Once a stage-discharge relation is established it can be used to estimate the discharge from stage heights measurement without the necessity of measuring discharge continuously. This relation is represented by a rating curve and can be expressed by the equation (Herschy, 1993):

$$Q = C\,(h + a)^d \qquad \text{Eq. 3.3}$$

where Q is the discharge, h is the stage height, a is the stage of zero flow (datum correction) and C and d are constants that depend on the control.

The control describes the channel characteristics at the gauging station and is defined by the geometry, bed slope, roughness of the bed and banks, sinuosity, stability and vegetation cover (Herschy, 1993; Wiche and Holmes, 2016). Eq. 3.3 is the basic function to describe the stage-discharge relationship, but a range of functions exist that are applied to specific cross-section geometries, such as triangular (Eq. 3.4), rectangular (Eq. 3.5) and parabolic controls (Eq. 3.6) (Le Coz, 2014):

$$Q = C_t \sqrt{2g} \tan\left(\frac{v}{2}\right)(h-b)^c \qquad \text{Eq. 3.4}$$

$$Q = C_t \sqrt{2g}\, \frac{B_p}{H_p}(h-b)^c \qquad \text{Eq. 3.5}$$

$$Q = C_t \sqrt{2g}\, B_w(h-b)^c \qquad \text{Eq. 3.6}$$

Where B_w = spillway width (m), B_p = width of the parabola (m), H_p = height of the parabola (m), v = triangle opening angle (°), C_t = discharge coefficient, g = gravitational acceleration, c = exponent for rectangular/parabolic/triangular cross-section, and b is the offset. The flow is zero if the water level falls below this offset.

The statistical nature of the rating curve implies that the curve is only valid for the conditions of the underlying stage and discharge samples. A new curve needs to be fitted if the control changes, e.g. due to erosion and sedimentation processes. It is necessary to fit different curves for the vegetation periods for controls with great seasonal variation of the vegetation cover (Kennedy, 1983; Coxon *et al.*, 2015). Ideally, a gauging site has a stable control and is unaffected by seasonal variations. However, controls change with time which requires continuous sampling (Herschy, 1993).

For gauging sites with irregular controls, it is required to fit multiple curves for each section of the control (Coxon *et al.*, 2015). For example, a river in which water drains in a triangular channel during low and medium flows but during high flows, the water enters a wide flat flood plain requires the development of separate rating curves; one for the low to medium flow conditions and one that is activated when the stage exceeds the bank-full stage.

The different stage-discharge relations for different stage heights requires the sampling over the full range of flows because the samples contain no information about the stage-discharge relationship above the highest and below the lowest samples. However, extreme flows occur less frequently by definition and are more difficult or impossible to sample. Therefore, fewer samples are available for the extremes of the flow spectrum which manifests in the larger uncertainty ranges compared to the medium flow range uncertainty (McMillan *et al.*, 2012). In practice, the samples do not cover the full range

of flows and rating curves are therefore extrapolated beyond the lowest and highest samples. As a guideline, the United States Geological Service (USGS) never extrapolates above twice the maximum observation (Wiche and Holmes, 2016). However, the extrapolation of rating curves is common practice in developing countries where site access and logistics make it difficult to measure during high flow conditions (Clarke, 1999).

3.1.4.2 Uncertainty sources

The uncertainty in the observed discharge results from errors in the measurements and the parameterisation of the statistical model. The measurement uncertainty comprises the errors in stage and discharge measurements and depends on various factors such as the measurement technique, velocity, number of verticals, and the duration of the measurements (ISO, 2007; Shrestha and Simonovic, 2010). The uncertainty in the stage measurements dominates for low flows and is comparatively small compared to the discharge measurement uncertainty which dominates for high flows (Clarke, 1999; Scanlon *et al.*, 2008; McMahon and Peel, 2019). The parameter uncertainty relates to the parameterisation of the statistical models and results from an imperfect approximation of the true stage-discharge relationship, the limited number of samples and the extrapolation beyond the sampled minimum and maximum (Kiang *et al.*, 2018). The structural uncertainty describes the fraction of uncertainty that results from the physical factors of the control that influences the stage-discharge relationship such as unsteady flow and backwater effects but also temporal changes in the channel cross-section due to morphological processes, changes in vegetation and ice formation (Kiang *et al.*, 2018).

3.1.4.3 Methods to estimate the discharge uncertainty

An increasing focus was set on the development of methods to estimate the uncertainty of stage-discharge rating curve uncertainty in the past 2½ decades. These methods use different approaches to estimate different aspects of the uncertainty. Traditional statistical approaches estimate the parameter uncertainty based on the residual variance from the curve or the variance of the parameter estimates (Coxon *et al.*, 2015). Examples are the approaches of Clarke (1999) which focuses on the uncertainty estimation of peak flows, and Petersen-Øverleir and Reitan (2005) which includes a methodology to detect rating curve shifts based on objective segmentation. Bayesian approaches are based on probability density functions that combine hydraulic knowledge to determine prior distributions of the model parameters and likelihood functions to account for the uncertainty in the individual gaugings into a posterior distribution (Le Coz *et al.*, 2014; Coxon *et al.*, 2015). The uncertainty is then estimated based on the percentiles of the posterior distribution. One example is the BaRatin model (Le Coz *et al.*, 2014) which estimates the uncertainty from the measurements and the model parameterisation. Alternative approaches utilize different methods such as fuzzy regression (Shrestha and Simonovic, 2010), variographic analysis (Jalbert *et al.*, 2011) or non-parametric regression (Coxon *et al.*, 2015) to estimate the rating curve uncertainty. All methods vary in the way that uncertainty is estimated, the considered sources of uncertainty, the temporal and spatial scale of their application, river types, data requirements and the output (Kiang et al., 2018). Therefore, the selection of an appropriate method depends on the study area, research aim and data availability.

3.2 The Flood Frequency Analysis

A hydrological model is applied to the Karnali catchment to predict the flood discharge into the fan for probabilistic climate scenarios. However, flood hazards are shaped by the individual characteristics of the flood-triggering rainfall event and the antecedent conditions which hampers a direct comparison of individual flood events (Bérubé *et al.*, 2022; Reaney, 2022). The Flood Frequency Analysis (FFA) is a statistical method that standardises flood hazards by relating the flood magnitude (flood discharge) with the flood frequency (occurrence probability) from a record of flood events (Hosking and Wallis, 1997). The FFA, therefore, enables the comparison of flood hazards predicted for different climatic conditions (e.g. past and projected climatic conditions) or different datasets (e.g. simulated and observed flood records).

3.2.1 The concept

Flood frequency analysis (FFA) utilises extreme value statistics to establish a statistical relationship between flood magnitude and exceedance probability from a record of flood discharges. These flood discharges must be independent of each other (i.e. one discharge per event) (WMO, 1989). This relationship is described by a flood frequency curve and the flood magnitude for a specific exceedance probability can then be derived from this curve. Flood frequency curves can be compared in space or time to investigate how the magnitudes of a given exceedance probability differ, or to quantify the differences in the exceedance probabilities for a given magnitude. Therefore, the FFA can be used to compare the change in the flood magnitude between the baseline and projected climates for a given exceedance probability.

The basic principle of the FFA is the relationship between flow rate and exceedance probability. The exceedance probability can be derived from the distribution of flows, particularly from the percentiles (also: non-exceedance probability) P that describe the fraction floods in the record that are below a given flow rate. The Annual Exceedance Probability (AEP) is defined as 1 – P and the return period as 1 / (1 – P). For example, assuming that the flow at the 80th percentile of a record of 100 years is 1000 m^3/s, the AEP is 0.8 and, hence, the return period is five years. This means that every year there is a 20% probability that flows ≥ 1000 m^3/s occur and hence these flows are exceeded statistically once every five years.

The quality of the FFA depends on the ability of an Extreme Value Distribution (EVD) to describe the statistical relationship between the flow magnitude and the frequency. A range of different approaches exists that differ regarding the sampling methods, methods to fit the distribution to the data, and the distribution itself.

3.2.2 Sampling methods

Every application of a FFA requires a record of flood discharges. However, only a fraction of a hydrograph is relevant for the analysis and the two most commonly applied models to extract the relevant information from the flow record are the Annual Maximum Flow model (AMAX) and Peaks-Over-Threshold model (POT) (also referred to as partial duration series) (Parkes, 2015). The AMAX model is a block sampling approach that divides the record into equally sized blocks of one year (or hydrological year) and extracts the maximum flow of each block. This model has the advantage that floods of different years can be considered independent (except for events during the turn of the year). On the other hand, information is lost if two floods occur within the same year, and non-flood flows are misclassified as floods during years without floods (Mangini *et al.*, 2018). The POT approach overcomes this limitation by keeping all flows above the defined flood level. This threshold is the bankfull flow (or stage) above which the floodplain is inundated. Hence the flood record is not restricted to a single value per year and does not contain any below-flood-level flows. However, this approach is sensitive towards the definition of the flood threshold and requires the validation that peaks are independent and not part of the same event (e.g. by defining a minimum number of days between the peaks) (WMO, 1989; Mangini *et al.*, 2018). The choice of sampling model affects the predicted magnitudes for the more frequent events (<10 years), but the differences decrease with the frequency (WMO, 2008; Shaw, 2005).

3.2.3 Extreme Value Distributions

The relationship between the magnitude and frequency is simulated from Extreme Value Distributions (EVD). Several distributions are suitable for the determination of the flood frequency and these distributions are classified based on the number of parameters that describe the distribution; the Gumbel and Weibull distributions have two parameters; the generalised extreme value, generalised logistic and Pearson Type III distributions have three parameters; the kappa and log Pearson Type III distributions have four parameters and; the Wakeby distribution has five parameters. More distributions exist, but the stated distributions are the most common for flood frequency analysis with AMAX datasets (Parkes, 2015). It is worth noting that the FFA evolved from extreme value statistics but not all distributions (i.e. Kappa, Wakeby) belong to the class of extreme value distributions (WMO, 2008). Nonetheless, for the sake of simplicity, all distributions that are commonly applied in the FFA are referred to as EVD in this thesis. For POT, the generalized Pareto distribution is most commonly chosen (Parkes, 2015). A more detailed description of the distributions is provided by Rao and Hamed (2000) Rao & Hamed (1999) and WMO (2008).

The application of the EVD involves the process of fitting the statistical distribution to the data. This fitting is the process of identifying the parameter values that best describe the observed distribution and, thus, produces the lowest differences between the simulated distribution (the cumulative distribution function of the EVD) and the observed distribution (the empirical distribution function of the flow record). A common approach for fitting distributions is the Maximum-Likelihood estimator approach which selects the parameters that are most likely to generate the observed data (Miura, 2011). However, this method is based on the large sample theory which conflicts with the low sample sizes of flood flows (Hosking *et al.*, 1985). A more common method for fitting the EVD is the method of moments. Moments are measures that describe the location, scale and shape of probability distributions, in particular the expected value (1st moment), variance (2nd moment), skewness (3rd moment) and kurtosis (4th moment) (Hosking and Wallis, 1997). However, the estimation of higher moments is biased which can, in some cases, result in unreliable estimations about the shape of the distribution. These biases are reduced in the L-Moments approach (Hosking, 1990) which is less sensitive towards outliers and provides more robust estimates for a wider range of shapes of the distribution (Hosking and Wallis, 1997). This method is widely used in FFA because it performs well for low sample sizes (Reiss and Thomas, 2007; WMO, 2008).

The EVD selection process aims to identify and select the fitted distribution that best describes the observed relationship between flood magnitude and frequency. The suitability of the distributions can be evaluated by their ability to reproduce those features of the data that are of importance for the application (Hosking and Wallis, 1997). The properties of a distribution that are most relevant for the FFA are the upper bound, the upper tail and the shape of the body of a distribution:

- The upper bound defines a reasonable upper limit of the distribution. Physical quantities, such as river flows, have an upper limit (which might not be known) beyond which these quantities are deemed physically impossible, and it is sometimes argued that bounded distributions should be used (Hosking and Wallis, 1997). However, this is only a concern for the extrapolation, and physically impossible values are only predicted for very high return periods (> 10,000 years) that are irrelevant for most applications. (Hosking and Wallis (1997) argue that it is, therefore, not required to select distributions that allow the definition of an upper bound for the FFA.
- The upper tail of the distribution contains the largest flood discharges which are naturally of the highest interest in the FFA (e.g. 99% of the flow range is below the flow with the 1% Annual Exceedance Probability). The tail weight defines the behaviour of the frequency curve with the increasing discharge and determines the rate at which quantiles increase when extrapolated beyond the range of the data (Hosking and Wallis, 1997).
- The shape of the body of the distribution is relevant for river systems where floods are caused by different processes (e.g. snow melt, heavy rainfall) and separate distributions can be applied for different parts of the data (e.g. separate distributions for snowmelt-induced floods and heavy rainfall-induced floods) (Hosking and Wallis, 1997).

3.2.4 Evaluation of the performance

The selection of the distribution depends on the catchment properties and flow regimes and while some distributions are recommended for specific sites, it is generally impossible to determine the most suitable distribution before the application (Hosking and Wallis, 1997; Parkes, 2015). Therefore, it is recommended to test multiple distributions and select the distribution that best describes the relationship between the flood discharge and AEP. This requires a procedure that evaluates the performance (also goodness-of-fit) of a distribution.

Several performance measures are deemed appropriate for evaluating the ability of a cumulative distribution function (in this case the EVD) to capture the empiric (observed) distribution. However, to this point, there is no consensus about which measure is most appropriate for the FFA (Heo *et al.*, 2013). The proposed measures can be classified as analytical or graphical measures. Graphical measures use visualisation techniques to compare the fitted distribution and the data. Commonly used graphical techniques include quantile-quantile plots and probability plot tests (Hosking and Wallis, 1997; Heo *et al.*, 2008; Parkes, 2015). Analytical measures evaluate the performance from statistical analysis and provide comparable, objective statistical values. Empirical distribution tests evaluate the correlation between the simulated (EVD) and empiric (observed) Cumulative Density Functions (CDF) (Heo *et al.*, 2013). Examples of such tests are the chi-squared test, the Cramer-von-Mises test, and the Kolmogorov-Smirnov test. These tests evaluate the performance across the range of the distribution, while the upper tail of the distribution is of particular importance for the FFA. The Anderson-Darling (AD) test (Anderson and Darling, 1952) emphasises the tails of the distribution by incorporating weights at the tails. The AD test was modified by Ahmad *et al.* (1988) who introduced weighting functions to give greater weight to one side of the tails. Heo *et al.* (2013) modified the AD test to consider the effect of unknown shape parameters.

3.2.5 The sources of uncertainty

Flood frequency analysis is a data-driven method and the largest drivers of uncertainty relate to the input data. The main sources of uncertainty are model errors, sampling errors, measurement errors, and the non-stationarity of the hydrological system. The model error describes the differences between the modelled and true distributions because the flood record is not bound to converge on any mathematically defined distribution or combination of distributions (Parkes, 2015). However, the model error is small compared to the other sources of uncertainty and is usually not considered in uncertainty assessments (Kjeldsen *et al.*, 2014). The measurement error results from inaccuracies in the flood record and is one of the main sources of uncertainty (Kjeldsen *et al.*, 2014). For more details on estimating measurement errors, the reader is referred to the overview of stage-discharge rating curves (Section 3.1.4). The sampling error and the uncertainty related to the non-stationarity are described in the following sections.

3.2.5.1 Sampling error

The sampling error results from estimating the true flow-frequency relationship from a subset of flows and is a dominant source of uncertainty (Apel *et al.*, 2008; Parkes, 2015). The root of the sampling error is the limited record length in any river system. Flood Frequency Analysis (FFA) is concerned with extreme values and hence the sample is low, even for well-observed rivers with long record lengths. This means that the sample is only a small subset of the true distribution which inevitably results in deviations between the true distribution and the sampled distribution. Consider two flood frequency curves fitted for the period 1950 – 2000 and 1960 – 2010, respectively. The sampling error describes the deviations of the two curves that result from using different samples of the same true distribution. The deviations, and thus the sampling error, increase with a decrease in the record length (Kjeldsen *et al.*, 2014).

The record length (sample size) is a crucial factor in the FFA that limits the estimation of flood magnitude for larger return periods. The 1-in-100-years flood is the benchmark for flood management in many parts of the world (Dalrymple, 1960; USDA, 2014; Parkes and Demeritt, 2016; Nathan and Ball, 2019). The estimation of this event requires extrapolation beyond the sample given the relatively short record range for most river systems. For example, the magnitude of a 1-in-100-years event might be estimated from a record of 30 years which might or might not contain a 1-in-100-years event. Thus, the sampling error is related to the large internal variation in extreme flows. The length of

extrapolation which still provides reasonable estimates depends on this internal variation and varies between river systems (Li *et al.*, 2018).

The common approach for reducing the sampling error is to increase the sample size for the parameter fitting (WMO, 2008). This can be achieved by adding information from additional sources such as the reconstruction of historic floods (Parkes and Demeritt, 2016). The regional FFA adds observations of similar river systems that have a high degree of correlation to the sample to increase the sample size (Hosking and Wallis, 1997; Kjeldsen *et al.*, 2014). When conducting the FFA with simulated flows from hydrological models, the models can be applied with longer (i.e. 1000 years) synthetic climate datasets that mimic the distribution of the observed data (Kjeldsen *et al.*, 2014).

In the case that the sample size cannot be extended the sampling error can be estimated from statistical methods. Burn (2003) estimated the sampling error from bootstrapping using a balanced resampling procedure. A record of *n* observations is duplicated *B* times, concatenated, and *B* synthetic records are created by randomly selecting *n* samples from the concatenated record. The third approach quantifies the sampling error using Monte-Carlo Simulations (MCS). *B* synthetic records of length *n* are created by selecting the flow at *n* random (but uniform distribution between 0 - 1) locations (quantiles) of the empirical CDF and repeating this process *B* times. A frequency curve is then fitted to each of the *B* records (Kjeldsen *et al.*, 2014).

3.2.5.2 Non-Stationarity of flood flows

The Flood Frequency Analysis (FFA) assumes that the data is stationary. Stationary data is identically distributed which means that the distribution does not change over time (Koutsoyiannis and Montanari, 2007). Non-stationarity in FFA means that the distribution of the flood record changes with time. Such changes are caused by changes in the hydrological system which can, for example, result from natural climate variability (e.g. in atmospheric circulation systems), anthropogenic climate change, land-use change, and human interventions (e.g. the construction of dams) (Pattison and Lane, 2012; Parkes, 2015; Bérubé *et al.*, 2022).

The assumption of stationarity is problematic for climate-change impact assessments because this implies a changing system and, thus, non-stationarity. The attribution of the observed climate change on flood discharge is challenged by the vast anthropogenic impact on the rivers in the past century, e.g. land use change, river engineering, and urbanisation (Douville *et al.*, 2021). However, the observed streamflow trends can only

be simulated in models when accounting for the anthropogenic radiative forcing (Douville *et al.*, 2021; Seneviratne *et al.*, 2021). Climate change is projected to alter flood flows globally whereas the response correlates with the projected emissions (Hirabayashi *et al.*, 2013, 2021; Dankers *et al.*, 2014; Seneviratne *et al.*, 2021). This indicates that the assumption of non-stationarity is not met for any kind of climate change impact assessment. Therefore, the violation of this assumption is inevitable for the objective in this research and this violation introduces uncertainty. The uncertainty could be reduced by lowering the record length so that the non-stationary signal perishes in the internal variation of flood flows. However, this reduction increases the sampling error and, therefore, the definition of the record length is a trade-off between the uncertainty introduced by the non-stationarity and the sampling error.

The modelling framework to predict the evolution of flood hazards consists of four research stages. Stage 1 and Stage 2 utilise models that are applied on the catchment scale and Stages 3 and 4 focus on predictions in the Karnali fan (floodplain scale). Stage 3 applies a morphodynamic model to predict the morphological evolution for the projected flood discharges. Stage 4 applies a hydrodynamic model for the projected flood magnitudes and topographies. Morphodynamic models often couple a hydrodynamic model to simulate the flow of water and a sediment transport model to predict sediment processes (erosion, transport, deposition) (Hardy, 2013). Therefore, hydrodynamic models are reviewed (Section 3.3) before the morphodynamic models (Section 3.4).

3.3 The simulation of water flow

Hydrodynamic models simulate the flow of water based on the assumption that fluid can be considered as a continuum (Lane, 1998). These models are based on the equations for the conservation of mass and the conservation of momentum. Due to their physical process description, these models are also referred to as physics-based or process-based models (Williams *et al.*, 2016). The physical process description is based on the Navier-Stokes equations for the conservation of mass and momentum which describe the motion of fluids (Ingham and Ma, 2005; Williams *et al.*, 2016). However, the simulation of this three-dimensional motion is overly complex for many applications and derivative equations reduce this complexity by reducing the dimensionality of the flow computations or by the level of process representation (Teng *et al.*, 2017).

3.3.1 One-dimensional models

One-dimensional models simulate the flow of water in the longitudinal direction neglecting the lateral and vertical motion, and are, thus, the simplest hydrodynamic models. This 1-D flow may be simulated from the 1-D Saint-Venant equations (SVE) for the conservation of mass and the conservation of momentum, approximations of the SVE which are simplifications of the 2-D Shallow-Water equations (which are in turn simplifications of the 3-D Navier-Stokes equations), the Muskingum-Cunge routing which uses a diffusion representation of the conservation of momentum, and from the Mannings equations (Miller, 1984; Ponce *et al.*, 1996; Teng *et al.*, 2017). These models assume the one-dimensionality of flow with a uniform velocity in the lateral and vertical directions and a horizontal water level in the lateral direction (Tayefi *et al.*, 2007).

The topography of the channel and floodplain constrains the flow and is, thus, an essential boundary condition for any hydrodynamic model. The topography is represented by a series of cross-sections through the channel and floodplain in 1-D models. The flow between the cross-section is simulated as the cross-section average. This approach generally produces satisfactory results for in-channel flows (Tayefi *et al.*, 2007). The variation in velocity is often larger for floodplain flows due to the more complex topography which cannot be captured by the cross-section averages of 1-D models (Tayefi *et al.*, 2007; Teng *et al.*, 2017). The lack of the lateral diffusion of the flood wave and the simplistic representation of the floodplain topography as a series of cross-sections results in the poor prediction of out-of-bank flows in settings with a

complex floodplain topography (Tayefi *et al.*, 2007; Williams *et al.*, 2016). These limitations led to the development of the coupling of 1-D models to simulate the channel flow, and 2-D models for the simulation of flow over the floodplain (Teng et al., 2017).

3.3.2 Two-dimensional models

Two-dimensional horizontal models simulate the flow of water in the longitudinal and lateral directions ignoring vertical motions from the full Shallow-Water Equations (SWE) which apply where the lateral extent is greater than the vertical extent, and, hence the water is shallow (Jirka and Uijttewaal, 2004; Williams *et al.*, 2016). These equations are obtained from depth-averaging the Navier-Stokes equations (Parkes, 2015):

Conservation of mass: $$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} + \frac{\partial (hv)}{\partial y} = 0$$ Eq. 3.7

Conservation of momentum: $$\frac{\partial hu}{\partial t} + \frac{\partial \, (gh^2/2 + hu^2)}{\partial x} + \frac{\partial huv}{\partial y} = gh \, (S_{0x} - \, S_{fx})$$ Eq. 3.8

$$\frac{\partial hv}{\partial t} + \frac{\partial huv}{\partial x} + \frac{\partial \, (gh^2/2 + hv^2)}{\partial y} = \, gh \, (S_{0y} - \, S_{fy})$$ Eq. 3.9

Where *x* and *y* are the two spatial dimensions, u and v are the depth-averaged velocities in the x and y directions respectively, S_{0x} and S_{0y} are the ground slopes in the x and y direction, S_{fx} and S_{fy} are the friction slopes in the x and y direction, *t* is time, *h* is the water depth, and *g* is the gravitational acceleration.

The full SWE are more complex than the 1-D SVE and some models use approximation of the full SWE to reduce this complexity by reducing the terms of the conservation of momentum equations (Neal *et al.*, 2012). The diffusion wave model maintains the pressure, ground slope and friction slope (3-term model) and omits the convective acceleration. The kinematic wave model also removes the pressure term and maintains the ground slope and friction slope (2-term model) (Teng *et al.*, 2017). The 2-D SWE (or approximations of the SWE) assume that the water depth is shallow in comparison to the longitudinal and lateral direction and hence these models cannot be applied to deeper systems such as reservoirs, lakes and estuaries (Teng *et al.*, 2017; USACE, 2020)

Bates *et al.* (2010) developed a simple 1-D inertial model which was derived from the SWE and is applied in the longitudinal and lateral direction to simulate 2-D flow at low computational costs (Bates *et al.*, 2010; Neal *et al.*, 2012; Coulthard *et al.*, 2013). These simplifications of the full SWE reduce the complexity and, thus, the computational

requirements while maintaining the predictive capabilities for many applications ((Neal *et al.*, 2012). However, the slope friction is related to Manning's equation (or Chezy equation) which introduces a roughness coefficient. This roughness coefficient compensates for the loss of physical complexity and these models are sensitive to the parameterisation of the roughness coefficient which adds parameter uncertainty (Horritt and Bates, 2002; Tayefi *et al.*, 2007; Bates *et al.*, 2010).

3.3.3 Three-dimensional models

Three-dimensional models simulate the motion of water in all three directions and are, thus, the most complex models. These models are based on the Navier-Stokes equations to describe the 3D motion of fluid but some models simulate the lateral and longitudinal flow using the 2D SWE and implement a quasi-3D extension to simulate the vertical velocity (Teng *et al.*, 2017). In general, these models provide a better representation of the hydrodynamic processes but require higher computation resources and are limited to the application on smaller spatial and temporal scales and are, thus, out of the scope for the application in the Karnali fan (USACE, 2020).

3.3.4 The selection of a model structure

Hydrodynamic models vary in their process description and predictive capabilities. The model complexity increases with the flow dimensionality from 1-D models to 3-D models. The complexity varies within each class based on the selection of the hydrodynamic equations. The computational requirements increase with model complexity which raises the question of how much complexity is required to simulate the flow of the braided Karnali River on the alluvial fan downstream of the mountain front.

The choice of the model dimensionality is determined by the characteristics of the simulated area and event. One-dimensional models provide a sufficient approximation of the flow in single-threaded channels (Tayefi *et al.*, 2007). These models assume that water in the floodplain flows parallel to the main channel and that the importance of the lateral motion is negligible for the research question. The single flow direction is predefined by the location and orientation of the cross-section which requires knowledge about the flow paths (Teng *et al.*, 2017; USACE, 2020). The floodplain flow is, thus, overly simplified in floodplains with a strong topographic effect on the flow (Tayefi *et al.*, 2007). One-dimensional models are overly simplistic for the application in the Karnali

fan for two reasons; i) the braided channel with the multiple-pathway and multi-directional flow cannot be adequately depicted by the one-directional flow (Williams *et al.*, 2016; USACE, 2020) and; ii) the floodplain flow is not parallel to the main channel for aggradational systems where the floodwater drains away from the river (Dixit, 2009; Sinha *et al.*, 2014).

Two-dimensional models predict the motion in the longitudinal and lateral direction and can, thus, simulate complex flow paths of braided river systems (Williams *et al.*, 2016). Furthermore, these models do not require prior knowledge of the flow paths which are predicted from the topographic data instead (USACE, 2020). These models are constrained by the availability of topographic data which need to capture the flow-affecting features with sufficient detail. However, the assumption of shallow water is met for braided river systems so 2D models have a sufficient complexity to simulate the flow of the Karnali River (Williams *et al.*, 2016).

The choice of the modelling equation depends on the flow characteristics of the simulated system. The approximations of the full SWE omit the acceleration terms which are important to predict dynamic flood waves such as dam outburst floods, levee branches and flash floods. Such flood waves are characterised by large changes in the velocity in time and space which requires the computation of the local acceleration (changes in velocity over time) and the convective acceleration (changes in velocity in space) (USACE, 2023). The acceleration is also important for rivers with a very flat slope because control of gravity and friction on the flow decreases (USACE, 2020, 2023). Hence, the full SWE need to be applied to dynamic flood waves and flat river systems. Furthermore, these equations are required for the simulation of transitions between subcritical and supercritical flows such as hydraulic jumps (Neal *et al.*, 2012).

The approximations of the full SWE can be applied for gradually-varied flows such as the riverine floods of the Karnali River. Neal *et al.* (2012) compared the predictions from the full SWE, a diffusive wave approximation, and the inertia model (Bates *et al.*, 2010) within the LISFLOOD-FP framework (Bates and De Roo, 2000; Bates *et al.*, 2010). The model differences were small in terms of velocity and depth so the full SWE are overly complex to simulate gradually-varied flow. The inertia model is the fastest and, thus, suited to simulate large ensembles or large areas at high resolution as long as the friction is not very low (Manning's $n \leq 0.03$). Furthermore, the modelling complexity had a lower impact on the results than model setup decisions such as the spatial resolution and the intervals of modelling outputs (Neal *et al.*, 2012).

3.4 The simulation of geomorphic changes

The field of morphodynamics investigates the topographic evolution of different drivers (e.g. water) on different spatial and temporal scales (Syvitski *et al.*, 2010). In this thesis, the term morphodynamics is used synonymously for fluvial morphodynamics which investigates the response of the morphology to river flows and the interactions between the river bed morphology, the sediment dynamics, and the flow (Syvitski *et al.*, 2010).

3.4.1 Introduction to morphodynamics

Morphodynamics involves a range of processes that interact and operate in a heterogeneous environment which changes constantly on different spatial and temporal scales. The river bed affects the flow field by causing turbulences which determine the dynamics of the sediment transport (Papanicolaou *et al.*, 2001, 2008; Hardy, 2013). The flow alters the river bed by depositing or eroding particles which induce changes to the flow field. The sediments can be transported as suspended load, or as bedload via saltation (consecutive hops), rolling, sliding, and creeping. Hence, sediments can be moved by different processes and these processes can occur simultaneously for sediments of different sizes (Hardy, 2013). However, the relationship between flow and sediment transport of different sizes is highly non-linear because larger sediments are heavier and require more force to be moved but are also more exposed to the flow than smaller sediments (Mosselman, 2012; Hardy, 2013). The complexity of the morphodynamic system is further increased by interactions between sediment particles (Papanicolaou *et al.*, 2008).

No modelling framework exists to describe the complex morphological system for all processes in all environments satisfactorily on all spatial and temporal scales. The imperfections of current modelling approaches arise from; i) the limited process understanding; ii) the non-linearities in the flow and sediment transport relationship; iii) the complexity of positive and negative feedbacks between these two; iv) the heterogeneity of the natural environment (e.g. bedrock and alluvial rivers); v) the wide range of fluvial behaviour (e.g. meandering and braided rivers) and; vi) the wide range of spatial scales (mm – thousands of km) and temporal scales (seconds – millennia) on which the processes operate (Hardy, 2013; Nicholas, 2013a; Ancey, 2020b). This is further challenged by the poor transferability of laboratory experiments to the natural environment and the lack of natural field studies for model verification and model

validation (Syvitski *et al.*, 2010; Mosselman, 2012; Hardy, 2013; Ancey, 2020b). Consequently, it is to this point not possible to predict bedload transport better than one order of magnitude (Ancey, 2020b).

3.4.2 Concepts in morphodynamic modelling

A range of morphodynamic models exists to simulate different environments (e.g. coasts, deltas, lagoons, rivers) on different scales (e.g. catchment, reach) (Syvitski *et al.*, 2010). This review focuses on models which simulate the morphodynamic behaviour of river systems on the reach scale. These models are based on the assumption that sediment transport is correlated with fluid flow and therefore couple flow models to simulate the motion of fluids with sediment transport models (Hardy, 2013). However, the models can vary widely in their representation of the processes and the environment. For example, the models differ in the type of flow (steady flow, unsteady flow), type of transport (bedload, suspended load), type of sediment (cohesive, non-cohesive), representation of sediment sizes (average grain size, multiple grain sizes), the modelled dimensions (1-D, 2-D or 3-D), and the modelling approach (Papanicolaou *et al.*, 2008; Williams *et al.*, 2016). Morphodynamic models can use either; i) a physics-based approach (also referred to as process-based approaches) which is based on classical continuum mechanics and aims to predict the morphodynamics from solving physical equations, or; ii) a cellular approach (also known as reduced-complexity or exploratory approaches) which use simple rule-based abstractions of the governing physics to represent the processes (Nicholas *et al.*, 2006; Hardy, 2013; Williams *et al.*, 2016).

3.4.2.1 Physics-based morphodynamic models

Physics-based models aim to predict the morphodynamic system from equations describing the physics of the system. For this, the models combine a hydrodynamic flow model which is based on the equations of continuity of mass and momentum (see Section 3.3) with a sediment transport model which is based on the equation for sediment continuity (Papanicolaou *et al.*, 2008). However, despite being physics-based, these models are not fully physical by reducing the complexity both within the flow model (e.g. the roughness coefficient – Section 3.3) and the sediment transport model).

Physics-based models simulate the flow and sediment transport processes in three consecutive steps (Spasojevic and Holly, 2008; Williams *et al.*, 2016). In the first step, the flow is predicted from a hydrodynamic model which provides the boundary condition or parameterisation for the sediment transport model (Hardy, 2013). In the second step, sediment transport processes (erosion, transport, deposition) are predicted by the sediment transport model. This simulation is the most complex one because the non-linear relationship between transport rate and flow strength is highly non-linear (Ferguson and Church, 2009; Hardy, 2013; Williams *et al.*, 2016). The bathymetry is then updated based on the predicted sediment transport processes in the third step.

The sediment transport model operates on the same spatial discretisation as the hydrodynamic model (1-D, 2-D or 3-D). One-dimensional models can predict the longitudinal morphological evolution but the cross-section averaged bed shear stress, the ignoration of secondary circulation and lateral sediment movement limit their applicability for braided river systems (Williams *et al.*, 2016). Two-dimensional models predict the lateral movement of water and sediments and have been able to predict the morphology of braided rivers (Kleinhans, 2010; Nicholas, 2013b; Williams *et al.*, 2016). However, model parameterisation remains a challenge for the application in natural river systems.

Three-dimensional models simulate the vertical fluid motion including turbulences which exhibit significant effects on sediment entrainment and provide better predictions than 2-D models (Papanicolaou *et al.*, 2001, 2008; Hardy, 2013; Williams *et al.*, 2016). These models can either use an Eulerian concept as 1-D and 2-D models which focuses on specific locations through which fluids and sediment pass, or a Langangean concept which focuses on the movement of fluid and sediment parcels and enables the tracking of particles (Hardy, 2013). Three-dimensional models provide the most complex representation of the morphodynamic system and are valuable tools to improve the system understanding but the higher computational costs limit the applicability to small spatial and temporal scales (Hardy, 2013; Williams *et al.*, 2016).

One main issue of morphodynamic modelling is that the morphodynamic system is not fully understood and, thus, cannot be fully replicated (Hardy, 2013). Furthermore, the predictive capability is reduced by modelling assumptions. For example, many models assume steady flow and thus ignore the effect of turbulent flow on sediment transport (Papanicolaou *et al.*, 2008; Hardy, 2013). Many models use the mean boundary shear stress which is insufficient to replicate the complex interactions of wake decay, boundary-layer development and topographically-induced acceleration and deceleration

(Hardy, 2013). In addition, the effect of vegetation on flow and sediment transport is not adequately considered in many models (Mosselman, 2012; Nicholas, 2013a).

3.4.2.2 Cellular models

Cellular models aim to describe the complex morphological system through interactions between elements without describing these interactions with detailed physical equations. The basic principle of these models is that the landform (e.g. the river and floodplain) is represented by a lattice of cells which interact with each other and that these interactions can be described by simple rules which are abstractions of the governing physics (Nicholas *et al.*, 2006). Thus, these models are two-dimensional and based on physics but simplify the system by neglecting or parameterising physical processes that are considered less important on larger temporal and spatial scales (Nicholas, 2013a). This simplification enables the simulation of the morphodynamic evolution at larger temporal and spatial scales compared to physics-based models (Coulthard *et al.*, 2013).

Cellular models share several similarities with physics-based models. As for physics-based models, these are based on the assumption that sediment transport is correlated with the flow. Hence these models incorporate routines to predict flow which provides the boundary condition for the prediction of the sediment transport. The simulation of flow and sediment transport is underpinned by the principle of mass conservation of sediment and water (Nicholas, 2013a). Furthermore, similar equations for the prediction of bedload transport may be incorporated. However, unlike physics-based models, these models do not use hydrodynamic models but flow routing schemes which neglect the principle of the conservation of momentum (Nicholas, 2013a).

This difference in flow routing is one key difference between cellular and physics-based models. Cellular models use simple rule-based routing schemes. The flow routing is either predefined or directed to all neighbouring cells whereas the distribution of water is determined from the local bed slope and, depending on the model, the local water surface elevation (Murray and Paola, 1994; Van De Wiel *et al.*, 2007; Coulthard *et al.*, 2013; Nicholas, 2013a). This simplistic flow representation results in an overconcentration of flows in local depressions and the locally false prediction of flow convergence and divergence (Thomas and Nicholas, 2002; Nicholas *et al.*, 2006).

Cellular models have, despite the simplified rule-based abstractions of physics, been able to replicate the morphological evolution of braided river systems. Murray and Paola (1994) replicate the highly non-linear relationship of bedload and flow and capture the spatial and temporal patterns of braided rivers. Ziliani *et al.* (2013) reproduce the macroscale evolution of a braided natural river in the Italian Alps. Examples of other cellular model applications are the investigation of the impact of vegetation on channel evolution (Murray and Paola, 2003; Ziliani and Surian, 2016), the evolution of upland rivers and alluvial fans (Coulthard *et al.*, 2002), the evolution of meandering rivers (Van De Wiel *et al.*, 2007), and the influence of aggradation and degradation on the channel morphology (Nicholas and Quine, 2007).

The simplistic nature of cellular models limits their predictive capability. These models cannot improve the understanding of the morphological processes due to the abstraction of the governing physics to simple rules (Hardy, 2013). A main criticism relates to the highly simplified flow models. These may produce unrealistic routing patterns that differ from hydrodynamic flow predictions and are sensitive to the grid structure (e.g. rectangular, hexagonal) (Nicholas, 2005). The flow algorithms do not redistribute momentum so the predictions are overly sensitive to the local bed slope whereas this sensitivity increases with the grid resolution (Nicholas *et al.*, 2006; Nicholas and Quine, 2007; Nicholas, 2013a; Williams *et al.*, 2016). Williams *et al.* (2016) found that the flow algorithm by Murray and Paola (1994) overestimated the branding for low flows and underestimated it for high flows and was inferior to a 2-D physics-based model.

Cellular models capture the complex balance of positive and negative feedback which control the morphodynamic systems and provide insights into the controls of the morphodynamic system despite their reduced complexity (Murray and Paola, 1994, 2003; Nicholas *et al.*, 2006; Nicholas and Quine, 2007). The main advantages of these models are their computational efficiency which enables simulations over large spatial and temporal scales, sensitivity analyses, and ensemble modelling (Nicholas, 2005, 2013a). Furthermore, several models have implemented improved flow routing algorithms.

The CAESAR-LISFLOOD (Coulthard *et al.*, 2013) model is a hybrid model which combines a physics-based flow model (Bates *et al.*, 2010) with a cellular sediment transport model. The updated flow model tackles one of the main weaknesses of cellular models and has significantly improved the performance for the application in braided river systems (Ziliani *et al.*, 2020).

3.4.3 Bedload transport equations

Morphodynamic models incorporate sediment transport models to simulate the sediment flux for the simulated flow. Different models are used for the simulation of suspended sediment and bedload transport. This overview of sediment transport models focuses on bedload as this dominates the morphodynamic evolution of the Karnali fan (Section 2.7).

Bedload transport equations are mathematical expressions that quantify the movement of coarse sediments transported as bedload. However, bedload sediment transport is complex as it occurs on different spatial and temporal scales and involves multiple processes that interact nonlinearly (Hardy, 2013; Ancey, 2020a). Consequently, bedload transport cannot be predicted accurately, with typical uncertainty being an order of magnitude, and no universally accepted bedload transport equation applies to all natural gravel bed rivers (Wainwright *et al.*, 2015; Ancey, 2020a).

Bedload transport models relate the sediment transport rate to the hydraulic conditions (e.g. discharge, bottom shear stress, stream power) of the modelled river system (Ancey, 2020a). This relationship is typically determined from laboratory observations, field observations, or theoretical considerations and is classified into empirical, stochastic, and mechanical models (Ancey, 2020a). The bedload transport models are developed for specific conditions (e.g. grain sizes, water, and sediment discharges) and, thus, the suitability of a bedload transport model depends on the hydraulic and morphological characteristics of the modelled river system (Wilcock and Crowe, 2003; Wainwright *et al.*, 2015).

The earliest quantitative formulation of a bedload equation was developed by du Boys (1879: in Ancey, 2020a):

$$q_s = X\, T_b (T_b - T_c) \qquad \text{Eq. 3.10}$$

Where q_s is the bedload transport rate per unit width, T_b is the bottom shear stress (Pa), T_c is the critical shear stress (Pa), and X is a material coefficient ($s^2\ m^{-1}$).

Du Boys model is an empirical model that introduced the concept that bedload transport is driven by excess shear stress and occurs if the bottom shear stress exceeds this critical bed shear stress (Ancey, 2020b). This concept was later formalised into the dimensionless Shields parameter (Shields, 1936) which is still used in bedload transport models.

The bedload model developed by Meyer-Peter and Müller (1948) refines the relationship between bedload transport and the hydraulic conditions by integrating the grain size of the sediment (represented by the median grain size D_{50}) and incorporating a hiding and exposure function of the grains (Deltares, 2023). The relationship was developed from flume experiments with high sediment discharge and hence is applicable for alluvial river sections in which the D_{50} ranges between 0.4 mm and 29 mm (Schwindt *et al.*, 2023). The bedload transport is calculated from the equation (Ancey, 2020b):

$$\Phi = 8\,(\Theta - \Theta_c^*)^{3/2} \qquad \text{Eq. 3.11}$$

$$\text{With } \Theta = \frac{pgR_h i}{g(p_p - p)d_x} = \frac{R_h i}{(s-1)D_{50}} \qquad \text{Eq. 3.12}$$

Where Φ is the dimensionless bedload transport rate (-), Θ_c^* = 0.047 is the dimensional critical shear stress (-), Θ is the Shields stress (-), p is the water density (kg m^{-3}), g is the gravitational acceleration (m s^{-2}), R_h is the hydraulic radius (m), i is the bed slope (%), p_p is the particle density (kg m^{-3}), p is the water density (kg m^{-3}), D_{50} is the median grain size (m) and s is the particle-to-water-density ratio (-).

One of the most accepted approaches for simulating the bedload of sand and gravel rivers is the Wilcock and Crowe (2003) model which calculates the bedload transport for different fractions of the grain sizes. It replaces the Shields parameter with a reference shear stress at which a small but constant sediment transport occurs (Parker *et al.*, 1982b, 1982a; Wilcock and Crowe, 2003). It quantifies the hiding and exposure effects of gravel transport as a function of the sand fraction in the riverbed (Wilcock and Crowe, 2003; Schwindt *et al.*, 2023). This empirical model is based on flume experiments of a wide range of grain sizes, flow conditions, and sediment discharges and applies to sand and gravel beds. The bedload transport rate Φ_i (-) of the grain fraction *i* is calculated as (Wilcock and Crowe, 2003):

$$\Phi_i = \begin{cases} 0.002\,\phi^{7.5}, & \phi < 1.35 \\ 14\left(1 - \frac{0.894}{\phi^{0.5}}\right)^{4.5}, & \phi \geq 1.35 \end{cases} \qquad \text{Eq. 3.13}$$

$$\text{With } \phi = \frac{\tau}{\tau_{ri}}$$

Where τ is the shear stress (Pa) and τ_{ri} is the reference shear stress for fraction *i* (Pa).

This model is a surface-based model that requires no information about the subsurface grain sizes. Furthermore, it does not require the calibration to the reach-specific conditions (Wilcock and Crowe, 2003; Deltares, 2023). The separate quantification for

multiple grain size classes enables the simulation of reaches with highly variable grain size distributions (e.g. downstream fining).

3.4.4 The predictive capabilities of morphodynamic models

Morphodynamic models are challenged to predict the complex morphodynamic system of natural real-world rivers as indicated in 3.4.1. This limited predictive capability arises, to a large degree, from the sediment transport models which Hardy (2013) describes as an incomplete science because of the poor process understanding and the complexity introduced by the heterogeneity of natural environments. The fundamental issues of sediment transport modelling are introduced briefly in the following section through the example of bedload sediment transport which is the governing transport mode in braided river systems (Murray and Paola, 1994; Williams *et al.*, 2016).

The incomplete knowledge about the sediment processes is compensated for by empirical relationships which reduce the complexity of the model. Current sediment transport models are based on physics by describing mass conservation as a continuity principle for the bed morphology but lack a real physical description (Hardy, 2013). Sediment transport models commonly rely on empirical relationships which relate sediment transport to the flow but are overly simplistic to simulate the complex, dynamic interactions between the bedload, riverbed and flow (Ancey, 2020a). Sediment transport involves many interrelated processes which vary nonlinearly due to the heterogenous environment, and occur on various spatial and temporal scales(Hardy, 2013; Ancey, 2020b). This results in large variations in bedload transport rates even in controlled laboratory experiments which cannot be replicated by simple empirical relations (Ancey, 2020b). The use of empirical relationships constricts the application to systems with similar conditions (e.g. grain sizes, transport rates) as those for which the relationship was derived (Wilcock and Crowe, 2003).

The predictive capability of sediment transport processes is further induced by the inability to quantify the effect of the near-bed turbulence. The near-bed turbulence exhibits a strong effect on sediment transport (Papanicolaou *et al.*, 2001, 2008). This turbulence is omitted in the steady state assumption of many flow models (Hardy, 2013). Furthermore, the use of velocity profiles for the determination of mean boundary shear stress leads to an inaccurate description of the effect of the bed roughness on the flow characteristics (Hardy, 2013).

Morphodynamic models combine flow and sediment transport models so that flow and sediment transport processes are simulated separately. This structure of separately simulated flow and sediment transport implies that the interactions of the riverbed and the flow field can not be predicted but only be approximated (Papanicolaou *et al.*, 2008).

Recent developments have improved the physical foundation of morphodynamic models (e.g. by improving the description of turbulent forces). In particular, Lagrangian schemes which enable the tracing of particles are promising tools to deepen the process understanding and improve the parameterisation of less complex models (Hardy, 2013). However, the fact that bedload transport cannot be predicted better than within one order of magnitude underlines the current limitations of morphological modelling (Ancey, 2020b).

3.4.4 The simulation of natural systems

The previous section showed that the capabilities of morphological models for the prediction of the morphological evolution of rivers are limited due to the limited process understanding, the complexity of the system, and structural model deficits. However, all environmental models are simplifications of the modelled system and reduce the process complexity by using parameters, and average processes in time and space. The lower predictive capability of morphodynamic models also arises from the lack of observations on adequate temporal and spatial scales that can be used to determine the boundary conditions, and to verify (the correct solution of the equations), calibrate (the identification of parameter values), and validate (the evaluation of the performance of the calibrated model) the model.

Morphodynamic models are sensitive towards the boundary conditions which are difficult to determine on adequate temporal and spatial scales. The required data comprises the topography, the inflow of water and sediments to the modelling domain, and the grain sizes of the sediment inflow and the riverbed. These datasets are scarce in most river systems so the determination of the boundary conditions is a common problem for the application of morphodynamic models in natural river systems (Papanicolaou *et al.*, 2008).

This data scarcity results from measurement difficulties, the involved spatial and temporal scales, and the heterogeneous environment. For example, bedload transport rates vary greatly in time and hence the measured rate is highly sensitive towards the measurement interval, which is further aggravated by the technical difficulties of

measuring bedload, and the lack of a universally accepted definition of the bedload transport rate (Ancey and Pascal, 2020).

Topographic data can be obtained at high resolution from remote sensing techniques but such datasets lack the description of the channels which is essential for morphodynamic modelling (Ziliani *et al.*, 2013). The field measurements of the riverbed properties (e.g. topography, surface and subsurface grain size distributions) are only applicable at limited spatial and temporal resolution (e.g. cross-sections) and need to be extrapolated to the unsampled areas (Papanicolaou *et al.*, 2008; Ziliani *et al.*, 2013). The issues of extrapolating spatially limited measurements are illustrated by the large variation of sediment grain sizes in the lateral (e.g. riverbed and river bank) and longitudinal directions (upstream, downstream) (Quick *et al.*, 2019; Dingle *et al.*, 2020b).

These challenges to measuring data at sufficient temporal and spatial scales also account for datasets which can be used for model calibration and validation. Examples of useful datasets to assess and optimise the model performance are sediment flux at the downstream outlet or repetitive topographic measurements (Ziliani *et al.*, 2013, 2020; Williams *et al.*, 2016). These measurements are commonly constrained to shorter periods (events – years) which stands in contrast to the millennia scale on which some morphodynamic processes occur (Nicholas, 2005).

The model calibration and validation are further complicated by the incomplete process description and the model's sensitivity towards the boundary conditions. Even for the hypothetical case that an ideal dataset exists it could not be used for a direct (e.g. pixel-to-pixel) calibration and validation. Model deficits and the uncertainty in the boundary condition cause deviations between the simulated and observed changes even if the model is a behavioural representation of the system (Nicholas, 2005; Mosselman, 2012; Hardy, 2013; Williams *et al.*, 2016). In other words, the model may predict the correct morphodynamic patterns (e.g. development of bars and pools) but not at the observed locations which complicates the model validation.

Most studies use indirect comparisons or a combination of direct and indirect comparisons of simulated and observed or simulated by a more complex model for the calibration and validation of the model performance. Direct comparisons include pixel-to-pixel performance indices (e.g. wet area, vegetated area, active channel area, morphological change, discharge per unit width) or sediment discharge at the outlet of the modelling domain (Thomas and Nicholas, 2002; Nicholas *et al.*, 2006; Ziliani *et al.*, 2013, 2020; Williams *et al.*, 2016). Indirect comparisons evaluate morphological relationships (e.g. bar perimeter and bar area, the width-depth ratio of channels),

channel characteristics (e.g. mean active channel width, mean wet area width, mean braided index) or total changes (e.g. erosion and deposition) (Nicholas *et al.*, 2006; Kleinhans *et al.*, 2008; Ziliani *et al.*, 2013, 2020; Williams *et al.*, 2016).

The data scarcity in conjunction with the sensitivity towards the boundary conditions, the non-linearity in the system, and the discrepancy between simulated and observed behaviour are fundamental challenges because they limit the model verification. This means that at the current stage, it is even a challenge to test whether a model solves the equations correctly. Mosselman (2012) suggests the establishment of hypothetical, laboratory and field cases that can be used to verify models in a more standardized and comparable way.

3.4.5 Summary

Stage 1 (Chapter 4) establishes a hydrological model to replicate the hydrological system of the Karnali River for the observed climatic and hydrological conditions which is the foundation for the quantification of the climate change impact on flood hazards in Stage 2. Empirical models are unsuitable for predicting changes because these are only applicable under stable conditions so either a physics-based or conceptual model needs to be applied. These models must be calibrated to the specific conditions of the catchment which requires information on the catchment behaviour, and performance indices to evaluate the fit between simulated and observed behaviour, whereas a combination of indices provides a more robust evaluation. Uncertainty in hydrological modelling arises from the simplification of the complex hydrological system in the model structure (structural uncertainty), errors in the calibration data which, for catchment discharge, can be quantified from stage-discharge rating curves (measurement uncertainty), and parameter uncertainty resulting from the imperfect parametrisation of the model. The uncertainty can be estimated using the GLUE framework which maintains all behavioural parameter sets and, thus, provides an ensemble of predictions (Section 3.1).

In Stage 2 (Chapter 5) probabilistic climate projections are applied to the hydrological model to predict the potential flood discharge for different climate scenarios. The change in the projected flood hazards can be quantified from a Flood Frequency Analysis (FFA) which predicts the flood magnitudes and frequencies from a flood record. This flood record can be obtained from the flow record using the Annual Maximum Flow model (AMAX) which does not require information about flood thresholds (bankfull discharge

or stage). The optimal Extreme Value Distribution (EVD) of the FFA is unknown before fitting the distributions to the flood record so multiple EVD should be evaluated using test statistics. The Anderson-Darling test evaluates the fit between the observed and simulated distributions with a focus on the tails of the distributions which affect the performance of floods with higher magnitude. The main sources of uncertainty in the FFA are the sampling error which can be quantified from MCS sampling or bootstrapping approaches, the measurement error which can be estimated by using an ensemble of flood records, and the non-stationarity of the data which relates to changes in the hydrological state of the catchment (Section 3.2).

Stage 3 (Chapter 6) applies a morphodynamic model to predict the fan evolution. Fluvial morphodynamics describes the complex system of flow and topography with positive and negative feedback and non-linear interactions on different spatial and temporal scales. Different approaches simulate this system with varying complexity. Physical-based models have a stronger physical basis but are computationally expensive and restricted to applications on smaller temporal and spatial scales. Cellular models are applicable on larger temporal and spatial scales but are based on weaker process descriptions. However, all models are limited in their predictive capabilities which results from the incomplete knowledge of the morphodynamic system, but also of the lack of data for the model verification, calibration, validation, and for the definition of the boundary conditions. Hence, the uncertainty of the prediction of the morphological evolution of fluvial systems arises from incomplete knowledge about the system as such, and the incomplete knowledge of the modelled environment (Section 3.4).

In Stage 4 (Chapter 7), a hydrodynamic model is applied to predict the spatial flood hazard patterns for the projected topographies and flood magnitudes. Hydrodynamic models differ in their spatial representation of the flow processes from 1-D (longitudinal flow), 2-D (longitudinal and lateral flow), and 3-D (longitudinal, lateral and vertical flow). The simulation of lateral flow is required for the simulation of braided river systems which restricts the model choice to 2-D models (3-D models are too complex for the spatial and temporal scales). Two-dimensional models differ in their process description whereas lower complexity models are not less accurate but restricted to fewer flow conditions (Section 3.3).

4 The simulation of the catchment hydrology of the Karnali River

This chapter presents the first stage of the environmental modelling framework which is the reproduction of the hydrological system of the mountainous catchment up to the mountain outlet from hydrological modelling to meet the first objective O1. The established modelling ensemble is the foundation to predict the potential flood discharge, flood magnitudes and frequencies for the projected climates in Stage 2 (Chapter 5).

The catchment hydrology determines the flood water that is released into the Karnali fan. The flood hydrographs are controlled by the interaction of rainfall pattern and the catchment topographic form, and by the antecedent hydrological conditions in the catchment (e.g. soil moisture) (Berthet *et al.*, 2009; Beven, 2012; Nathan and Ball, 2019; Reaney, 2022). While riverine floods in Central Himalayan catchments are caused by heavy and large-scale rainfall events, baseflow, snowmelt and glacier melt also contribute water which has to be predicted by the hydrological model (Lutz *et al.*, 2014; Nepal *et al.*, 2014; MacClune *et al.*, 2015; Wijngaard *et al.*, 2017; DHM, 2018).

This chapter presents an introduction to assessed modelling approaches (Section 4.1), the methodology of the selected modelling approaches (Section 4.2), the utilised datasets (Section 4.3), the modelling results (Section 4.4) and the discussion of these results (Section 4.5). The final section concludes this chapter.

4.1 The model selection

Hydrological simulation models represent the flows of water within the landscape and river system to predict the catchment's behaviour. It is, therefore, possible to apply such models with data of climate change projections to predict the future streamflow. Hydrological models differ in their process representation (empirical, conceptual, physics-based), spatial domain (lumped, distributed), time domain (event-based, continuous), data requirements, and parametrisation (Devi *et al.*, 2015). For this study, the model selection is based on the following criteria:

i) A physical process representation is required since empirical models are based on the assumption of stationary (environmental conditions are stable) which excludes the impact assessment of changing climatic conditions (Pechlivanidis *et al.*, 2011);
ii) A continuous model is preferred because event-based models are sensitive to the antecedent conditions which are unknown for the projected flood events in stage 2 (USDA, 2014; Nathan and Ball, 2019);
iii) The ability to simulate glacier melt and snowmelt since these are important runoff components that are significantly affected by climate change (Immerzeel *et al.*, 2012; Lutz *et al.*, 2014; Nepal, 2016);
iv) Low data requirement as the location is data poor and the definition of the boundary conditions is dependent on satellite imagery;
v) Computational efficiency to predict the hydrological uncertainty and simulate climate ensembles in Stage 2.

Several continuous models were assessed against these criteria. The Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström, 1976; Lindström *et al.*, 1997) is a conceptual model with moderate data requirements (e.g. daily temperature, precipitation, potential evapotranspiration) that includes algorithms to simulate glacier melt processes (Seibert *et al.*, 2018). This model has a coarse spatial resolution that does not capture the small-scale heterogeneity of the environmental conditions in the catchment and is therefore disregarded.

The Soil and Water Assessment Tool (SWAT) (Arnold *et al.*, 1993) and J2000 (Krause, 2001) models aggregate areas with similar environmental conditions (e.g. elevation, slope, aspect, soil, land cover, etc.) to Hydrological Response Units (HRU) enabling a more realistic representation of the spatial variability. Both models include snowmelt and glacier melt routines using a degree-day-factor (DDF) (Nepal, 2012; Omani *et al.*, 2017).

The glacier representation in both models assumes an infinite glacier storage which can result in the overprediction of glacier melt for glaciers that are projected to decrease or vanish during the study period (Immerzeel *et al.*, 2012, 2013).

The Spatial Processes in HYdrology (SPHY) model (Terink *et al.*, 2015a) is selected because it is characterized by moderate data requirements and an advanced description of snowmelt and glacier melt processes, matching the criteria set out above. The decisive argument for SPHY was the distributed representation of the environmental conditions to capture the heterogeneity of Himalayan catchments, and the model's ability to simulate the retreat of glaciers (Khanal *et al.*, 2021). This representation is an important asset compared to models with static glacier extent since the glacier melt depends on the glacier mass (increasing melt for decreasing mass) and hence the importance of the glacier dynamics increases with the length of the modelled period (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014). Furthermore, it is an open-source model that can be adapted to the catchment and objective.

4.2 Methods

The hydrological model SPHY is a conceptual model that simulates the hydrological processes as a series of storages considering the underlying physics. This model is introduced in the next section. The methods to calibrate the model to the hydrology of the Karnali catchment are presented in Section 4.2.2. This calibration aims to maximise the agreement between the simulated and observed catchment behaviour. The discharge at the mountain outlet is an essential indicator of the catchment behaviour because it is the sum of all upstream processes. However, discharge records may contain substantial uncertainty which propagates to the model during the calibration. Therefore, the discharge uncertainty is estimated and the method for this estimation is presented in Section 4.2.3.

4.2.1 Spatial Processes in Hydrology (SPHY)

SPHY (Terink *et al.*, 2015a) is a raster-based, conceptual cryospheric-hydrological model that simulates the hydrological processes at daily resolution. It is written in the Python programming language and uses the PCRaster modelling framework (Karssenberg *et al.*, 2001; Karssenberg, 2002). The catchment's hydrology is simulated conceptually as a series of water fluxes and storages which enables the computationally efficient quantification of hydrological processes (Devi *et al.*, 2015). The fluxes and storage changes are simulated as averages of the cells of the modelling grid. A sub-grid routine is implemented for glacial processes; grid cells can be fully or partly covered by glaciers and the melt, runoff and percolation processes are only simulated for the glaciated cell fraction.

4.2.1.1 Model structure

The model is structured as a collection of storages that store the water and fluxes that connect these storages. The conceptual layout of the model structure is presented in Figure 4.1. Precipitation is separated into snow and rainfall based on a temperature threshold parameter (*tcrit*); if the temperature in a cell is below the threshold all the precipitation falls as snow, if not as rain. Precipitation (rain and snow) that falls in cells covered by snow or glaciers is added to the snow or glacier storage. If the cell is covered by vegetation, the precipitation is added to the interception storage where it is either

stored until the next time step or lost as evapotranspiration. Surplus rainfall that exceeds the interception storage capacity (effective precipitation) is passed on to the soil storage.

The soil storage is divided into three separate layers that represent the root zone, subzone and groundwater zone. The effective precipitation can either infiltrate the root zone or is routed downstream as surface runoff. The water in the root zone can be lost as evapotranspiration, routed downstream as lateral flow or percolate into the subzone where it either percolates further to the groundwater layer or rises back up to the root zone. Glaciated areas have no root zone and subzone layers so meltwater drains as runoff or percolates to the groundwater layer.

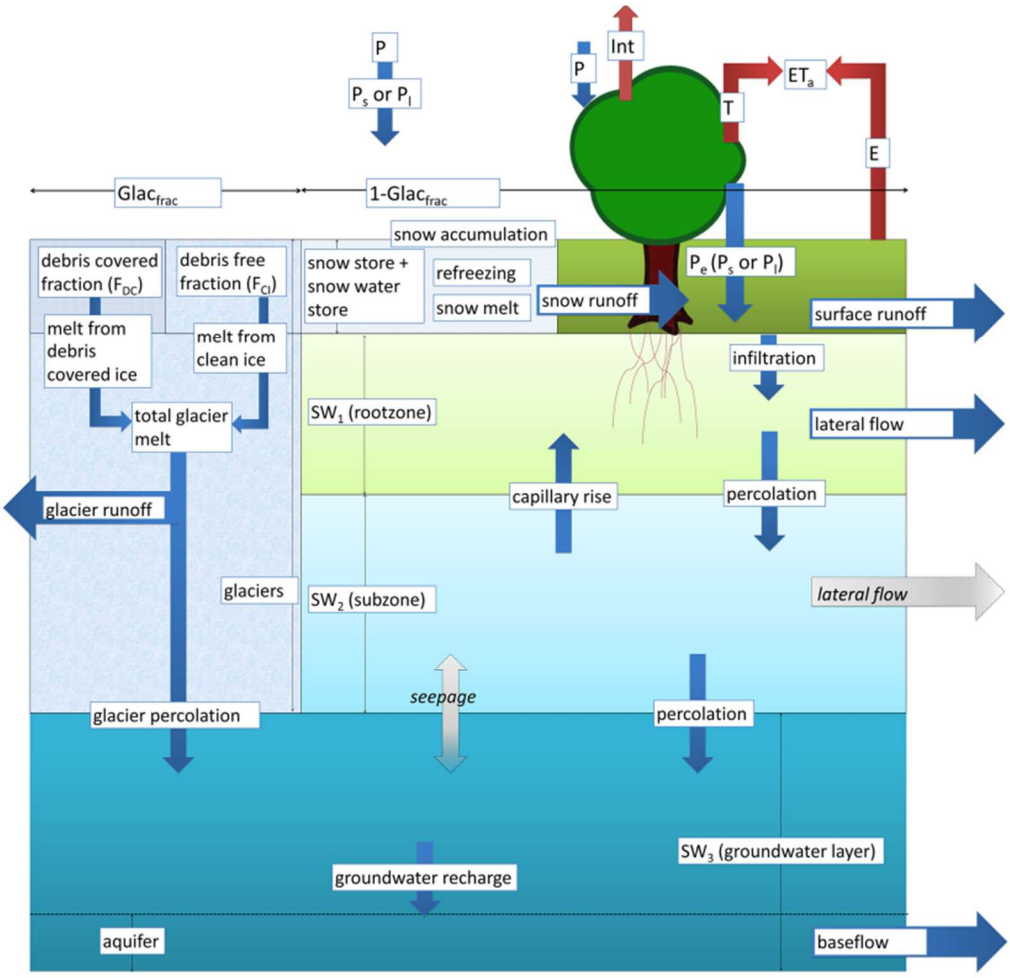


Figure 4.1: The conceptual layout of the SPHY model. The grey fluxes are only simulated if the groundwater module is not used. P is the precipitation; P_s is the snowfall; P_l, is the rainfall; P_e, is the effective precipitation; Int is the intercepted precipitation, T is temperature; ET_a, is the actual evapotranspiration; $Glac_{frac}$ is the glaciated fraction of the cell; SW_1 is the upper soil water storage; SW_2 is the lower soil water storage, and SW_3 is the groundwater layer. This figure is taken from Terink et al. *(2015, Figure 2).*

The model distinguishes four runoff components that differ in their origin, path and timing. The fastest runoff component is the rainfall-runoff (surface runoff, lateral flow). The slowest runoff is the baseflow, which ensures river flows during dry seasons. Snow and glacier runoff originate from melt processes, whereas glacier meltwater does not require precipitation.

4.2.1.2 The process representation in the model

This section describes the equations that represent the hydrological processes and calculate the fluxes between the storages and the water that leaves the model. The model uses parameters that require the adjustment to the catchment-specific conditions and are highlighted in italic.

Evapotranspiration

SPHY calculates three evapotranspiration rates to simulate how much water is lost from the system by evapotranspiration (Terink *et al.*, 2015a).

The reference evapotranspiration (ET_{ref}) describes the evapotranspiration of the climatic conditions that would occur on a reference surface with an unlimited water supply (Allen *et al.*, 1998). This rate is calculated by the modified Hargreaves equation (Hargreaves and Samani, 1985):

$$ET_{ref} = 0.0023 * 0.408 * Ra \left(T_{avg} + 17.8\right) * TD^{0.5} \quad \text{Eq. 4.1}$$

where T_{avg} is the average daily temperature (°C), Ra is the extraterrestrial radiation (MJ m^{-2} day^{-1}) which is calculated from the day of the year and the latitude, and the daily temperature range TD (°C).

The potential evapotranspiration (ET_{pot}) describes the evapotranspiration rate for a specific vegetation type under ideal conditions (e.g. no water stress, pests, diseases, etc.). In SPHY, ET_{pot} is estimated by introducing a crop coefficient (*Kc*) after Allen *et al.*, (1998) which determines the difference in the evapotranspiration rates between the specific vegetation and the ET_{ref} (Terink *et al.*, 2015a):

$$ET_{pot} = ET_{ref} * Kc \quad \text{Eq. 4.2}$$

In most cases, the actual evapotranspiration (ET_{act}) is lower than ET_{pot} because several limiting factors reduce the evapotranspiration rate. SPHY considers the reduction of

evapotranspiration rates that are caused by water excess and water shortage stresses by introducing a reduction parameter that is determined by the water content in the root zone layer (Terink *et al.*, 2015a):

$$ET_{act} = ET_{pot} * ET_{wet} * ET_{dry} \quad \text{Eq. 4.3}$$

The reduction parameter (ET_{wet}) can only be set to 0 and 1; if the soil is saturated, the parameter is set to 0 and no water is lost to evapotranspiration. The ET_{dry} parameter describes the reduction in evapotranspiration caused by water shortage and can take any value from 0-1. It is calculated using the Feddes equation (Feddes et al., 1978 in: Terink *et al.*, 2015) which assumes a linear decline in rootwater uptake if water pressure drops below a critical value (pF):

$$ET_{dry,t} = \frac{SW_{1,t} - SW_{1,pF4.2}}{SW_{1,pF3} - SW_{1,pF4.2}} \quad \text{Eq.4.4}$$

The critical value (pF) can be obtained from the soil water retention curve (pF curve) and is a conversion of the suction force (H) (Terink *et al.*, 2015a):

$$pF = log_{10}(-H), \quad \text{Eq. 4.5}$$

where $pF_{4.2}$ represents the critical value at the permanent wilting point and pF_3 is the critical value at the start of the linear decline in rootwater uptake (Terink *et al.*, 2015a).

Snowmelt

Snow processes are simulated by implementing the dynamic snow storage model by (Kokkonen *et al.*, 2006) using a degree-day-factor approach (Hock, 2003). The snow storage is adjusted for each time step by simulating melt, refreeze and runoff processes (Terink *et al.*, 2015a).

The precipitation either falls entirely as snow or rain depending on the air temperature at a particular time step and grid cell. Precipitation falls as snow and is added to the snow storage if the air temperature (T_{avg}) is below a calibrated threshold (*tcrit)* (Terink *et al.*, 2015a).

The snowmelt is temperature-dependent and occurs if the air temperature exceeds 0°. In this case, the potential snowmelt is the product of the air temperature and a calibrated degree-day factor (DDF) s*DDF*. The actual snowmelt of day t ($A_{act,t}$) equals the potential snowmelt but the upper limit is the snow storage of the previous day (SS_{t-1}). The snow storage of the current day (SS_t) is then updated (Terink *et al.*, 2015a):

$$SS_t = \begin{cases} SS_{t-1} + P_{s,t} + SSW_{t-1}, & T_{avg,t} < 0 \\ SS_{t-1} + P_{s,t} - A_{act,t}, & T_{avg,t} \geq 0 \end{cases}$$ Eq. 4.6

where SSW_{t-1} is the meltwater stored in the snowpack from the previous time step. The amount of liquid water stored in the snowpack depends on a calibrated storage capacity parameter (*snowsc)* that describes the total water equivalent of snowmelt (mm) that can freeze per mm water equivalent of snow. The liquid water content of the snowpack is calculated using equations 4.7 and 4.8 (Terink *et al.*, 2015a):

$$SS_{Wmax} = SSC * SS_t$$ Eq. 4.7

$$SSW_t = \begin{cases} SSW_{t-1} + P_{l,t} + A_{act,t}, & T_{avg,t} \geq 0 \\ 0, & T_{avg,t} < 0 \end{cases}$$ Eq. 4.8

where $P_{l,t}$ is the precipitation that falls as rain. If the temperatures are below 0°, SSW_t is zero.

Runoff from snowmelt (SRo) occurs when T_{avg} is above 0 and no more meltwater is available to refreeze (Terink *et al.*, 2015a):

$$SRo_t = \begin{cases} A_{act,t} + P_{l,t} - \Delta SSW, & T_{avg,t} \geq 0 \\ 0, & T_{avg,t} < 0 \end{cases}$$ Eq. 4.9

Glacier melt

The glacier melt processes are, like the snow melt processes, simulated using the DDF approach, whereas different factors for glaciers with and without debris cover can be provided. The daily melt (A_{gl}) for a specific glacier type (clean ice, debris-covered) depends on the air temperature, the type-specific degree day factor (for clean ice glaciers and $gDDF_{db}$ for debris-covered glaciers) (°C) and the fraction of the cell that is covered by glaciers ($Glac_{frac}$) (Terink *et al.*, 2015a):

$$A_{gl} = \begin{cases} T_{avg,t} * DDF_{gl} * Glac_{frac}, & T_{avg,t} \geq 0 \\ 0, & T_{avg,t} < 0 \end{cases}$$ Eq. 4.10

The glacier meltwater drains the glaciers either as surface runoff or percolates to the groundwater. The path of the meltwater is defined by the *glacF* parameter which represents the fraction of water that drains as surface runoff on day t ($G_{ro,t}$). The remaining fraction of the meltwater ($G_{perc,t}$) is added to the groundwater layer in the soil module (Terink *et al.*, 2015a):

$$G_{ro,t} = A_{gl} * glacF \quad \text{Eq. 4.11}$$

$$G_{perc,t} = A_{gl} * (1 - GlacROF) \quad \text{Eq. 4.12}$$

The glacier evolution is simulated for each glacier by using a mass-conserving ice redistribution approach (Khanal *et al.*, 2021). This redistribution is based on the mass balance of the individual glaciers (I) calculated from the total accumulated snow (SnowS) and the total melt generated from the glaciers (GM) at the end of the hydrological year (Khanal *et al.*, 2021):

$$I_{n,j} = SnowS_{n,j} - GM_{n,j} \quad \text{Eq. 4.13}$$

Where n and j are the glacier and the unique grid cell identification.

The accumulated snow in the accumulation zone is transformed into ice and redistributed to the ablation zone if the total melt exceeds the snow accumulation ($I_{n,j} < 0$) (Khanal *et al.*, 2021):

$$Vred_{n,j} = \begin{cases} 0 & , j \epsilon B_{n,j} \\ \sum_{j \epsilon B_{n,j}} I_{n,j} * \frac{Vini_{n,j}}{\sum_{j \epsilon A_{n,j}} Vini_{n,j}} & , j \epsilon A_{n,j} \end{cases} \quad \text{Eq. 4.14}$$

Where, A is the part of the glacier with a negative imbalance, and B is the one with a positive imbalance. Vini is the initial total volume of ice and hence the redistribution scales with the initial ice volume (Khanal et al., 2021).

Soil water processes

The soil module is the most complex in terms of the pathways that water can take and greatly affects the shape and timing of the simulated hydrograph (Terink *et al.*, 2015a). The upper layer SW_1 (root zone) interacts with the atmosphere, receiving water from precipitation or delivering water as evapotranspiration. The water is either stored in the upper layer, percolates to the lower layer SW_2 (subzone) or is routed downstream as overland flow or lateral flow. In the second layer SW_2, water can either rise back up to SW_1 through the capillary rise, or percolate further to the groundwater layer SW_3, but no water is lost to evapotranspiration or lateral flow. The capillary rise is limited by a calibration parameter (*capRise*). The groundwater layer SW_3 is fed by glacier melt contributions and water percolating from SW_2. The water is either stored in the layer or drains the cell as baseflow, but no water can rise back up to the second layer SW_2 (Terink *et al.*, 2015a).

The percolation from SW_1 to SW_2 is calculated following a storage routing methodology that is also implemented in the SWAT model (Neitsch *et al.*, 2011). Percolation only occurs if the lower layer is not saturated and only the surplus water $W_{l,exec}$ that exceeds the field capacity is available for percolation. However, only a fraction $W_{l,perc}$ of the surplus water percolates depending on the percolation travel time $TT_{perc,\,l}$ (days) which depends on the saturated hydraulic conductivity $K_{sat,l}$ (mm day^{-1}), the field capacity $SW_{l,fc}$ (mm) and the saturated water content $SW_{l,sat}$ (mm) (Terink *et al.*, 2015a):

$$W_{l,perc} = W_{l,exec} * (1 - exp\frac{-1}{TTperc,l}). \qquad \text{Eq. 4.15}$$

$$TT_{perc,l} = \frac{SW_{l,sat} - SW_{l,fc}}{K_{sat,l}} \quad . \qquad \text{Eq. 4.16}$$

This routine is implemented to calculate the amount of water that percolates to the lower layer. Water that percolates to the groundwater layer SW_3 will eventually reach the shallow aquifer from where it drains as baseflow. The groundwater recharge rate of day t is calculated using an exponential decay weighing function (Venetis, 1969 in Terink *et al.*, 2015) to account for the slower nature of groundwater recharge (Terink *et al.*, 2015a):

$$\mathrm{Gchrg_t} = \left(1 - \exp^{\frac{-1}{\delta \mathrm{gw}}}\right) * \mathrm{W_{2,perc}} + \exp^{\frac{-1}{\delta \mathrm{gw}}} * \mathrm{Gchrg_{t-1}}, \qquad \text{Eq. 4.17}$$

where δgw is the delay time (d) which can be treated as a calibration parameter (*deltaGW*), $\mathrm{Gchrg_{t-1}}$ the groundwater recharge on the previous day (mm) and $\mathrm{W_{2,perc}}$ (mm) the amount of water that percolates from SW_2 to SW_3 (mm).

Runoff components

The overland flow RO comprises saturation excess flow (RO_{sat}) and infiltration excess flow (RO_{inf}). The saturation excess flow begins if the water stored in SW_1 exceeds the saturated water content of the layer $SW_{1,sat}$ (mm) A second routine enables the simulation of infiltration excess runoff if the infiltration capacity Inf_{cap} is exceeded. For this, three calibration parameters are required that describe i) the rainfall fraction of daily rainfall that occurs during the hour of the most intense rainfall (*alphaInf*); ii) a coefficient that determines the reduction speed of the infiltration capacity ($labda$) and; iii) a factor which the effective saturated hydraulic conductivity *kEff* (Terink *et al.*, 2015a):

$$RO = RO_{sat} + RO_{inf} \qquad \text{Eq. 4.18}$$

$$RO_{sat} = \begin{cases} SW_1 - SW_{1,sat} & if\ SW_1 > SW_{1,sat} \\ 0 & if\ SW_1 \leq SW_{1,sat} \end{cases} \qquad \text{Eq. 4.19}$$

$$RO_{inf} = \begin{cases} \frac{P_l - (alphaInf * P_l - labda)^2}{(alphaInf * P_l)} & if\ Inf_{cap} < \text{alphaInf} * P_l \\ 0 & if\ Inf_{cap} \geq \text{alphaInf} * P_l \end{cases}$$ Eq. 4.20

$$Inf_{cap} = \frac{kEff * Ksat_{root}}{24} * (1 + \frac{SW_{1,sat} - SW_1}{SW_{1,sat}})^{\lambda},$$ Eq. 4.21

where $Ksat_{root}$ describes the saturated hydraulic conductivity of SW_1.

The lateral flow of each cell ($LF_{l,\ hill}$) (mm) in the root zone SW_1 and subzone SW_2 is calculated according to Sloan and Moore (1984 in: Terink *et al.*, 2015) as a function of available water and flow velocity. The amount of water available for drainage is determined by the soil properties and only the excess water ($W_{l,exec}$) (mm) exceeding the field capacity is available for lateral drainage (Terink *et al.*, 2015a):

$$LF_{l,hill} = W_{l,excfrac} * v_{lat,l}$$ Eq. 4.22

$$W_{l,execfrac} = \frac{W_{l,exec}}{SW_{l,sat} - SW_{l,fc}}$$ Eq. 4.23

$$v_{lat,l} = K_{sat,l} * slp,$$ Eq. 4.24

where $K_{sat,l}$ is the saturated hydraulic conductivity of soil layer *l* (SW_1, SW_2), $SW_{l,\ fc}$ is the field capacity of soil layer *l,* and *slp* is the slope (increase in elevation per unit distance) of the hill. The units of $v_{lat,l}$ and $K_{sat,l}$ are mm day^{-1}.

The lateral flow is slower than surface runoff and only a fraction of the available water drains on the same day. The lateral flow entering the channel on day t (LF_l) (mm) is calculated based on Neitsch *et al.* (2011) using the travel time ($TT_{lag,l}$) (days), which depends on the field capacity ($SW_{l,fc}$) (mm), the saturated water content ($SW_{l,sat}$) (mm) and the hydraulic conductivity of the soil ($K_{sat,l}$) (mm day^{-1}) (Terink *et al.*, 2015a):

$$LF_l = \left(LF_{l,hill} + LF_{l,hill,t-1}\right) * \left(1 - \exp\left[\frac{-1}{TT_{lag,l}}\right]\right)$$ Eq. 4.25

$$TT_{lag,l} = \frac{SW_{l,sat} - SW_{l,fc}}{K_{sat,l}}.$$ Eq. 4.26

The baseflow into the channel on day t $BF_{3,t}$ (mm) from SW_3 is calculated based on the steady-state response of groundwater flow to recharge and water table fluctuations which result in a non-steady response of groundwater to periodic groundwater recharge after Smedema & Rycroft (1983 in Terink *et al.*, 2015). Two parameters are implemented to adjust the baseflow simulation to the catchment characteristics; i) a threshold (*baseThr*) (mm) below which no baseflow occurs and; ii) a recession coefficient (*alphaGW*) (-) that relates the baseflow response to changes in groundwater recharge

and determines the speed with which the baseflow drains into the channels (Terink *et al.*, 2015a):

$$BF_{3,t} = \begin{cases} 0 & if\ SW_3 \leq baseThr \\ BF_{t-1} * \ exp^{-alphaGW} + Gchrg_t * \left(1 - \ exp^{alphaGW}\right) & if\ SW_3 \ > baseThr \end{cases}$$ Eq. 4.27

Routing

The cell-specific total runoff (QTot) (mm) is the sum of the rainfall-runoff (RRo) (mm) (surface runoff Ro + lateral flow LF_1 from SW_1), snow runoff (SRo) (mm), glacier runoff (Gro) (mm), and baseflow (BF) (baseflow BF_3 from SW_3) (mm).

This cell-specific total runoff at day t ($QTot_t$) (mm) is routed through a simple routing scheme to obtain the river discharge ($Qrout_t$) ($m^3\ s^{-1}$). It is routed through the downstream cells to the catchment outlet based on the flow direction (F_{dir}) derived from a DEM using the D8 algorithm (O'Callaghan and Mark, 1984). The cell-specific runoff is accumulated for all upstream cells and the travel time of the water is simulated using a recession coefficient *recessCoef* that delays the runoff which would otherwise reach the outlet within one day (Terink *et al.*, 2015a):

$$QTot_t^* = \frac{QTot_t * 0.001 * A}{24 * 360}$$ Eq. 4.28

$$Q_{accu,t} = accuflux(F_{dir}, QTot_t^*)$$ Eq. 4.29

$$Q_{rout,t} = (1 - kxrecessCoef) * Q_{accu,t} + recessCoef * \ Q_{rout,t-1},$$ Eq. 4.30

where A represents the grid cell area (m^2).

This approach is an efficient way to simulate water transport in channels without relying on solving the St. Venant or Manning equations which have higher demands towards data and computational resources (Terink *et al.*, 2015a).

Precipitation correction

A precipitation correction factor is introduced which scales the daily precipitation input to account for biases in the precipitation data (Lutz *et al.*, 2014; Nepal *et al.*, 2014). The precipitation correction is used as a calibration parameter (*precipFactor*).

4.2.2 Model calibration

The model calibration is conducted to fit the model to the conditions of the Karnali catchment by identifying the model parameter combinations that result in the best match of simulated and observed catchment behaviour. A Regional Sensitivity Analysis (RSA) is conducted in the first stage to identify insensitive parameters. These parameters are removed from the further calibration process which reduces the required sample size of the calibration. The calibration is conducted within a GLUE framework to account for model equifinality and estimate the uncertainty in the model predictions.

4.2.2.1 Regional Sensitivity Analysis

In the first stage of the calibration process, the model is run with 1,500 different combinations of 21 parameters sampled using Latin Hypercube Sampling (LHS). An overview of these parameters is presented in Table 4.1. The Regional Sensitivity Analysis (RSA) is conducted for multiple performance measures that evaluate different aspects of the hydrograph, namely:

i) the Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) of the discharge at the catchment outlet because of its emphasis on high flows (Gupta *et al.*, 2009);
ii) the coefficient of determination (R^2) to evaluate the correlation between simulated and observed discharge;
iii) the percentage BIAS (PBIAS) of the simulated and observed discharge ($PBIAS_{runoff}$);
iv) the PBIAS of the discharge during the winter months ($PBIAS_{winter}$);
v) the PBIAS of the simulated and reference snow extent ($PBIAS_{snow}$);
vi) the PBIAS of the simulated and reference annual actual evapotranspiration ($PBIAS_{actET}$).

The performance measure $PBIAS_{winter}$ was introduced because it was observed that many parameter sets with high efficiencies vastly underestimate the low flows during the winter months which indicates a poor representation of the catchment's hydrology. All parameters that are classified as sensitive for either of the performance measures are maintained in the further calibration process. A single value is used for the insensitive parameters. This value is obtained from the parameter set with the highest NSE.

The actual evapotranspiration is calibrated using the scaling factor *kcFactor*. The kc values of the individual factors are determined from the comparison with remote sensing estimates in initial model runs. These kc values are then multiplied with the *kcFactor* in the calibration.

Table 4.1: The parameter and their range used for the Regional Sensitivity Analysis (RSA).

Module	Parameter	Description	Unit	min	max
Glacier	gDDF_cl	Degree-day factor debris-free glaciers	mm ◦C^{-1} day^{-1}	0.5	9
	gDDF_db	Degree-day factor debris-covered glaciers	mm ◦C^{-1} day^{-1}	0.5	9
	glacF	Fraction of glacier melt that becomes glacier runoff	-	0.2	0.7
Groundwater	alphaGW	Baseflow recession coefficient	day^{-1}	0.01	0.07
	deltaGW	Groundwater recharge delay time	day	15	50
	gwSat	Saturated soil water content of the groundwater layer	mm	250	600
	h_gw	Initial groundwater table height	m	0.8	2.2
	yieldGW	Specific yield of the groundwater storage	m/m	0.1	0.5
	baseThr	The threshold for baseflow to occur	mm	50	250
Routing	recessCoef	Flow recession coefficient	-	0.7	0.99
Snow	sDDF	Degree-day factor for snow	mm ◦C^{-1} day^{-1}	0.5	9
	tcrit	Temperature threshold for precipitation to fall as snow	◦C	-3	3
	snowsc	Water storage capacity of snow pack	mm mm^{-1}	0.1	0.7
Soil	alphaInf	Fraction of daily precipitation that occurs during the hour of most intense rainfall	-	0.1	0.5
	lambda	Infiltration coefficient that affects the speed of infiltration capacity reduction	-	0.25	0.8
	kEff	Effective saturated hydraulic conductivity	-	0.1	0.5
	rootDepth	Thickness of root zone	mm	350	800
	soilDepth	Thickness of subzone	mm	350	1000
	capRise	Maximum capillary rise from subzone to root zone	mm day^{-1}	5	20
Correction	precipFactor	Precipitation correction factor	-	1.2	1.45
	kcFactor	Crop coefficient multiplication factor	-	0.5	1.5

4.2.2.2 Sampling design

The 11 parameters identified as sensitive are calibrated using the GLUE framework. The parameters are sampled using the LHS design to reduce the sample size to 10% of the size of a random sampling generator. However, the required sample size is unfeasible even within an LHS framework which is illustrated by the following example: the optimal sample size of 11 parameters and 10 parameter values is 10^{11} x 0.1 = 10 Billion (see 3.1.2.2). The computation and post-processing of a single model run requires ~2h so the calibration would take > 2 Million years in CPU time, which is beyond the scope of this project.

Most studies which apply complex, (semi-) distributed models use considerably lower sampling sizes. Shen *et al.* (2012) used a sampling size of 10,000 for 20 parameters with LHS; Nepal *et al.* (2014) used a sample size of 1,600 for 16 parameters with Monte-Carlo-Sampling (MCS), and Pearson (2020) used a sampling size of 5,000 for 12 parameters with LHS.

In this research, a sample size of 10,000 is deemed appropriate for the 11 sensitive parameters with LHS using a uniform probability distribution. However, the sample size selection is a tradeoff between the sampling density and available computational resources and while it orients on the sampling strategy of previous studies, the choice is somewhat subjective as GLUE lacks formal guidelines for the parameter sampling (Beven and Binley, 2014).

4.2.2.3 Performance measures

Performance measures evaluate the fit between simulated and observed catchment behaviour and are used for the identification of behavioural parameter sets. This study uses a multi-criteria approach that includes one graphical technique, one dimensionless technique, and one error index as proposed by Legates and McCabe (1999). The choice of appropriate evaluation techniques depends on the objective and availability of calibration data. In this case, the objective is the estimation of the evolution of flood hazard frequency and magnitude so that the focus of the hydrological modelling is on the simulation of flood peaks. Therefore, the selection of evaluation techniques is based on the ability to depict the performance during high flows. The NSE is chosen as a dimensionless evaluation technique because of its sensitivity towards high flows. The PBIAS is calculated for flows ≥ 5000 m^3/s to gain information about discharge volume differences during high flows ($PBIAS_{high}$). A modified version of eGLUE ($eGLUE_{mod}$) (Liu

et al., 2009) is applied to account for the uncertainty in the observed discharge. The original eGLUE quantifies the number of days for which the simulated discharge is within the uncertainty intervals of the observed discharge. The modified $eGLUE_{mod}$ describes the mean discharge of all days for which the simulated discharge is within the uncertainty intervals of the observed discharge. This adjustment is made to prevent the low flow conditions which prevail during nine months from dominating this performance measure.

4.2.2.4 Selection of behavioural parameter sets

The selection of behavioural parameter sets contains two stages, the identification of unbehavioural parameter sets, and the selection of the best-performing parameter sets. The thresholds for the identification and exclusion of unbehaviorual parameter sets are obtained and adjusted to account for the modelling time step and data quality based on Moriasi *et al.* (2007, 2015).

In the first stage, the unbehavioural parameter sets are identified from multiple performance measures. The parameter sets are evaluated in regards to the snow dynamics and annual evapotranspiration rates and parameter sets with larger PBIAS of ±30% in snow extent and annual ETA are removed. The inspection of the hydrographs reveals that a fraction of parameter sets with high efficiencies vastly underpredict the discharge in the winter months which indicates a poor representation of groundwater processes. The PBIAS for winter months is calculated and all parameter sets that deviate more than ±30% are eliminated from the selection.

In this study, the selection of behavioural models uses a combined selection approach that combines threshold and ranking selection approaches. This is a pragmatic choice to prevent one criterion from dominating the others due to a discrepancy in the number of classified behavioural parameter sets. The threshold for NSE is 0.7 and ± 25% for $PBIAS_{high}$. These thresholds are based on the recommendations of Moriasi *et al.* (2007, 2015) and are adjusted to account for the uncertainty in the discharge observations. No threshold is defined for $eGLUE_{mod}$ due to the lack of reference thresholds and interpretability. For each performance measure, the n highest performing parameter sets are maintained, whereas n is the number of behavioural parameter sets of the performance measure with the lowest number of behavioural parameter sets.

4.2.2.5 Model setup

The SPHY model (Version 3.0) is run at daily time steps for the period 2002 – 2006 as the calibration period, and 2007 – 2015 as the validation period. The year 2001 is the warm-up period to fill the water storages and is not considered in the evaluation. The spatial resolution of the model is 500 x 500m resulting in ~ 182.000 modelling cells. The model version is V3.0 and the infiltration, glacier, groundwater, and snow modules are used.

The calibration is conducted on the Hamilton Supercomputer of Durham University. Minor adjustments in the model code were conducted by Dr S. Reaney to adapt it to the Linux environment. The precipitation correction factor was added to the source code. It multiplies the precipitation input rasters with the precipitation correction factor when loaded into the memory.

4.2.3 Estimation of the discharge uncertainty

4.2.3.1 Bayesian Rating Curve Framework

The Bayesian Rating Curve Framework (BaRatin) (Le Coz *et al.*, 2014) is used to estimate the uncertainty of the discharge time series. The method is selected because:

i. the uncertainty estimates include the parametric, structural, and observational uncertainty sources;
ii. the model output is presented as confidence intervals and matches the eGLUE format;
iii. the uncertainty is estimated based on hydraulic knowledge of the gauging site which makes the method less uncertain when extrapolating compared to other Bayesian approaches (Kiang *et al.*, 2018);
iv. the uncertainty ranges can be provided for the hydraulic knowledge and measurements which is very useful since the gauging site could not be visited during this study.

In BaRatin, the estimation of the stage-discharge relationship and the associated uncertainty is based on the hydraulic knowledge about the controls at the gauging site (Le Coz *et al.*, 2014). This information determines the stage-discharge relationship and, thus, the model parameterisation. The relationship between stage and discharge can be represented by a segmented regression in the case that the gauging site is characterized by multiple controls that alter the stage-discharge relationship at different stage levels.

The estimation of measurement uncertainty is based on the assumption that stage measurement errors are negligible and that discharge measurement errors follow a Gaussian distribution with mean zero and known standard deviation (Le Coz *et al.*, 2014). The uncertainty of each discharge gauge can be specified as a percentage. A remnant error is implemented to account for the structural uncertainty in the rating curve. It is assumed that this error follows a Gaussian distribution with mean zero and unknown standard deviation and is estimated from the gaugings (Le Coz *et al.*, 2014). The prior distributions of the rating curve parameters and the specified measurement uncertainty are combined using Markov Chain Monte Carlo (MCMC) simulations (n = 1000) and hence a large set of rating curves is created with different values for the parameters and observations. The remnant error is added to each rating curve and the total uncertainty interval is then determined from the 5 and 95 percentiles of rating curves.

4.2.3.2 Rating-curve parameterisation

Rating curve methods are based on the assumption of stable controls which means that the geometry of the channel (or subsets of the channel for channels with multiple controls) which determines the stage-discharge relationship remains constant. Thus, multiple rating curves need to be established if one or more controls shift during the study period which requires information about the timing of shifts. Additionally, the BaRatin method requires information about the geometry of the controls to estimate a reasonable parameter range for the parameter fitting.

Investigation of temporal shifts

In this study, temporal shifts in the stage-discharge relation are identified from a combination of graphical and statistical methods. In the first step, a log-log transformation is conducted on the stage and discharge measurements to convert the logarithmic stage-discharge relationship to a linear relationship (Herschy, 1993). This relationship is plotted and it is investigated whether for certain stage ranges (hereafter referred to as stage groups) the variances in discharge are larger compared to other stage ranges. A higher variance in one of the stage groups indicates that the active control (for the given stage group) shifts more frequently or stronger compared to the controls that determine the stage-discharge relation for the other stage groups.

The unequal variance t-test (also Welch's t-test) (Welch, 1947) is conducted for each of the stage groups to identify the timing of the shift. This test compares the means between two independent populations without assuming equal variances to investigate if both populations are similar (evidence for no shift) or different (evidence for shift) (Ahad and Yahaya, 2014). The unequal variance t-test is chosen over the student's t-test and Mann-Whitney U test because it provides robust evidence for populations with equal variance as well as populations with unequal variance (Ruxton, 2006). However, Welch's t-test assumes that the populations are normally distributed. It cannot be verified that this assumption is met because of the low sample size (see section 4.3.3.3).

The unequal variance t-test requires a measure that contains information about the stage and discharge. Since the log-log transformed stage-discharge relationship follows a linear trend, this measure can be obtained by dividing log(discharge) and log(stage). The test is then conducted for the populations of this measure of subsequent years. If the null hypothesis is rejected with high confidence (p-value ≤ 0.05) it indicates that the

distributions of the measure of the two groups (years) are different and, hence, that the stage-discharge relationship has shifted.

Prior hydraulic knowledge

The hydraulic information is required to identify the number of controls and the parameter range of each control. This information comprises the activation stage, width, slope and Strickler coefficient of each control. In the ideal case, detailed knowledge of the gauging site from a field survey or the expertise of the gauging site manager would be used to define the controls and their initial parameter range. In the case of the gauging station at the mountain outlet, no official information is available, and no field survey could have been conducted. Therefore, the information is derived from literature and remote sensing products. The derivation of the rating curve parameter from remote sensing products is more uncertain compared to field surveys. However, the information is used to provide a first rough estimation of the parameter range and not the exact parameter values (Le Coz, 2014; Le Coz *et al.*, 2014). The software allows the definition of uncertainty ranges for the prior hydraulic information and hence higher uncertainty ranges account for the lower accuracy of the remote-sensing-based information.

The identification of the different controls is based on a cross-section profile of a DEM derived from World-View imagery (Figure 4.2). This DEM does not depict the topography of those parts of the riverbed which are covered by water during the acquisition of the images. This part is defined at the low flow control for which an activation stage of 0 ± 2 m is assumed (the activation stage can be negative depending on the calibration of the stage). The medium control is activated at the median stage height of 2.9 m in the first week of March (the week of the image acquisition). The activation stages of the high flow and flood flow controls are measured from the cross-section profile.

A rectangular geometry is chosen for each control which seems most in line with other rating-curve applications of comparable cross-sections (Le Coz, 2014). The stage-discharge relationship of rectangular channels is calculated as (Le Coz, 2014) ;

$$Q(h) = a\,(h-b)^c \; for \; h > k \qquad \text{Eq. 4.31}$$

$$a = K_s B_w \sqrt{S} \qquad \text{Eq. 4.32}$$

where Q = discharge (m^3/s), h = stage (m), k = activation stage (m), K_s = Strickler coefficient (m/s), B_w = Channel width (m) and S = channel slope. The exponent c is 5/3 for rectangular channels, and the parameter b does not require specification in the BaRatin framework.

The width (B_w) is measured at the upper bound of each control but excludes the width of the next lower control. The slope is measured from the DEM and averaged over 2 km to remove the local noise. The Strickler coefficient is the inverted Manning's n and is obtained from Chow (1959); the value for mountain streams with gravel, cobbles and few boulders is selected for all three controls, which seems most reasonable because the gravel-sand transition occurs further downstream (Dingle *et al.*, 2020b). The BaRatin parameterisation is summarised in Table 4.2.

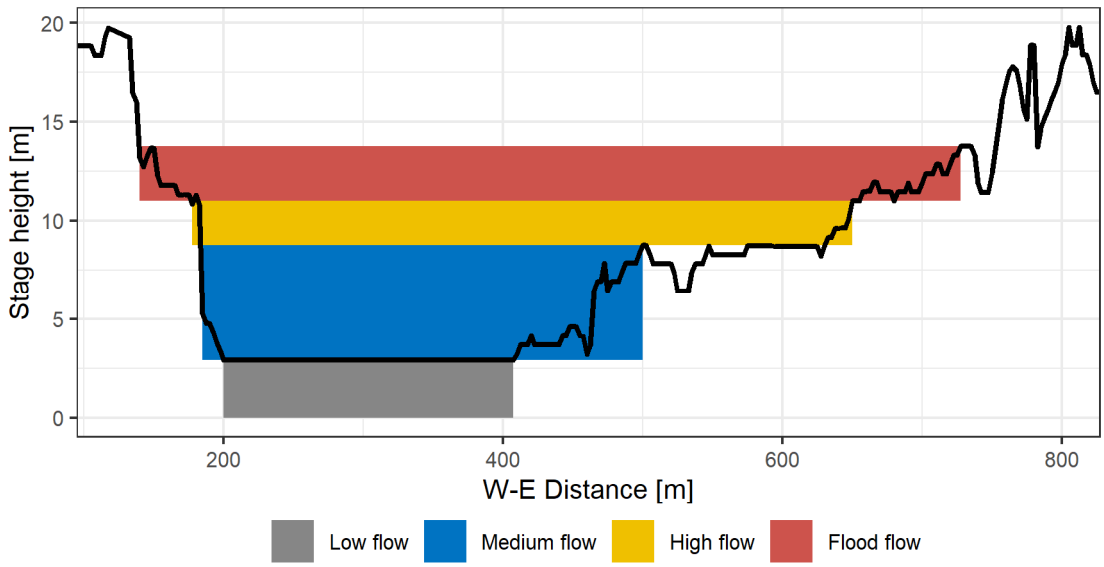


Figure 4.2. Controls of the Rating Curves. The solid line is the cross-section profile of the Gauging station which was derived from the WV-2 DEM (Section 6.2.5) and the rectangular areas indicate the identified controls.

Table 4.2: The parameterisation of BaRatin.

Control	Geometry	Width [m]	Activation stage [m]	Slope [m/m]	Strickler-coefficient [m s^{-1}]
Low flow	rectangular	207 ± 20	0.0 ± 2	0.001 ± 0.0005	25 ± 12
Medium flow	rectangular	107 ± 20	2.9 ± 2	0.001 ± 0.0005	25 ± 12
High flow	rectangular	157 ± 20	8.8 ± 2	0.001 ± 0.0005	25 ± 12
Flood flow	rectangular	115 ± 20	11.0 ± 2	0.001 ± 0.0005	25 ± 12

4.3 The utilised datasets

The datasets used for the hydrological model fall into three categories; static data which does not change throughout the modelling period (Section 4.3.1), climate forcing data that describes the daily atmospheric conditions and changes with time (Section 4.3.2), and calibration data is not fed into the model but used to calibrate the model and validate its performance (Section 4.3.3). The datasets and their sources are summarized in Table 4.3.

All spatial datasets are reprojected to the modelling coordinate system (CRS) WGS84 / UTM Zone 44N (EPSG: 32644), converted to the raster format, resampled to the modelling resolution (500 x 500m), and aligned to the model grid in the preprocessing stage. The data sources and individual processing steps are described in this chapter.

Table 4.3: The datasets used for the hydrological modelling.

Variable	Dataset	Usage	Source
Topography	HydroSHED V 1.0	Watershed delineation, temperature downscaling, Flow routing	Lehner *et al.*, 2008
Soil	HiHydroSoils	Soil saturated water content, field capacity, permanent wilting point of the root zone and subzone layers	De Boer, 2016
Geology	Physiographic Divisions of Nepal	Demarcation of zones with different geological properties	ICIMOD, 2011
Vegetation	Globcover 2009 V2.3	Demarcation of land use zones for actET simulation	Arino *et al.*, 2012
Glacier	Randolph Glacier Inventory 6.0	Glacier delineation	RGI Consortium, 2017
	Debris-covered glaciers	Classification of debris-covered glaciers	Scherler *et al.*, 2018
	Ice thickness	Glacier depth	Farinotti *et al.*, 2019
Snow	MOYDGL06	8-day snow cover composites	Muhammad and Thapa, 2020
Temperature	WFDEI	Minimum, maximum and mean daily air temperature	Weedon *et al.*, 2011, 2014
Precipitation	GPM IMERG Final Precipitation L3 1 Month V006	Monthly precipitation maps	Huffman *et al.*, 2019
	Gauged precipitation (daily)	Temporal disaggregation of monthly precipitation maps	DHM
Actual Evapo-transpiration	MODIS 16A3GF V006	Model calibration	Running *et al.*, 2019
Discharge	Gauged discharge (daily)	Model calibration	DHM
	DHM stage-discharge observations	Uncertainty estimation of daily discharge	DHM

4.3.1 Static datasets

Static datasets are spatial maps that represent the environmental conditions and remain similar over the full modelling period. In reality, environmental conditions are dynamic and change over time. For example, the geomorphological processes (i.e. landslides) alter the topography and human activities change the land cover and vegetation. However, the modelling period covers 13 years, and it is assumed that the effect of these changes is negligible for the modelling period (e.g. Lutz *et al.*, 2014; Nepal *et al.*, 2014; Dhami *et al.*, 2018; Chandel and Ghosh, 2021; Khanal *et al.*, 2021).

The topography is represented by the HydroSHED V 1.0 DEM (Lehner *et al.*, 2008). This Digital Elevation Model (DEM) is a derivative of the Shuttle Radar Topography Mission (SRTM) DEM that is optimized for hydrological applications (Figure 4.3). It has a spatial resolution of ~90 m and represents the topography of the year 2000. The dataset is used to delineate the watershed boundaries, the runoff routing and the downscaling of the temperature data.

The Globcover 2009 V2.3 dataset (Arino *et al.*, 2012) is used to represent the spatial distribution of different land cover types in the catchment. This dataset is based on images collected from January – December 2009 by the Merris sensor on board the ENVISAT satellite and has a resolution of 300 m. The overall accuracy of the dataset reaches 67.5% (Bontemps *et al.*, 2011) but varies with the landcover class; classes that are classified with high confidence (i.e. croplands, permanent snow and ice, water) cover 36% of the catchment, while uncertain classes (sparse vegetation, herbaceous vegetation) cover 21%. The land cover raster is presented in Figure 4.4.

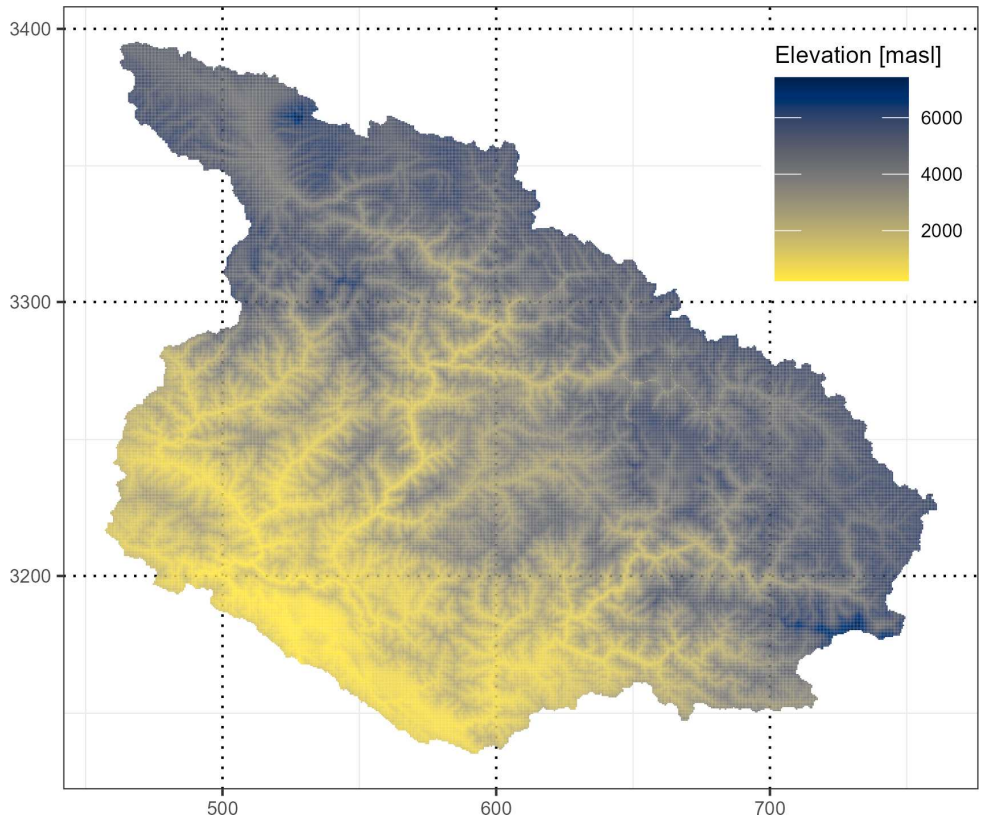


Figure 4.3: The HydroSHED DEM (Lehner et al.*, 2008) which represents the topography in the hydrological modelling. The coordinates are in WGS 84 / UTM Zone 44N (km)*

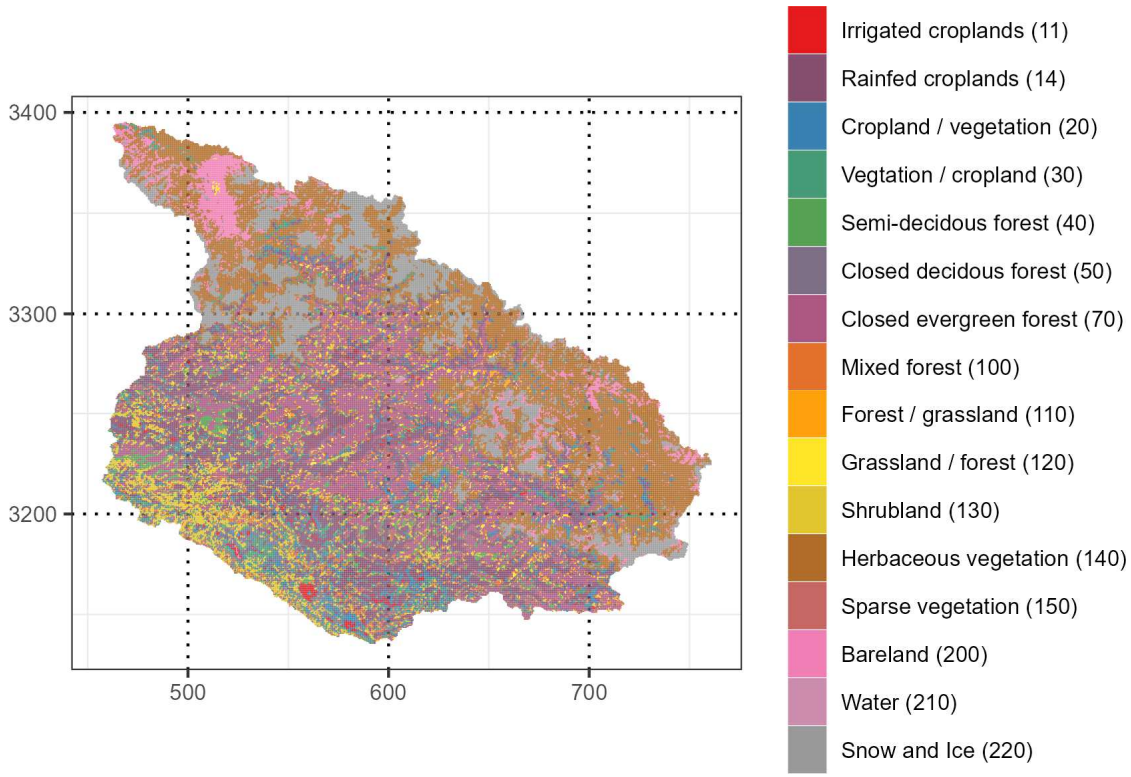


Figure 4.4: The Globcover dataset (Arino et al.*, 2012) which represents the land cover in the hydrological modelling. The class names were simplified and the Globcover ID is provided in the brackets for identification. The coordinates are in WGS 84 / UTM Zone 44N (km).*

Soil information is derived from the HiHydroSoil database (De Boer, 2016). This dataset has a resolution of 900 m and is based on the SoilsGrid1km (Hengl *et al.*, 2014), and is gap-filled with data from the Harmonized World Soil Database (HWSD). The HiHydroSoil database converts the soil properties of the gap-filled SoilsGrid1km data to hydraulic soil properties using the Mulan Van Genuchten model (Van Genuchten, 1980). The SPHY-Preprocessor (Terink *et al.*, 2015b) is used to derive maps of the wilting point and permanent wilting point (only root zone SW_1) (Figure 4.5), and the field capacity, saturated water content, saturated hydraulic conductivity (separate maps for root zone SW_1 and subzone SW_2) (Figure 4.6), and There is a lack of information on the depth of the root zone and subzone layers so that the depth of the layers is calibrated.

The SPHY model requires information on the groundwater storage capacity. Despite the availability of data on the spatial distribution of geological layers in the Himalayas, there is a lack of information about the storage capacity of these layers. Nepal *et al.* (2014) used the physiographic divisions of the Himalayas to estimate the spatial distribution of groundwater storage capacities. Even though the physiographic divisions do not reflect the local variability in geologic conditions, the large-scale differences are captured. The storage capacity reduces with elevation in the northern direction. In this work, the Physiographic Divisions of Nepal dataset (ICIMOD, 2011) is used (Figure 2.1). The storage capacities were estimated in personal communication with Dr. S. Nepal and are 1650 mm (Siwalik range), 1350 mm (Middle Hills) and 1050 (High Hills, High Mountains).

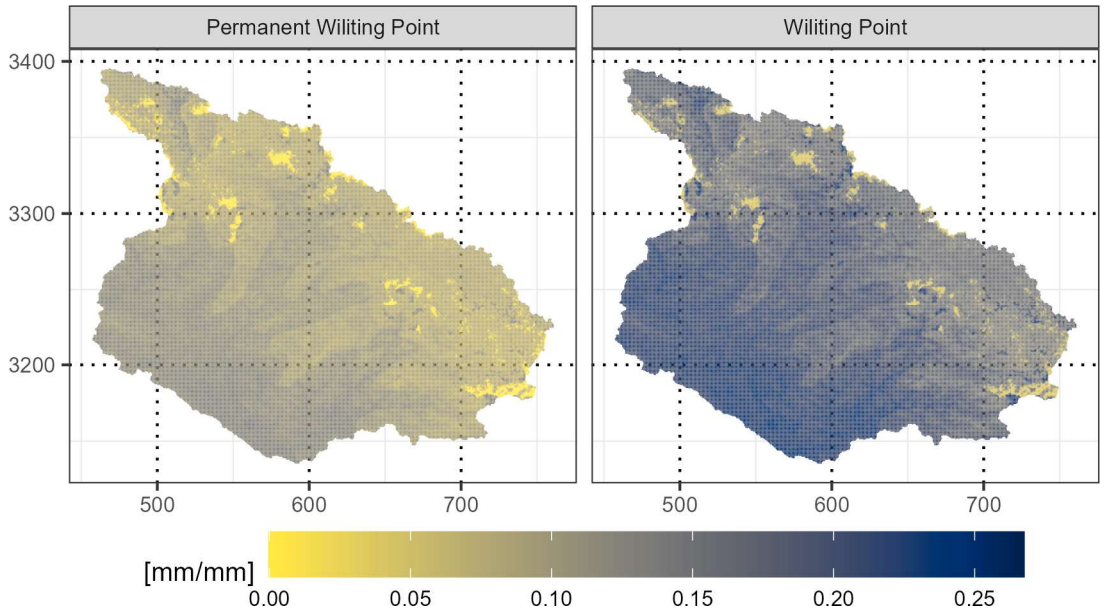


Figure 4.5: The permanent wilting point and the wilting point of the root zone layer (De Boer, 2016). The coordinates are in WGS 84 / UTM Zone 44N (km).

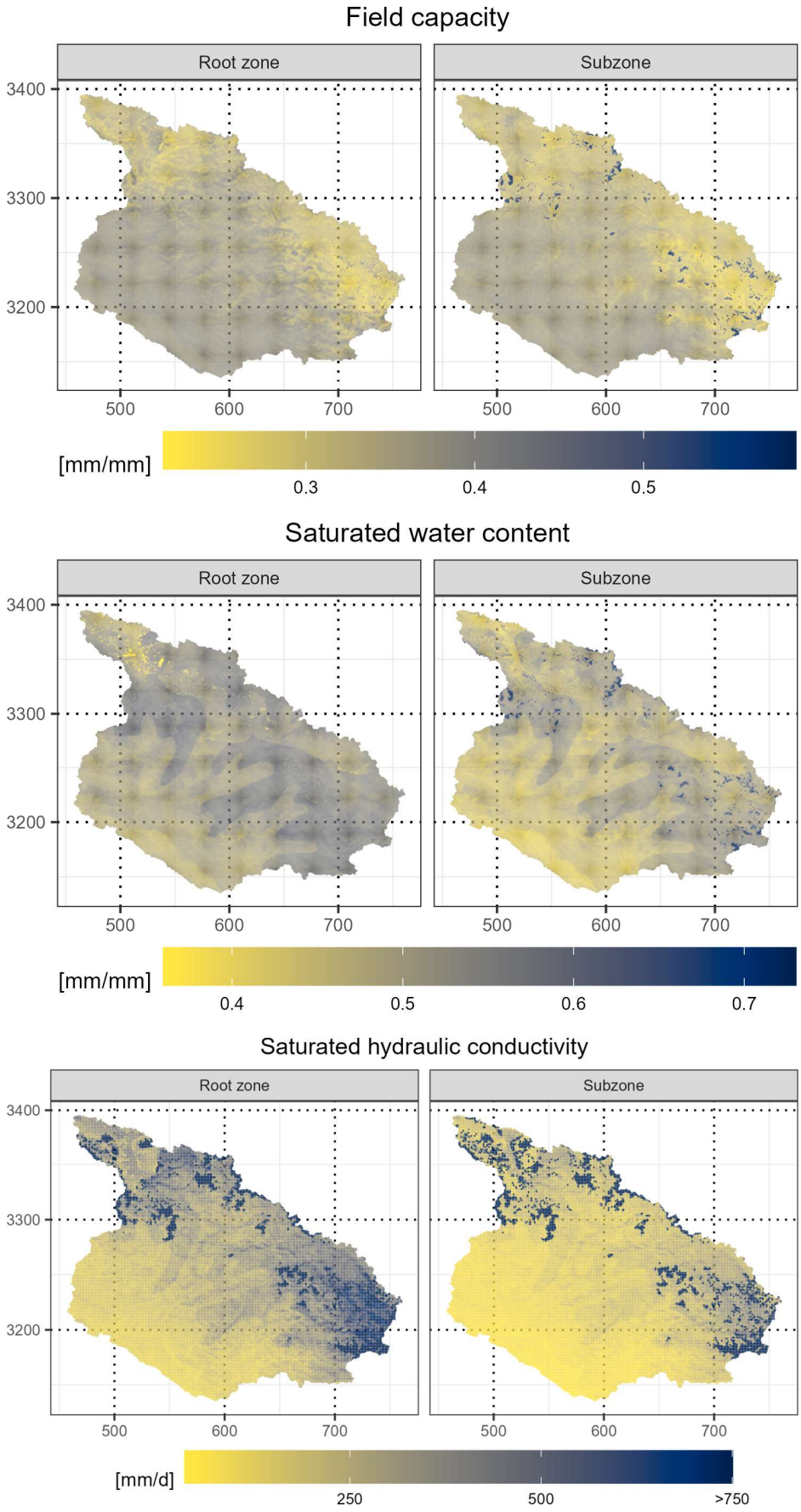


Figure 4.6: Soil properties for the root zone and subzone layers (De Boer, 2016). The coordinates are in WGS 84 / UTM Zone 44N (km).

Three datasets are used as input for the glacier module; the Randolph Glacier Inventory 6.0 (RGI Consortium, 2017) dataset is used for the delineation of glaciers (Figure 4.7). The outlines of supraglacial debris cover for RGI 6.0 (Scherler *et al.*, 2018) are used to classify the glaciers into clean ice and debris-covered glaciers. The ice thickness is represented by the modelled ice thickness distribution for RGI 6.0 (Farinotti *et al.*, 2019). The fraction of a modelling cell that is covered by glaciers is calculated during the data preparations. In SPHY a glaciated modelling cell can either be of type clean ice or debris-covered. Therefore, the glacier type for each glaciated modelling cell is classified based on the dominating glacier type. The glacier elevation for each glaciated cell is obtained from the RGI 6.0 dataset (median elevation). In the case that the median RGI 6.0 elevation is higher than the DEM elevation, the minimum RGI 6.0 elevation is used.

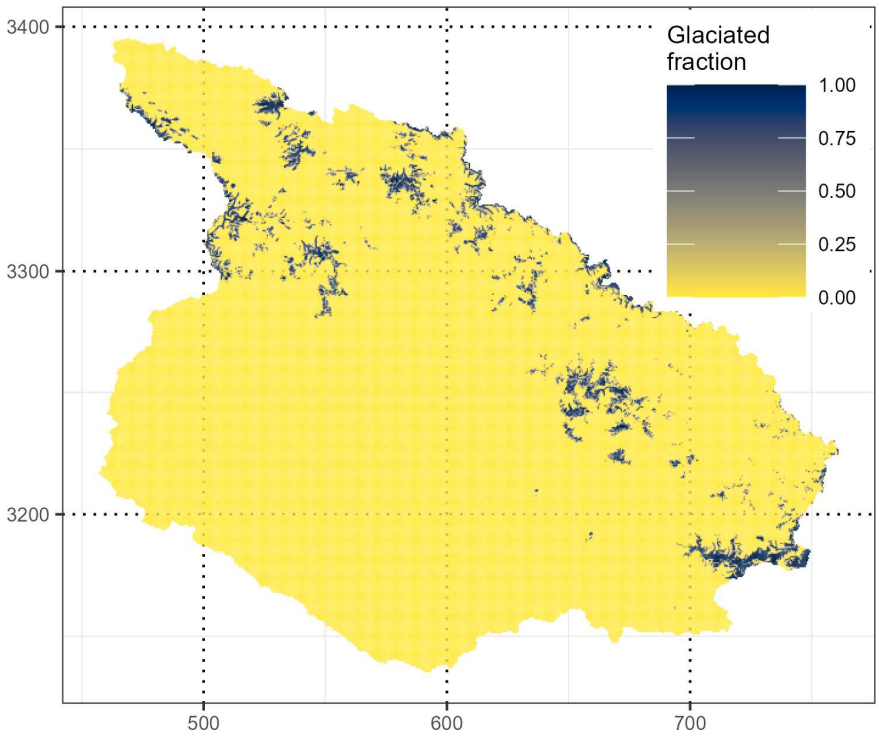


Figure 4.7: Glaciated areas in the hydrological modelling (RGI Consortium, 2017). The value indicates the fraction of the grid cell that is glaciated. The coordinates are in WGS 84 / UTM Zone 44N (1000 km).

4.3.2 Climate forcing data

The climate forcing data represents the atmospheric conditions of the modelling period at the daily resolution. This data comprises the minimum, maximum and average daily air temperature and the daily precipitation. Gridded products for temperature and precipitation are used in this study due to the low density and concentration along valleys of the climate station network. The gridded products can be grouped into three categories: i) interpolated datasets that are based on ground observations; ii) satellite-based datasets, and; iii) datasets that are generated by merging reanalysis data and ground-based observations (Lutz and Immerzeel, 2016). The main criteria for the dataset selection are the spatial resolution because of the high variability of atmospheric conditions on the local scale in the Himalayas (Collier and Immerzeel, 2015) and the accuracy of the product. A reanalysis product is used to represent the temperature, and a combination of satellite-based datasets and ground observation is used as precipitation input.

4.3.2.1 Temperature forcing data

Interpolated temperature datasets lack observations at high elevations and lack accuracy in mountainous regions (Winiger *et al.*, 2005; Lutz and Immerzeel, 2016; Pan *et al.*, 2019). Comparisons of observed temperatures and temperatures retrieved from remote sensing indicate that satellite-based products are useful data sources for global analyses (Sobrino *et al.*, 2020). However, these products lack data for pixels with cloud cover, and hence large data gaps are inherent during the monsoon season. Therefore, this research uses reanalysis data for the temperature forcing. The WATCH forcing data methodology applied to ERA-Interim dataset (WFDEI) (Weedon *et al.*, 2011, 2014) depicts the temperatures with low biases and is, therefore, chosen in this research (Lutz and Immerzeel, 2016; Bhattacharya *et al.*, 2020).

The basis for WFDEI is the ERA-Interim data which is bias-corrected and elevation-corrected using monthly average and diurnal temperature data of the Climate Research Unit (CRU) gridded station observation products CRU TS 3.1 for 1979-2009 and CRU TS 3.21 for 2010-2012 (Weedon *et al.*, 2014). The WDFEI dataset has a spatial resolution of 0.5 X 0.5° and is downscaled to the modelling resolution using a lapse rate. This lapse rate is derived from gauges within and in proximity to the catchment using linear regression modelling.

The lapse rate Is different for dry and wet adiabatic conditions which vary in time and space (Immerzeel *et al.*, 2014; Lutz and Immerzeel, 2016). The orographic rainfalls are, therefore, expected to cause elevation-dependent differences in the lapse rates. However, the highest temperature gauge is located at ~3,000 masl which prevents the determination of separate lapse rates for elevation bands. Therefore, the lapse rate is applied to the total catchment. Both annual and seasonal lapse rates, which account for the seasonal variation in the cooling, are calculated (Figure 4.8).

These lapse rates are used to downscale the temperature data from the WFDEI resolution to the modelling resolution. This downscaled temperature is then compared with the temperature gauges for validation (Figure 4.9). This comparison reveals that the temperatures at higher elevations which are downscaled by the winter lapse rate systematically underestimate the observed temperatures. The data downscaled by the annual lapse rate does not contain a systematic bias and, hence, the annual lapse rate of -0.0054 ºC/m is used for the spatial downscaling. This lapse rate is identical to the one used by Dhami *et al.* (2018) for the Karnali catchment.

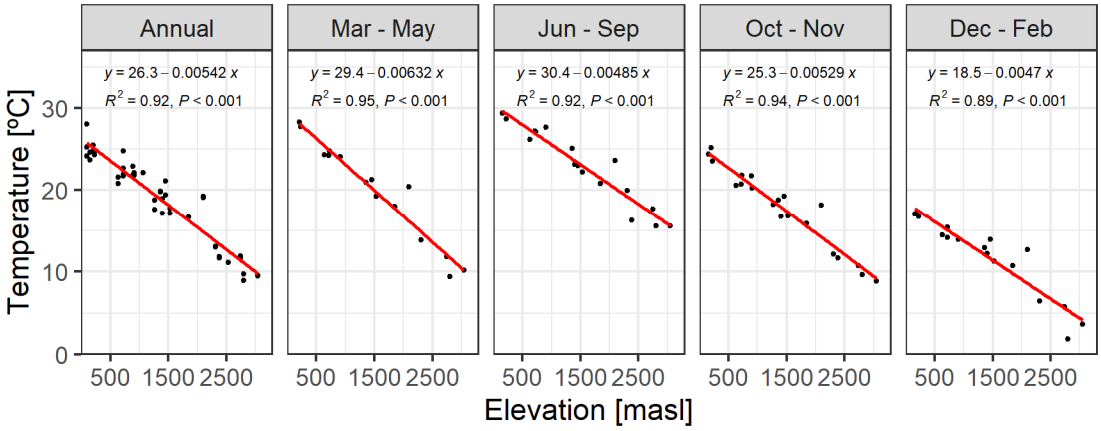


Figure 4.8: Lapse rates calculated from gauges (DHM) in the period 1998 – 2016. The red lines indicate the linear regression model of the lapse rates. The x in the regression equations is the temperature decrease (ºC/m).

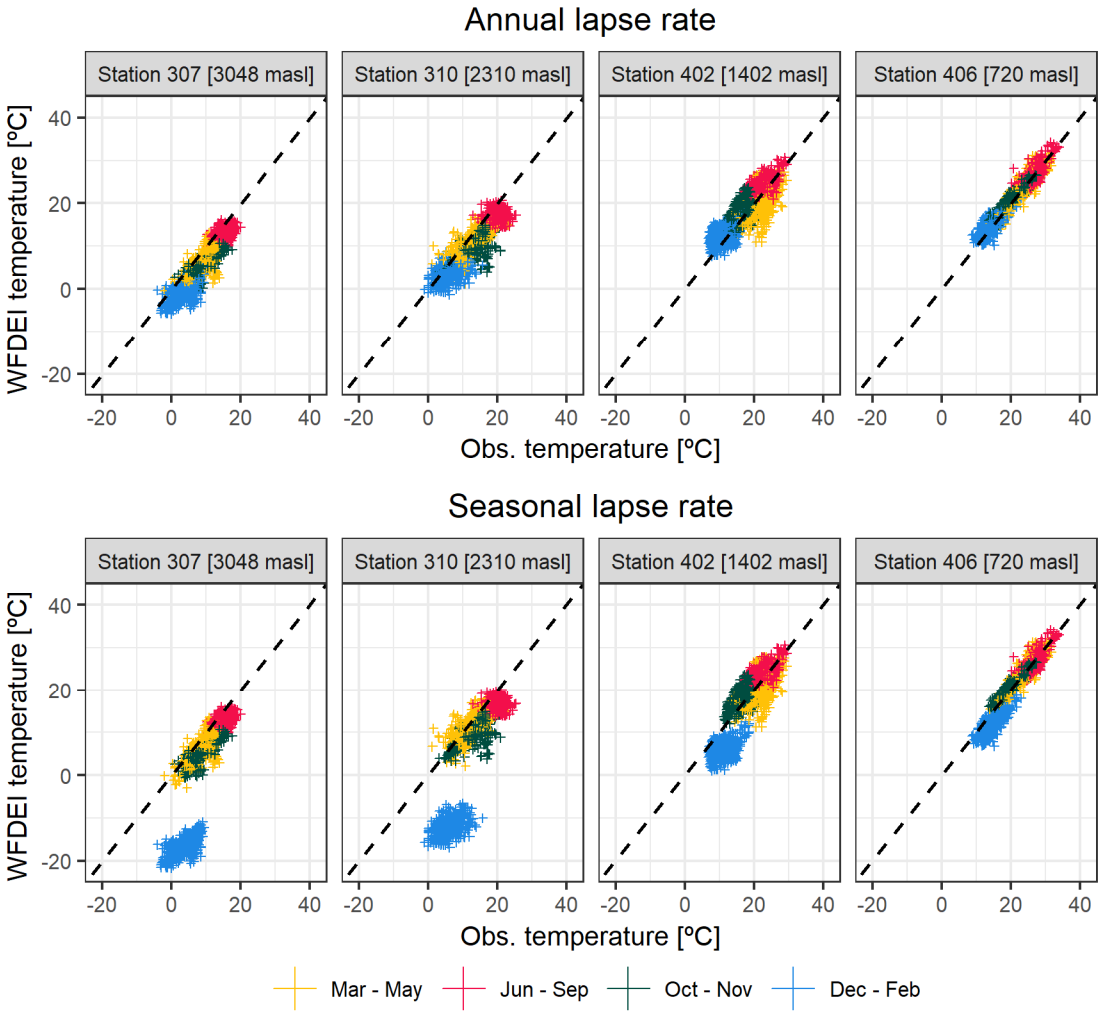


Figure 4.9: Elevation corrected WFDEI temperatures and observed temperatures of selected gauges (DHM) at daily resolution for the years 2001 – 2006. The dashed line is the line of the perfect fit. The elevation correction is conducted using the annual lapse and seasonal lapse rates presented in Figure 4.8.

4.3.2.2 Precipitation forcing data

As for the interpolated temperature datasets, the interpolated precipitation datasets suffer from observation gaps along the hillslopes and mountain ridges and hence these datasets are more robust in the southern regions of the study area. The central and northern parts of the catchment with elevations ≥ 3000 masl lack observations and hence the precipitation is more uncertain. Satellite-based precipitation products have large spatial coverage and are useful to improve the understanding of spatiotemporal precipitation patterns in data-scarce regions (Arias-Hidalgo *et al.*, 2013). However, the accuracy of such products decreases with the temporal resolution and the daily products may include large biases (Arias-Hidalgo *et al.*, 2013; Bhardwaj *et al.*, 2017; Pearson *et al.*, 2022). Furthermore, precipitation estimates are more accurate in oceanic and flat environments than in areas with complex topographies which is due to the complex convection mechanics that are not captured in the remote sensing imagery and the sparser gauge network to identify and correct biases (Arias-Hidalgo *et al.*, 2013; Collier and Immerzeel, 2015; Sunilkumar *et al.*, 2019). Precipitation of climate reanalysis products also may exhibit large biases in mountainous regions which may lead to poor performance of hydrological models if forced with this data (Bhattacharya *et al.*, 2020). It is therefore not clear which precipitation product captures the spatiotemporal precipitation patterns at the daily resolution in the catchment best.

Several gridded precipitation products are compared with gauged precipitation to asses their potential for application in hydrological modelling (Figure 4.10). These evaluated datasets are APHRODITE-2 (Yatagai *et al.*, 2012) and CPC Global Unified Precipitation (National Center for Atmospheric Research, 2023) (interpolated gauge-based datasets), TRMM 3B42_Daily (Huffman *et al.*, 2016), GPM Level 3 IMERG Daily (GPM_3IMERGDF) (Huffman *et al.*, 2019) and PERSIANN-CDR (Ashouri *et al.*, 2015) (all satellite-based precipitation estimates), and the reanalysis product WFDEI (Weedon *et al.*, 2011, 2014).

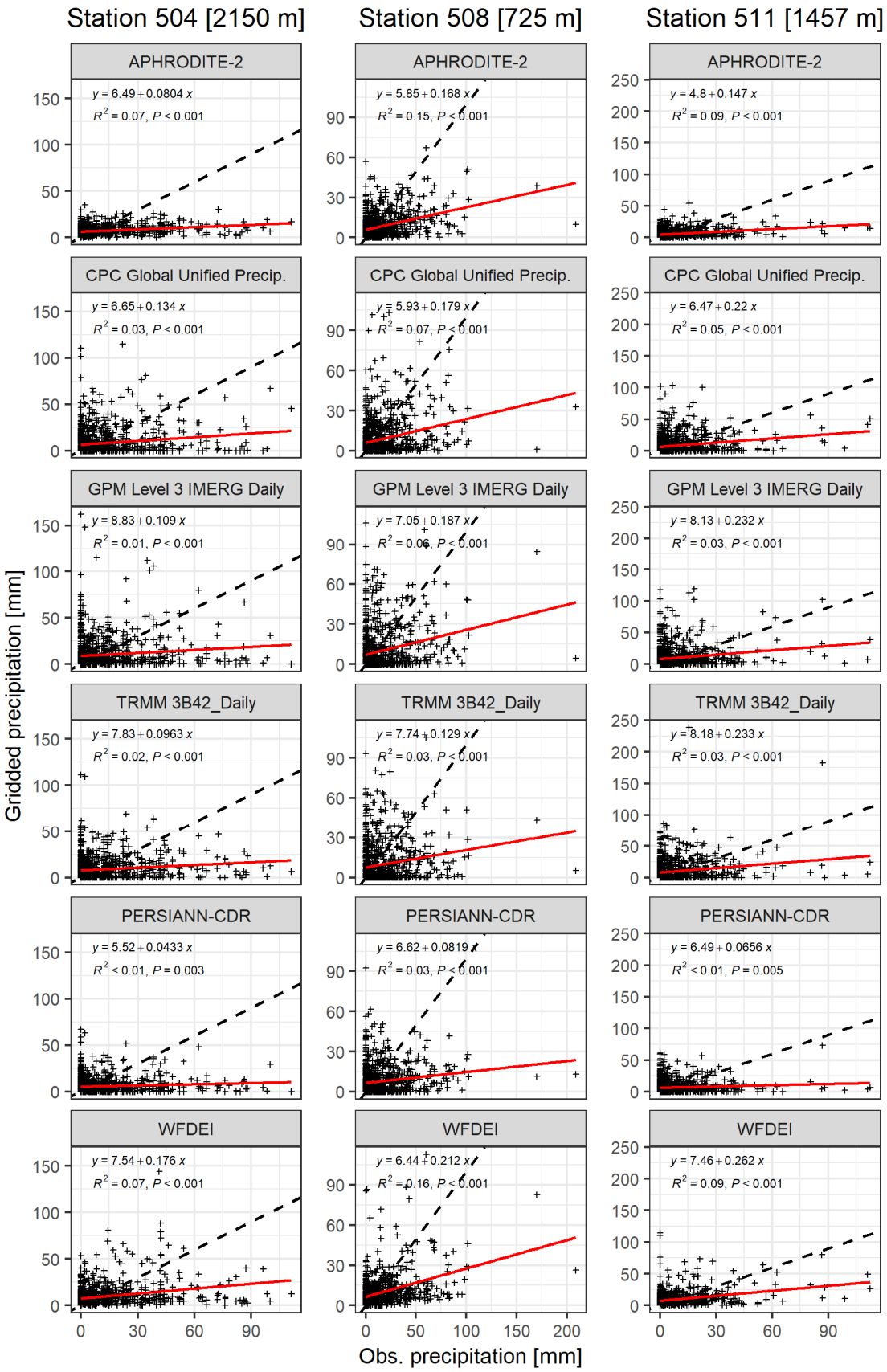


Figure 4.10: Comparison of gridded daily precipitation and DHM observations for three stations in the Karnali River catchment. The investigated period is 2000 - 2009. The red line indicates the linear regression and the black dashed line indicates the line of the perfect fit.

None of these products depicts the observed precipitation at any rainfall gauge (Figure 4.10). The observed precipitation extremes are systematically underestimated by the precipitation products. Furthermore, the precipitation extremes are often estimated on days with no or low observed precipitation. The poor performance of all products can partly be explained by comparing gridded products with point observations at the stations. However, the degree of non-correlation indicates that none of the products reflects the daily variations of precipitation in the study area sufficiently, which agrees with the findings of Bhardwaj *et al.* (2017) for the Garhwal Himalaya and Pearson *et al.* (2022) for the East Rapti catchment in the Central Himalayas. Therefore, none of the products is of sufficient quality for the application in hydrological modelling.

A gridded precipitation product is combined with gauged precipitation to generate the precipitation input for the hydrological modelling. The accuracy of satellite precipitation estimates increases with decreasing temporal resolution (Huffman *et al.*, 2010; Arias-Hidalgo *et al.*, 2013). A gridded monthly product is used to represent the spatial precipitation patterns and this dataset is downscaled to the daily resolution using precipitation gauges of the Department of Hydrology and Meteorology Nepal (DHM) following the temporal disaggregation approach of Arias-Hidalgo *et al.* (2013). This approach was also applied by Pearson *et al.* (2022) in the Himalayan East Rapti catchment. In this study, the monthly GPM precipitation product GPM IMERG Final Precipitation L3 1 Month V006 (Huffman *et al.*, 2019) is selected because it has the highest spatial resolution (0.1° x 0.1°).

The temporal disaggregation combines the spatial patterns of the giddded precipitation estimate and the temporal pattern of the precipitation gauges. To obtain the temporal patterns, the observations of the gauges are interpolated using the Inverse Distance Weighting (IDW) method. The obtained daily rasters are summed to the monthly raster. The daily fraction of the monthly rainfall is then obtained by dividing the daily rasters by the monthly raster. The monthly gridded precipitation of the satellite estimate is then multiplied by the daily fractions to disaggregate the monthly precipitation estimates. With this approach, the temporal variation observed in the DHM gauges is transferred to the GPM data. The gauge network is not dense enough to depict the local variations in the precipitation patterns, but the large-scale patterns (e.g. westward-moving monsoon rain fields) should be reflected in the gauge data.

The disaggregated satellite estimate precipitation correlates significantly better with the gauged precipitation than the daily products (Figure 4.11). The disaggregated GPM data is not bias-corrected in this study because there is no information about the bias in vast unobserved regions of the catchment. Therefore, it is assumed that the spatial variations are better reflected by the GPM data than by the gauge network. However, the precipitation correction factor implemented in SPHY accounts for under-, or overestimations in the GPM data without changing the spatial patterns of the GPM precipitation.

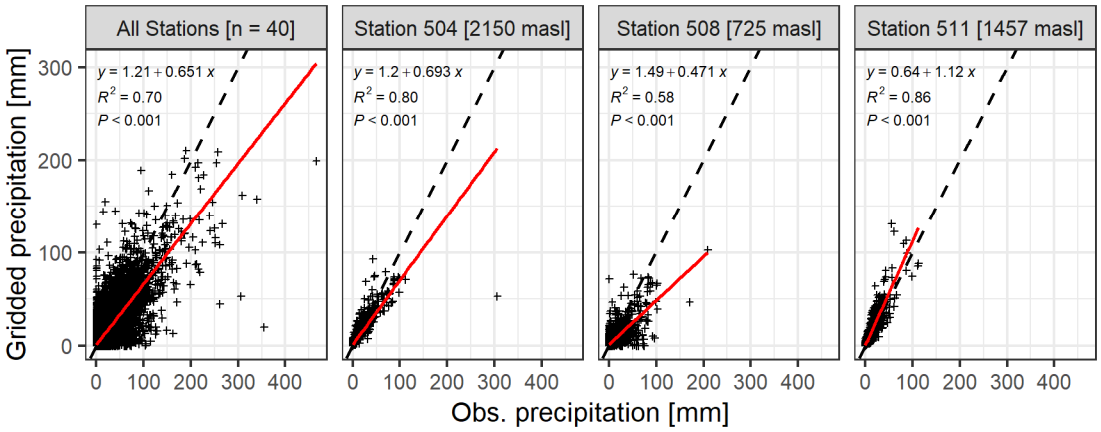


Figure 4.11: Comparison of the disaggregated monthly GPM precipitation and the observed daily precipitation for all 40 stations (plot 1) and selected stations (plots 2-4). The red lines indicate linear regression models and the black dashed lines indicate the line of the perfect fit. The comparison for each of the 40 stations is presented in the Appendices 4.1 and 4.2.

4.3.3 Calibration data

A multi-criteria approach to calibrate different aspects of the model with different datasets is more robust than the calibration with a single type of data. Three different datasets are used in this study to calibrate different modules. These datasets describe the actual evapotranspiration to calibrate the evapotranspiration rates (Section 4.3.3.1), the snow cover to calibrate the snow module (Section 4.3.3.2), and the discharge to calibrate the remaining modules that simulate the rainfall-runoff conversion (Section 4.3.3.3).

4.3.3.1 Actual evapotranspiration estimates

There is no continuous information on the evapotranspiration rates available for the Karnali basin. Therefore, the gridded satellite estimate product Moderate Resolution Imaging Spectroradiometer (MODIS) MOD16A3GF V006 dataset (Running *et al.*, 2019) with a spatial resolution of 500 m and a temporal resolution of one year is used for the calibration of the actual evapotranspiration rate (ET_{act}). Two versions exists which are based on the imagery of the MODIS Terra and MODIS Aqua satellites, respectively. The MODIS Terra product is chosen because this satellite was launched earlier and covers the complete modelling period. This product combines remotely sensed data such as land cover, Leaf Area Index (LAI), the fraction of absorbed photosynthetically active radiation (FPAR) and albedo with meteorological data of reanalysis products such as air pressure, air temperature, humidity and radiation to estimate ET_{act} based on the Penman-Monteith equation (Monteith, 1965). The spatial resolution of the dataset equals the spatial resolution of the underlying satellite product (500 x 500 m). However, the resolution of the underlying reanalysis products is coarser (1°) and hence the spatial heterogeneity of ET_{act} is only partly depicted at the grid scale (Sullivan *et al.*, 2019).

Cloud cover leads to an increase in the albedo and a decrease in FPAR and LAI which decreases the product accuracy. The yearly product uses gap-filling algorithms to reduce the errors that are caused by cloud cover. Running *et al.* (2019) compared the daily ET_{act} with ground observations for two sites and found that the average mean absolute error lies within the accuracy range of the observations. However, Long et al. (2014; in Sullivan *et al.*, 2019) found that satellite-based ET retrievals underestimate ET_{act} in wet conditions and overestimate them in dry conditions. Sullivan *et al.* (2019) found that the MODIS products underestimate ET_{act} during the months with the highest ET_{act} in the Southern Great Planes in the U.S. which leads to an underestimation of

yearly ET_{act} of 38%. There is no information about the accuracy of the product in the Himalayas, but ET_{act} is likely underestimated due to the frequent cloud cover during the monsoon season.

4.3.3.2 Snow cover

The simulated snow extent is compared with snow extent observations for the calibration of the snow module. Snow features such as snow coverage and thickness are difficult to measure in the field, especially in remote areas in the mountains (Muhammad and Thapa, 2019, 2020). Remote sensing products overcome this limitation and provide continuous snow observations on a global scale. The snow cover products derived from the MODIS satellite provide daily records (MOD10A1) at 500 m resolution dating back to the year 2000 (Riggs *et al.*, 2019).

Satellite-based snow products underestimate the snow coverage in areas with dense cloud cover (Muhammad and Thapa, 2020). This effect is reduced in the 8-day composite product (MOD10A2) in which the snow cover represents every pixel that is classified as snow for at least one day. However, underestimations due to persistent cloud cover persist especially during the winter and monsoon seasons (Hall and Riggs, 2007; Muhammad and Thapa, 2020). Furthermore, thin snow layers during the early and late snow seasons are only depicted with low accuracy by the MODIS sensors (Hall and Riggs, 2007). While the snow cover is underestimated in certain conditions, the spatial resolution and the large zenith angle result in an overestimation of snow cover, especially in the outer cells of the sensing swath (Li *et al.*, 2016).

Despite these limitations, an accuracy of 93% is observed for the daily product (MOD10A1) in the Rio Grande region in the U.S., whereas the accuracy depends on the land cover, seasonality, and topography (Hall and Riggs, 2007). For the complex topography of the Hindu Kush Himalayas (HKH), the accuracy of the composite product is 77% and therefore considerably lower and the greatest source of uncertainty is related to the large zenith angle (Muhammad and Thapa, 2020).

Muhammad & Thapa (2020) propose a methodology that combines MODIS Terra and MODIS Aqua images to decrease the overestimations and applies spatial, temporal and seasonal filters to decrease the underestimations. Following this approach, the accuracy of the 8-day composite is increased by 10% to 87% in the mountains of the HKH. Two improved MODIS derivatives are published for the high mountains in Asia; a daily product (MOY10A1GL006; Muhammad and Thapa, 2019) and an 8-day composite (MOYDGL06, Muhammad and Thapa, 2020).

The snow extent of these MODIS-based products varies in absolute amounts and temporal dynamics (Figure 4.12). The daily MOD10A1 snow cover is highly variable and for some days during the winter and pre-monsoon season, no snow is observed which is unrealistic. Furthermore, the snow extent varies by several thousands of kilometres within a few days which indicates the strong noise in this product. The temporal variation in the snow extent of the improved daily product (MOY10A1GL006) is lower and appears more reasonable.

The snow cover is larger in the 8-day composites than in the daily products. The MODIS 8-day composite shows higher snow cover extents throughout the year compared to the improved MODIS product. This difference is explained by the overestimations due to the large zenith angle. In this study, the MOYDGL06 dataset is used because this data – in light of the absence of field observations to validate the datasets – is least affected by noise (e.g. cloud cover) and represents the snow conditions of the study area best. This dataset covers the period 2002 – 2018 and glaciated areas are removed.

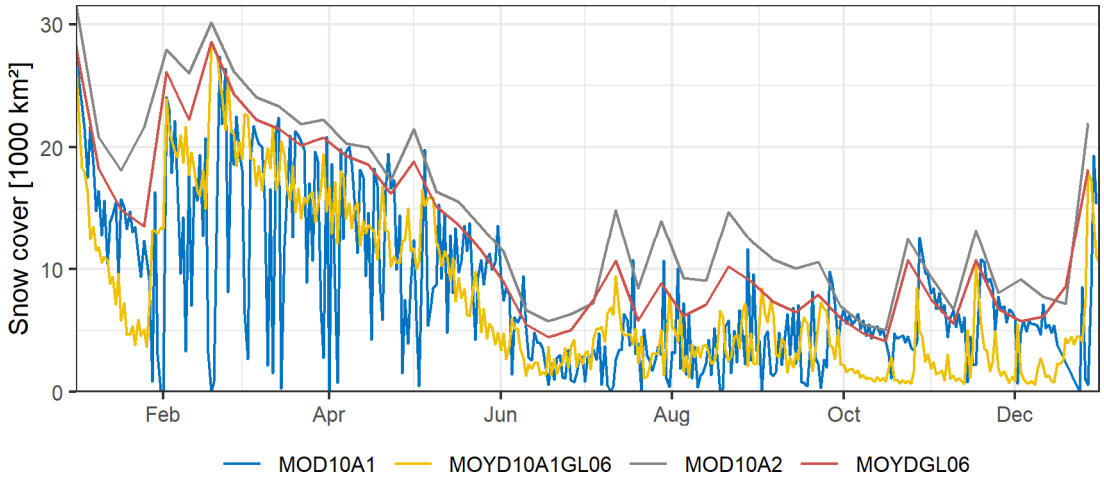


Figure 4.12: Comparison of the satellite-based snow extent estimation over the Karnali river for the year 2003. The compared products are the daily products MOD10A1 (Hall et al., 2016) and MOY10A1GL006 (Muhammad & Thapa, 2020), and the 8-day composites MOD10A2 (Hall & Riggs, 2021) and MOYDGL06 (Muhammad & Thapa, 2019).

4.3.3.3 Discharge data

Discharge data is available for three stations in the Karnali catchment; Station 280 is the discharge of the Karnali River at the mountain outlet (catchment size 45,600 km²); Station 265 is the Chera river which drains a subbasin of 7,100 km² in the Eastern parts of the catchment; and Station 251 records the West Seti River which drains a small subbasin in the Eastern middle hills of the mountain (157 km²). The discharge data is the mean daily discharge and covers the period 2001 – 2015 (Stations 280 and 265), and 2001 – 2014 (Station 251). The data was obtained from the Department of Hydrology and Meteorology Nepal (DHM). It was generated from stage observations at 08:00, 12:00 and 16:00 for each day using stage-discharge rating curves by DHM. The records are characterised by large flow ranges whereas the maximum flow exceeds the median flow by the factors 28 at station 280, 25 at station 265, and 55 at station 251 (Figure 4.13).

Stage-discharge measurements are available for station 280 at the mountain outlet. This dataset comprises 410 observations that are measured at the same point in time and used by DHM to generate rating curves and convert the continuous stage measurements into discharge rates. The dataset was provided by DHM and is used to estimate the uncertainty associated with the daily discharge rates for this station. The observations date back to the year 1963 to the year 2013. However, stage-discharge relations frequently change and hence only the 73 samples since 2000 are considered hereafter. The average number of samples per year is 5.2; the maximum number of samples were collected in 2004 (11 samples) and in the years 2005 and 2007 no samples were collected (Figure 4.14).

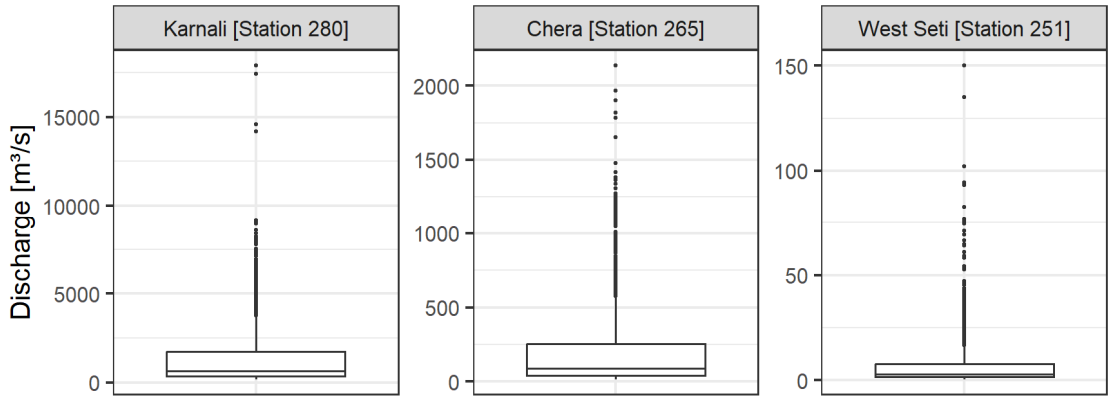


Figure 4.13: Boxplots of the observed discharge at three stations in the Karnali River. Station 280 is the Karnali River catchment at the mountain outlet in Chisapani (45,500 km²); Station 265 is the Chera subbasin (7,100 km²); and Station 251 is the West Seti subbasin (160 km²). The records are in the period 2001 – 2015 for Stations 280 and 265, and 2001 – 2014 for Station 251.

The stage-discharge samples concentrate on low to medium flow conditions. The mean discharge rate is 1,020 m^3/s and the highest discharge of 5,530 m^3/s was measured at a stage of 10.75 m on the 2nd August 2000. This observation captures a high-flow event that exceeds the warning level (10 m) and the danger level (10.7 m) (MacClune *et al.*, 2015) (Figure 4.14). The remaining observations are more than 2 m below the warning level and cover low to medium flow ranges. The highest observation after 2004 was taken in 2008 at a stage of ~8 m and the highest observation after 2009 was taken in 2012 at a stage of 4.9 m. Hence no information about the stage-discharge relationship during high flows is available for recent years. Two outliers are identified at stages ~4.5 m and 8 m which have both corresponding discharge measurements of < 500 m^3/s.

The low sampling density, the concentration on low and medium flows, and the presence of outliers indicate the uncertainty in the discharge data and the confidence interval estimation which are both derived from these observations.

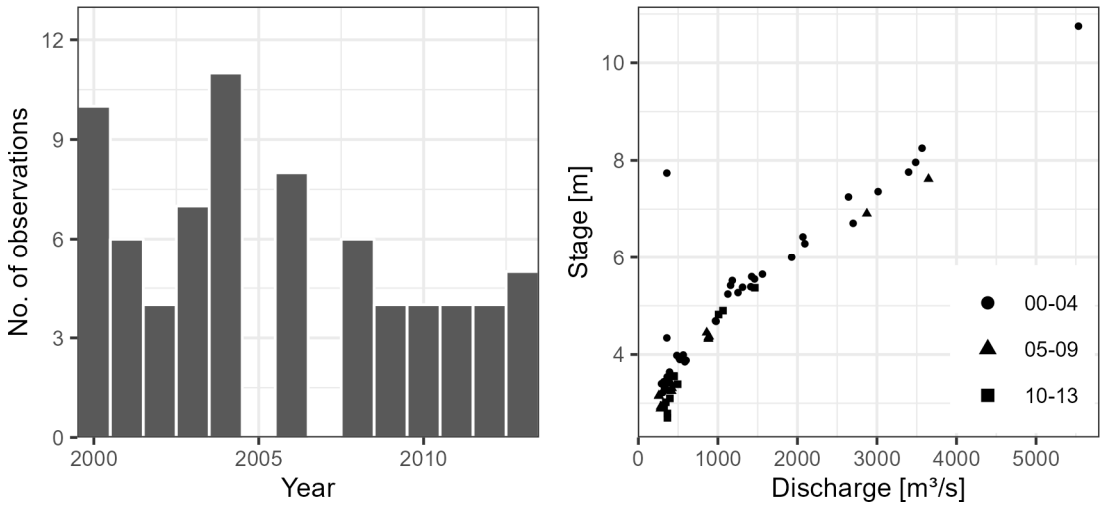


Figure 4.14: The number of stage-discharge samples (left) and the stage-discharge measurements (right). The samples are taken at the mountain outlet of the Karnali River in Chisapani by DHM.

4.4 Results: Predicting the historical catchment hydrology

The hydrological model SPHY is used to simulate the hydrological system of the Karnali River catchment until the mountain outlet for the period 2002 – 2015. The model is calibrated using a multi-criteria approach that includes the discharge, actual evapotranspiration, and snow extent. The uncertainty in the discharge observations is considered in the calibration process. Therefore, the quantification of this uncertainty is presented in the first Section 4.4.1 before the model calibration (Section 4.4.2) and validation (Section 4.4.3). The runoff composition provides information about the hydrological system and is presented in Section 4.4.4. The last Section 4.4.5 then presents the predictions of the flood events.

4.4.1 Estimation of the discharge uncertainty

The uncertainty in the discharge data is estimated from stage-discharge observations at the mountain outlet using the BaRatin model. This requires the detection of shifts in the stage-discharge relationship which result in changes in the cross-section geometry (Section 4.4.1.1). The uncertainty is then estimated from an ensemble of rating curves which are fitted for the periods with stable control (section 4.4.1.2).

4.4.1.1 Temporal analysis of the stage-discharge relationship

In the first step, a log-log transformation is conducted on the stage and discharge measurements to convert the logarithmic stage-discharge relationship to a linear relationship (Herschy, 1993). The variation in the measurements is larger for lower flows < 665 m^3/s (< 6.5 log discharge) than above this threshold (Figure. 4.15). This indicates that controls which are activated at higher stage levels shift less frequently and are, thus, more stable. This is reasonable because these controls are activated less frequently.

Therefore, the identification of temporal changes in the stage-discharge relationship is conducted for the lower and higher flows separately. Since the log-log transformed stage-discharge relationship follows a linear trend, it is possible to divide log(discharge) by log(stage) to obtain a measure that indicates the stage-discharge relationship. The unequal variance t-test (also Welch-test) is performed on this measure of different years. A shift in the stage-discharge relationship is indicated by a rejection of the null hypothesis with high confidence (p-value ≤ 0.05). Four shifts are identified for the period 2000 – 2013. Three shifts are detected for the low-flow range and one for the high-flow range (Table. 4.4).

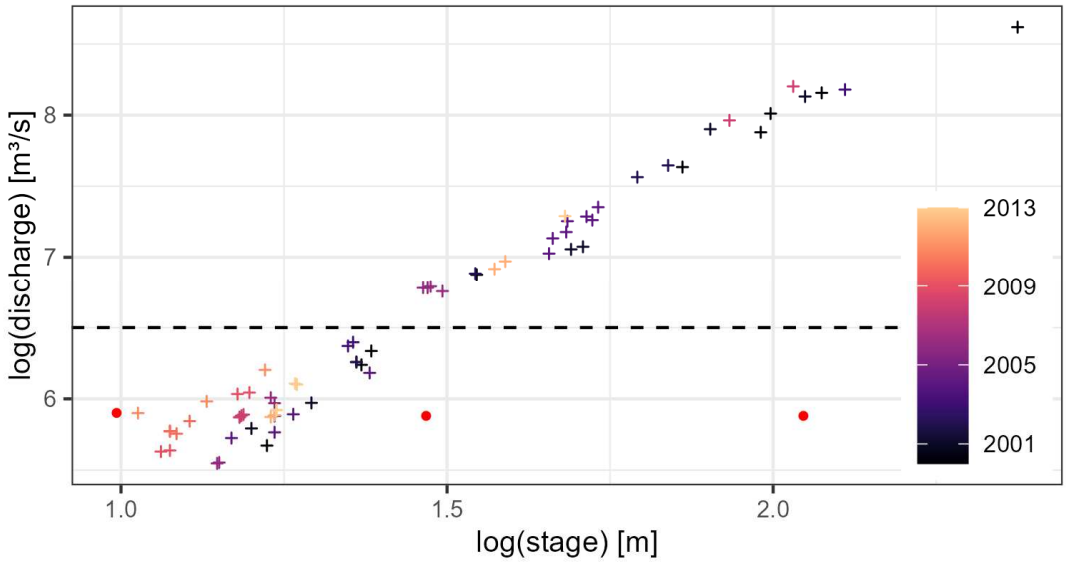


Figure 4.15: Logarithmic stage-discharge relationship for the observations between 2000 - 2013. The dashed line divides the samples into lower flows and higher flows for which temporal shifts are investigated separately. The red dots indicate outliers that are removed from the processing.

Table 4.4: Unequal variance test statistics of groups that are classified as different in their distributions of log(stage) / log(discharge).

Group 1	Group 2	t-value	p-value	Flow
2000 - 2004	2005 - 2008	-6.3	<0.05	low flow (< 650 m^3/s)
2005 - 2008	2009 - 2011	-7.7	<0.05	low flow (< 650 m^3/s)
2009 - 2011	2012 - 2013	12	<0.05	low flow (< 650 m^3/s)
2009 - 2011	2012 - 2013	12	<0.05	low flow (< 650 m^3/s)
2000 - 2004	2006 - 2013	-3.1	<0.05	high flow (≥ 650 m^3/s)

Four stage-discharge datasets are prepared based on the identified temporal shifts in the control. These datasets combine lower flow and higher flow observations of different periods to maintain as many of the sparse high flow observations as possible. These four datasets are:

i) 2000 – 2004 (low flow) and 2000 – 2004 (high flow);
ii) 2005 – 2008 (low flow) and 2006 – 2013 (high flow);
iii) 2009 – 2011 (low flow) and 2006 – 2013 (high flow);
iv) 2012 – 2013 (low flow) and 2006 – 2013 (high flow).

It is worth mentioning that the Welch test does not identify any temporal shifts except between 2004 and 2006 (no measurements were conducted in 2005) for the high flow control. This does not mean that no shifts occurred, just that there is no statistical evidence for a shift which may also be caused by the low sampling density after 2004 after which only 30% of the samples were collected.

Preliminary rating curves are generated for the four datasets and all observations that deviate more than 10% from the simulated value are considered outliers and removed from the datasets (McMahon and Peel, 2019). The observations are classified as outliers and removed from the rating curve generation (Figure 4.15).

4.4.1.2 Curve fitting and uncertainty estimation

The uncertainty in the discharge data is estimated from an ensemble of rating curves (n = 1,000). These rating curves are fitted to each of the four identified stage-discharge datasets and are presented in Figure 4.16.

The median predictions of the rating curve ensembles of RC2 and RC4 are similar. The discharge predicted by the ensemble median of RC1 is lower compared to the other rating curve ensembles above a stage of 3.25 m (300 – 400 m^3/s) (Figure 4.16 A). The largest difference is observed for RC3 which predicts larger discharges than the other rating curves above a stage of 8 m (3,600 – 3900 m^3/s). The slope of RC3 is steeper than the ones of the other ensembles, particularly for stages ≥ 8 m, and hence higher discharge rates are predicted and this difference increases with the stage. RC 3 predicts a discharge of 8,400 m^3/s at the warning level at stage 10.7 m, whereas this is with 7,300 – 7,700 m^3/s lower for the remaining rating curves (median predictions).

The prediction interval (the predictions at the 2.5th and 97.5th percentiles of the rating curve ensembles) are similar for RC2, RC3, and RC4 but are lower for RC1 for all stages (Figure 4.14 A). The absolute difference in the predicted discharges between RC1 and RC2-4 increases with the stage. At the danger level (stage 10.75 m), the lower boundary ($P_{2.5}$) of RC1 is 600 – 800 m^3/s lower than the ones of RC2-4, and the upper boundary ($P_{97.5}$) is 1,100 – 1,200 m^3/s lower than the ones of RC2-4. This difference increases to 2,000 – 2,500 m^3/s and 3,000 – 3,500 m^3/s at stage 17.5 for $P_{2.5}$ and $P_{97.5}$, respectively. The predictions of the upper boundary of RC1 are similar to the median predictions of RC3 above stages ≥ 8 m.

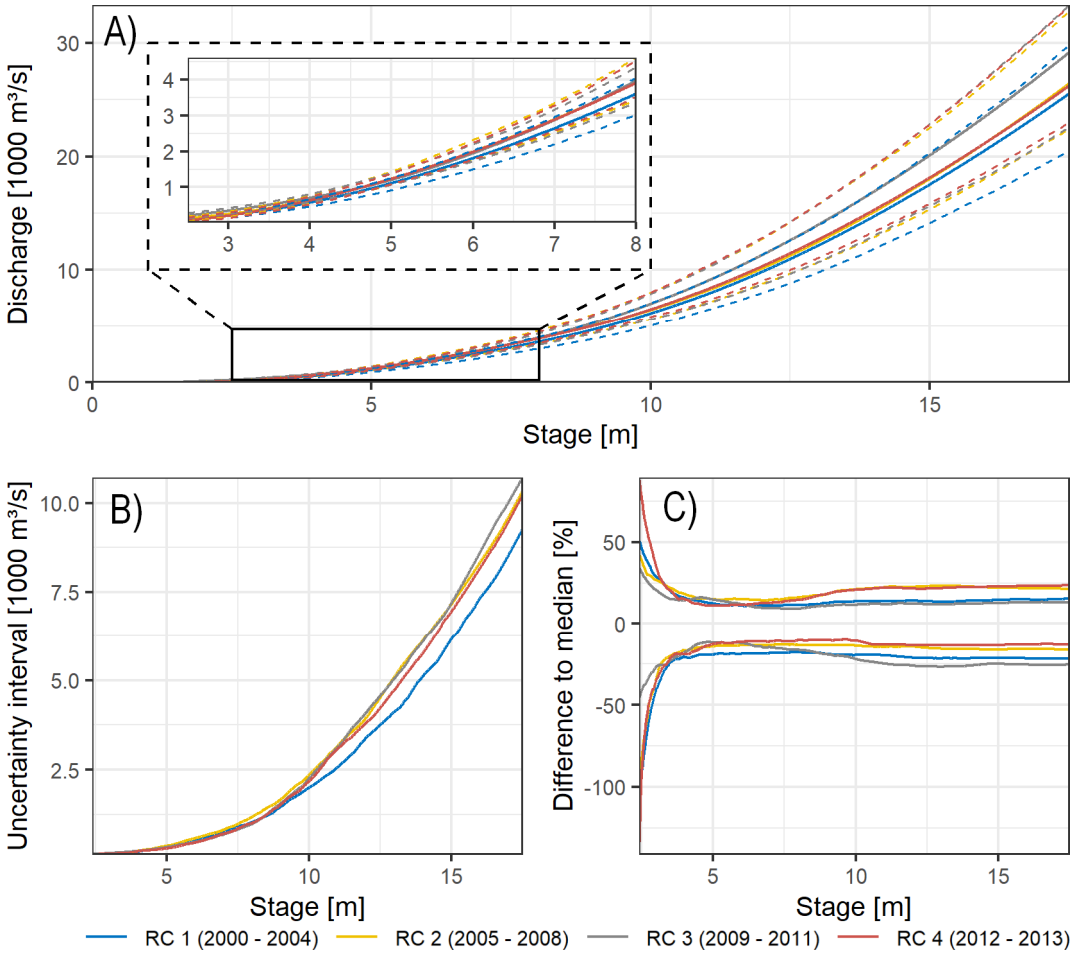


Figure 4.16: A) The predictions of the rating curve ensembles. The solid lines indicate the median and the dashed lines the $P_{2.5}$ and $P_{97.5}$ confidence intervals (n = 1000). B: Absolute range of the predicted discharges ($P_{97.5} – P_{2.5}$) with stage. C: The %-difference between the $P_{2.5}$ and the median predictions (negative values), and between the $P_{97.5}$ and median predictions (positive values).

The percentage difference between the $P_{2.5}$ and $P_{97.5}$ and the median rating curves is highest for extremely low flows (stage < 2.5 m) and then decreases to ± 25 – 30% for the medium to high flow range (stage 5 – 10 m). The uncertainty range increases for flood flows (> 10 m) to ± 35 – 40% but it remains stable with the increase in the stage. The %-range is similar for all rating curves but for RC1 and RC3 the uncertainty leans towards the lower bound whereas it leans to the higher bounds for RC2 and RC4 (Figure 4.16 C). However, the absolute uncertainty range (m^3/s) is similar for RC2 – RC4 for stages ≥ 10 m but is lower for RC1 (Figure 4.16 B).

The discharge uncertainty (the predictions at $P_{2.5}$ and $P_{97.5}$ of the rating curve ensembles) is transferred to the discharge observations. However, the stage data is not available for the complete modelling periods. Therefore, the observed discharge is related to the median discharge of the rating curve predictions to transfer the discharge uncertainty. The median discharge of the rating curves is consistently lower than the DHM discharge observations and this difference increases with the flow rate (Figure 4.17). This indicates that the rating curve used by DHM is steeper than the median BaRatin one, and that uncertainty estimation is distorted by the uncertainty transfer using the discharge observations.

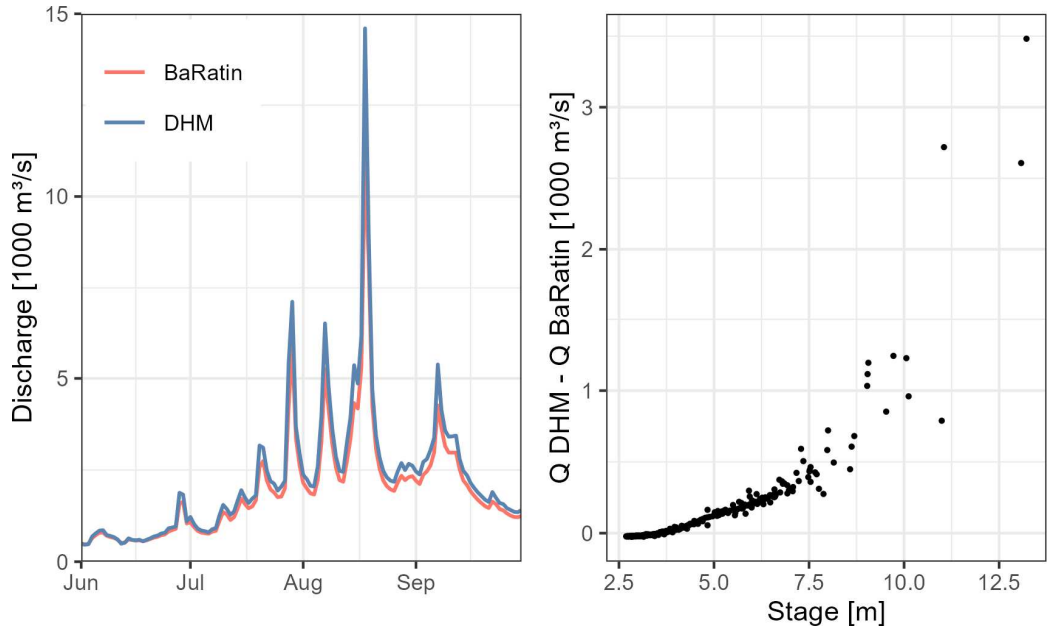


Figure 4.17: Comparison of the observed discharge (DHM) and the estimated discharge from the median rating curve prediction (BaRatin). Left: Hydrograph of the monsoon season 2009. Right: Difference between both discharge datasets with the stage observations of the corresponding date.

4.4.2 The calibration of the hydrological model

The SPHY model is calibrated for the period 2002 – 2006 within a GLUE framework to account for the model equifinality and quantify the uncertainty of the predictions. These predictions are validated for the period 2007 – 2015 to assess the performance of the model in reproducing the temporal variations of different hydrological variables. The results of the calibration process are presented in the next section including the identification of sensitive parameters and the selection of behavioural parameter sets. Section 4.4.2.2 validates the hydrological simulations from the comparison of the simulated and observed catchment behaviour.

4.4.2.1 The calibration process

A hybrid calibration approach which combines manual and automated calibration procedures is applied in the initial calibration phase to reduce the parameter ranges for further calibration. In the manual calibration, a sensitivity analysis is conducted where a single parameter value is varied to gain an understanding of the effects of the parameter on the simulated hydrographs. Afterwards, an iterative automated calibration is conducted with a Regional Sensitivity Analysis (RSA) to identify sensitive parameters which are maintained in the further calibration process and identify the parameter space of behavioural model realisations.

Identification of sensitive parameters

The RSA is conducted for 21 parameters and 1,500 parameter combinations to assess the sensitivity and eliminate insensitive parameters from the further calibration process. Each parameter set is classified into behavioural and non-behavioural models based on five different performance measures (NSE, R^2, $PBIAS_{runoff}$, $PBIAS_{snow}$, $PBIAS_{actET}$). The cumulative frequency curves of the parameter values of the behavioural and non-behavioural classes are then plotted. The sensitivity of a parameter increases with the split between the curves. The thresholds for the classification into behavioural and non-behavioural models were pragmatically chosen based on balanced group sizes to enhance the interpretability of the sensitivity plots.

The RSA identifies 13 parameters (parameters are highlighted by italic fonts) as sensitive to one or more performance measures (Figure 4.18). The parameter sensitivity depends on the performance measure; the groundwater parameters *alphaGW*, *deltaGW*, and *gwSat* control the amount and timing of the baseflow and are sensitive towards NSE and R^2. The parameters *kEff* and *lambda* affect the infiltration capacity and, thus, the generation of surface runoff and are also classified as sensitive. However, *kEff* is sensitive towards NSE, R^2, $PBIAS_{runoff}$, and $PBIAS_{actET}$, whereas *lambda* only influences the performance of R^2. The *alphaInf* parameter determines the fraction of daily rainfall that falls during the hour of most intense rainfall and, thus, affects the amount of overland flow, and is sensitive towards R^2 and $PBIAS_{runoff}$. The *recessCoef* parameter determines the speed of the transport in the channel network and is highly sensitive toward NSE and R2. The precipitation-correction-factor *precipFactor* affects how much water enters the catchment and is sensitive towards the NSE and $PBIAS_{runoff}$. The scaling factor of the crop coefficients (*kcFactor)* affects the actual evapotranspiration and is highly sensitive towards $PBIAS_{actET}$.

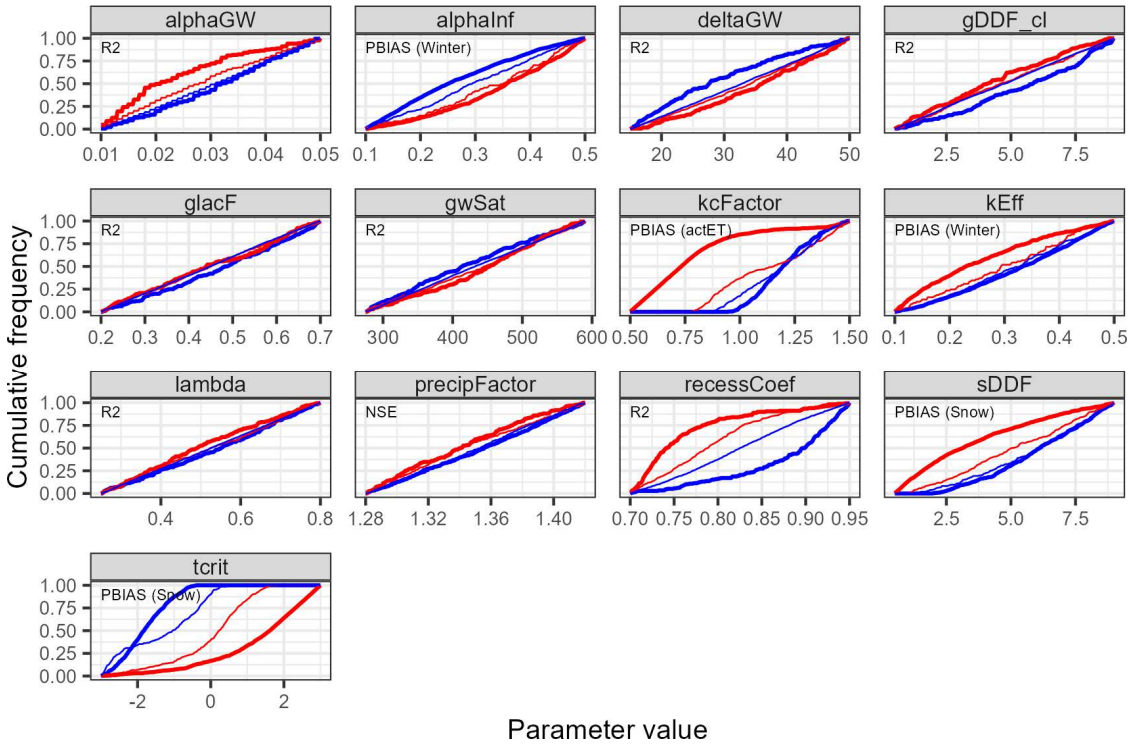


Figure 4.18: Regional Sensitivity Analysis (RSA) of the parameters identified as sensitive to at least one performance measure. The red line indicates the cumulative frequency of the non-behavioural realisations and the blue line is the cumulative frequency of the behavioural simulations. The performance measure of the RSA is shown in the top-left corner of each parameter. The RSA of all parameters and performance measures is included in the Appendices 4.3 – 4.6.

The RSA also classifies four parameters of the snow and glacier modules as sensitive. The Degree-Day-Factor (DDF) for clean ice glaciers (*gDDF_cl*) controls the melt rates and is sensitive towards the NSE and R^2. However, the DDF for debris-covered glaciers (*gDDF_db*) is not classified as sensitive which can be explained by the smaller fraction of this glacier type. The *glacF* factor controls the amount of meltwater that percolates to the groundwater storage and is slightly sensitive towards R^2. The DDF of the snow module (*sDDF*) is highly sensitive towards $PBIAS_{snow}$ but the RSA also indicates a moderate sensitivity towards $PBIAS_{actET}$. The *tcrit* parameter defines the temperature threshold for snowfall and is highly sensitive towards $PBIAS_{snow}$, and moderately sensitive towards R^2.

The remaining 11 parameters are classified as insensitive from the RSA and are, thus, excluded from the further calibration process in the GLUE framework.

Generalized Likelihood Uncertainty Estimation

The final model calibration is conducted in a Generalized Likelihood Uncertainty Estimation (GLUE) framework with 10,000 parameter sets for the 13 sensitive parameters. The selection of behavioural parameter sets is conducted in three steps: firstly, the removal of non-behavioural parameter sets based on $PBIAS_{winter}$, $PBIAS_{snow}$, and $PBIAS_{actET}$; secondly, the identification of behavioural parameter sets based on the observed discharge at the catchment outlet using NSE, $PBIAS_{high}$, and $eGLUE_{mod}$ and; the selection of the highest *n* parameter sets of each performance measure, whereas *n* is the number of behavioural parameter sets of the efficiency with the lowest number of behavioural parameter sets in the final step (see Section 4.2.2.4).

In the first step, 2,462 of 10,000 parameter sets are classified as non-behavioural and removed from the selection process. This parameter removal is based on the performance of $PBIAS_{winter}$ because all parameter sets are within the ± 30% interval of the $PBIAS_{snow}$ and $PBIAS_{actET}$ (Figure 4.19). The analysis found that 6,138 of the remaining parameter sets exceed the NSE threshold (0.75). There are 25 parameter sets within the threshold boundary of $PBIAS_{high}$ (± 25%), and hence the 25 highest parameter sets of each of NSE, $eGLUE_{mod}$, and $PBIAS_{high}$ are selected for the simulations. This results in the selection of 64 parameter sets after the removal of duplicates. The parameter ranges of these 64 models are presented in Table 4.5.

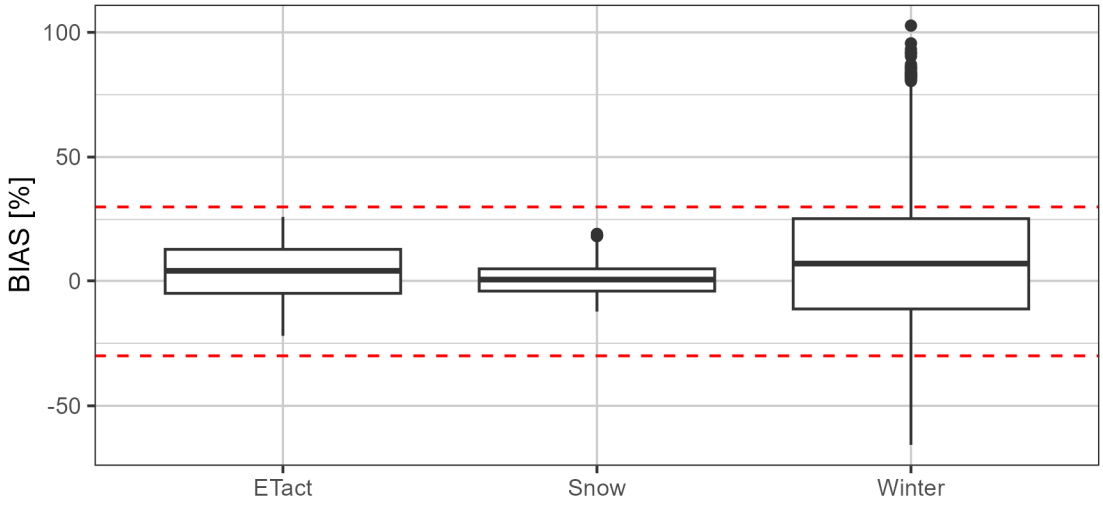


Figure 4.19: Efficiencies of all parameter sets (n = 10,000) regarding the performance measures $PBIAS_{winter}$, $PBIAS_{snow}$, and $PBIAS_{actET}$. The red lines indicate the 30% threshold for the classification as a non-behavioural parameter set.

Table 4.5: The parameter range of the sensitive parameters of the ensemble of the 64 behavioural parameter sets.

Module	Parameter	Description	Unit	Parameter value		
				Min	Max	Median
Glacier	gDDF_cl	Degree-day factor debris-free glaciers	mm ºC^{-1} day^{-1}	1.55	8.90	6.71
	gDDF_db	Degree-day factor debris-covered glaciers	mm ºC^{-1} day^{-1}	1.80	8.94	6.89
	glacF	Fraction of glacier melt that becomes glacier runoff	-	0.31	0.60	0.49
Ground-water	alphaGW	Baseflow recession coefficient	day^{-1}	0.01	0.04	0.03
	deltaGW	Groundwater recharge delay time	day	20	46	26
	gwSat	Saturated soil water content of groundwater layer	mm	325	500	390
	h_gw	Initial groundwater table height	m	1.10	2.10	1.60
Routing	recessCoef	Flow recession coefficient	-	0.75	0.95	0.84
Snow	sDDF	Degree-day factor for snow	mm ºC^{-1} day^{-1}	2.32	8.97	5.72
	tcrit	Temperature threshold for precipitation to fall as snow	ºC	-3.5	-0.6	-2.4
Soil	alphaInf	Fraction of daily precipitation that occurs during the hour of most intense rainfall	-	0.16	0.40	0.34
	lambda	Infiltration coefficient that affects the speed of infiltration capacity reduction	-	0.26	0.79	0.44
	kEff	Effective saturated hydraulic conductivity	-	0.20	0.50	0.24
Correction factors	precipFactor	Precipitation correction factor	-	1.35	1.42	1.40
	kcFactor	Crop coefficient multiplication factor	-	0.80	1.19	0.85

4.4.3 The validation of the hydrological simulations

The performance of the hydrological ensemble is evaluated by comparing the simulated catchment behaviour with the observed one. This evaluation includes the discharge at the catchment outlet and two subbasins, the snow extent, and the actual evapotranspiration (ET_{act}).

4.4.3.1 Catchment discharge

The simulated discharge at the catchment outlet reproduces the observed seasonality during the calibration and validation periods as indicated by the hydrograph comparison (Figure 4.20). The seasonal dynamics are captured with low flows during the Non-monsoon season (Oct – May) and high flows during the monsoon season (Jun – Sep). The timings of the rising and recession limbs are predicted well by the ensemble. This high agreement of simulated and observed discharge rates is also indicated by the high NSE and R^2 efficiencies; the NSE of the ensemble members ranges between 0.75 – 0.87 and 0.71 – 0.85 for the calibration period (2002 – 2006) and the validation period (2007 – 2015), respectively. The median NSE of the ensembles is 0.85 and 0.82, for both periods respectively. The R^2 of the ensemble ranges between 0.77 – 0.87 and 0.80 – 0.86 and the median R^2 is 0.85 and 0.84 for the calibration and validation period, respectively (Table 4.6).

The visual inspection of the hydrographs (Figure 4.20) indicates some systematic deviations between the simulated and observed discharge. The discharge peaks in the first six months of the year are overestimated consistently throughout the calibration and validation periods by all ensemble members. This includes the first month of the monsoon season (Jun) and hence the peaks of the rising limb are consistently overestimated. These overestimations lead to positive PBIAS values for most months in the first half of the year (Figure 4.21). The PBIAS for May is 10% in the calibration period and 23% in the validation period. The highest monthly PBIAS is simulated in June and reaches 58% in the calibration period and 47% in the validation period. The BIAS decreases afterwards and the discharge is underestimated in August by 21% in the calibration period but reproduced well in the validation period (PBIAS: -3%). The PBIAS is low (within ± 10%) at the end of the monsoon season in September. The post-monsoon season (Oct – Nov) discharge of the recession limb is overestimated in both periods, and this gradually declines and turns into an underestimation at the start of the winter season in December.

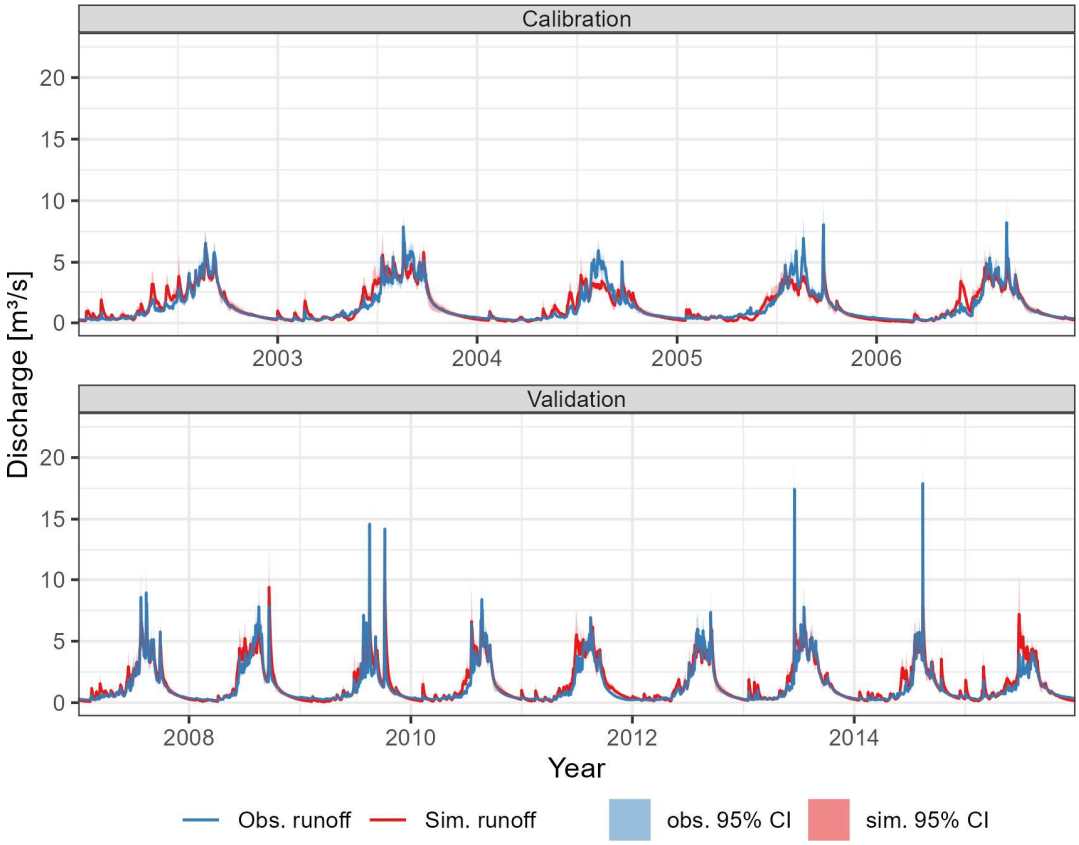


Figure 4.20: Hydrographs of the simulated and observed discharge at the catchment outlet. The solid lines indicate the observed discharge and the ensemble-median predictions (red). The shaded areas indicate the 95th confidence intervals of the observations and the simulations.

Table 4.6: Statistics of the performance measures of the 64 behavioural parameter sets for the calibration period (2002 – 2006) and the validation period (2007 – 2015). The P_{10} and P_{90} are the efficiencies of the 10th and 90th percentile of the ensemble.

		Min	P_{10}	Median	P_{90}	Max
NSE	Calibration	0.75	0.78	0.85	0.86	0.87
	Validation	0.71	0.743	0.82	0.84	0.85
R^2	Calibration	0.77	0.79	0.85	0.86	0.87
	Validation	0.80	0.82	0.84	0.85	0.86
PBIAS	Calibration	-6.8	-1.84	3.15	6.57	9.6
	Validation	4.2	8.82	13.9	17.21	20.2
eGLUE	Calibration	18.8	28.5	38.6	47.6	51.7
	Validation	22.3	28.7	35.6	44.9	47.5
$eGLUE_{mod}$	Calibration	318	419	533	582	600
	Validation	408	470	540	599	620

The performance of the hydrological predictions varies between the years. Overall, the timing and shape of the monsoon season hydrographs are depicted well but for the years 2008, 2011 and 2015 the flows in the early monsoon season are overestimated (Figure 4.20). On the contrary, the monsoon season flows of the years 2004 and 2005 are vastly underestimated. The transition from high to low flows in the post-monsoon season is well timed for all years except 2011 where the lag between simulated and observed recession is around 10 days. In the following months, the discharge is consistently overestimated, which is only observed for this particular period. For all other years, the low flow during winter (Dec – Feb) is underestimated except for the discharge peaks. This underestimation is most prevalent in the winters of 2005 – 2006, 2008 – 2009, and 2015.

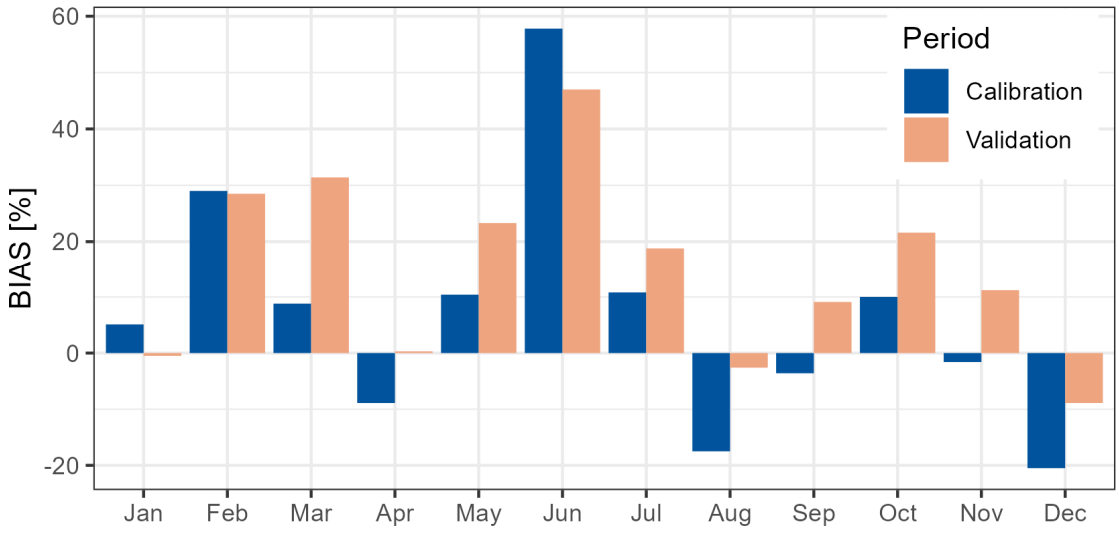


Figure 4.21: The monthly PBIAS of the simulated discharge (median) and the discharge observations for the calibration and validation periods at the catchment outlet.

The variation of the predictions of river flow between the calibration and validation periods is low for the model ensemble. This manifests in the very similar median NSE and R^2 performance measures which only deviate by 0.03 and 0.01 between the calibration and validation periods (Table 4.6). However, several differences are observed; the simulated and observed discharge sums are almost similar and only deviate by 3% (PBIAS) for the calibration period. These differences increase in the validation period to an overall overestimation of 14% (Table 4.6). On the contrary, high flows are underestimated whereas the underestimations decrease during the validation period. The $eGLUE_{mod}$ performance increases slightly in the validation period, while the eGLUE performance decreases by 3%. Therefore, the increase in $eGLUE_{mod}$ results from higher flows for the days where the simulated discharge falls within the observed confidence intervals while the number of such days decreases which is indicated by the decrease in eGLUE (Table 4.6).

The ensemble variation is higher for the upper percentiles and the low-flow seasons. The ensemble-median discharge is 3,150 m^3/s in the monsoon season, the $Q_{2.5}$ discharge is 84% of the median discharge, and the $Q_{97.5}$ discharge is 118% of the median discharge (Table 4.7). The variation of the upper percentiles is higher for all seasons except for the post-monsoon season where it is 25% higher and the lower percentiles 28% lower. The variation relative to the median flow decreases with the flow rate from 84 – 118% during the monsoon season to 63 – 164% during the winter. However, the absolute variation in the ensemble is lowest during the low flow seasons in the winter and pre-monsoon seasons.

Table 4.7: The seasonal prediction range of the modelling ensemble. The uncertainty interval of the observed discharge is presented for comparison in the brackets. Q_n is the discharge predicted for the n^{th} percentile of the ensemble.

	Median discharge [m^3/s]	**$Q_{2.5}$ relative to Q_{50} [%]**	**$Q_{97.5}$ relative to Q_{50} [%]**	**Range [%]**
Pre-monsoon	546 (481)	73 (78)	143 (123)	70 (45)
Monsoon	3,150 (2,917)	84 (86)	118 (115)	34 (29)
Post-monsoon	1,110 (974)	72 (84)	125 (116)	53 (32)
Winter	365 (355)	63 (77)	164 (126)	101 (49)

4.4.3.2 Discharge of the Chera and West Seti subbasins

The model ensemble also depicts the general hydrological dynamics in the subbasins of the small West Seti (157 km^2) and the medium-sized Chera (7,076 km^2) (Figure 4.22). As for the catchment outlet, the rising limb is overestimated for both subbasins for the calibration and validation period. In the West Seti subbasin, the PBIAS of May is 75% in the calibration period while it is -1% for the validation period (Figure 4.23). In June, the PBIAS of the validation period exceeds the one of the calibration period by 64% and reaches 116%. For the Chera subbasin, the monthly PBIAS increases throughout the first half of the year, reaching the peak of 68% in June. In the following months, the hydrographs show a contrary behaviour between the subbasins; the monsoon season flows are overestimated with a decreasing trend towards the post-monsoon season and the low flows in winter are underestimated in the West Seti subbasin. In the Chera subbasin, the monsoon season flows are increasingly underestimated until the start of the post-monsoon season, in which the flows are overestimated (Figure 4.23).

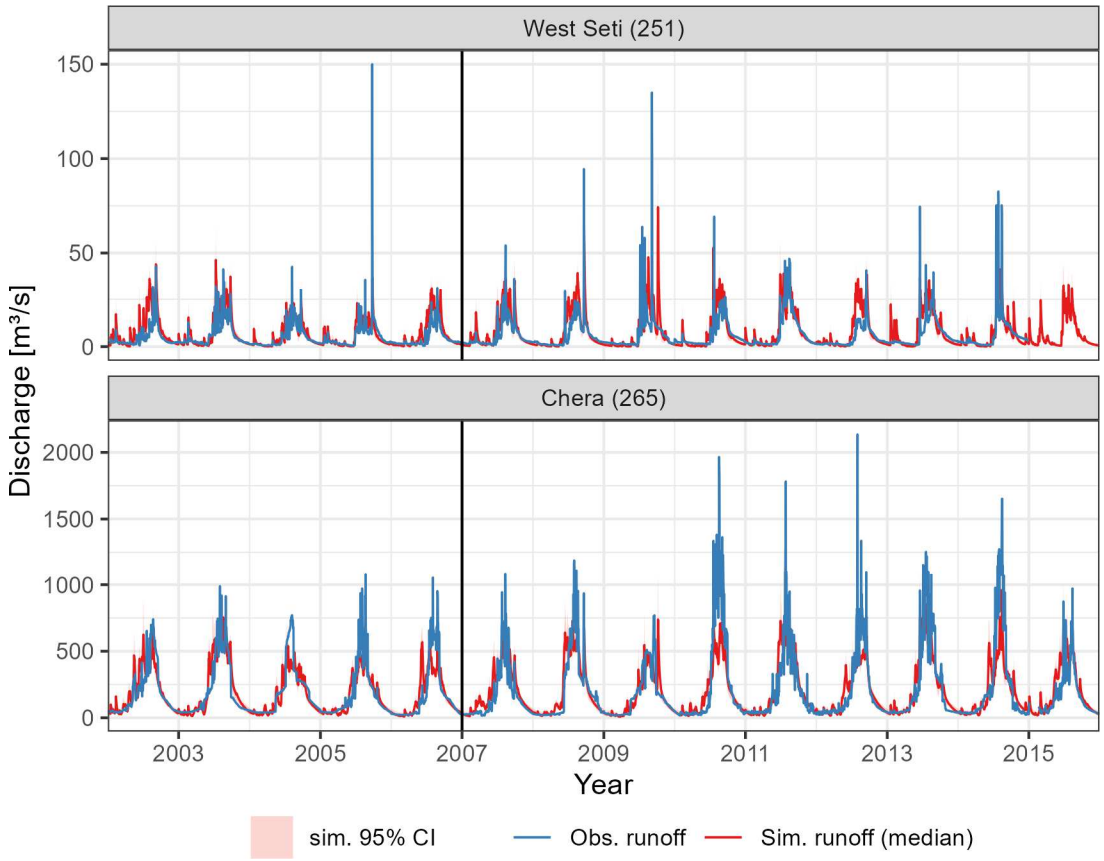


Figure 4.22: Hydrographs of the West Seti and Chera subbasins. The solid lines indicate the observed runoff (blue) and the median simulated runoff (red). The shaded area indicates the prediction interval of the $P_{2.5}$ and $P_{97.5}$ of the ensemble, and the vertical black line indicates the break between the calibration and validation period. The discharge observations of the West Seti subbasin are not available for the year 2015. The number in the bracket is the gauge identification of DHM.

The performance measures for the Chera and West Seti subbasins are presented in Table 4.8. The model ensemble performs better for the larger Chera subbasin. The median R^2 is 0.70 and 0.68 in the calibration and validation periods for the Chera, and the median R^2 is 0.51 in both periods for the West Seti. The NSE indicates a good performance of the ensemble for Chera in both periods (0.70 and 0.65), whereas it is considerably lower for the West Seti (0.17 and 0.3). The overall PBIAS lies within the ± 5% interval for the Chera basin, while the discharge is overestimated by 41% and 25% in the West Seti subbasin for the calibration and validation period, respectively.

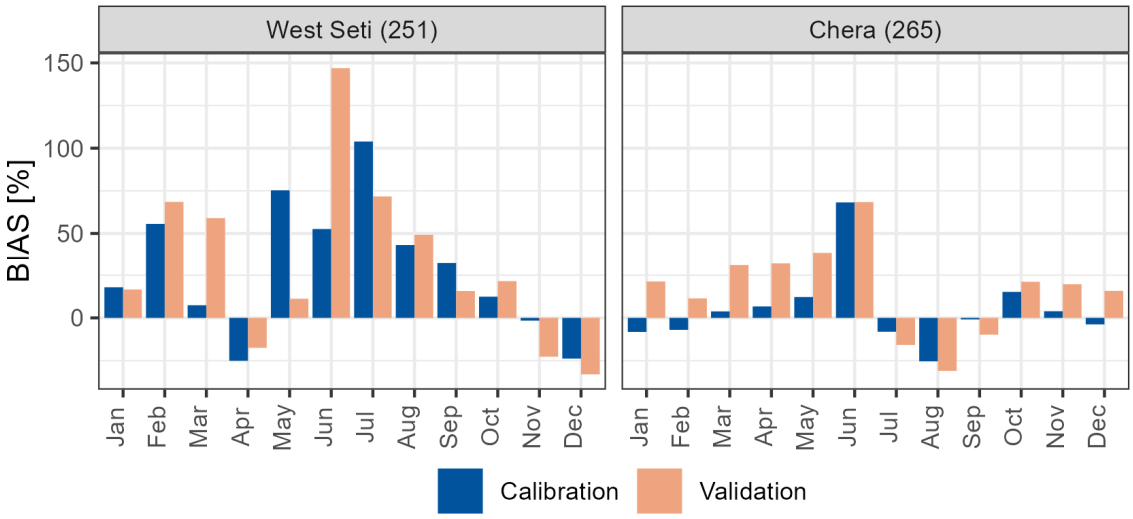


Figure 4.23: Monthly PBIAS of the simulated (median) and the observed discharge for the West Seti and Chera subbasins.

Table 4.8: Efficiencies of the discharge for the West Seti and Chera subbasins.

Subbasin	Efficiency	Period	Min	P_{10}	Median	P_{90}	Max
West Seti	NSE	Calibration	-0.16	-0.1	0.16	0.27	0.32
		Validation	-0.01	0.06	0.3	0.41	0.44
	R^2	Calibration	0.47	0.48	0.51	0.54	0.56
		Validation	0.45	0.46	0.51	0.53	0.55
	PBIAS	Calibration	24.6	35.09	40.8	45.35	46.8
		Validation	13.1	20.75	25.35	29.04	30.3
Chera	NSE	Calibration	0.59	0.64	0.7	0.74	0.76
		Validation	0.57	0.61	0.65	0.68	0.69
	R^2	Calibration	0.6	0.65	0.7	0.75	0.77
		Validation	0.58	0.62	0.68	0.7	0.72
	PBIAS	Calibration	-8.9	-6.24	-0.25	5.27	8.9
		Validation	-11.7	-8.37	-2.85	1.31	5.5

4.4.3.3 The validation of the snow dynamics

The snow module is calibrated and validated against the 8-day snow extent derived from the improved MODIS dataset (hereafter referred to as the observed snow extent) (Muhammad and Thapa, 2020). Grid cells with a snow depth < 5 mm are excluded from the analysis to remove very shallow and infrequent snow cover (Khanal *et al.*, 2021). The snow extent is only an approximation for the agreement of observed and simulated snow dynamics because it lacks snow depth information and is, thus, just a 2-D approximation of the 3-D snow storage.

The snow seasonality is captured by the ensemble throughout the modelling period as indicated by the median R^2 of 0.63 (Table. 4.9). The median PBIAS is -3% for both calibration and validation period, which indicates a very good agreement. The R^2 of the ensemble ranges between 0.58 and 0.65 which indicates that the ensemble models predict a similar seasonality. The PBIAS of the ensemble members ranges between -11% and +12%. Hence, some members underpredict the observed snow extent while other members overpredict it. These performance measures suggest that the differences in the temporal dynamics between the ensemble members are small compared to the variation in the spatial snow extent. Furthermore, only small differences are observed between the calibration and validation periods. The median R^2 is 0.63 in both periods but the lowest R^2 increases from 0.58 to 0.61. The median PBIAS decreases slightly from -3.2% to -3.0%.

The good agreement between the simulated and observed snow extent seasonality is also indicated by the snow extent time series (Figure 4.24) and the monthly PBIAS (Figure 4.26). The maximum snow extent in the winter (Dec – Feb) is well depicted although it is underestimated in 2005 and overestimated in 2012. The timing of the transitional periods of snowmelt (Apr – Jun) and snow accumulation (Sep – Dec) are reproduced by the model. However, the model tends to accumulate snow earlier than observed which leads to positive PBIAS rates in November and December (33 – 47%). The snow extent during the melting season in May and June is overestimated by 13 – 17% (Figure 4.25).

The hydrological ensemble systematically underpredicts the snow extent during the monsoon season (Figures 4.24 and 4.25). The mean observed monsoon season snow extent (calibration and validation period) ranges between 6,400 km^2 and 7,400 km^2. The model underestimates the snow extent by up to 90% and in August only 700 (± 300) km^2 are covered with snow. The observed snow cover is lowest in June but it is variable throughout the season due to the precipitation falling as snow at higher elevations. This

expansion of the monsoon snow extant is not depicted by the modelling ensemble which predicts no expansion until September – October.

Table 4.9: Statistics of the performance of the simulated 8-Day snow extent.

	Period	Min	P_{10}	Median	P_{90}	Max
PBIAS	Calibration	-11.1	-8.2	-3.2	6.7	11.4
	Validation	-11.0	-8.0	-3.0	7.1	11.9
R^2	Calibration	0.58	0.61	0.63	0.65	0.65
	Validation	0.61	0.62	0.63	0.64	0.65

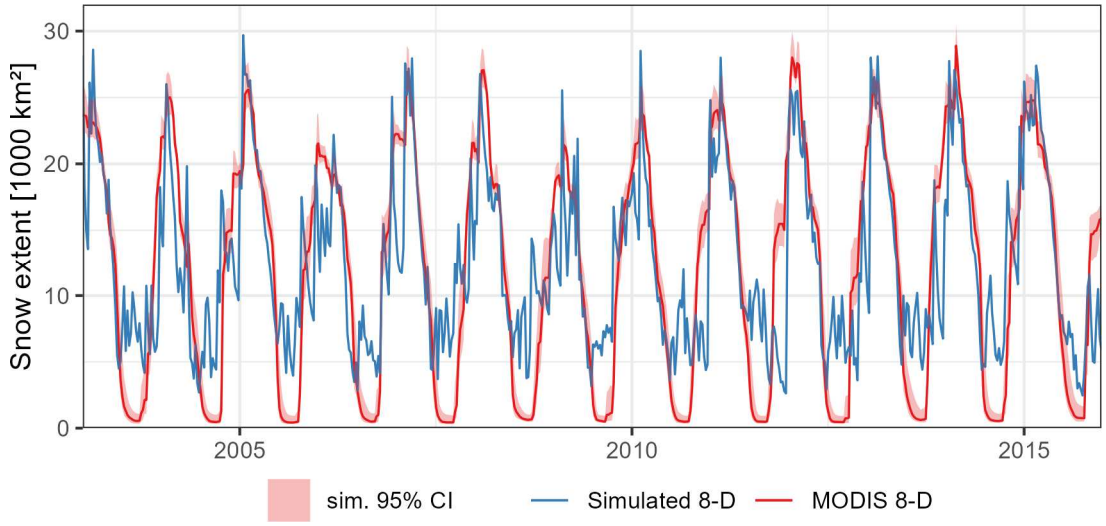


Figure 4.24: Time series of the simulated maximum 8-day snow extent and improved MODIS 8-day snow extent (Muhammad and Thapa, 2020). The shaded area indicates the interval of the $P_{2.5}$ and $P_{97.5}$ of the ensemble predictions.

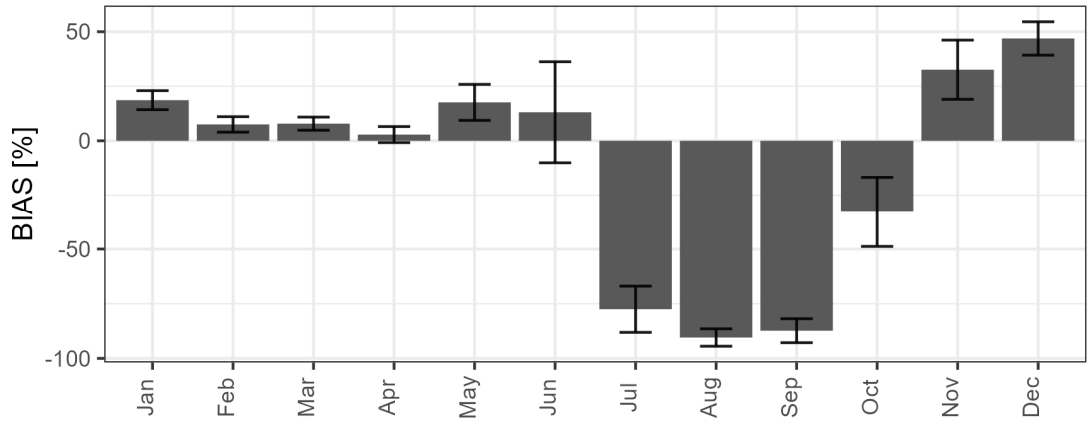


Figure 4.25: Monthly mean PBIAS of the simulated 8-day snow extent. The error bars indicate the standard deviation of the PBIAS arising from differences between the years and the ensemble members.

4.4.3.4 Actual evapotranspiration

The actual evapotranspiration ET_{act} of the modelling ensemble is compared with the ET_{act} estimated from satellite-based annual estimates (MOD16A3GF V006) (Running *et al.*, 2019) which is hereafter referred to as the observed ET_{act}.

The ensemble median underpredicts catchment-mean ET_{act} throughout the modelling period but for most years this underprediction is moderate. The mean annual ET_{act} of the modelling period is 573 mm/m^2 for the ensemble-median while the observed one is 653 mm/m^2. The PBIAS is -11% and -13% in the calibration and validation periods respectively (Table 4.10). The underpredictions of the modelling ensemble are low to moderate for most years (3 – 15%) but particularly for the years 2005, 2014, and 2015 it is underpredicted by 22 – 31% (Figure 4.26). However, the ensemble members at the upper prediction interval overpredict ET_{act} moderately by up to 10% (Table 4.10).

The ET_{act} is underpredicted for most of the land cover classes (Figure 4.27). The highest ET_{act} is simulated and observed for the water class which is almost double than the catchment-mean ET_{act}. Water is also the only class for which the simulated ET_{act} exceeds the observed ET_{act} (median predictions). Further classes with high ET_{act} rates are snow and ice, herbaceous vegetation, and bareland. Sparse vegetation has very low ET_{act} rates in both, the simulated and observed products. The standard deviation is similar between MODIS and SPHY for most classes which indicates that the annual variation is higher than the ensemble variation (Figure 4.27). However, this does not account for the classes with the highest ET_{act} rates for which the ensemble variation is larger than the interannual variation of the MODIS ET_{act}. The largest differences between the simulated and observed ET_{act} rates are observed for the deciduous forest with -13% and -21%, and for irrigated croplands with -12% and -22% in the calibration and validation periods, respectively.

Table 4.10: PBIAS of the annual actual evapotranspiration averaged over the catchment. The reference dataset is MOD16A3GF V006 (Running et al., 2019a). The PBIAS for each land cover class is presented in Appendix 4.7.

Period	Min	P_{10}	Median	P_{90}	Max
Calibration	-20	-15	-11	-6	10
Validation	-22	-18	-13	-8	7

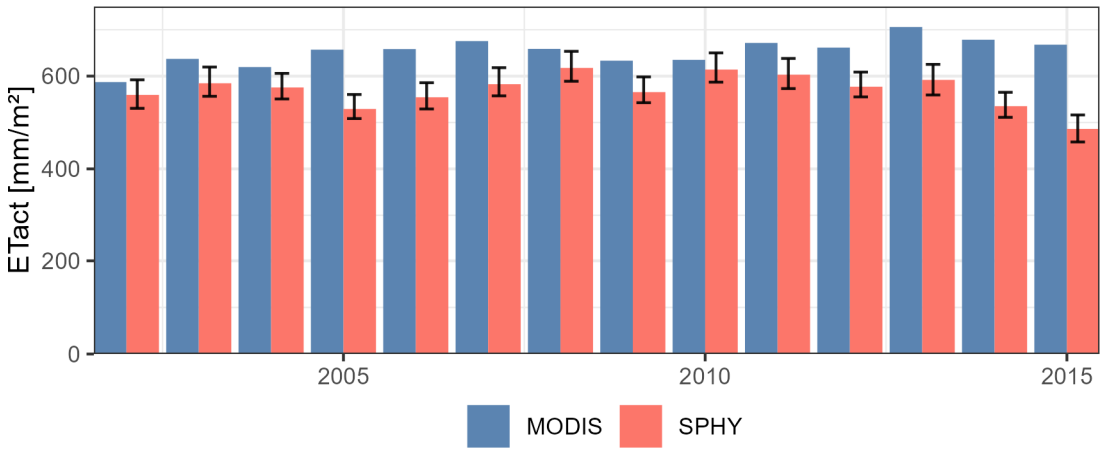


Figure 4.26: Catchment-mean annual actual evapotranspiration rates of the hydrological simulations (SPHY) and satellite estimated (MODIS) (Running et al.*, 2019). The bars show the ensemble median and the error bars indicate the standard deviation of the hydrological ensemble.*

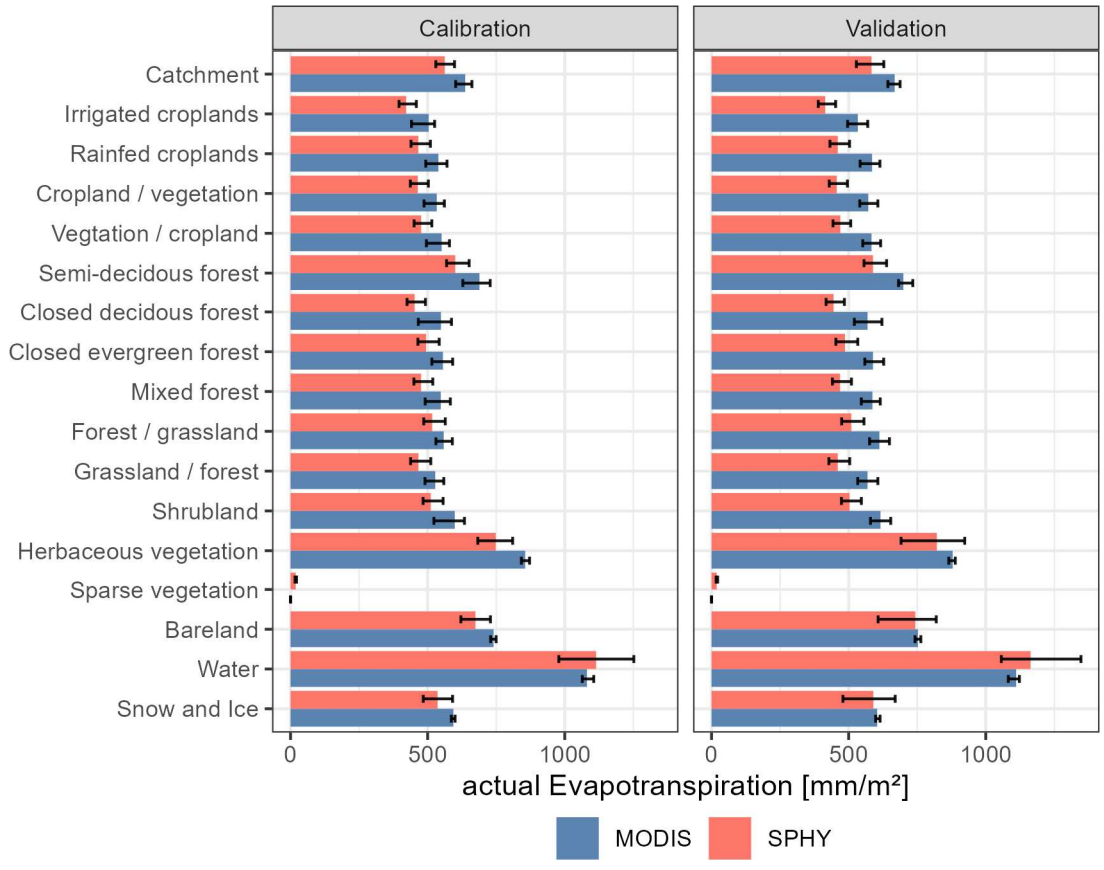


Figure 4.27: Actual evapotranspiration rates for each land cover class of the hydrological simulations (SPHY) and satellite estimated (MODIS) (Running et al.*, 2019). The bars are the annual mean (MODIS) and the annual mean of the median predictions (SPHY). The error bars indicate the standard deviation of the annual variation (MODIS), and the standard deviation of the ensemble predictions and the annual variation (SPHY).*

4.4.4 The runoff composition

The runoff composition provides information about the hydrological system of the Karnali River and the annual and seasonal variations. It provides information about the catchment hydrology which can be used to validate the model performance qualitatively but may also be used to guide water management. This runoff composition deviates between the parameter sets. The knowledge about the hydrology of the Karnali River system is limited and hence each parameter set is an equally likely representation of this system.

The first section 4.4.4.1 presents the overall runoff composition of the calibration and validation periods. It further assesses the impact of the parameters on the composition to gain a better understanding of the model behaviour regarding parameterisation. The later sections 4.4.4.2 and 4.4.4.3 present the annual and monthly composition to gain a better understanding of the effect of non-linearities on flows, and drivers of the flow seasonality.

4.4.4.1 The runoff composition of the calibration and validation periods

Rainfall-runoff dominates the hydrological regime of the Karnali River as it contributes more to the discharge than the other components combined (mean predictions[1]). The mean rainfall-runoff contribution to the total discharge is 57% in the calibration period and 59% in the validation period (Figure 4.28). Baseflow contributes around one-quarter of the total runoff (24% in the calibration period and 23% in the validation period). It is followed by snowmelt with 15% in the calibration period and 14% in the validation period. Glacier melt contributes the least to the total runoff with 4% in the calibration period and 3% in the validation period.

[1] The ensemble median is commonly used in this to describe the predictions of the hydrological ensemble because it is less affected by outliers. However, the ensemble median contributions of the different runoff components do not necessarily add up to 100%, and hence the ensemble mean is used to describe the runoff composition.

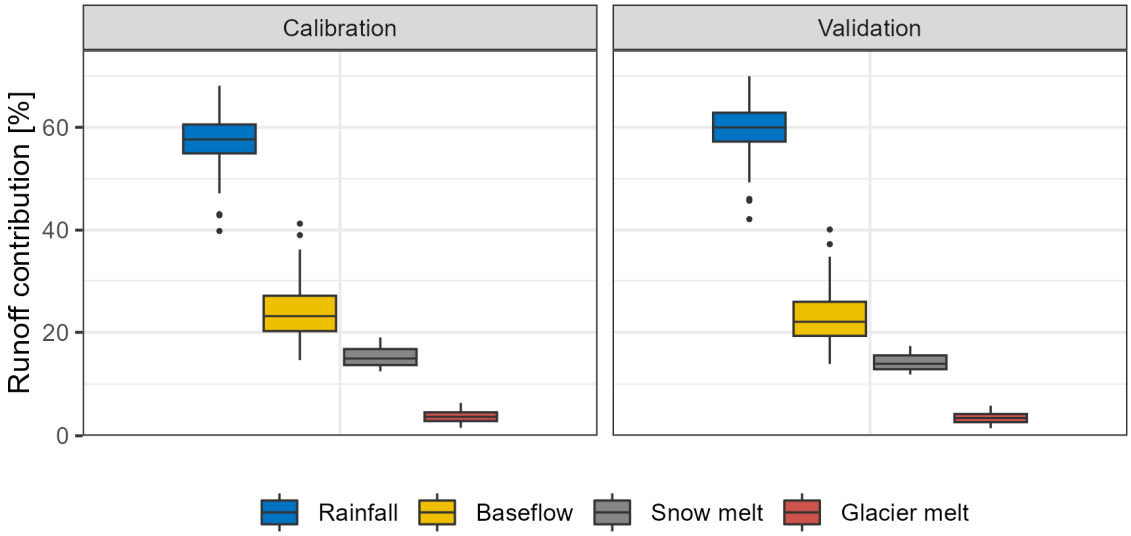


Figure 4.28: The composition of the total discharge leaving the catchment for the calibration period (2002 - 2006) and the validation period (2007 - 2015).

The ensemble members differ mainly in their allocation of water to rainfall-runoff and baseflow (Figure 4.28). The rainfall-runoff contribution ranges from 40 – 70% and the baseflow contribution ranges between 15 – 41%. The ensemble members which simulate the lowest rainfall-runoff contribution simulate the highest baseflow contribution, and the members predicting very high rainfall-runoff contributions (≥ 67%) predict the lowest baseflow contributions (≤ 17%). In this light, the ensemble variation of the snowmelt (12 – 19%) and glacier melt runoff (1 – 6%) is low.

The ensemble variation is driven by differences in the rainfall-runoff and baseflow allocation, particularly by differences in those parameters that control the infiltration of rainfall. The relationship between the parameter value and the runoff contribution is presented in Figure 4.29. The ensemble members that vary in the rainfall-runoff and baseflow allocations differ distinctively in the *alphaInf* parameter which describes the fraction of the daily rainfall that occurs during the hour of most intense rainfall whereas the contributions increase with the parameter value.

High baseflow contributions are associated with the *kEff* parameter which defines the saturated hydraulic conductivity in the soil, and the *deltaGW* parameter which controls the groundwater recharge delay. The baseflow contributions increase with increasing *kEff* and decreasing *deltaGW* parameter values (Figure 4.29).

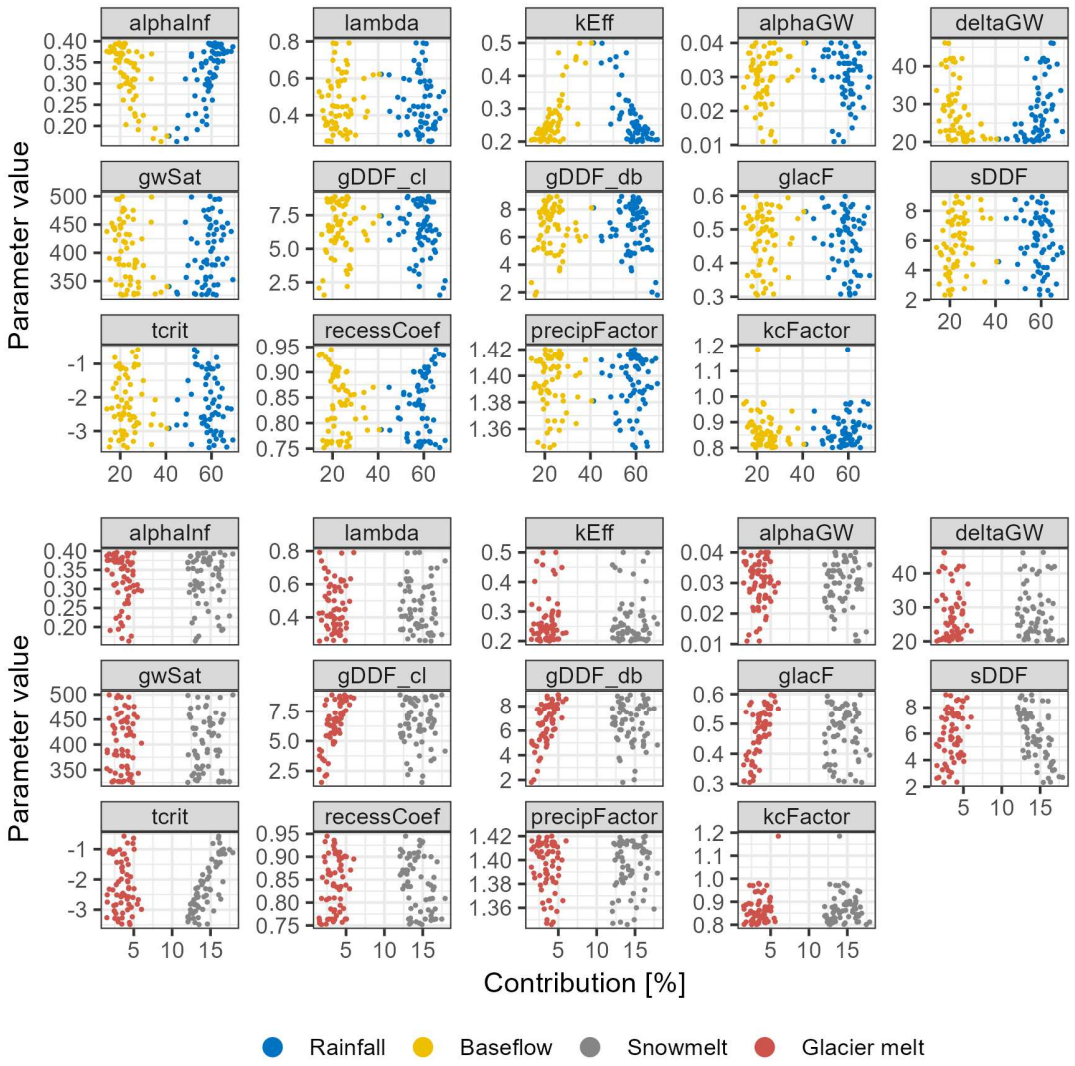


Figure 4.29: The relationship between the parameter value and the runoff contribution (total contribution over the simulation period 2002 – 2015 in percent).

The recession coefficient *recesCoef* which determines the speed of transport in the channel network also affects the rainfall-runoff and baseflow allocation because high parameter values (≥ 0.9) predict high rainfall-runoff contributions (Figure 4.29). However, members with lower parameter values (< 0.9) predict low and high contributions which indicates that this parameter is not the dominant control except for the upper parameter range. The ensemble members with the highest rainfall-runoff contribution were selected by the NSE and $PBIAS_{high}$ performance measures.

The snowmelt contribution ranges from 12 – 19%. The ensemble differences are caused by the *sDDF* parameter (negative relationship between value and contribution), and the *tcrit* parameter (positive relationship between value and contribution). The glacier melt contribution over the simulation period ranges from 1 – 6%, whereas higher glacier Degree-Day-Factor values (*gDDF_cl*, *gDDF_db*) result in larger glacier melt

contributions. The *glacF* parameter describes the fraction of meltwater which drains as surface runoff and higher values facilitate higher glacier melt contributions. However, the ensemble differences in the snowmelt and glacier melt contributions are small compared to the variation in the rainfall-runoff and baseflow contributions.

4.4.4.2 Yearly runoff composition

The variations in the annual flow composition are characterised by differences in the rainfall-runoff and snowmelt contributions while the glacier melt contribution is consistent throughout the years (Figure 4.30). The highest ensemble-mean rainfall-runoff contribution of 68% is simulated in 2009, and for this year the snowmelt contribution reaches its minimum of 7%. Conversely, the lowest rainfall-runoff contribution (52%) is simulated for the year 2015 with the highest snowmelt contribution (25%). The baseflow and glacier melt contributions are consistent throughout the simulation period with small variations between the years. The ensemble-mean baseflow contribution ranges between 20 – 26% and the glacier melt contribution between 3 – 5%.

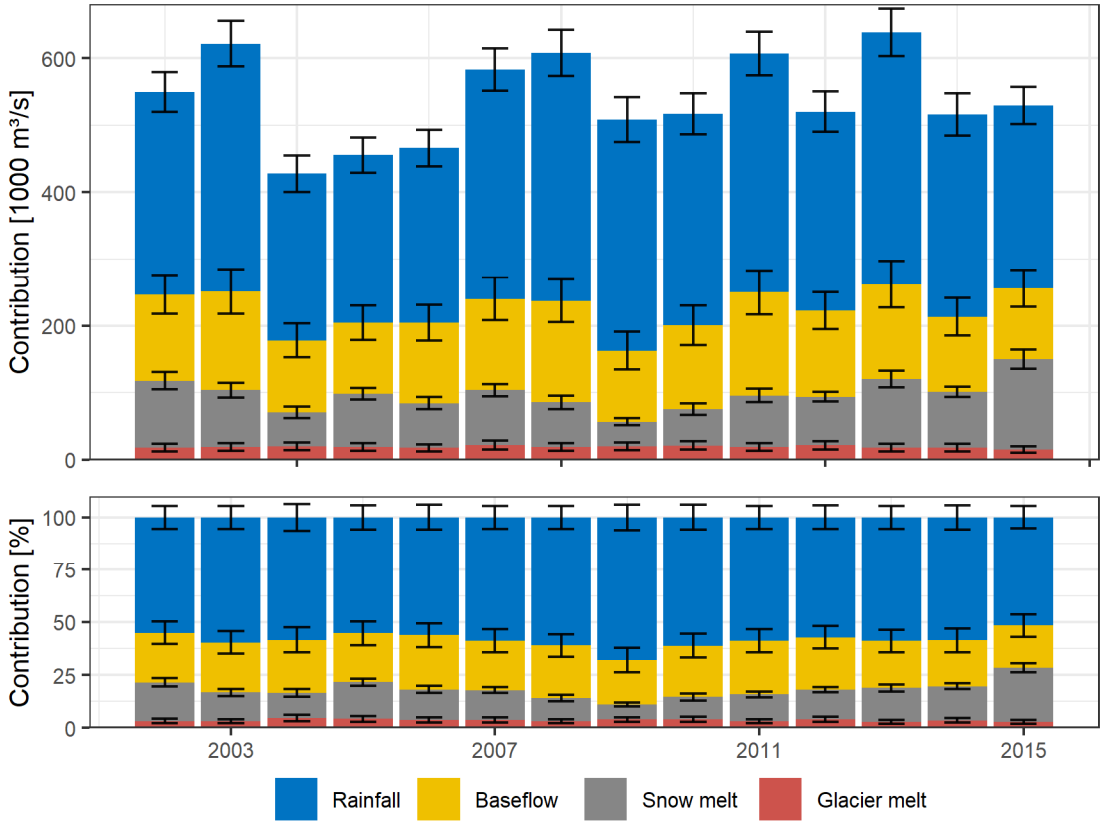


Figure 4.30: Mean annual runoff composition of the modelling ensemble. The error bars indicate the standard deviation of the composition of the ensemble members.

The prediction range of the ensemble is stable throughout the simulation period as indicated by the low variations of the standard deviation. This annual standard deviation of the contributions decreases with the mean contribution; it is largest for rainfall-runoff (5.4 – 6.3%), followed by baseflow (5.2 – 6.0%), snowmelt (0.9 – 2.2%), and glacier melt (0.9 – 1.4%). This low variation in the annual standard deviations indicates the consistent behaviour of the ensemble members regarding the variation in the climatic boundary conditions.

4.4.4.3 Monthly runoff composition

The monthly runoff composition is shaped by the seasonality of the climate, whereas the monsoon season (Jun – Sep) is dominated by rainfall-runoff, the post-monsoon (Oct – Nov) and winter (Dec – Feb) seasons by baseflow, and the pre-monsoon season (Mar – May) by snowmelt (Figure 4.31).

Most of the runoff during the monsoon season originates from rainfall-runoff. In the early monsoon season (Jun), rainfall-runoff and snowmelt are equally important and contribute 95% to the monthly discharge. The largest absolute snowmelt rates occur during this month. Rainfall-runoff then increases in importance during July and August contributing more than three-quarters to the monthly discharge. During these months, the groundwater storage refill and the baseflow contributions increase. The importance of baseflow increases further in the late monsoon season (Sep) as the rainfall-runoff slowly decreases. Glacier melt occurs only during the monsoon season and the highest melt rates are predicted for August.

The monthly discharge in the non-monsoon season (Oct – May) is dominated by snowmelt and baseflow (Figure 4.31). The baseflow gains importance in the post-monsoon season and early winter (Oct – Dec) and the contributions increase from 60% in October to 95% in December, whereas rainfall-runoff contributes the remaining fraction. Baseflow is the dominant runoff source in January (53%) but the importance of the rainfall-runoff increases as the westerlies bring some precipitation. It becomes the most important source in February (65%) which is the first month with notable snowmelt contributions (12%). Snowmelt then consistently gains importance in the pre-monsoon season from 44% in March to 70% in May. The baseflow contributions decrease as the groundwater storages deplete with time to the monsoon season, and it only contributes 5% at the start of the following monsoon season (Jun).

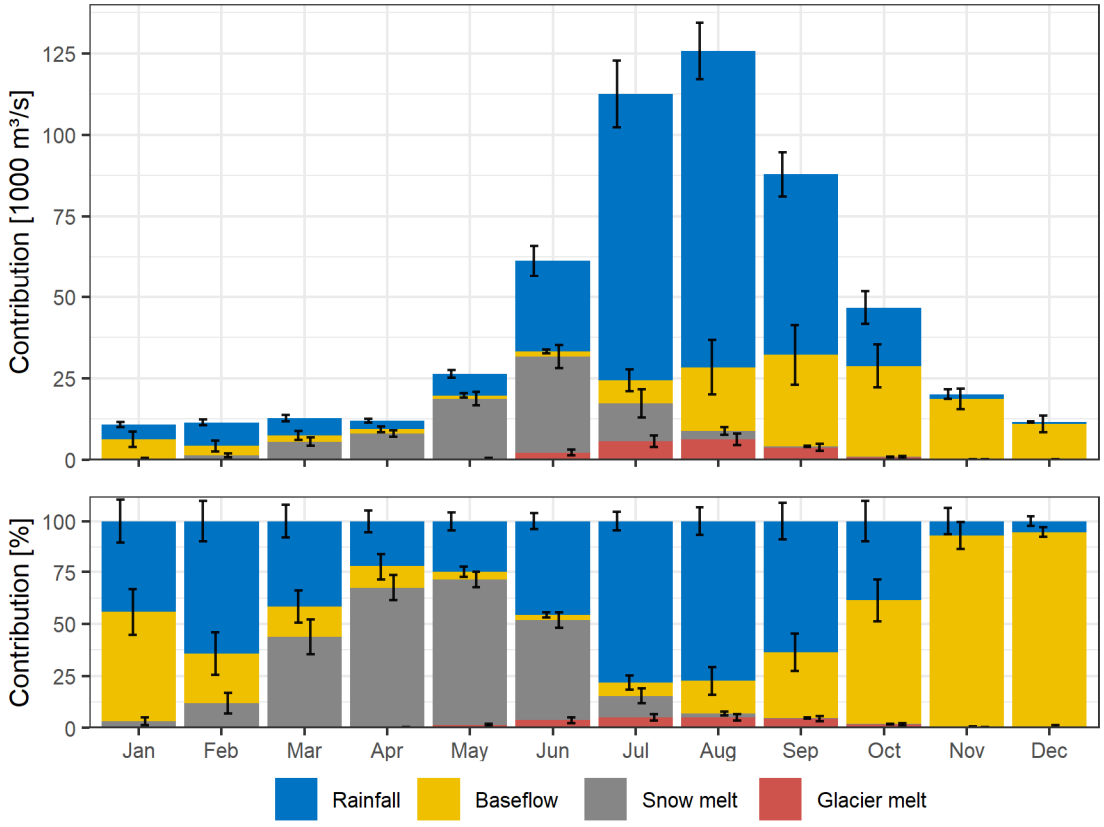


Figure 4.31: Mean monthly runoff contribution of the modelling ensemble (2002 – 2015). The error bars indicate the standard deviation caused by the annual variation and the ensemble variation.

The ensemble variation is largest for the rainfall-runoff and baseflow components (Figure 4.31). The standard deviation results from the ensemble variation and the annual climate variations. The standard deviations of the composition (%) are highest for the baseflow and rainfall-runoff, particularly during the transition from high to low flows (Sep – Oct), and the beginning of the year (Jan – Mar). The standard deviation of the absolute contributions is highest in June to September for rainfall-runoff, and August to October for baseflow. For snowmelt, the relative deviations are highest in March and Apr, and the absolute deviations are highest in June and July.

4.4.5 The simulation of the historic flood events

The modelling ensemble tends to underestimate the flood discharge throughout the modelling period (Figure 4.32). This underestimation increases with the magnitude of the flow. Flood flows are classified as flows ≥ 5,500 m^3/s which is a rough estimate of the danger level at stage 10.7m. These flows are underestimated by 25% in the calibration period and 16% in the validation period. The Annual Maximum flows (AMAX) are underestimated by 34% in the calibration period and by 40% in the validation period in which the higher AMAX flows occurred. For several days, flows in the range 5,500 – 8,000 m^3/s are overestimated in the validation period. However, 89% of the flood flows are underestimated.

The AMAX events are underestimated throughout the period, whereas the confidence intervals overlap during three events with lower magnitude (6,000 – 7,000 m^3/s) (Figure 4.33). The year 2008 is the only year for which the AMAX flow is overestimated but it is delayed by one day. The flows of the three largest floods (2009, 2013, 2014) are underestimated by over 50%; the observed flood flows range between 14,500 – 18,000 m^3/s and the simulated ones between 6,700 – 7,800 m^3/s. The simulated and observed flood hydrographs of these large events differ not only in magnitude but also in duration. The observed flood hydrographs (≥ 10,000 m^3/s) are sharp and with a quick recession and the discharge two days after the flood is similar to the discharge on the day before the flood. The simulated hydrographs have a lower magnitude but a longer duration, and it takes two additional days to generate the same cumulative discharge as is observed from the day before the flood event to the day past the flood event (Figure 4.34).

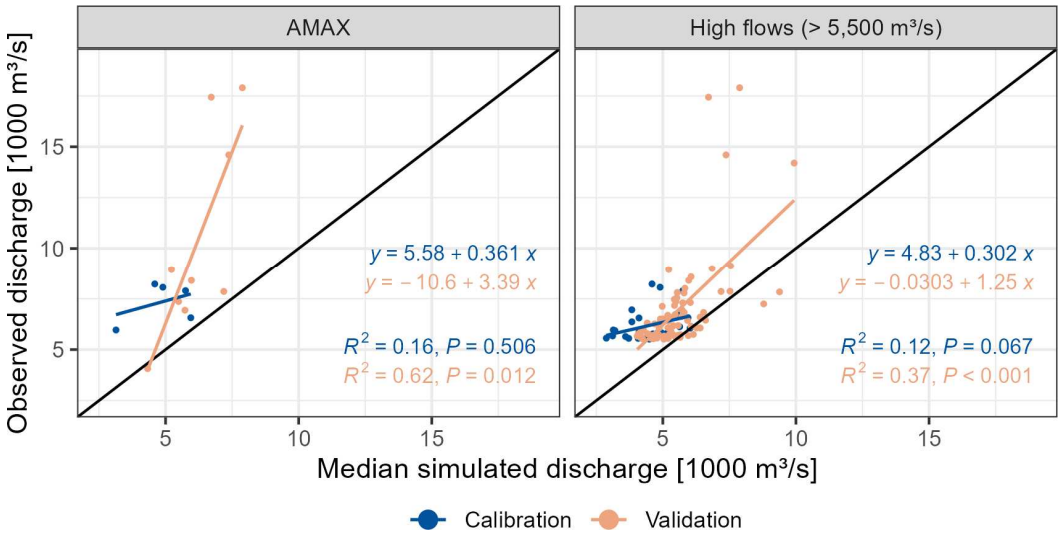


Figure 4.32: Comparison of simulated and observed annual maximum flows (left) and high flows (> 5,500 m^3/s) (right). The coloured lines indicate linear regression models of the simulated and observed flood flows, and the black line indicates the line of the perfect fit. The observations are obtained from DHM.

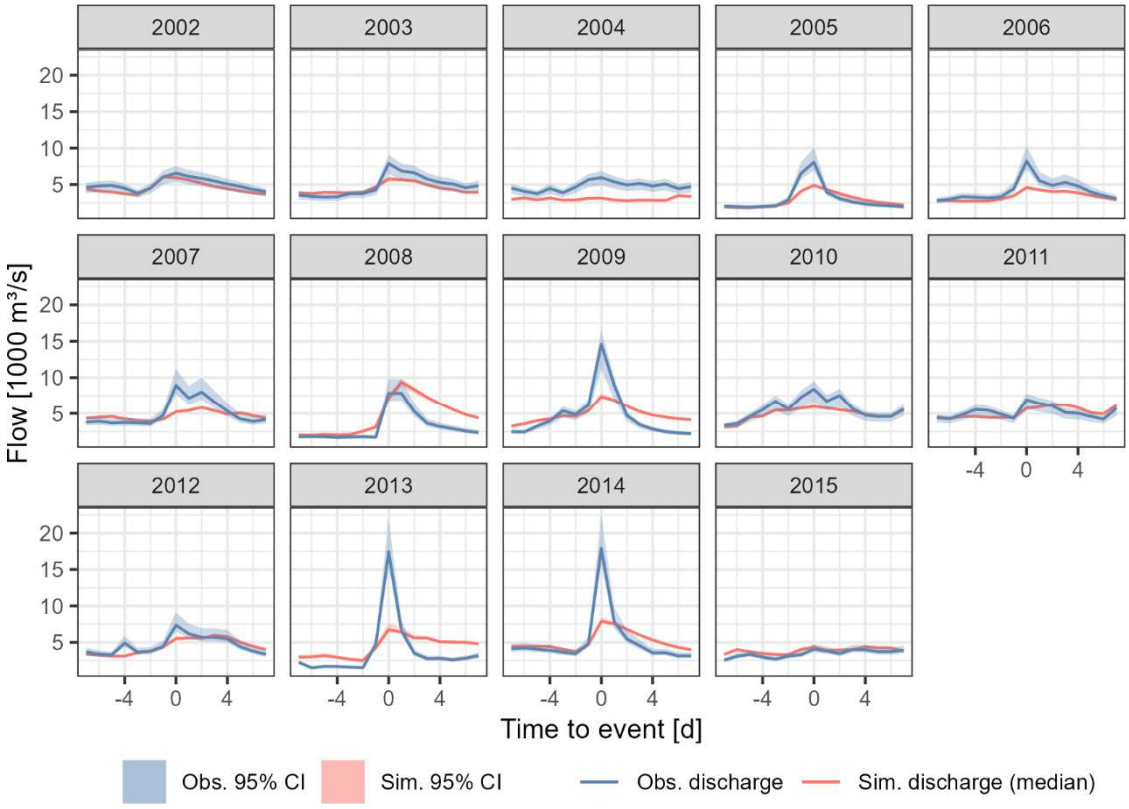


Figure 4.33: Observed and simulated hydrographs of the annual maximum flow events. The solid lines indicate the median simulated discharge and the observed discharge (DHM). The shaded areas indicate the prediction intervals ($Q_{97.5}$ – $Q_{2.5}$) of the BaRatin uncertainty estimation and the hydrological ensemble.

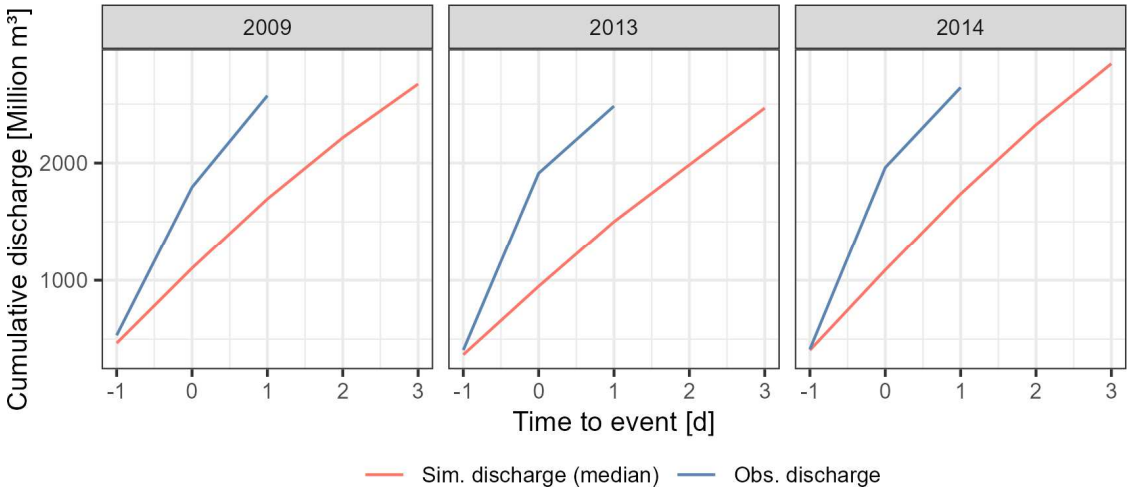


Figure 4.34: Cumulative discharge of the three largest observed flood events. Day 0 refers to the day of the flood event.

4.4.5.1 Flood flow composition

The rainfall-runoff is the most important source of flood water and the contributions are higher than on the monthly and yearly scale. It contributes 85% (median predictions) to the AMAX event which is considerably higher than the contribution to the overall flow of 59% (Table 4.11). Consequently, the contribution of the other runoff components decreases; the baseflow contribution is 7% (overall composition: 22%), the snowmelt contribution is 0.8% (overall composition: 14%), and the glacier melt contribution decreases to 3% (overall composition: 4%) (median predictions). However, the baseflow and snowmelt components are important contributors for individual AMAX events and ensemble members with contributions of up to 40% and 30%, respectively. The range of contributions is larger for the AMAX flows compared to the overall composition (Table 4.11); e.g. the rainfall-runoff prediction range is 40% (56 – 96%) for the AMAX events compared to 28% (41 – 69%) in the total simulation period. This larger prediction range is caused by differences between the AMAX events and the different parameterization of the ensemble members.

Table 4.11: The composition of the flood flows. The ensemble describes the variation in the composition arising from the hydrological ensemble and the different hydro-meteorological conditions of the flood events. The variation of the hydrological ensemble is removed in the ensemble mean and the variation arises from the different hydro-climatological conditions of the flood events. The overall simulation period describes the composition of all flows in the total simulation period (2002 – 2015) and is included for comparison.

	Component	**Min [%]**	**P_{10} [%]**	**Median [%]**	**P_{90} [%]**	**Max [%]**
Ensemble (AMAX)	Rainfall-runoff	56	74	85	91	96
	Baseflow	0	1	7	20	40
	Snowmelt	0	0	1	16	30
	Glacier melt	0	1	3	5	8
Ensemble mean (AMAX)	Rainfall-runoff	76	79	83	90	90
	Baseflow	1	3	9	14	17
	Snowmelt	0	0	2	10	22
	Glacier melt	1	2	3	4	4
Ensemble (total simulation period)	Rainfall-runoff	41	51	59	65	69
	Baseflow	14	18	22	30	41
	Snowmelt	12	12	14	16	18
	Glacier melt	1	2	4	5	6

The influence of the event characteristics is indicated by the varying composition of the ensemble mean which solely results from differences between the AMAX events. The rainfall-runoff contribution ranges from 76% in 2015 to 90% in 2009, baseflow contributes between 01 – 17%, snowmelt between 0 – 22%, and glacier melt between 1 – 4% (Table 4.11 and Figure 4.35). The AMAX composition is not, or at least not solely, controlled by the flow magnitude; i.e. the 2010 and 2015 events are of comparable magnitude but for the 2010 event rainfall-runoff contributes 90% of the flood flow, compared to 76% for the 2015 event. The reduced rainfall-runoff contribution is compensated by higher snowmelt which contributes 22% in the 2015 event and to 3% in the 2010 event (median predictions) (Figure 4.35).

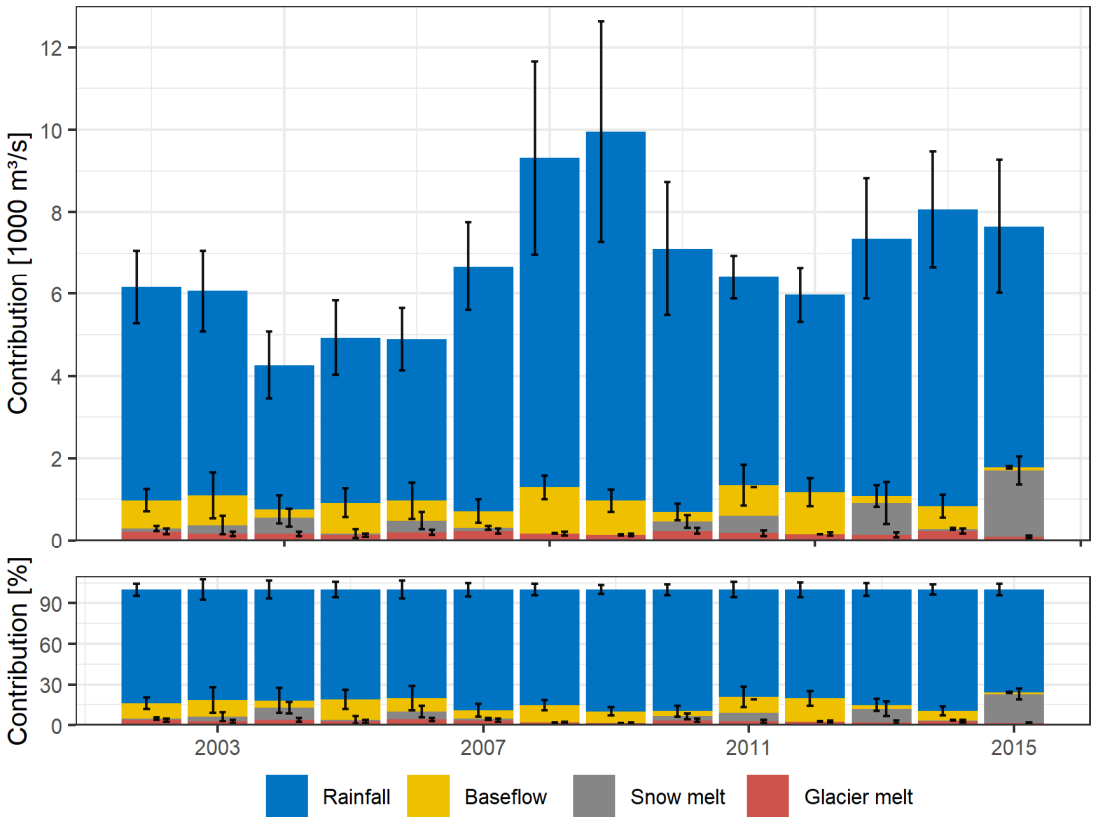


Figure 4.35: Composition of the AMAX events. The error bars indicate the standard deviation of the hydrological ensemble.

These differences in the AMAX composition relate to the timing of the AMAX event (Figure 4.36). Rainfall-runoff is the dominant runoff source throughout the monsoon season without clear temporal trends. However, the contributions of the other sources are affected by the timing. The snowmelt contribution is higher for earlier events but is not relevant after week 30 (Mid-July). Glacier melt contributions increase until week 34 (Mid-August) and gradually decline afterwards. Baseflow contributions are negligible until week 30 (Mid-July) but increase as the monsoon season progresses.

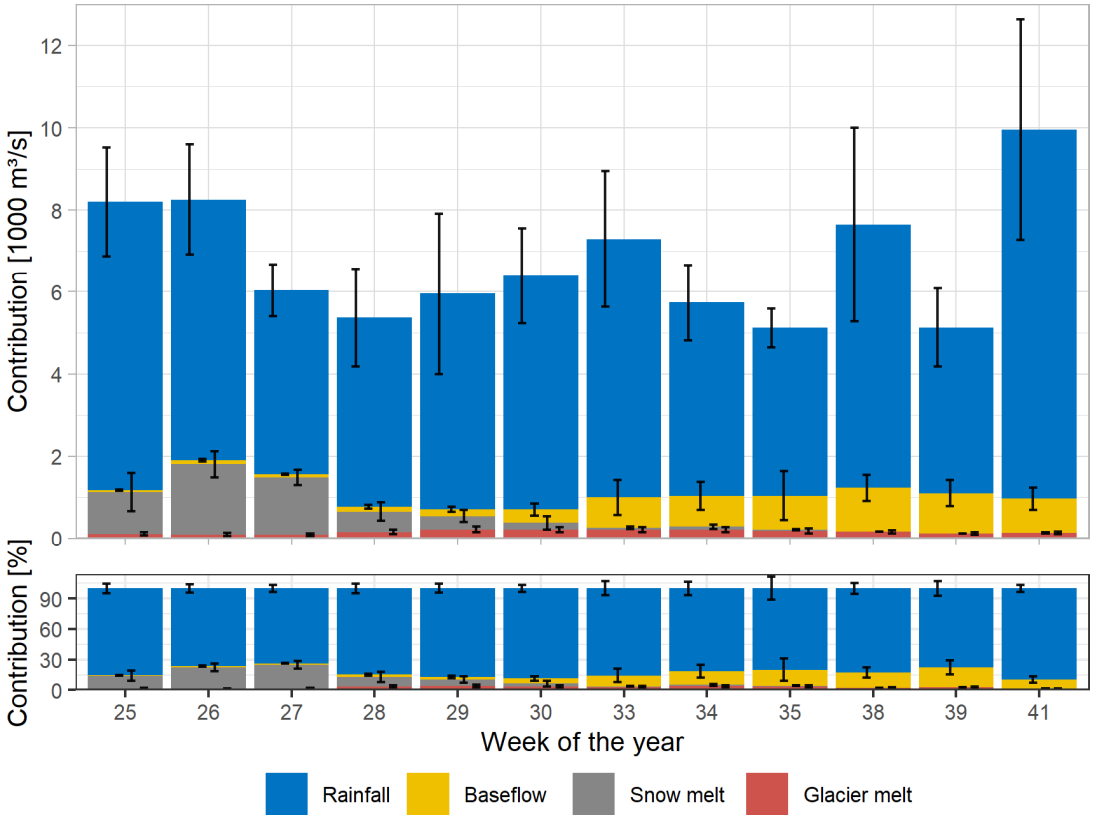


Figure 4.36: Mean AMAX composition aggregated by the week of the year during which the AMAX event occurred. The error bars indicate the standard deviation arising from the hydrological ensemble and the hydro-meteorological conditions of the events.

4.4.5.2 Ensemble variation

The contribution range is larger for the ensemble than the ensemble mean which indicates that the variation in the runoff composition is affected by the hydro-meteorological conditions of the flood event but also by the model parametrisation (Table 4.11). The relationship between model parameterisation and runoff composition is shown in Figure 4.37.

The *alphaInf* and *kEff* parameters influence the AMAX contribution of the rainfall-runoff and baseflow contributions whereas the rainfall-runoff contribution increases with the *alphaInf* value and decreases with the *kEff* value as also observed for the runoff composition of the total simulation period (Figure 4.29 and 4.37). Furthermore, lower *alphaGW* values and higher *deltaGW* values increase the probability of higher rainfall-runoff and lower baseflow contributions. The precipitation correction factor does not affect the rainfall-runoff contribution.

For the snowmelt contribution, a clear relation between the *sDDF* parameter and the snowmelt contribution is observed similar to the overall composition. However, unlike for the overall contribution, no clear relationship between the *tcrit* parameter and snowmelt is observed for the AMAX contribution. Furthermore, higher *precipFactor* values increase the likelihood of larger snowmelt contributions. The AMAX glacier melt contribution is influenced by the glacier DDFs, and the *glacF* parameter which defines the fraction of glacier melt that percolates to the groundwater.

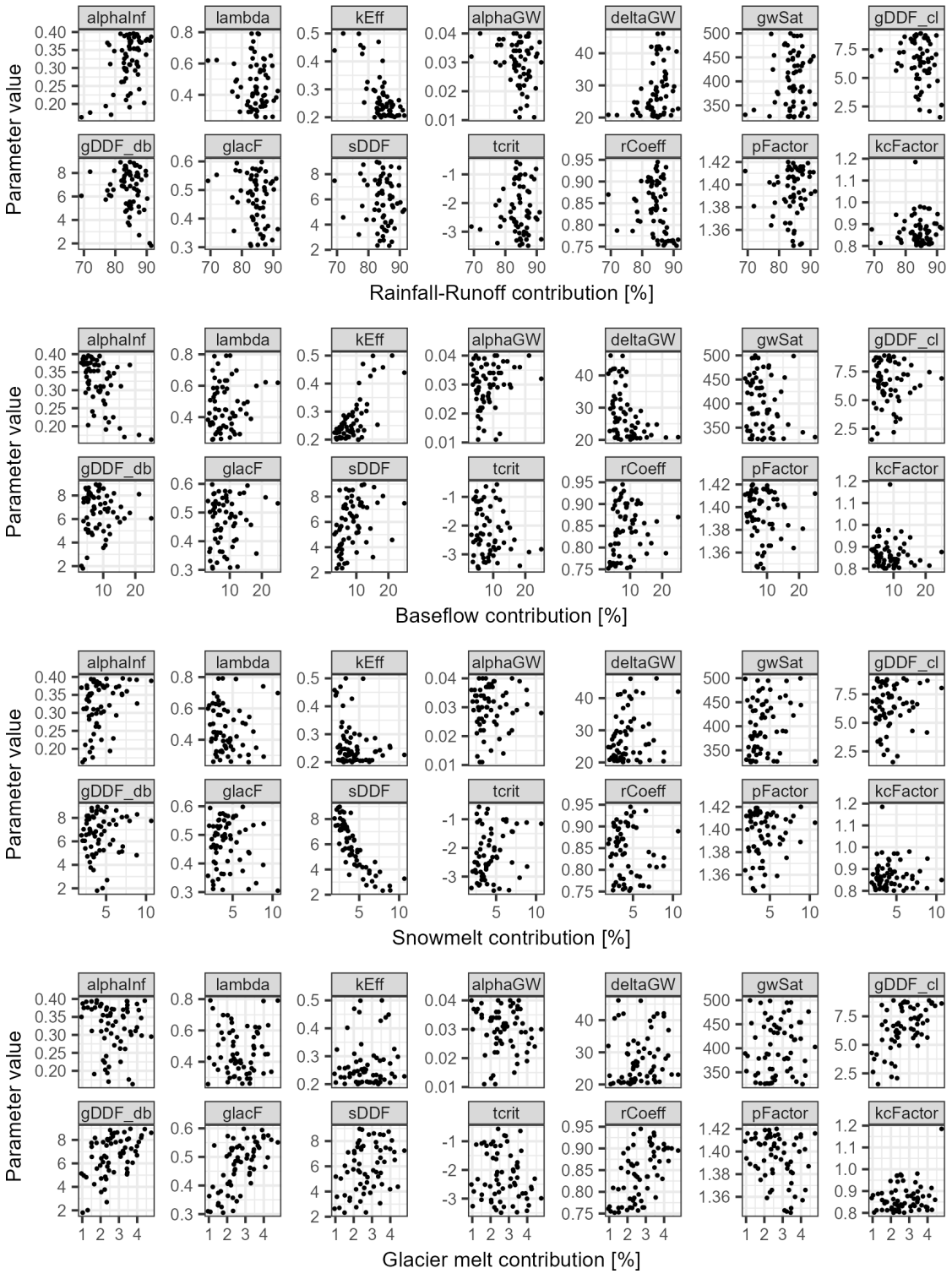


Figure 4.37: Comparison of the parameter values and the AMAX runoff contribution (mean of all AMAX events) of the 64 ensemble members. The rCoeff is the recession coefficient and the pFactor the precipitation correction factor.

The model ensemble produces more similar results for events with higher rainfall runoff contribution (Figure 4.38). The standard deviations of the individual AMAX events indicate the ensemble variations in the runoff contribution resulting solely from the parameterization of the ensemble members. A positive relationship between the ensemble mean contribution and the standard deviation is observed for the glacier melt contribution. The snowmelt and baseflow components show a weaker positive relation. This means that the ensemble variation increases with the importance of these sources. However, a negative relationship is predicted for the rainfall-runoff component. Hence, the ensemble members tend to produce more similar results with increasing rainfall-runoff contribution. This indicates that the model complexity reduces because the infiltration excess runoff increases and hence more water drains as surface runoff and less water drains through the soil and groundwater which are controlled by more parameters.

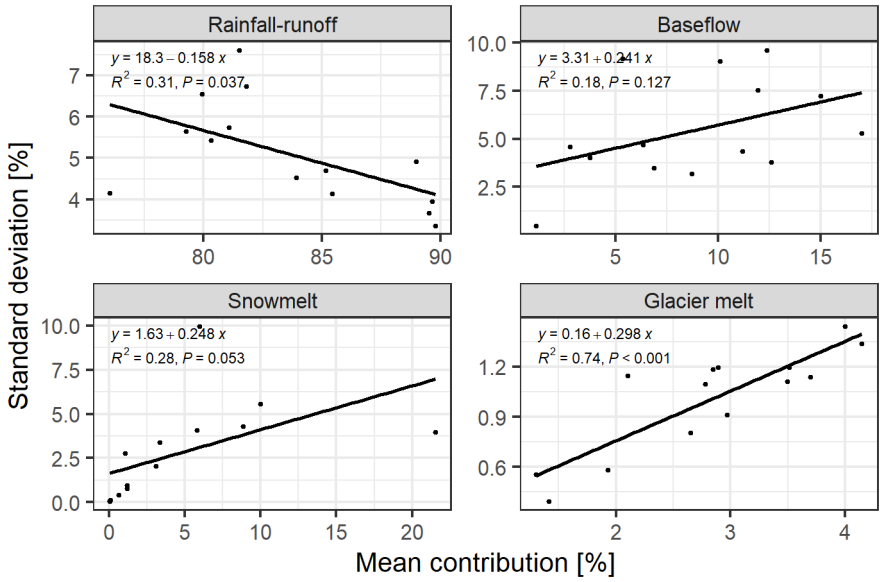


Figure 4.38: The relationship between the ensemble mean and the ensemble variation (indicated by the standard deviation) of the contribution for each AMAX event.

4.5 Discussion of the hydrological predictions

The inherent imperfection of hydrological models arises from the incomplete knowledge of catchment hydrology (the modelled system), the incomplete knowledge about the catchment behaviour (the modelled catchment), and the simplification of the hydrological system by aggregating the processes in time and space which requires the integration of parameters that cannot be measured and, therefore, need to be calibrated to replicate the catchment behaviour (Gupta *et al.*, 2005). However, different parameter combinations can produce similar results so that it is impossible to determine a single best parameter set, which is further complicated by the lack of information about the internal state of the catchment's hydrology (Beven and Freer, 2001; Gupta *et al.*, 2005; Odoni, 2007). The ensemble modelling accounts for this uncertainty whereas every ensemble member represents an equally likely representation of the hydrological system (Beven and Binley, 1992; Beven, 2012; Beven and Binley, 2014). Due to the lack of information about the internal state of a catchment's hydrology, the classification of behavioural models can only be approximated but any behavioural model should have three characteristics:

i) the input-state-output behaviour is consistent with the observations;
ii) the model predictions are accurate (small bias) and precise (small prediction range):
iii) the model behaviour is consistent with the current hydrological understanding of the catchment (Gupta *et al.*, 2005).

The simulations of the Karnali catchment are discussed in light of these characteristics to evaluate whether the ensemble predictions are behavioural representations of the catchment hydrology of the Karnali River. The following section discusses the simulated hydrological regime which provides information about the overall model performance (Section 4.5.1), followed by the discussion of the runoff composition which provides information about the internal state (Section 4.5.2). These sections are the foundation for the assessment of the model's ability to predict flood hazards (Section 4.5.3).

4.5.1 The predicted hydrological regime of the Karnali River

The hydrological ensemble reproduces the observed hydrological regime of the Karnali River. The catchment discharge at the mountain outlet aggregates all hydrological processes in the mountainous catchment and is therefore an important indicator of the model performance. The great variation of flows in Himalayan River systems with high flows during the monsoon season and low flows in the non-monsoon season poses a challenge for hydrological models (Nepal *et al.*, 2014). However, the established hydrological ensemble captures these seasonal dynamics of the Karnali River well as indicated by the comparison of the observed and simulated hydrographs (Figure 4.20). The transitional periods in the pre-monsoon and post-monsoon seasons are well timed by the ensemble. This good depiction of the seasonality is underlined by the high coefficient of determination (R^2) which is classified as good even for the member with the lowest efficiency of 0.77 (Table 4.6) (Moriasi *et al.*, 2015b).

The monsoon season flows are, in general, well replicated as indicated by the high Nash-Sutcliffe efficiencies that reach 0.85 in the calibration period and 0.82 in the validation period (Table 4.6) (median predictions). The differences in the efficiencies between the calibration and validation period are small which indicates that the ensemble depicts the hydrological system of the Karnali River and is not overfitted to the calibration data. This is also true for the R2 which is 0.85 and 0.84 for the calibration and validation periods, respectively (median predictions).

The NSE and R^2 are similar to the efficiencies obtained in other studies of the three major river systems in Nepal. These reported efficiencies range between 0.84 – 0.87 (NSE) and 0.85 – 0.88 (R^2) for the Koshi River (Nepal *et al.*, 2014), 0.56 – 0.78 (NSE) and 0.63 – 0.84 (R^2) for the Gandak River (Bhattarai *et al.*, 2018), and 0.83 – 0.85 (NSE) and 0.84 – 0.89 (R^2) for the Kanali River (Dahal *et al.*, 2020; Pandey *et al.*, 2020). The performance of all members is classified as good to very good regarding the NSE, and good regarding the R^2 (Moriasi *et al.*, 2015b). The median PBIAS of the discharge is 3% for the calibration period and 13% for the validation which is classified as very good and satisfactory (Moriasi *et al.*, 2015b). These high efficiencies, the small differences between the calibration and validation period, and the visual hydrograph comparison indicate that the model ensemble is a behavioural representation of the hydrological system of the Karnali River.

The good model performance is also indicated by the comparison of the hydrographs of the West Seti and Chera subbasins. The performance of the Chera subbasin is satisfactory regarding the NSE, R^2 and PBIAS efficiencies (Table 4.8) (Moriasi *et al.*, 2015b). These efficiencies are not satisfactory for the West Seti subbasin. However, this subbasin covers < 0.5% (157 km^2) of the catchment. The hydrograph comparison indicates that the temporal dynamics are well captured for the West Seti subbasin which indicates that the model depicts the hydrologic conditions of the Karnali River on a small spatial scale (Figure 4.22).

Despite the good overall performance, several systematic biases are observed from the comparison of the simulated and observed hydrographs at the catchment outlet, namely the overestimation of peak flows of the rising limb in the pre-monsoon season and in the early monsoon (May – Jun), the overestimation of peak flows outside the monsoon season, and the underestimation of flood flows.

The higher simulated flow of the rising limbs is linked to the overestimation of the snowmelt runoff indicated by the underestimation of the snow extent during the monsoon season (Figure 4.24). The snow extent is only an indicator of snow storage because it does not provide information about the snow volume and, thus, the water stored as snow. However, the high contributions of snowmelt (50 - 70% from April – June) indicate that a fraction of the overprediction is related to the overestimated snowmelt (Figure 4.31). However, the seasonality of snow melt and snow accumulation is well timed.

The overprediction of the non-monsoon season peak flows is related to the use of global (in time) model parameters that cannot represent the seasonal variability of the climatic conditions, particularly the precipitation. The Indian monsoon system dominates the climate in the summer months bringing frequent and intense rainfall events. In the winter and pre-monsoon season, the Westerlies cause low-intensity precipitation events (Gautam and Acharya, 2012; Krishnan *et al.*, 2019). However, both precipitation systems are simulated with a global parameter set which was calibrated with an emphasis on the monsoon season. The consistent overprediction of discharge peaks outside the monsoon season particularly between Jan – May indicates that the model parameterisation does not represent these low-intensity westerly rainfalls. The *alphaInf* parameter defines the fraction of daily rainfall during the hour of most intense rainfall and is implemented to simulate infiltration excess runoff. This global parameter cannot reflect the different characteristics of the monsoon and non-monsoon season precipitation which leads to a fast rainfall-runoff conversion and consequently the overprediction of runoff.

The overpredicted non-monsoon season peak flows are likely to also be affected by the global precipitation correction factor which is unlikely to reflect the precipitation bias of both, the monsoon and Westerly systems. The application of this factor may lead to the overestimation of the precipitation outside the monsoon season. Furthermore, this parameter interacts with the *alphaInf* parameter. This parameter combination favours, the overestimation of surface runoff and, consequently, leads to an underestimation of groundwater recharge. The baseflow is the main contributor to runoff in the post-monsoon season and winter until the start of the snowmelt and the groundwater storage depletes with time from the monsoon season (Nepal, 2012). This behaviour of overpredicting non-monsoon seasonal peak flows is also occurring in the simulations of (Pandey *et al.*, 2020) which emphasises the challenges of representing the diverse hydrologic conditions in Himalayan watersheds stated by (Nepal *et al.*, 2014).

4.5.2 Runoff composition

The model ensemble reproduces the catchment discharge of the Karnali River well. However, the catchment discharge is the sum of the upstream hydrological models and hence the ensemble could predict the correct output for the wrong reasons (i.e. model equifinality). The representation of the internal catchment behaviour can be investigated by comparing the runoff composition with our current understanding of the hydrology of Central Himalayan catchments. The runoff contributions in Himalayan catchments are not well understood and hence it resorts to modelled datasets which suffer from similar uncertainties (Bookhagen and Burbank, 2010; Andermann *et al.*, 2012b; Bookhagen, 2012). Nonetheless, the comparison of the simulated runoff composition with the ones in other studies provides information about whether the obtained results are within reason.

4.5.2.1 Overall contribution

The overall composition throughout the simulation period agrees with the current understanding of the hydrology in Central Himalayan catchments. Rainfall-runoff dominates the composition contributing between 57 – 59% in the calibration and validation periods (median predictions) (Figure 4.28). This is consistent with the current understanding that monsoon season rainfall dominates the runoff in these river systems (Bookhagen and Burbank, 2010; Bookhagen, 2012; Scott *et al.*, 2019). (Lutz *et al.*, 2014) simulated the runoff composition of the Upper Ganga River to which the Karnali contributes. In their study, the rainfall-runoff contribution is 66% and, thus, slightly higher than in this study, and the baseflow contribution is 14% which is 10% lower than in this study. However, groundwater storages have a significant impact on retarding the rainfall in the Central Himalayas and the role of groundwater is not well understood in this region (Andermann et al., 2012; Scott et al., 2019). The simulated baseflow of 24% is reasonable in this light and considering the large uncertainties. The overall baseflow contribution agrees with the 20% contribution simulated for the central Himalayan Dudh Koshi River (Nepal *et al.*, 2014).

The melt contribution (the combined snowmelt and glacier melt) is 17- 19% and agrees with the 20% simulated for the Upper Ganga River (Lutz *et al.*, 2014). The snowmelt contribution of 14 – 15% falls within the predictions of other studies for the Karnali River which range between 7 – 21% (Bookhagen and Burbank, 2010; Andermann *et al.*, 2012b; Dhami *et al.*, 2018). The snowmelt contribution varies strongly between the years

and hence this discrepancy could result from the different study periods (Figure 4.30). The predicted snowmelt contribution is reasonable because it falls within the range of contributions provided in the literature and does not exceed the maximum snowmelt contributions of 20 – 25% in Central Himalayan catchments (Bookhagen and Burbank, 2010).

The predicted snowmelt during the monsoon season is overestimated as indicated by the consistently lower monsoon season snow extent compared to the MODIS estimates, and the overestimated rising limb (Figures 4.20 and 4.24). This could result from interactions between the temperature data and the melt parameters. The regionalisation of the temperature data was conducted with a global lapse rate and extrapolated beyond the highest gauge at ~3,050 masl. However, it is established that the lapse rate varies with elevation (Immerzeel *et al.*, 2012). It is plausible that the global temperature-dependent melt parameters are too high for the regionalised high mountain temperatures. This might be further aggravated by the use of an annual lapse rate which does not account for the wet-adiabatic temperature cooling during the monsoon at higher elevations.

Nonetheless, the comparison with studies of other central Himalayan River systems suggests that the overall runoff composition is reasonable considering the unique catchment characteristics, the temporal variability of the composition, and the reliance on modelling approaches for the quantification of runoff sources.

4.5.2.2 Seasonality of the runoff composition

The seasonal dynamics of the runoff components are well reflected by the modelling ensemble. The rainfall-runoff dominates during the monsoon season and contributes 60 – 80% in July – September and snowmelt is the largest contributor in June (Figure 4.31). This is consistent with the compositions of the Ghagara River (Lutz *et al.*, 2014), and the Dudh Koshi River (Nepal *et al.*, 2014). The baseflow is highest during the transition from the monsoon season to the post-monsoon season (Sep – Oct) and is the main runoff source until the end of winter (Nov – Feb) after which it declines as the groundwater storages deplete which agrees with our knowledge of Central Himalayan catchment hydrology (Andermann *et al.*, 2012b; Lutz *et al.*, 2014; Nepal *et al.*, 2014).

The snowmelt seasonality is well depicted but the snow storage might deplete too early. The highest absolute snowmelt contributions are simulated for the months May to July which agrees with Dhami *et al.* (2018) for the Karnali River, Lutz *et al.* (2014) for the Ghagara River, and (Immerzeel *et al.*, 2013) for the mountainous Langtang River. However, the length of the snowmelt season varies in the literature; Dhami *et al.* (2018) simulate contributions ≥ 10% throughout the non-monsoon season (Oct – May) for the Karnali; Lutz *et al.* (2014) simulate notable contributions in the period January – September for the larger Gharaga River system; and Bhattarai *et al.* (2018) predict snowmelt between February and August in the neighbouring Gandak River. The longer melting season in several studies indicates that the simulated melting season may be underestimated and the snow storage may deplete too early.

The underestimation of the snow extent between July and October provides evidence that the snow storage deplete too early (Figure 4.24). Therefore, it is likely that the snowmelt contribution is overestimated during June and July and underestimated afterwards which is supported by the higher snowmelt rates in Lutz *et al.* (2014) and Dhami *et al.* (2018). However, the importance of snowmelt declines after June – July in every study and hence the effect of these potential underestimations on the monsoon hydrograph is low.

The simulated glacier melt season is shorter than in comparable studies. Glacier melt occurs during the monsoon season (Jun – Sep), and small melt rates are simulated in October (Figure 4.31). The highest melt rates are simulated during July and August which agrees with other Central Himalayan catchments (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014; Nepal *et al.*, 2014). However, the melt season is shorter than for the Ghagara River (May – Nov) (Lutz *et al.*, 2014), the Dudh Koshi (Apr – Nov) (Nepal *et al.*, 2014), and Langtang (May – Oct) (Immerzeel *et al.*, 2013).

The shorter predicted glacier melt period is somewhat unexpected because the most plausible cause for the shorter melting season is the underestimation of high-elevation temperatures. This contradicts the snow melt simulations because the most plausible cause for the low snow extent during the monsoon season is the overestimation of high-elevation temperatures. Furthermore, the degree-day factor which positively controls the melt rate is 0.6 °C higher for the glacier module than for the snow module. The ensemble mean degree-day factors are comparable with the ones used by Lutz *et al.* (2014) and Dhami *et al.* (2018) and hence these differences are not attributed to an unreasonable parameterisation. However, the uncertainty about the actual glacier melt rates in the region is large, the total melt rates reasonable, and the shorter melt season has only

minor effects on the simulation results due to the dominance of rainfall on the monsoon season flows.

4.5.3 The simulation of the historic flood events

The hydrological modelling aims to predict changes in flood flows for the projected climates (Chapter 5) and hence the representation of flood peaks is an essential aspect of this study. However, the monsoon season flood peaks are consistently underestimated and these underestimations increase with the flow magnitude to > 50% for the largest flood events (Figures 4.32 and 4.43). There are four potential causes for this underprediction:

i) the model parameterisation that does not reflect the hydrology during heavy rainfall events;
ii) deficits in the process representation that limit the model's ability to reproduce the sharp flood peaks of the Karnali River;
iii) the underestimation of heavy precipitation and hence the underestimation of the water in the catchment;
iv) the overestimation of the observed flood hydrograph so that the model unknowingly reproduces the flood hydrograph.

The parametric error of the model can relate to the runoff generation that converts rainfall on the hillslopes and transports it to the channel network, and/or the runoff routing that transports the water in the channel network to the mountain outlet (Beven, 2012). In the SPHY model, the runoff routing is described by the *recessCoeff* parameter that controls the travel time within the channel network. The flow velocity changes with flow rate and the global *recessCoeff* parameter cannot account for this non-linearity (Beven, 2012). The runoff generation processes depend on the antecedent conditions (the initial wetness) and the rainfall characteristics and are, thus, also non-linear which cannot be represented by the global parameters, e.g the *alphaInf* parameter that controls the infiltration excess runoff. The observed cumulative flood flow (AMAX ± 1 day) of the three largest events is reached after three simulation days (Figure 4.35). This indicates that the underestimation of flood flows is affected by deficits in the model's process representation which hamper the reproduction of the rapid rainfall-runoff conversion of the Karnali River to the full extent.

These deficits of the process representation may arise from the aggregation of the hydrological processes to the daily resolution. The Karnali River is, despite its large catchment area, highly responsive and the flood peak is reached within 24 hours of the rainfall event (MacClune *et al.*, 2015; Shrestha *et al.*, 2015a). It is likely that the aggregation of the flood generation and routing processes to the daily resolution oversimplifies the system and causes an underprediction of the flood peaks. However, the daily time step is constrained by the hydro-meteorological data (precipitation, temperature, discharge) which is only available at the daily resolution. Furthermore, floods are usually simulated from event-based models that simulate single flood events (USDA, 2014; Nathan and Ball, 2019). However, these models require the calibration for single events and are sensitive to the antecedent conditions which are unknown for the past and projected climates (Berthet *et al.*, 2009; Beven, 2012; Nathan and Ball, 2019). Therefore, the application of a continuous model at a daily resolution is required in this research and the structural errors must be accepted as part of the research design. The prediction of flood flows is a challenge in hydrological modelling, particularly for the most extreme events and biases of 50% common (Huang *et al.*, 2017; Hirabayashi *et al.*, 2021).

The accurate representation of the flood-triggering rainfall is essential to reproduce the flood hydrograph. However, precipitation data is particularly uncertain in mountainous areas with steep elevation precipitation gradients and coarse gauge networks (Winiger *et al.*, 2005; McMillan *et al.*, 2012; Lutz and Immerzeel, 2016). Furthermore, the uncertainty is larger for convective rainfall events which trigger the floods in the Karnali River (McMillan *et al.*, 2012). The rainfall intensity and its spatial distribution are important variables for hydrological simulations. The global rainfall volume over the catchment determines the amount of water in the hydrological system (assuming the evapotranspiration is negligible during heavy rainfall events) and, thus, determines the runoff volume and the spatial patterns of the rainfall strongly influence the timing of the hydrograph (Gupta *et al.*, 2005; Huang *et al.*, 2017; Reaney, 2022).

The rainfall intensity and the spatial variations are highly uncertain in the study area. The gauge network is sparse and concentrates along valleys and the floodplain so that large areas of the study area are unobserved. A monthly gridded precipitation product was used to overcome this limitation. The poor quality of every investigated gridded product at the daily resolution indicates the high uncertainty of gridded precipitation in the catchment. The monthly dataset was downscaled to the daily resolution based on the sparse gauge network and hence both, the volume and spatial variation are uncertain. However, the model ensemble matches the observed flood volume but within a longer

duration which indicates that rainfall volume with the applied precipitation correction factor is adequate. The timing of the hydrograph can be affected by both, errors in the model structure and the spatial rainfall patterns.

The difference between the simulated and observed flood hydrographs may also result from errors in the observed discharge record. The uncertainty in the discharge estimation of high and flood flows from stage-discharge rating curves is usually between ± 15 – 40% (Westerberg *et al.*, 2011; McMillan *et al.*, 2012; Coxon *et al.*, 2015). These reported uncertainty ranges agree with the predicted uncertainty range of ± 30 – 40% in this study (Figure 4.16). However, the uncertainty varies between gauging stations and the discharge record at the catchment outlet is particularly uncertain because of the lack of observation at high flows. The discharge of the three largest floods in 2009, 2013, and 2014 is 250 – 300% larger than the maximum stage-discharge measurement after 1990, and 400 – 500% larger than the largest recent measurement after 2008. These extrapolations are well beyond the USGS guidelines that do not extrapolate beyond 100% of the max. measurement (Wiche and Holmes, 2016). Furthermore, the stage-discharge relationship may also be altered by changes in the geometry of the riverbed and these potential changes may not captured by the coarse temporal resolution of the samples.

The observed high flows could be, in principle, underestimated by the stage-discharge conversion but the comparison with other studies suggests an overestimation. Two studies conducted hydrological simulations of the Karnali with overlapping periods; Dahal *et al.* (2020) simulated the monthly discharge between 1990 - 2005 using the SWAT model and point precipitation data from DHM, and Pandey *et al.* (2020) simulated the daily discharge between 1995 – 2009 using the SWAT model and use a combination of the gridded TRMM product and point DHM observations. Both studies consistently underestimate the high flows like in this study. The SWAT model is a continuous model which suffers from similar structural deficits as the SPHY model and hence this underestimation could be caused by structural model errors. However, all three simulations underestimate the monsoon season flow during the years 2004 and 2005. This is particularly evident in the SPHY simulations of this study, and the monthly SWAT simulations which match the monsoon season flow of the previous years well. The consistent underprediction in three models which use different precipitation datasets suggests that the discharge record may overestimate the actual discharge, at least during some years which may be caused by errors or changes in the instrumentation or shifts in the stage-discharge relationship.

The consistent underprediction of the model ensemble results most probably from the combination of uncertainties in the hydrological model and the discharge data. The consistent underestimation of more frequent flood events by the ensemble indicates structural and parametric errors which can be aggravated by uncertainties in the heavy rainfall data. However, it is likely that the observed peaks are overestimated and that the true peaks fall within the observed and simulated ones. The prediction of peak flows are a common challenge in hydrological modelling and the bias in the prediction of the Karnali flood events falls within the bias of comparative studies (Huang *et al.*, 2017; Wijngaard *et al.*, 2017; Hirabayashi *et al.*, 2021).

4.6 Conclusions

The objective of this chapter was to establish a hydrological model that replicates the hydrological system of the mountainous catchment and reproduces the observed discharge at the mountain outlet. The established modelling ensemble is a behavioural representation that reflects our knowledge of the hydrology of the Karnali River. The simulated input-state-output is consistent with the observations as indicated by the high efficiencies of the modelling ensemble. The evaluation consists of multiple evaluation criteria and hence this conclusion is based on a robust foundation which considers the seasonality of flows and the differences between observations and simulations regarding the discharge, snow dynamics and actual evapotranspiration.

The prediction range of the ensemble is accurate as indicated by the good to very good performance regarding the R^2 and NSE and satisfactory to very good performance regarding the PBIAS for all ensemble members (Moriasi *et al.*, 2015b). The small differences in the efficiencies between the calibration period and the validation period indicate that the modelling ensemble reproduces the catchment hydrology and is not overfitted to the calibration data. The main variation within the ensemble relates to the different distribution of rainfall into rainfall-runoff and baseflow. The seasonality of the composition is consistent with other simulations and reflects the current knowledge of the hydrology in Central Himalayan catchments.

The ensemble has some systematic inaccuracies despite the good overall performance, namely the overprediction of the non-monsoon season runoff spikes, the underestimation of the flood discharge, and the underestimation of snow cover in the monsoon season. These systematic errors are affected by parametric uncertainties and deficits in the process representation of the model. In general, the application of hydrological models in these catchments is challenging because the models use global parameters that cannot represent the diverse seasonality of the hydrological processes. This parametric uncertainty manifests in the overprediction of the non-monsoon season spikes which is caused by the overestimation of the infiltration-excess runoff due to the *alphaInf* parameter and the precipitation correction factor which cannot reflect the different characteristics of the monsoon and Westerly precipitation. The underestimation of flood flows arises from the simulation of sub-daily processes on a daily scale which reduces the model's ability to reproduce the fast rainfall-runoff conversion and routing to the full extent.

The limited data availability and quality pose another challenge for the application of hydrological models in the Himalayas. This mainly accounts for the climatic boundary conditions and the discharge data. The highest temperature station in the catchment is located at 3,050 masl and hence it is not possible to determine the temperature lapse rate for the higher elevations in which melt processes are important features of the hydrology. The underestimation of the monsoon season snow extent is most likely caused by inaccuracies in the temperatures of the high mountainous parts of the catchment and consequently the miscalibration of temperature-dependent melt parameters.

The lack of precipitation data in the high mountainous parts of the catchment translates into errors in the gridded precipitation product which uses in-situ observation for the enhancement of the quality. The station network is coarse and concentrates along valleys and in the plain. Therefore, orographic rainfalls and the local variations cannot be considered in the downscaling of the monthly precipitation product which is likely to reduce the simulation of peak flows. The calibrated precipitation correction factor has performed well as indicated by the simulated cumulative discharge of flood events which matches the observed one with two days delay. This is a global parameter which was calibrated with emphasis on the monsoon season. However, it is unlikely that the precipitation bias is similar during the monsoon and non-monsoon seasons and hence this correction factor potentially amplifies the overestimation of non-monsoon season runoff spikes.

The discharge data is very uncertain as indicated by the consistent underprediction of 2005 monsoon season flows in multiple independent studies using different precipitation products. This data is essential for the model calibration and hence errors in the discharge data propagate to the model. There is evidence that the discharge observations overestimate the flood flows and, therefore, the bias of the simulated flood flows may be lower as suggested by the comparison with the discharge observations.

The ensemble captures the catchment's hydrology well but better performance is limited by constraints in the model structure and the available data. It is suggested that the following recommendations can help to improve the predictive capability of hydrological models in Central Himalayan catchments;

- Monitoring of high mountainous temperatures: the lack of information on the temperatures in the mountains (≥ 3,000 masl) hampers the simulation of the hydrological processes in these areas and limits our understanding of the snow and glacier dynamics. It is acknowledged that these areas are inaccessible and the

resources of the Nepalese government for environmental monitoring are limited. However, the Himalayas are the water towers of Asia and provide water resources to billions of people downstream. Therefore, the monitoring should not be limited by the resources of Nepal and should be supported by the global community.

- Monitoring of precipitation along the southern Himalayan flank: The Indo-Gangetic plain is home to frequent and devastating floods. Monitoring the precipitation is essential to understand, reconstruct, and forecast these flood events to provide better flood management systems. The current observation network concentrates along valleys and plains. Further stations should be established at strategic locations along the slopes and ridges to gain a better understanding of orographic rainfall patterns. The Indo-Gangetic Plain is among the most flood-prone areas in terms of fatalities and affected population in the world and such investments would have a high impact.
- Discharge data: The discharge data is uncertain due to the dynamic river morphology and the large flow range. These characteristics make it very challenging to accurately observe the large Nepali River Systems. However, the large uncertainty is rarely considered in the modelling of these rivers. It is common to regard the discharge data as observations without acknowledging this uncertainty. It could be a simple and cost-effective measure to distribute the discharge data with confidence intervals, the stage-discharge rating curves, the stage-discharge measurements, and the cross-sections. This would raise awareness of the uncertainty in the data, improve the understanding of the model behaviour, and potentially improve the calibration of hydrological models.
- Dynamic parameter: The use of global parameters (in time) limits the predictive capability of modelling rivers with a strong seasonality of the hydro-meteorological conditions. This could be improved by implementing dynamic parameters which are calibrated for the specific characteristics of the seasons. This implementation is technically not complex but it would require a larger sample size in the calibration process and consequently more computational resources. Nonetheless, the effect of dynamic parameters on model performance should be explored for these rivers.
- Calibration of the lapse rate: The vast underestimation of the monsoon season snow cover indicates the poor quality of the extrapolation of the lapse rate in catchments affected by the monsoon circulation. Biases in the temperature data translate to the calibration of the parameters which control the snow and glacier dynamics. Therefore, it might provide better results to calibrate the lapse rate for elevation bands alongside these parameters.

- Coupling continuous and event-based models: The flood flows are underpredicted which is caused by the inability of the continuous model to reproduce the fast rainfall-runoff conversion to the full extent. The continuous model was chosen due to the lack of hydro-meteorological data on the sub-daily scale, the lack of information about the boundary conditions (e.g. soil moisture), and the complex hydrological system with snowmelt and glacier melt contributions. An event-based model can be coupled with the simulations of this continuous model to obtain the antecedent conditions and the snowmelt and glacier melt runoff. A statistical approach can be used to downscale the daily data to the sub-daily resolution. This coupling of continuous and event-based models may improve the predictions of flood flows and should be further explored.

The hydrological modelling complemented our knowledge of the hydrology of Central Himalayan River systems. Particularly, these simulations have contributed to a deeper understanding of the seasonal and annual variation of the runoff composition. The simulations have improved our understanding of the processes that govern floodwater generation and how these processes vary between the timing and intensity of the flood-triggering rainfall event. Furthermore, it has identified key parameters which have a high impact on the model's ability to reproduce monsoonal peak flows which may improve further modelling applications.

The established hydrological ensemble (64 parameter sets) is used in the next stage (Chapter 5 – Stage 2) to predict the climate change impact on catchment hydrology and flood flows. For this, the hydrological ensemble is applied with probabilistic climate projections. A Flood Frequency Analysis (FFA) is then conducted for the projected flood flows to quantify the changes in flood magnitudes and frequencies for different climate scenarios.

5 The prediction of the potential flood hazards on the catchment scale

This chapter focuses on Stage 2 in the modelling framework to predict the potential flood hazards for the projected climatic conditions on the catchment scale. This stage applies the calibrated hydrological modelling ensemble of Stage 1 (O1) and a Flood Frequency Analysis (FFA) to address the second and third objectives:

- O2: To predict the flood discharge at the mountain outlet for the projected climates from an ensemble of climate models until the end of the century using the hydrological model established in O1.
- O3: To quantify the flood frequencies and magnitudes at the mountain outlet for the flood discharge projected in O2.

The predicted flood discharge is then used to simulate the morphological evolution of the Karnali fan in Stage 3 (Chapter 6). The projected flood magnitudes from the FFA provide the inflow boundary conditions for the hydrodynamic simulations in Stage 4 to predict the spatial flood hazard characteristics (Chapter 7).

The first Section 5.1 introduces the problem context followed by the description of the probabilistic climate projections in Section 5.2. Section 5.3 describes the setup of the hydrological modelling and the FFA. The results of the simulations are presented in Section 5.4 and discussed in Section 5.5. Section 5.6 concludes the prediction of the potential flood hazards for the projected climates and provides recommendations to improve future research.

5.1 Introduction

Anthropogenic climate change alters the hydrological cycle and consequently impacts freshwater availability by changing the runoff composition and seasonality, and the flood hazards by changing the frequency and magnitude of flood flows (Douville *et al.*, 2021). However, the direction of changes (i.e. increase/decrease) and their drivers (e.g. rainfall, snowmelt, and glacier melt) vary in space and time which hampers the transferability of projected trends between catchments (Immerzeel *et al.*, 2013; Dankers *et al.*, 2014; Lutz *et al.*, 2014; Hirabayashi *et al.*, 2021).

Previous studies of the climate change impact on the hydrology of Central Himalayan catchments focus on the water availability and seasonality without quantifying the impact on flood flows (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014; Nepal, 2016; Shea and Immerzeel, 2016; Bajracharya *et al.*, 2018; Bhattarai *et al.*, 2018; Dahal *et al.*, 2020; Chandel and Ghosh, 2021; Khanal *et al.*, 2021). Several studies quantify the climate change impact on flood flows but these studies use coarse global climate projections and/or earlier generations of climate models (Hirabayashi *et al.*, 2013; Dankers *et al.*, 2014; Huang *et al.*, 2017; Pechlivanidis *et al.*, 2017; Wijngaard *et al.*, 2017; Hirabayashi *et al.*, 2021). However, the global projections are inferior in capturing the small-scale variability of the climate in the Himalayas (Krishnan *et al.*, 2019; Scott *et al.*, 2019).

This research updates the previous studies by using probabilistic climate projections of the latest generation of global climate models which are downscaled and bias-corrected to the regional scale to better represent the small-scale climate variation. It uses an ensemble of 13 climate projections of the Coupled Model Intercomparison Project 6 (CMIP6) which were downscaled and bias-corrected by Mishra *et al.* (2020). Quality control is conducted to identify and remove those ensemble members that do not capture the climatic characteristics in the Karnali catchment (Section 5.2).

The CMIP6 climate scenarios combine the CMIP5 anthropogenic radiative forcing scenarios with newly added Shared Socioeconomic Pathways (SSPs) which describe alternative evolutions of future societies in the absence of climate change or climate policies (O'Neill *et al.*, 2016). These SSPs enable the consideration of other socio-economic variables which affect the climate (e.g. land use change). The descriptor of the scenario combines the SSP and the anthropogenic radiative forcing. For example, the SSP126 scenario combines SSP1 (economic growth and a shift to renewable energy, investments in health and education, well-functioning institutions) with an

emission scenario which results in an anthropogenic radiative forcing of 2.6 W m^{-2} at the end of the century (O'Neill *et al.*, 2016).

The medium-emission scenario SSP245 and the high-emission scenario SSP585 are selected for the simulation of the projected streamflow. These scenarios are selected because these scenarios continue those of CMIP5 which enables the comparison of the predictions with previous studies. The SSP245 scenario represents the medium range of future emission pathways and updates the RCP4.5 of CMIP5. The SSP2 assumes a central pathway in which past trends are projected into the future (O'Neill *et al.*, 2016). The SSP585 describes the upper end of the emission pathways and is the successor of RCP8.5. The SSP5 assumes high investments in health and education and the rapid growth of a fossil-based economy (O'Neill *et al.*, 2016). These scenarios cover the range of medium to high climatic changes. The SSP126 scenario represents the best-case scenario which would limit the global climate increase below 2ºC. However, the comparison of observed and projected climate change indicates that this scenario is less likely (Carvalho *et al.*, 2022; UN, 2023). Therefore this scenario is omitted in light of the limited computational resources

The SSP245 and SSP585 climate projections are used to predict the climate change impact on flood flows by combining a catchment-scale hydrological model and a flood frequency analysis. For this, the hydrological model that was established for the Karnali River in Chapter 4 is applied with climate projections of an ensemble of CMIP6 models to predict the flood discharge at the mountain outlet. The baseline scenarios of the past climatic conditions are simulated as a reference to quantify the change between the past and projected climates. A Flood Frequency Analysis (FFA) is then conducted with the simulated flood discharge (the annual maximum flows) to determine the flood hazard frequencies and magnitudes. The magnitudes of a given frequency (e.g. the 1-in-100-years event) is then compared between the scenarios to quantify the climate change impact on flood hazards.

5.2 The evaluation of the probabilistic climate projections

The hydrological modelling ensemble utilises the same static datasets (i.e. land cover, geology, soil, topography) of the historical simulations (Chapter 4) and the reader is referred to Section 4.3.1 for information about these datasets. The climate data is replaced with projections of the CMIP6 ensemble which is described in the following section.

The CMIP6 ensemble consists of the projected climates predicted from 13 Global Circulation Models (GCM) and contains the daily precipitation, and the daily minimum and maximum air temperatures. The spatial resolution of these GCMs ranges between 0.7º and 2.0º. The datasets have been downscaled to 0.25º for South Asia by Mishra *et al.* (2020). This downscaling included a bias correction procedure with empirical quantile mapping. The mean temperature which is required by the SPHY model is calculated as the mean of the daily minimum and maximum temperature.

The climate projections are analysed in the next section to gain an understanding of the projected changes which is relevant for the interpretation of the hydrological projections and to assess their quality and remove those projections that do not capture the climatic characteristics of the Karnali catchment in the baseline scenario. Section 5.2.2 presents the regionalisation of the CMIP6 data to the resolution of the hydrological model.

5.2.1 The analysis of the climate projections

The Mean Annual Air Temperature (MAAT) (30-year mean) is projected to increase with time and emissions (Figure 5.1). The ensemble-mean MAAT increases from 7.3 ºC at the end of the baseline scenario in 2014 to 9.8 ºC (SSP245) and 11.8 ºC (SSP585) at the end of the century. The variation between the ensemble members increases in time from 0.5 ºC (7.1 – 7.6 ºC) in 2014 to 1.5 ºC (9.1 – 10.6 ºC) for SSP245 and 1.9 ºC (11.0 – 12.9 ºC) for SSP585 in 2099. The temperatures projected by the CanESM5 model are notably higher than for the other ensemble members and reach 11.3 ºC and 14.6 ºC for the SSP245 and SSP585 scenarios, respectively.

The temperatures are projected to increase in every season whereas the projected change is higher in the winter (Dec – Feb) and the monsoon season (Jun – Sep) (Figure 5.2). The ensemble-mean temperatures are projected to increase by 3.0 ºC (SSP245) and 5.1 ºC (SSP585) in the winter and by 2.6 ºC (SSP245) and 4.5 ºC (SSP585) in the ,onsoon season. The temperatures of the pre-monsoon (Mar – May) and post-monsoon

(Oct – Nov) seasons are projected to increase by 2.1 – 2.2 ºC (SSP245) and 3.9 – 4.1 ºC (SSP585).

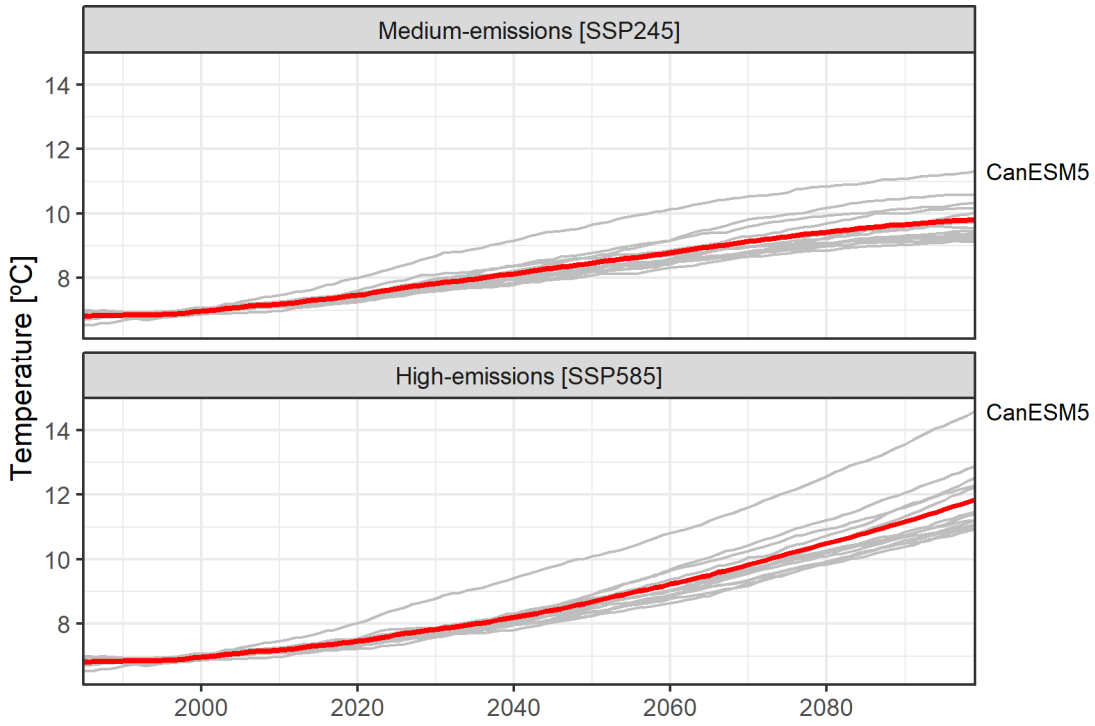


Figure 5.1: Time series of the catchment-average Mean Annual Air Temperature (MAAT) (30-year mean) for the CMIP6 ensemble members. The red line indicates the ensemble-mean MAAT, and the grey lines show the MAAT of the individual members. The period before the year 2014 is obtained from the baseline.

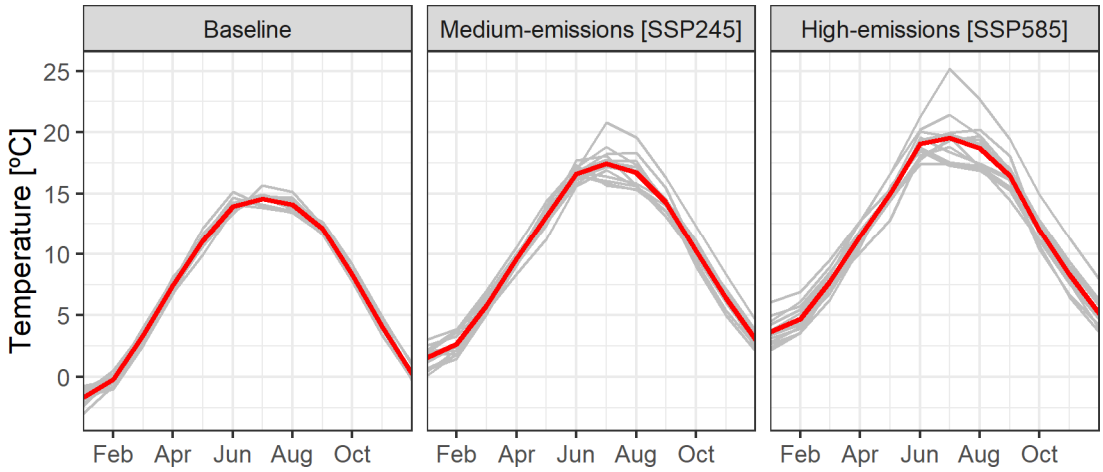


Figure 5.2: Mean monthly air temperature of the CMIP6 ensemble for the three scenarios. The periods are 1985 – 2014 for the baseline and 2070 – 2099 for the SSP245 and SSP585 scenarios. The red line indicates the ensemble mean, and the grey lines indicate the individual members. The temperatures represent the catchment mean.

The annual precipitation is projected to increase alongside the temperatures, whereas this increase scales with time and emissions (Figure 5.3). The 30-year mean annual precipitation (ensemble mean and catchment mean) is 1,030 mm at the end of the baseline period in 2014. This precipitation is projected to increase by 11% for SSP245 and 13% for SSP585 by the year 2065. The difference between both scenarios increases towards the end of the century with a moderate further increase of 4% to 1,190 mm for SSP245 and a high increase of 23% to 1,400 mm/a for SSP585.

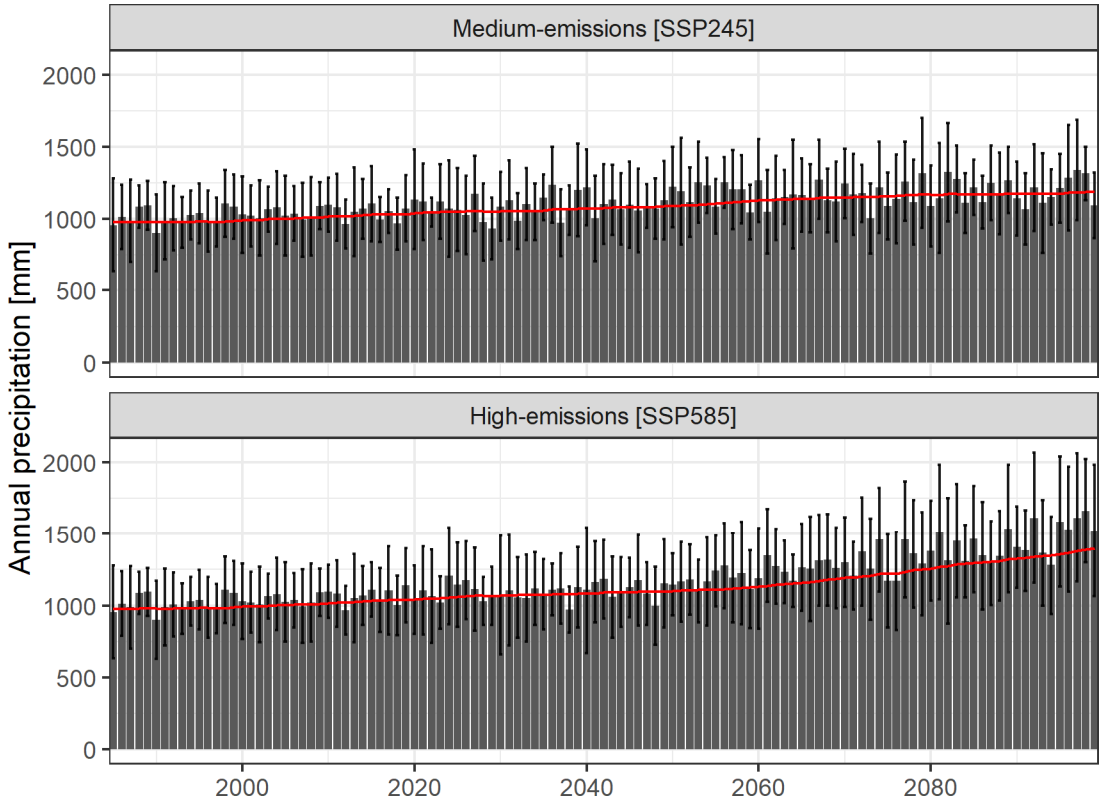


Figure 5.3: Time series of the catchment-mean annual precipitation. The bars show the ensemble-mean precipitation and the error bars the standard deviation of the ensemble members. The red line shows the 30-year mean annual precipitation of the ensemble. The period before the year 2014 is obtained from the baseline scenario.

The CanESM5 model does not depict the seasonality in the catchment of the Karnali River (Figure 5.4). The annual precipitation is dominated by the monsoon precipitation which accounts for 80% of the annual precipitation (Gautam and Acharya, 2012; Collier and Immerzeel, 2015; Krishnan *et al.*, 2019). This seasonality is captured by the CMIP6 members with an ensemble mean monsoon precipitation of 74% (60 – 86%) in the baseline. The only exception is the CanESM5 model which predicts only 18% of the annual precipitation during the monsoon season. Therefore, this model does not reproduce the precipitation seasonality and is removed from the ensemble. The remaining models are maintained as they reproduce the monsoon seasonality. These models predict an increasing contribution of the monsoon precipitation which increases to 78% (SSP245) and 79% (SSP585) (ensemble mean).

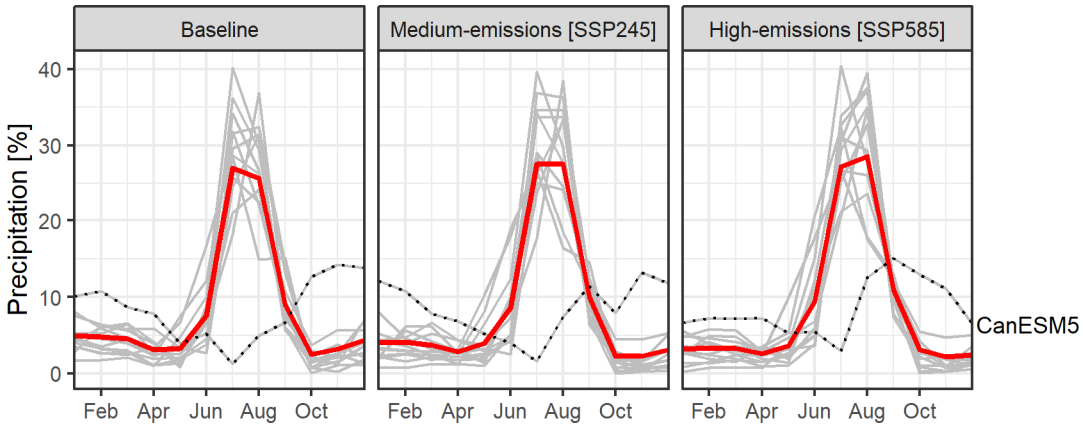


Figure 5.4: Mean monthly precipitation as a percentage of the annual precipitation for the baseline (1985 – 2014) and the projected scenarios (2070 – 2099). The red line shows the ensemble mean precipitation and the grey lines are the precipitation of the ensemble members. The dotted line shows the distribution of the CanESM5 model which is distinctively different to the other members. The precipitation was averaged (mean) over the catchment.

The projected increase in precipitation coincides with an increase in the number of heavy rainfall events (Figure 5.5). Events with catchment-mean precipitation rates ≥ 50 mm per day occur on average (ensemble-mean) 18 times in the baseline (1985 – 2014). The frequency of these events are projected to increase to 40 times (SSP245) and 73 times (SSP585) in the period 2070 – 2099. However, the variation between the ensemble members is large; several models predict a vast increase in the occurrence of these events (e.g. ACESS-ESM1-5; EC-Earth3, EC-Earth3-Veg) while others project a small increase (e.g. MPI-ESM1-2-HR) or a decrease (MPI-ESM1-2-LR). This increasing ensemble variation manifests in the standard deviation which increases from ±4 events in the baseline scenario to ±20 (SSP245) and ±47 (SSP585).

The trend of increasing occurrence and ensemble variations are also projected for events with catchment-mean precipitation rates ≥ 100 mm per day (Figure 5.5). The number of these events is projected to increase from 1±1 in the baseline to 4±2 and 11±9 in the SSP245 and SSP585 scenarios, respectively. The models which project a large increase in events ≥ 50 mm per day also project a large increase for events ≥ 100 mm per day. However, there are deviations for some models; e.g. IN-CM-4-8 projects a decrease in the number of ≥ 100 mm per day events, and MPI-ESM1-2-LR projects an increase in ≥ 100 mm per day events but a decrease in ≥ 50 mm per day events.

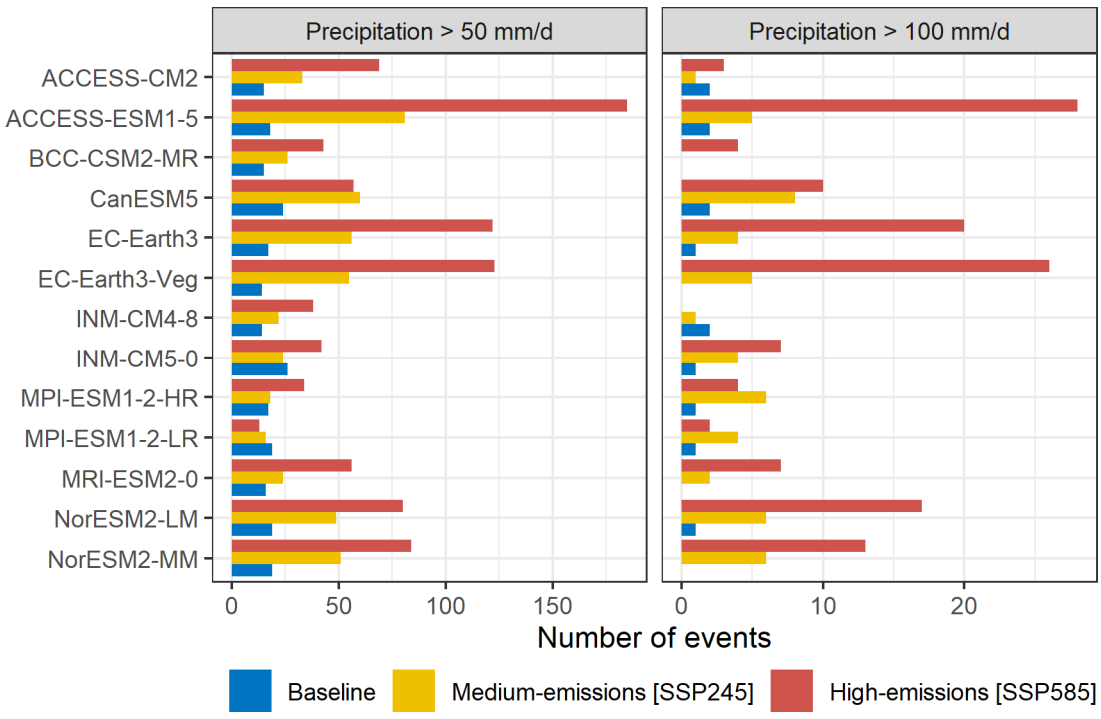


Figure 5.5: Total number of heavy rainfall events with catchment mean precipitation rates ≥ 50 mm per day (left) and ≥ 100 mm per day (right) for the baseline (1985 – 2014) and the projected scenarios (2070 – 2099) for each CMIP6-ensemble member. The precipitation rates refer to the catchment mean.

5.2.2 The regionalisation of the dataset

The CMIP6 climate data has a spatial resolution of 0.25 X 0.25º (~24 km at the latitude of the Karnali River catchment) which is too coarse for the hydrological simulations. The dataset is therefore downscaled to the modelling resolution of 500 X 500 m.

The temperature is regionalised using a lapse rate approach. The reference elevation of each CMIP6 grid cell is calculated by calculating the mean grid cell elevation from the HydroSHED SRTM DEM (see Chapter 4.3.1). The difference between the SPHY elevation and the reference elevation is then calculated in the second step. This elevation difference is then multiplied with a lapse rate of -0.0054 ºC/m (see Chapter 4.3.2) to obtain the temperature difference between the reference and modelling elevation. The no-data grid cells of the CMIP6 rasters are then filled by bilinear interpolation and the obtained raster is resampled to the modelling resolution by bilinear interpolation. The temperature difference is then added to the temperature raster to obtain the regionalised modelling temperature map. This procedure is the same as the regionalisation of the WFDEI temperatures in Chapter 4 with one difference; the reference elevation for the WFDEI regionalisation was based on the WFDEI reference elevation. However, such a dataset does not exist for the CMIP6 dataset because it was downscaled by a statistical approach (Mishra *et al.*, 2020). Therefore, it was assumed that the grid-cell mean HydroSHED DEM elevation is an appropriate descriptor of the reference elevation. An example of the original and regionalised temperature is shown in Figure 5.6.

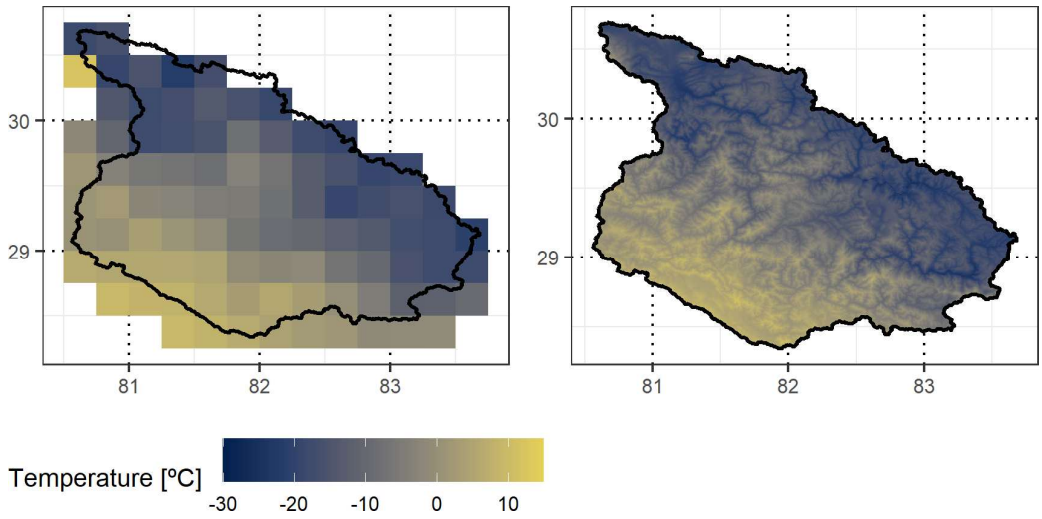


Figure 5.6: The original CMIP6 temperature of Mishra et al. (2020) (left) and the regionalised SPHY input map (right). These maps show the 2053-12-24 of the EC-Earth3 SSP245 scenario. The coordinate system is WGS84 (EPSG: 4326).

The precipitation is downscaled from statistical interpolation because there is no information on the precipitation characteristics at higher elevations (> 3,500 masl) and lapse rate estimations from glacier mass balance modelling are deemed out of the scope of this research (Immerzeel *et al.*, 2015). Instead, it is assumed that the bias-corrected CMIP6 data represents the large-scale orographic rainfall patterns reasonably and is regionalised without elevation correction. For this, the precipitation of the no-data grid cells along the northern border is estimated using bilinear interpolation, and this filled dataset is then downscaled to the modelling resolution using bilinear interpolation. An example of the original and regionalised precipitation is shown in Figure 5.7.

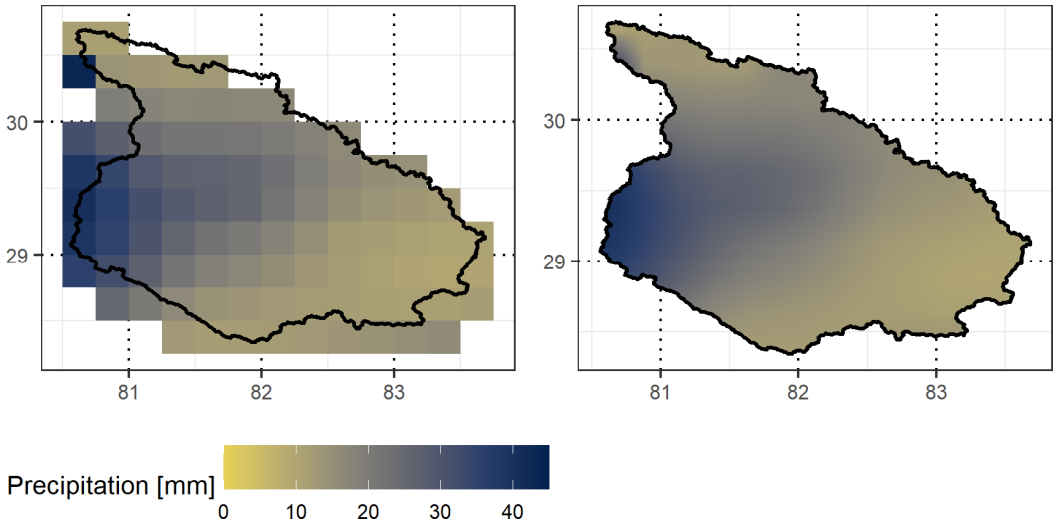


Figure 5.7: The original CMIP6 precipitation of Mishra et al. (2020) (left) and the regionalised SPHY input map (right). These maps show the 2053-07-22 of the EC-Earth3 SSP245 scenario. The coordinate system is WGS84 (EPSG: 4326).

5.3 Methods

The calibrated hydrological model (Chapter 4) is linked with a Flood Frequency Analysis (FFA) to predict the changes in flood hazard frequencies and magnitudes for the projected climates. The setup of the hydrological model is presented in Section 5.2.1 and a detailed description of the hydrological model was presented in Section 4.2.1. An overview of the FFA is presented in section 5.2.2.

5.3.1 The simulation of the projected flood flows

The hydrological model Spatial Processes in Hydrology SPHY (Terink *et al.*, 2015a) was calibrated and validated in Chapter 4 and 64 parameter sets were identified as behavioural representations of the Karnali River system. This ensemble of 64 hydrological models is applied with climate projections of the 12 CMIP6 models to simulate the hydrological response to these projected climates. A precipitation correction factor was implemented for the model calibration to account for the bias in the observed precipitation data. This factor is removed from the modelling because it was calibrated to a specific dataset with a specific bias, and it is questionable that each of the CMIP6 members has this specific bias. No other adjustments are made to the hydrological ensemble except for the removal of the precipitation correction factor and the replacement of the climate data with the CMIP6 projections.

The hydrological models are applied with climate projections of three different scenarios. The baseline scenario represents the past climatic conditions and is applied for 40 years from 1975 – 2014 with the years 1971 – 1974 as the warm-up period and is used as the reference to quantify the changes of the future scenarios. The medium-emission scenario SSP245 and the high-emission scenario SSP585 are applied for the period 2020 – 2099 with the years 2015 – 2019 as the warm-up period. The CIMP6 ensemble comprises 12 different models which are applied to each of the 64 hydrological parameter sets. Therefore, the prediction ensemble comprises 768 (12 x 64) models for each scenario.

The hydrological model was validated against the observed catchment behaviour in Chapter 4, and it was established that it is a behavioural representation of the hydrological system of the Karnali River. It was calibrated (and validated) against specific datasets. However, the parameter transfer to the CMIP6 climate data might result in a shift in the simulated behaviour. The simulations of the projected climates, therefore,

require some sort of validation. A direct comparison with observations is not possible because the baseline scenarios do not represent actual events (e.g. the 2013 Karnali flood). The comparison of observed and simulated distributions is hampered by the short record length (≤14 years) in conjunction with the large internal variability of the climatic variables. Therefore, an indirect model evaluation is conducted by comparing the baseline simulations with the historical simulations of Chapter 4 to understand the implications of the parameter transfer for the hydrological simulations.

The hydrological model is applied at the daily time step and has a spatial resolution of 500 x 500m. The simulations are conducted on the High-Performance Computing (HPC) facility at Durham University.

5.3.2 The simulation of the flood magnitudes and frequencies

The flood discharge describes the amount of water that is released into the floodplain at the individual events. The FFA is applied with these flood discharges to quantify the differences between the climate scenarios. The FFA is a statistical method to determine the flood magnitudes and frequencies from a flow record. This method standardises flood events which enables the comparison of different flood events of similar frequencies. For example, the climate change impact on flood hazards can be quantified by comparing the magnitudes of 1-in-100-year events of the past climatic conditions (the baseline) and the projected scenarios (the medium- and high-emission scenarios).

The FFA is a data-driven method which requires a record of flood events. A comprehensive review of the different methods and limitations of the FFA was provided in section 2.3. The extraction of the flood flows requires the definition of flood events which is most commonly conducted by the Annual Maximum Flow (AMAX) or Peak-over-Threshold (POT) approaches (WMO, 1989). The AMAX approach classifies the maximum flow of each year as a flood event and this approach is chosen to avoid the uncertainty of defining a flood threshold. In this study, the flood record is derived from the simulated flow of the hydrological model. The simulated records are split into three periods each consisting of 40 years (and hence 40 flood events); the baseline period from 1975 – 2014 of the baseline scenario, the near-future from 2020 – 2059 of the SSP245 and SSP585 scenarios, and the far-future from 2060 – 2099 of the SSP245 and SSP585 scenarios.

One of the main uncertainty sources of FFAs is the sampling error that relates to the deviation between the sampling distribution and the true distribution of flood magnitudes. The sampling error is reduced by increasing the record length from the climate reference period of 30 years to 40 years. The sampling error uncertainty is estimated using the bootstrapping method proposed by Burn (2003) where the 40-year record is duplicated 999 times, concatenated and randomly divided into 999 40-year records. A further source of uncertainty is related to the errors of the flood flows. This study uses simulated flood discharges predicted by an ensemble of 768 model realisations for each scenario. These ensemble predictions are used to estimate the measurement uncertainty. The estimation of sampling error and measurement error uncertainties results in the fitting of 767,232 flood frequency curves (12 CMIP6 models X 64 hydrological models X 999 bootstrapping samples).

The FFA uses Extreme Value Distributions (EVD) to simulate the distribution of the flood record. However, it is impossible to determine which EVD estimates the empirical distribution of the flow record best. Therefore, the following EVD are fitted to the simulated records (not the bootstrapped records) using the L-Moments parameter estimation method as recommended by Hosking and Wallis (1997); the Gumbel distribution (GUM), Weibull distribution (WEI), Generalised Extreme Value distribution (GEV), Generalised Logistic distribution (GLO), Pearson Type III distribution (PT3), Generalized Pareto distribution (GPA), Exponential distribution (EXP), and Wakeby distribution (WAK). The evaluation is conducted on the Anderson-Darling and Kolmogorov-Smirnov tests since the application of graphical evaluation techniques is unfeasible for the number of flood frequency curves. The distribution with the best median test statistics for all fitted frequency curves is selected for the final FFA including the sampling error uncertainty estimation.

5.4 Results: Predicting the projected flood events

This section presents the results of the simulated projections of the flood discharges, flood magnitudes, and flood frequencies. Since the hydrological model was calibrated from the observed hydro-meteorological conditions (Chapter 4), the parameter transfer from the observed data to the simulated CMIP6 data might change the behaviour of the hydrological model. Therefore, the first Section, 5.4.1, compares the CMIP6 baseline scenario (hereafter baseline simulations) with the observations and the historical simulations of Chapter 4 (hereafter historical simulations) to investigate the validity of the hydrological model. The catchment response to the projected climates is then presented in section 5.4.2. These two sections provide information about the model validity and set the foundation for the predicted flood discharges in Section 5.4.3. These flood flows are the basis for the FFA and the predicted flood frequencies and magnitudes are presented in the final Section 5.4.4.

5.4.1 The comparison of the baseline and historical simulations

The historical simulations are compared with the baseline ones to investigate changes in the model behaviour which is introduced by the application of the CMIP6 data and the parameter transfer between the climate datasets. River discharge has a large internal variability which leads to deviations between the historical simulations with a 14-year record and the baseline simulations with a 40-year record (Giuntoli *et al.*, 2018; Martel *et al.*, 2018; Bérubé *et al.*, 2022). However, strong and systematic differences indicate shifts in this behaviour hence the comparison of historic and baseline simulations provides insights into the reliability of the predictions of the CMIP6 simulations. The runoff composition is the product of the different paths which water takes through the hydrological system and is used as the indicator to compare the model behaviour of the historical and baseline simulations. Furthermore, the total discharge is compared with the observed discharge record of the respective period 1975 – 2014.

5.4.1.1 Catchment discharge

The baseline simulations predict lower annual discharge than observed and predicted by the historical simulations (Figure 5.8). The percentage difference between the median annual discharge of the baseline simulations (38 Billion m^3) is -13% for the median observed annual discharge (43 Billion m^3) and -19% for the median historical annual discharge (46 Billion m^3). Therefore, the baseline simulations produce less discharge as observed and as simulated with the historical climate data. This lower annual discharge is predicted for each CMIP6-baseline member whereas two members (NorESM2-LM, ACCESS-CM2) generate higher median annual discharge (40 - 41 Billion m^3) than the other members (37 – 38 Billion m^3) (Figure 5.9). However, the lowest and highest annual discharges are predicted for the baseline simulations (10 and 67 Billion m^3). This large prediction range is due to the large sample size that is composed of 12 CMIP6 models, the ensemble of 64 SPHY models, and the record length of 40 years (n = 30,720). This is considerably larger than the Historic simulations with 64 SPHY models for 13 years (n = 832), and the 40 years of the observed annual flows (n = 40).

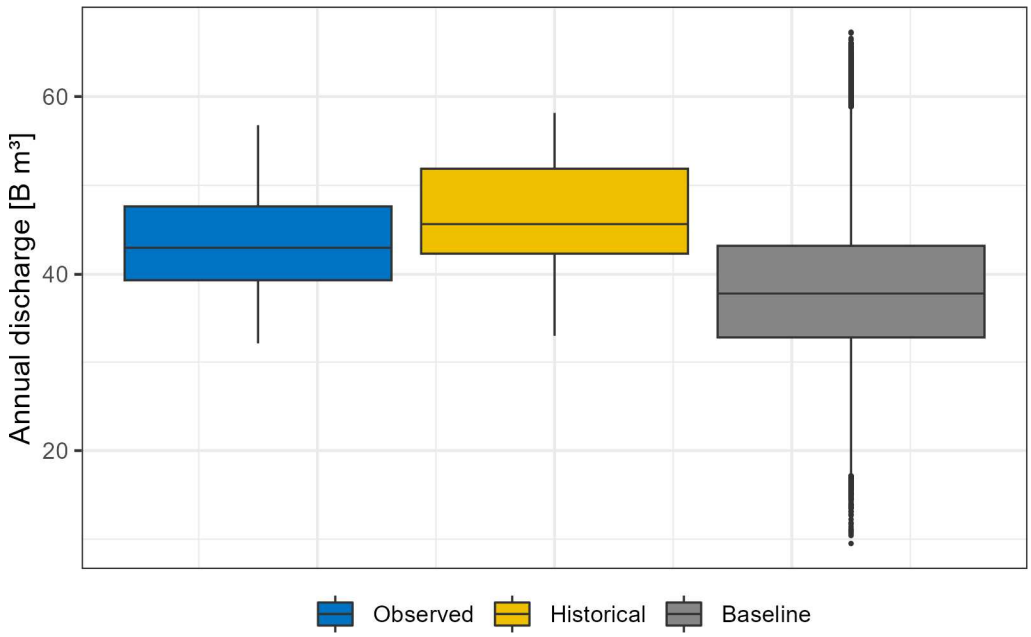


Figure 5.8: Comparison of the annual discharge at the catchment outlet of the observations (1975 - 2014), the historical simulations (2002 - 2015), and the baseline simulations (1975 - 2014). The ensemble sizes of the datasets are 40, 832, and 30,720 respectively.

The lower discharge of the baseline simulations results from the deficit of the monsoon season discharge (Figure 5.10). The ensemble median baseline discharge is higher than the observed one for most months throughout the year. The percentage difference between the baseline simulations and the observed discharge is 12 – 17% in the winter (Dec – Feb). It increases to 21 – 27% in the pre-monsoon season (Mar – May) and declines to 6% at the beginning of the monsoon season in June. The surplus turns into a deficit which increases from -26% in July to -43% in September. The highest absolute deficit is predicted for August (-3.2 Billion m^3). The discharge differences transition from under- to overpredictions during the post-monsoon season in October (-10%) and November (+14%). The discharge deficit (ensemble median) of the monsoon season is -7.6 Billion m^3 and exceeds the non-monsoon season surplus of 1.4 Billion m^3 by a factor of 5.5.

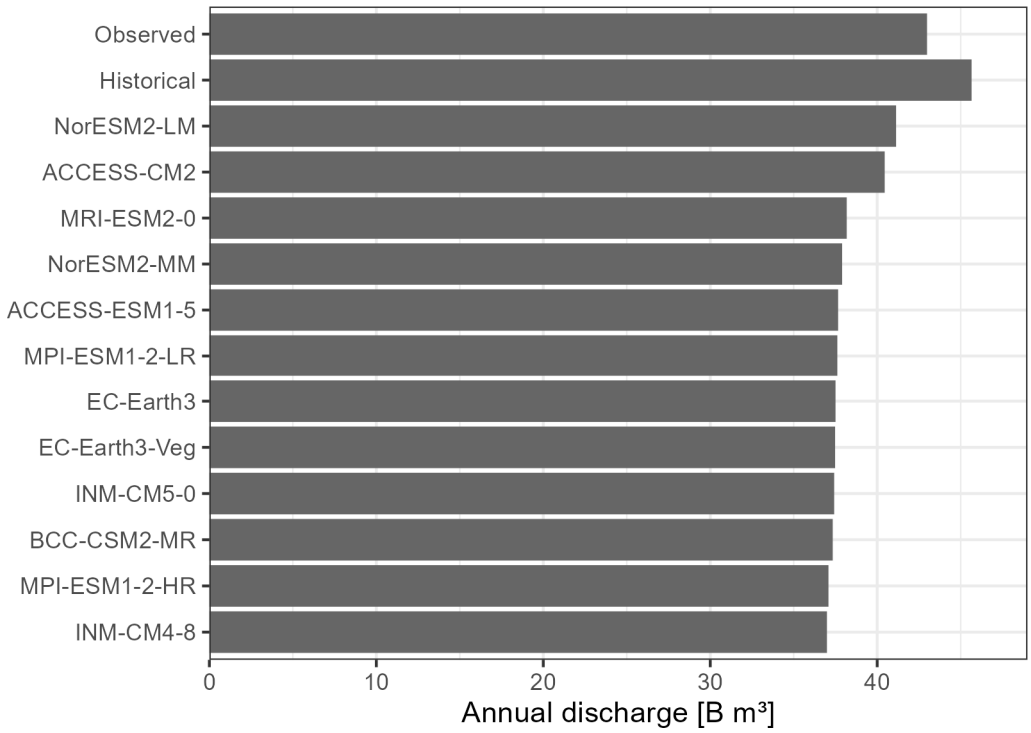


Figure 5.9: Comparison of the median annual discharge of the baseline ensemble members, the historical simulations and the observations.

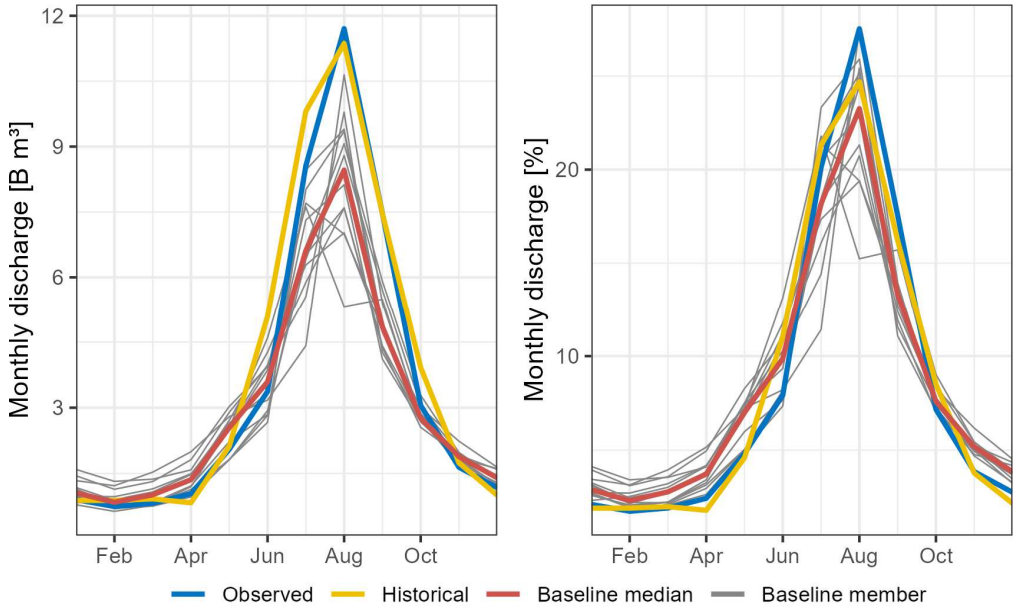


Figure 5.10: Comparison of the median monthly discharge at the mountain outlet of the observations (1975 - 2014), the historic simulations (2002 - 2015), and the baseline simulations (1975 - 2014).

These trends of overpredicted non-monsoon season discharge and underpredicted monsoon season discharge are consistent for each baseline ensemble member (Figure 5.11). The median observed monsoon season discharge accounts for 73% of the annual discharge. The highest baseline monsoon season contribution is simulated for the ACCESS-ESM1-5 (70%) and EC-Earth3-Veg members (67%) and the lowest one for the NorESM2-LM and MRI-ESM2-0 members (59%). The two members with the highest median annual discharge (ACCESS-CM2, NorESM2-LM) are among the members with the lowest monsoon season contribution. Hence, their high annual discharge is not primarily driven by the monsoon season precipitation but is also affected by higher precipitation in the non-monsoon season.

The baseline ensemble predicts the seasonality with high discharge during the monsoon season and low flows during the winter and the transition between these seasons well despite the systematic over- and underpredictions. This manifests in the high coefficient of determination (R^2) between simulated and observed median monthly discharge for each ensemble member (Figure 5.12). The INM-CM4-8 member has the lowest R^2 which is with 0.78 still a reasonable representation of the runoff seasonality. The R^2 of the other members is between 0.88 and 0.99. Hence, more than 88% of the monthly discharge variation is reflected in the prediction of the ensemble members. The EC-Earth3 and EC-Earth3-Veg members reflect the runoff seasonality exceptionally well as indicated by the R^2 values of 0.99.

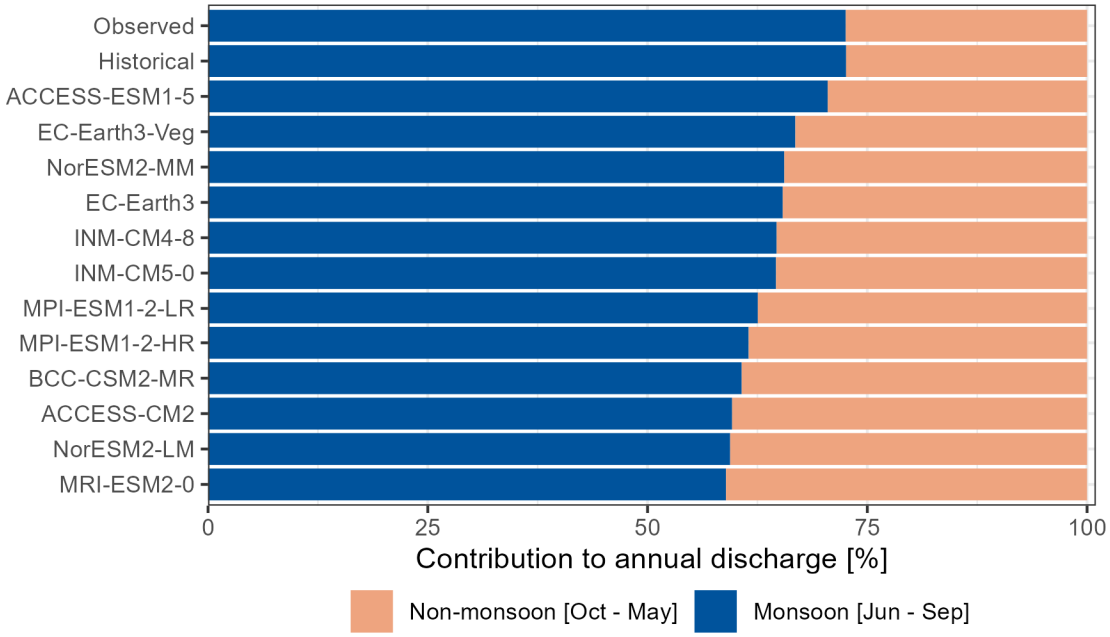


Figure 5.11: Median seasonal contribution to the annual discharge for the baseline ensemble members, the historical simulations and the discharge observations.

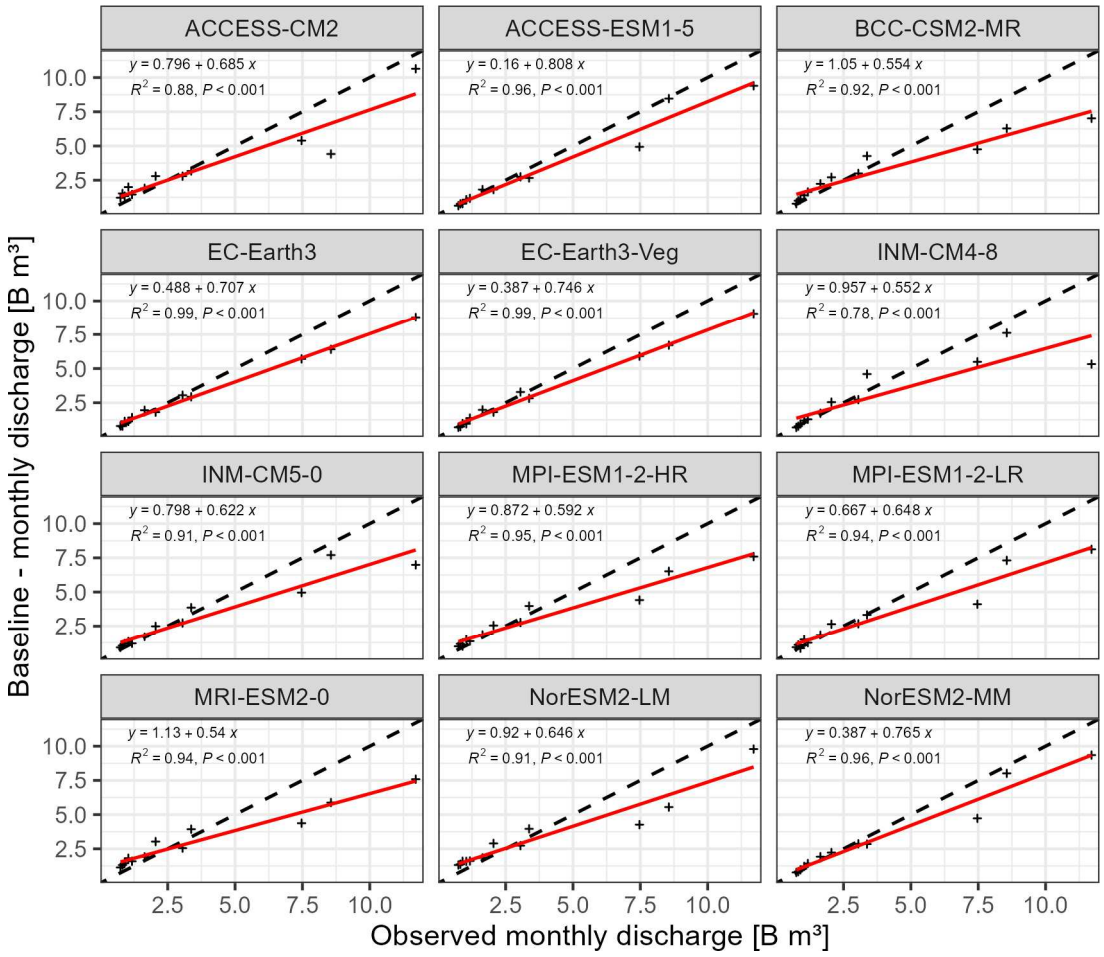


Figure 5.12: Comparison of the observed and simulated (baseline scenario) median monthly discharge. The red line shows the linear regression model and the dashed line shows the line of the perfect fit.

5.4.1.2 Runoff composition

The previous section verified that the baseline simulations represent the seasonality in the Karnali River system. This discharge is the sum of the hydrological processes in the upstream catchment and, as such, an unreasonable quantification of these processes could potentially add up to a behavioural discharge representation Furthermore, different model parameterisations with different internal behaviour may produce similar outputs (model equifinality). This section investigates whether the parameter transfer caused changes in the ensemble behaviour by analysing the runoff composition.

The lower annual discharge predicted for the baseline simulations results from lower rainfall-runoff contributions (Figure 5.13). The mean annual contribution[2] of this component is 27 Billion m^3 in the historical simulations and 15 Billion m^3 in the baseline simulations. Consequently, the rainfall-runoff loses importance for the annual budget of the baseline simulations with a mean contribution of 40% compared to the 59% contribution in the historical simulations. Higher glacier melt compensates in part for the lower rainfall-runoff. This component contributes 1.7 Billion m^3 (4%) in the historical simulations but is an important runoff source in the baseline simulations with contributions of 7.6 Billion m^3 (20%). The baseflow runoff is similar for both simulations with contributions of 11 Billion m^3. However, the relative importance increases from 23% in the historical simulations to 28% in the baseline ones due to the lower total annual discharge. Snowmelt runoff reduces from 7 Billion m^3 (14%) to 4 Billion m^3 (11%) in the baseline simulations (mean predictions).

The larger prediction range in the baseline simulations is expected due to their larger ensemble size (768 vs 64) and longer record length (40 years vs 14 years). The rather small differences in the snowmelt contributions could result from these differences in the record length because the historical simulations are more prone to be affected by internal climate variability. However, the differences in rainfall-runoff and glacier melt are too large to be only caused by internal climate variability.

[2] The median is commonly used in this thesis to aggregate the hydrological ensemble. However, the median runoff contribution does not add up to 100% which is why the mean is used to describe the runoff composition.

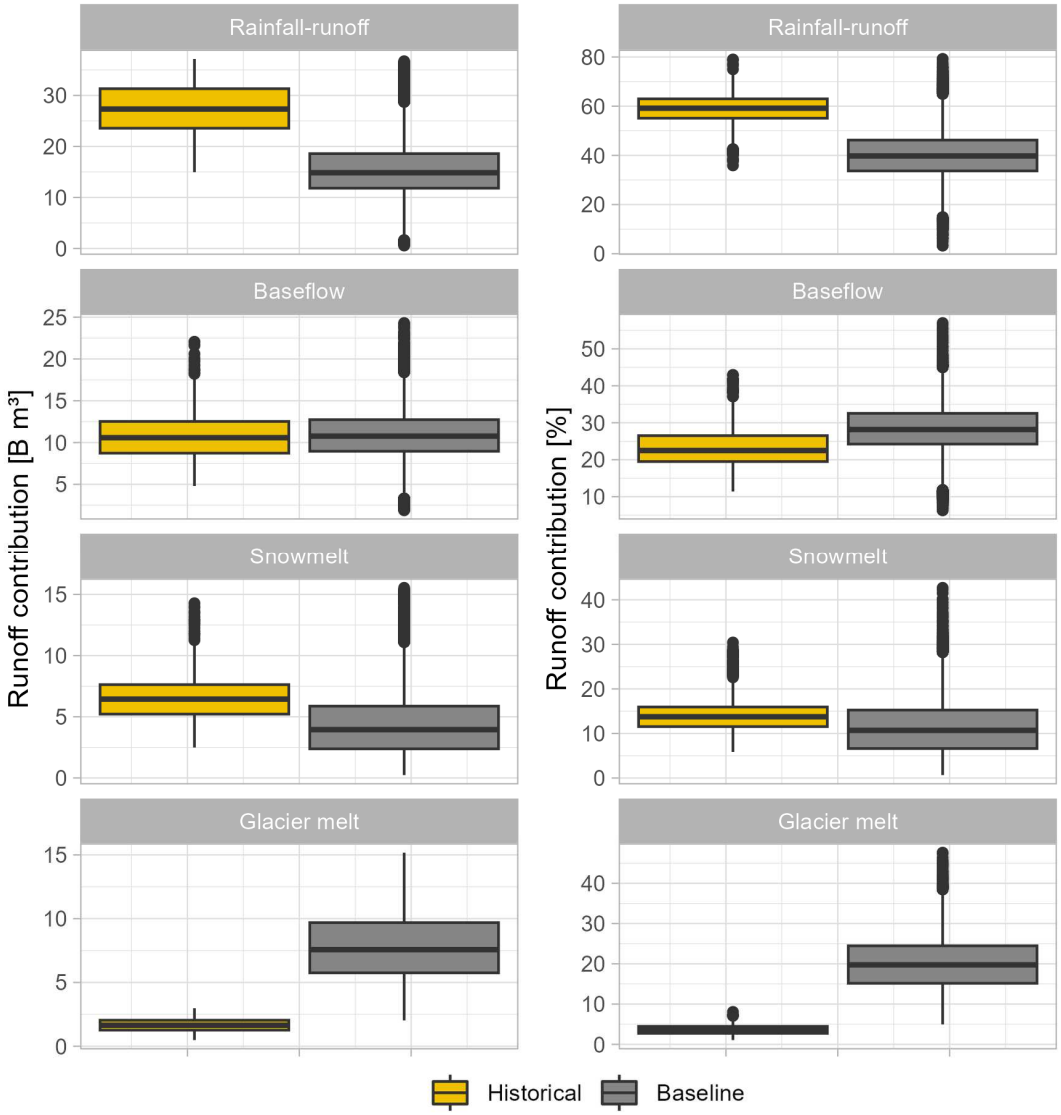


Figure 5.13: Annual runoff composition of the historical simulations (2002 - 2015) and the baseline simulations (1975 - 2014).

Most rainfall-runoff drains during the monsoon season, and hence the highest rainfall-runoff deficit of the baseline simulations is generated during this season (Figure 5.14). This deficit is 1.8 Billion m^3 at the onset of the monsoon season in June and increases to 3.6 Billion m^3 in July. Afterwards, it decreases to 2.9 Billion m^3 and 2.6 Billion m^3 in August and September, respectively. However, while the largest absolute difference is simulated for the peak monsoon season (Jul – Aug), the percentage-difference is larger for the early monsoon season in June (119%) and late monsoon season in September (77%). The total difference in the simulated monsoon season rainfall-runoff (10.9 Billion m^3) is larger than the total runoff deficit (9.3 Billion m^3).

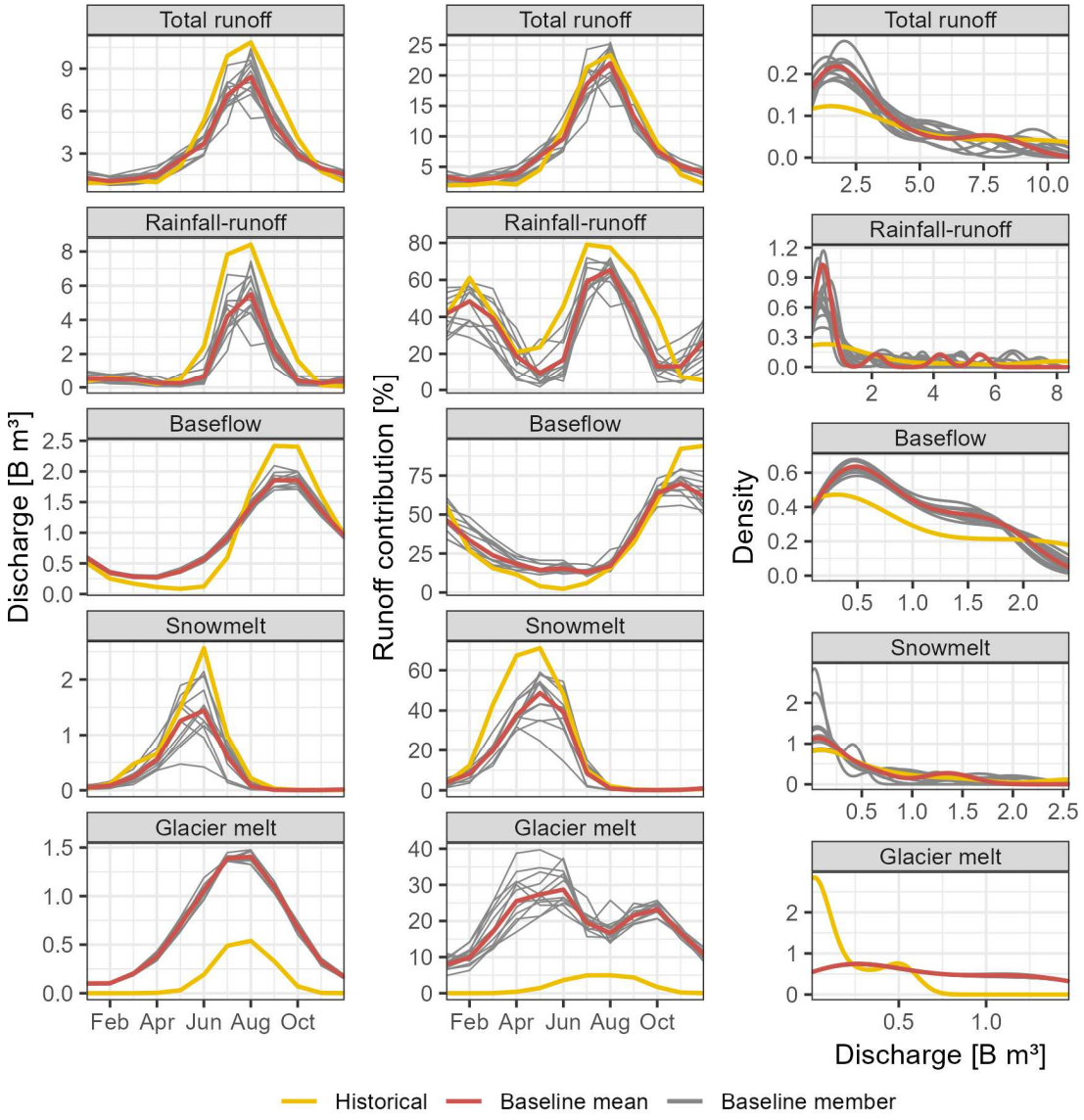


Figure 5.14: Mean monthly runoff (left column); b) Mean monthly runoff contribution (centre column); c) Density curves of the mean monthly runoff (right). The yellow line is the mean monthly runoff of the historical simulations (2002 – 2014), the red line is the monthly mean of the baseline ensemble (1975 – 2014), and the grey lines are the monthly mean of the baseline members (1975 – 2014). The runoff contribution of the total runoff refers to the mean monthly contribution to the annual discharge while the runoff contribution of the components refers to the contribution to the monthly discharge.

The timing and duration of the monsoon season differ between baseline and historical simulations, which is particularly prevalent in the rainfall-runoff distribution (Figure 5.14). The monsoon season starts earlier and lasts longer in the historical simulations. The rainfall-runoff contributes 46% to the monthly discharge in June in the historic simulations whereas it contributes 17% in the baseline simulations which indicates that the monsoon season starts later in the CMIP6 simulations. Furthermore, the CMIP6 data has lower late monsoon rainfalls which is indicated by the lower rainfall-runoff

contribution in October in the baseline simulations (13%) compared to the historical simulations (40%). However, the rainfall-runoff is lower in the baseline simulations throughout the monsoon season and this deficit is with -3.6 Billion m^3 highest in July and decreases to -2.6 Billion m^3 in September (mean predictions).

The lower rainfall contribution of the baseline simulations is, in part, compensated by higher baseflow and glacier melt contributions (Figure 5.14). The baseflow contributions are 0.4 Billion m^3 and 0.3 Billion m^3 higher in June and July but 0.3 Billion m^3 and 0.6 Billion m^3 lower in August and September and, thus, adding to the total discharge deficit. Contrary, glacier melt contributions are considerably larger throughout the year in the baseline simulations. This difference ranges between 0.8 – 0.9 Billion m^3 in the monsoon season. Therefore, glacier melt is more important in the baseline simulations and contributes 17 – 29% to the monthly discharge compared to the 4 – 5% contribution of the historical simulations. The baseline snowmelt contributions are lower throughout the monsoon season, particularly at the peak of the melting season in June for which this difference is 1.1 Billion m^3 (mean predictions).

The higher baseline discharge during the non-monsoon season is driven by different contributors. In October, the lower rainfall-runoff causes a deficit which is reduced by the higher glacier melt contributions. The baseline discharge is higher from November until the start of the next monsoon season. In November, this difference is 0.2 Billion m^3 which is caused by higher rainfall-runoff and glacier melt rates, whereas the 0.2 Billion m^3 lower baseflow reduces the runoff differences by 50%. In December and Januarty, the surplus (0.4 – 0.5 Billion m^3) is caused by higher rainfall-runoff (0.2 – 0.4 Billion m^3). In February and March, the discharge of both simulations converge and the contributions are similar. The baseline simulations predict higher discharge in April and May and this is driven by higher glacier melt and, to a lower degree, higher baseflow contributions. The snowmelt component predicts lower contributions in the baseline simulations throughout the year.

The seasonality with low discharge in the non-monsoon season and high discharge during the monsoon season is more pronounced in the historical simulations which is also prevalent in the density curves of the monthly discharge (Figure 5.14). A density curve describes the frequency of samples in the value range, whereas the area under the curve equals all samples. A higher density indicates a higher frequency of samples for the corresponding value. All density curves have the maximum in the lower 25% quantile of the runoff range due to the longer duration of the non-monsoon season (Oct – May). However, the density curves of the historical simulations are flatter than the baseline ones indicating a more pronounced runoff seasonality. The larger density of flows above the 75% quantile in the historical simulations emphasises the higher

monsoon season runoff of the historical simulations. This is the case for the total discharge and all components except glacier melt, where the density curve of the baseline simulation is flatter, indicating the higher baseline glacier melt rates.

The baseline ensemble reproduces the seasonality of the Karnali river system well and agrees with the historical simulations despite the less pronounced seasonality (Figure 5.15). The rainfall-runoff remains, although on a lower level, the most important runoff source with similar seasonal patterns (Figures 5.13 and 5.14). This is also valid for the baseflow as the second-largest runoff contributor in both simulations. Snowmelt loses importance and is the smallest contributor to the total runoff in the baseline simulations. However, its seasonality is well depicted and it remains the dominant runoff source in the pre-monsoon season. The largest difference between the baseline and historical simulations is the glacier melt which becomes more important than snowmelt in the baseline simulations. However, the contribution increase from 4% in the historical simulations to 20% in the baseline ones is also caused by the lower total runoff which is driven by the reduction of rainfall-runoff. Nevertheless, both simulations predict a similar seasonality for the total runoff and each component. This is illustrated by the strong linear relationship betwwen the monthly runoff rates predicted for the historical and baseline simulations (Figure 5.15). The coefficient of determination (R^2) ranges between 0.94 – 0.98 except for glacier melt. The R^2 is lower due to the extended melt season but remains on a high level (0.86). The high agreement of the monthly runoff rates indicates that both simulations predict a similar hydrological system.

All CMIP6 ensemble members predict similar glacier melt runoff, and the variation of the baseflow runoff is also comparatively low during the monsoon season (Figure 5.16). The variation is larger for the snowmelt component and largest for rainfall-runoff. The highest rainfall-runoff is predicted for the ACCESS-ESM1-5 member. This member predicts the highest mean monsoon season runoff of 27 B m^3 despite predicting the lowest snowmelt contributions (0.6 B m^3). The lowest mean monsoon season runoff of 22 B m^3 is predicted for the MRI-ESM2-0 and BCC-CSM2-MR models.

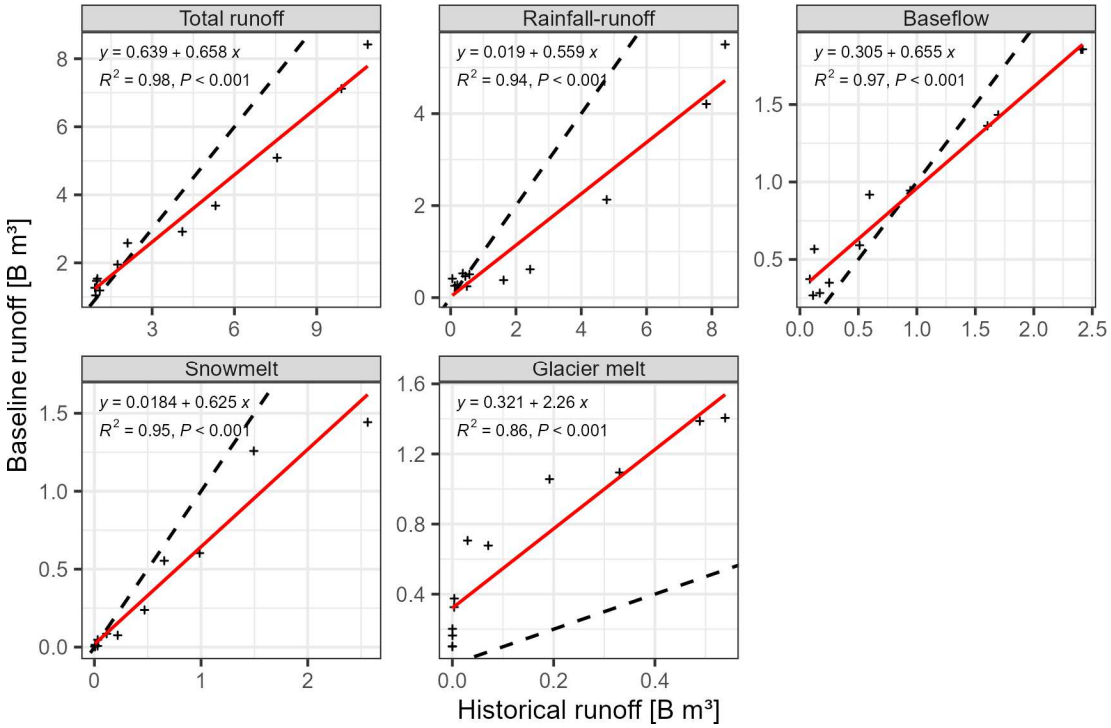


Figure 5.15: Comparison of the mean monthly runoff of the historical simulations (x-axis) and baseline ensemble (y-axis) for the total runoff and the runoff components. The red line indicates the linear regression and the black dashed line is the line of perfect fit. The comparison of each baseline ensemble member and the historical simulations is presented in Appendix 5.1.

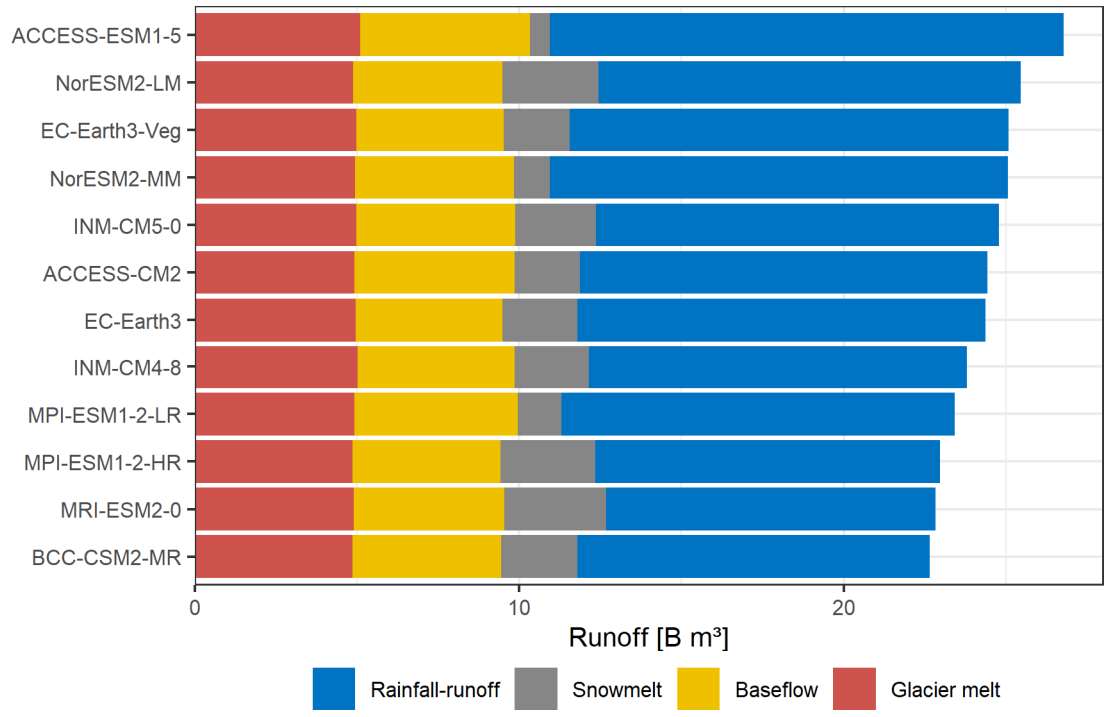


Figure 5.16: Mean monsoon season runoff composition of the baseline ensemble members (1975 - 2014).

5.4.1.3 The flood events of the baseline simulations

This section evaluates the predictions of the flood discharge defined as the Annual Maximum flows (AMAX) of the baseline simulations and compares it with the ones of the historical simulations and the observed flood record. Both simulations consist of an ensemble of 64 hydrological models. This ensemble is applied to the climate data of 12 CMIP6 models for the baseline simulations and hence the ensemble size of the baseline is larger. Furthermore, the compared record lengths, and thus the number of floods, differ between the observations and baseline simulations (1975 – 2014, 40 events), and the historical simulations (2002 – 2015, 14 events) and hence the historical flood flows have a lower sample size and are, thus, more prone to outliers.

Both simulations predict similar flood discharges of rare events but underpredict the observed ones. The baseline AMAX P_{50} (see Box 5.1) has a magnitude of 5,400 m^3/s which is lower than the corresponding historical flood discharge (6,200 m^3/s) and the observed one (7,400 m^3/s) (Figure 5.17 and Table 5.1). The difference between the baseline and historical flood discharge predictions decreases with the AMAX percentile and converges around the AMAX P_{80} which describes flood flows that statistically occur once every 5 years. The baseline simulations predict higher flood discharges ≥ AMAX P_{95} (median predictions).

The simulated flood discharge of both simulations is lower than the observed flood discharge (Figure 5.17 and Table 5.1). This difference is -31% (baseline) and -18% (historical) for the AMAX P_{50}. The observed flow duration curve increases stronger than the simulated ones and hence the differences increase and reach -43% (baseline) and -55% (historical) for the highest flood discharge (AMAX P_{100}) which is 9,900 m^3/s for the historical simulations,11,300 m^3/s for the baseline simulations and 17,400 m^3/s for the observations (median predictions).

Box 5.1: Definition of the AMAX percentiles.

AMAX percentiles:
The flood discharge (classified as the AMAX flows) of the ensemble members is aggregated (median) by the percentile of the flood discharge which is referred to as the AMAX percentile. This AMAX percentile is used for the aggregation of the ensemble because it is not reasonable to aggregate it by the year since the hydrological models might react differently to the climate input. The individual ensemble members may predict the highest (or lowest or medium) flood discharge in different years and hence events of different frequencies would be combined. Instead, the flood flows are aggregated by their frequency. These AMAX percentiles refer to the percentile of the flood discharge (e.g. the highest flood flow predicted by the ensemble members) and do not refer to the percentile of the ensemble predictions.

For example, the AMAX P_{100} is the ensemble median of the highest flood discharge predicted by the individual members. The AMAX P_{95} is the flood discharge which is exceeded by 5% of the events. The flood record contains 40 AMAX events for the CMIP6 simulations and hence this AMAX P_{95} is the median of the 2nd highest flood prediction of each ensemble member and is hence predicted statistically once every 20 years.

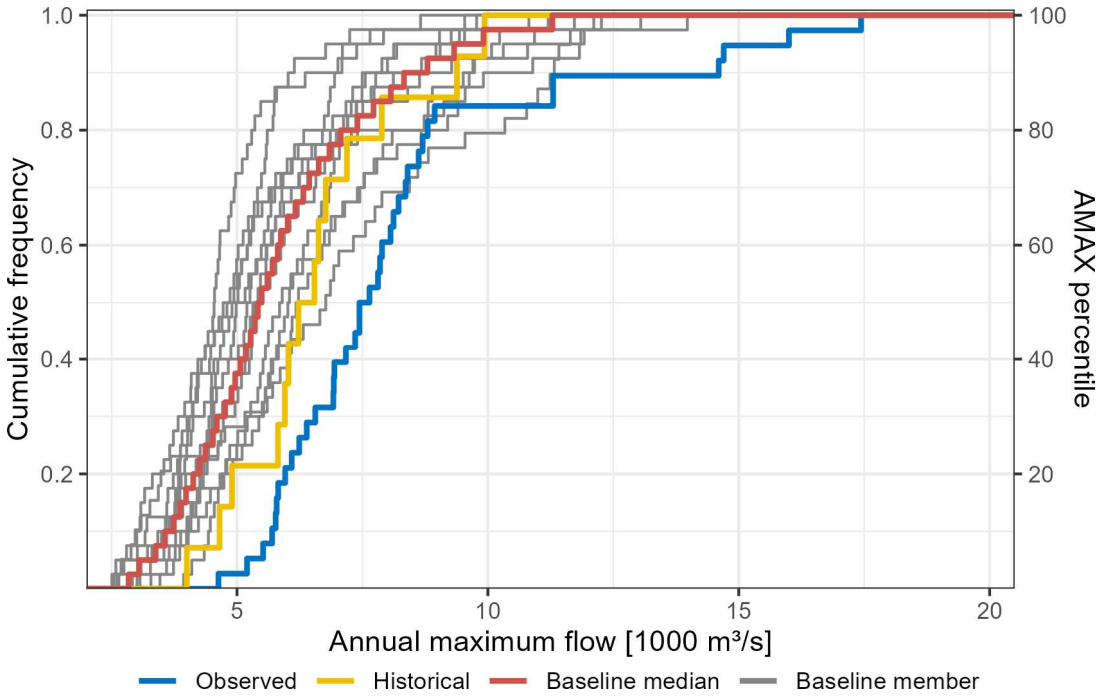


Figure 5.17: The cumulative frequency curves of the flood (AMAX) flows. The hydrological ensembles are aggregated (median) by the AMAX percentile. The record length is 40 years except for the historical simulations for which it is 14 years.

Table 5.1: Median AMAX discharge for selected AMAX percentiles. The P_{50} is exceeded statistically once every two years, the P_{80} once every 5 years, the P_{95} once every 20 years and the P_{100} is the highest flow of the record.

	AMAX [1,000 m^3/s]				Difference to observed AMAX [%]			
	P_{50}	P_{80}	P_{95}	P_{100}	P_{50}	P_{80}	P_{95}	P_{100}
Baseline	5.4	7.1	9.3	11.3	-31	-21	-45	-43
Historical	6.2	7.2	9.4	9.9	-18	-19	-44	-55
Observed	7.4	8.7	14.7	17.4	-	-	-	-

The timing of the flood events differs between the baseline and historical simulations and the observations but the monsoon seasonality is similar for the three datasets. The flood events occur earlier in the baseline simulations than in the historical ones. The median week of occurrence is week 31 (beginning of August) in the baseline simulations and week 33 in the historical simulations and the observed AMAX (Figure 5.18). The different shape of the historical density curve is explained by the smaller sample size of the flood record. The comparison of the baseline and observed density curves indicate a temporal shift, whereas the flood season starts and ends two weeks earlier for the baseline simulations. Furthermore, the curve of the baseline simulation reaches a plateau between weeks 30 and 34, while the observed one steadily increases until week 34 after which it drops. However, despite the temporal shift and the extended peak period, both curves follow similar trends and indicate a similar seasonal dynamic.

Several flood events are predicted outside the monsoon season (weeks 22 – 40) whereas all observed flood events fall within the monsoon season (Figure 5.18). These non-monsoon season events account for 7.5% of all events for the baseline simulations and 6% for the historical simulations. The historical non-monsoon season events are predicted for October but not outside the period June - October. The baseline simulations also predict some events before June and after October. The baseline non-monsoon season events have a median rank of 17 (out of 40) which indicates that the events are of below-average magnitude, whereas flows in the post-monsoon season are higher (median rank: 28) compared to the pre-monsoon season (median rank: 14). This is reasonable because of the wetter antecedent condition after the monsoon season and the higher baseflow rates.

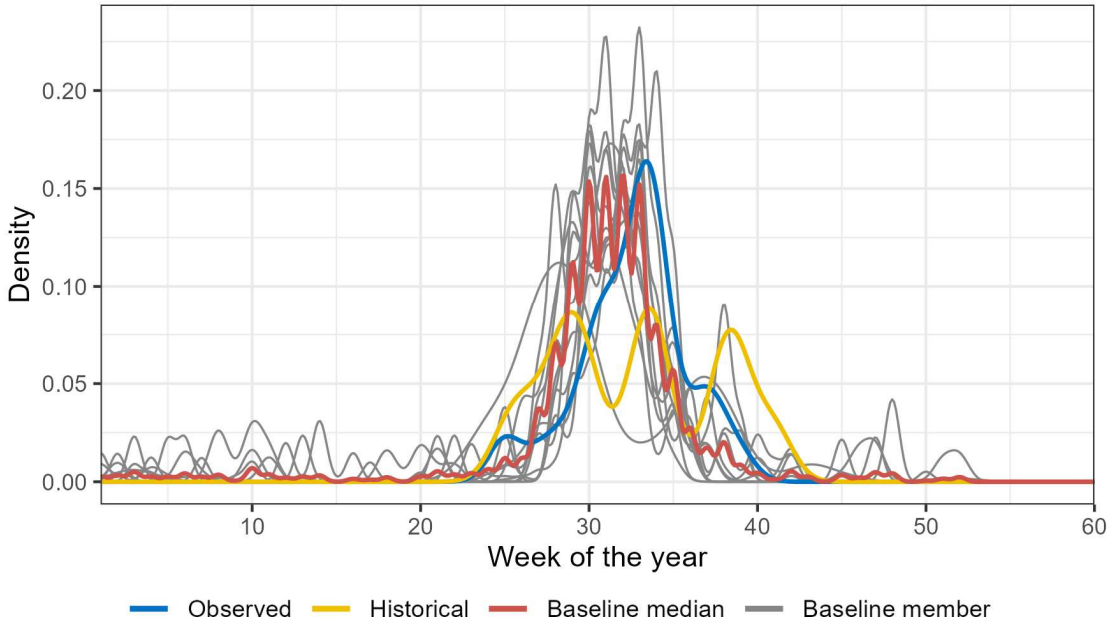


Figure 5.18: Timing of the Annual Maximum flows. The sample sizes of the records are different due to differences in the record length and ensemble size. The sample sizes are 40 (1x40) for the observations, 869 (14 x 64) for the baseline, 2560 (40 x 64) for the baseline members, and 30,720 (40 x 768) for the baseline median.

The composition of the flood events is similar between the historical and baseline simulations except for the glacier melt component. Most flood events are dominated by rainfall-runoff in both simulations. The mean rainfall-runoff contribution to all events is 79% and 84% in the baseline and historical simulations, respectively (Figure 5.19). The predicted baseflow contribution is consistent in both simulations (8 – 9%). Both models also agree on the low importance of the snowmelt contribution (2 – 5%). One notable difference is, as for the annual and seasonal flows, the higher glacier melt rates of the baseline simulations. It contributes 10% (the second largest contribution) in the baseline simulations whereas it is the least important source in the historical simulations (3%). However, the glacier melt contributions decrease with increasing flow and for the highest events the predicted contribution reduces to 5% for the baseline simulations. The largest differences in the composition of the maximum events are predicted for the baseflow component which contributes 11% in the historical simulations and 4% in the baseline ones. However, this difference results from the timing of the event rather than systematic differences.

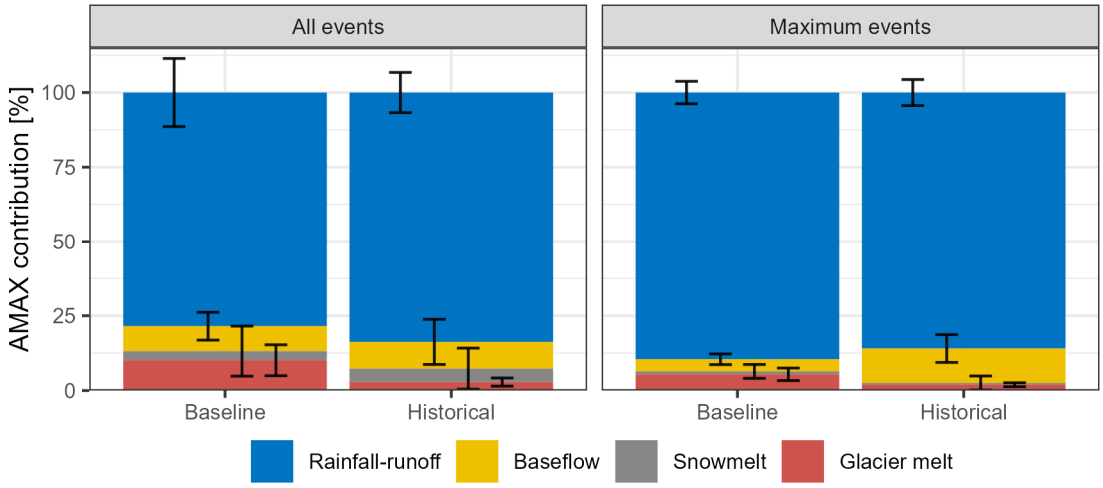


Figure 5.19: Ensemble mean composition of the simulated flood events. The left plot is the composition of all events, and the right plot is the composition of the flood events with the highest magnitude. The error bars indicate the standard deviation in the ensemble predictions.

The rainfall event is the primary control of the flood flow composition. This controlling effect of the rainfall characteristics is illustrated by the relationship between the composition and the flood discharge (Figure 5.20). The rainfall-runoff contribution increases with flow magnitude and, consequently, the other sources lose importance. The rainfall-runoff is the dominant source for above-median flood discharge and only for a few events in the baseline simulations, the contribution is below 50%. All events above the 80th percentile are composed of ≥ 50% rainfall-runoff. In the historical simulations, this is true for all events. The composition ≥ 90th percentile is strongly dominated by the rainfall-runoff but for a few events higher snowmelt contributions lead to lower rainfall-runoff contributions of 70 – 75%. However, for the rarer flood flows (≥ 95th percentile), the rainfall-runoff increases to ≥ 85%. High contributions of the other components are primarily simulated for below-average flood events. It can, therefore, be argued that these events are shaped by the absence of a heavy rainfall event during the respective year and that the higher contributions result from the misclassification of the flood event from the AMAX approach and not from high contributions itself.

The composition of the flood events is furthermore characterised by the timing of the event (Figure 5.20). The highest snowmelt contributions are predicted during the weeks 22 – 30 (June – Mid-July), higher glacier melt contributions for the mid-monsoon season during the weeks 27 – 36 (July and August), and higher baseflow contributions from week 33 onwards (Mid-August). The composition of the flood event, thus, follows the flow seasonality. However, rainfall-runoff is the dominant source throughout the flood season in both simulations. Snowmelt can be the most important runoff source for earlier floods in the baseline simulations. However, while the timing affects the flood composition the primary driver is the intensity of the rainfall event.

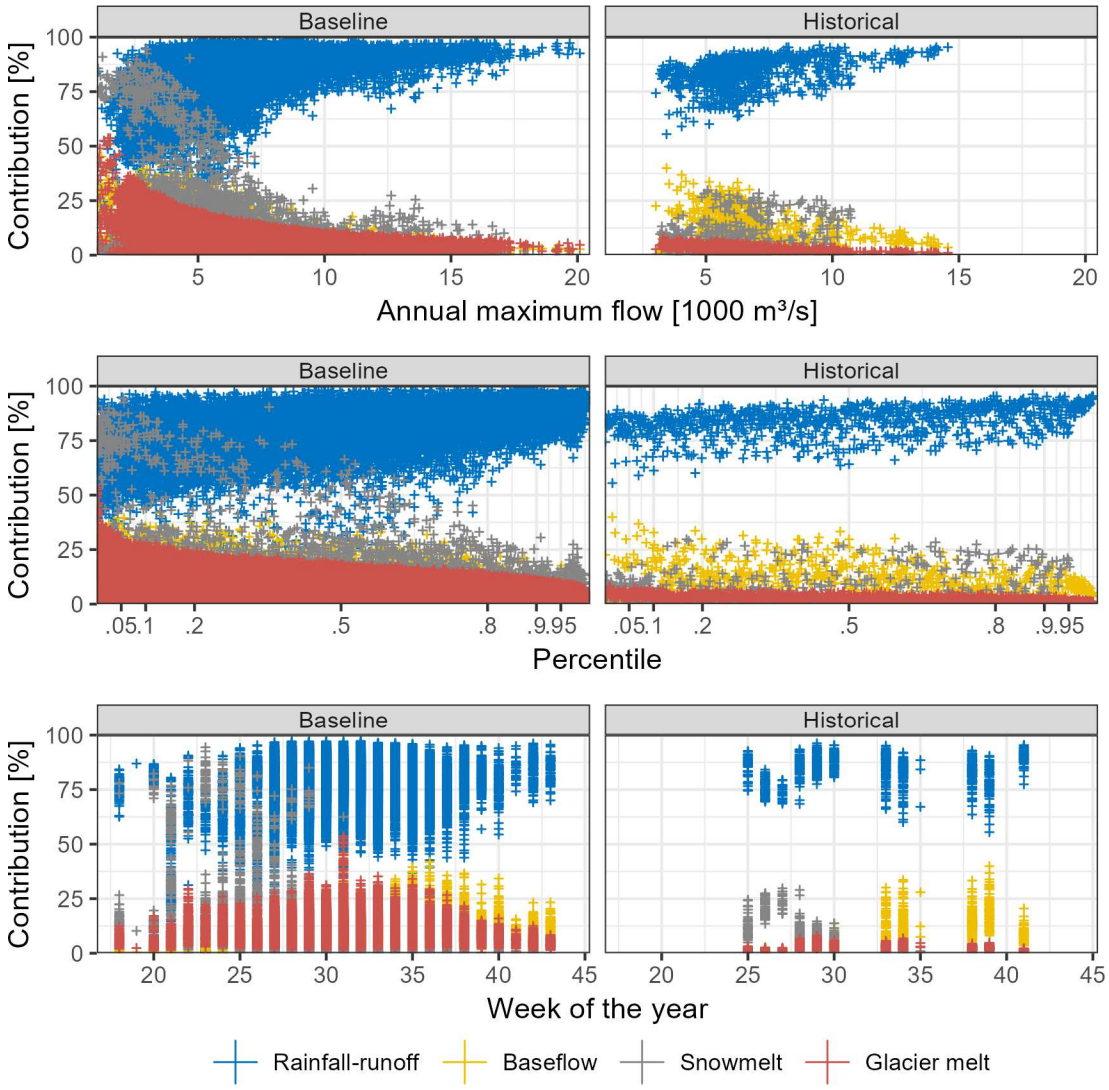


Figure 5.20: The relationship between runoff composition and flow magnitude (top), the flow quantile (middle), and the timing (bottom). The baseline flood record consists of 30,720 events (12 CMIP6 models x 64 hydrological models x 40 years), and the historical one of 896 events (64 hydrological models x 14 years).

The variation of the predicted flood flows between the CMIP6 members increases with the magnitude of the flood discharge. The members can be classified into three groups according to the prediction of the most extreme floods AMAX P_{100} (Table 5.2). The highest AMAX P_{100} are predicted by the NorESM2-LM and ACCESS-ESM1-5 members which predict flows ranging between 13,000 – 14,000 m^3/s. They underpredict the observed maximum flood (17,400 m^3/s) by 20 – 25%. The second group (MRI-ESM2-0, INM-CM5-0, NorESM2-MM, EC-Earth3-Veg, EC-Earth3, MPI-ESM1-2-HR, MPI-ESM1-2-LR) predict medium flows between 10,800 and 12,300 m^3/s. The last group predicts considerably lower flows between 8,700 – 9,800 m^3/s and, thus underpredicts the highest observed flood by 44 – 50%. These members are BCC-CSM2-MR, ACCESS-CM2, and INM-CM4-8. The prediction range of the ensemble members reduces with the percentile and for the AMAX P_{50}, which approximates the 1-in-2-years event, this difference is 1,700 m^3/s compared to 5,300 m^3/s for the AMAX P_{100}.

Table 5.2: Flood predictions of the CMIP6 ensemble members. The hydrological ensemble is aggregated by the median. The non-monsoon columns indicate the median magnitude of the non-monsoon season floods and the fraction of non-monsoon season flood events.

	AMAX Percentile				**Non-monsoon**	
	P_{50}	P_{80}	P_{95}	P_{100}	Median [1,000 m^3/s]	Fraction [%]
NorESM2-LM	6.7	9.5	11.4	14.0	2.6	4
ACCESS-ESM1-5	5.8	8.7	10.8	13.0	5.6	3
MRI-ESM2-0	4.7	7.0	9.0	12.3	4.0	15
INM-CM5-0	4.9	7.4	11.8	12.1	6.1	9
NorESM2-MM	6.2	8.1	10.3	11.7	5.4	4
EC-Earth3-Veg	5.4	7.1	9.0	11.3	5.2	2
EC-Earth3	5.3	6.9	8.1	11.2	6.0	1
MPI-ESM1-2-HR	4.5	5.3	6.8	11.2	5.6	13
MPI-ESM1-2-LR	5.0	6.7	10.1	10.8	6.0	16
BCC-CSM2-MR	4.8	5.6	7.4	9.8	4.8	5
ACCESS-CM2	6.0	7.3	8.2	9.5	5.6	12
INM-CM4-8	5.2	6.3	7.1	8.7	6.6	7

One further difference between the baseline members is the frequency of the non-monsoon season flood events (Table 5.2). Four members (MPI-ESM1-2-LR, MRI-ESM2-0, MPI-ESM1-2-HR, ACCESS-CM2) predict > 10% of the flood events outside the monsoon season. The median discharge of these events is lower than the AMAX P_{50} for MRI-ESM2-0 and ACCESS-CM2 which indicates that these events occur during the monsoon seasons with comparatively low flows and might be a misclassification of floods using the AMAX approach. However, for the both MPI-ESM models, the median non-monsoon season discharge is larger than the AMAX P_{50} which indicates that these floods are not a misclassification but triggered by the climate predictions of the members. Therefore, these members are weaker in predicting the flood seasonality of the Karnali River system. The members which predict the highest flood magnitudes generally predict fewer events (< 5%) outside the monsoon season (NorESM2-LM, ACCESS-ESM1-5). The lowest number of non-monsoon season floods are predicted for the EC-Earth-3 and EC-Earth3-Veg models.

The baseline members also differ regarding the flow ratio between the more common and rare events (Figure 5.21). Particularly the MRI-ESM2-0 and MPI-ESM1-2-HR models predict low discharges for most of the AMAX percentile but these discharges increase strongly above the AMAX P_{95}. On the contrary, the ACCESS-CM2 model predicts high discharges for the lower AMAX percentiles but the increase with the percentile is less pronounced and hence the flow duration curve is flatter. The difference in the flood discharges between the AMAX P_{50} and the AMAX P_{100} is 3,500 m^3/s which is lower than for any other model. The NorESM2-LM, ACCESS-ESM1-5 and NorESM2-MM models predict comparatively high flood discharges across the full AMAX percentile range.

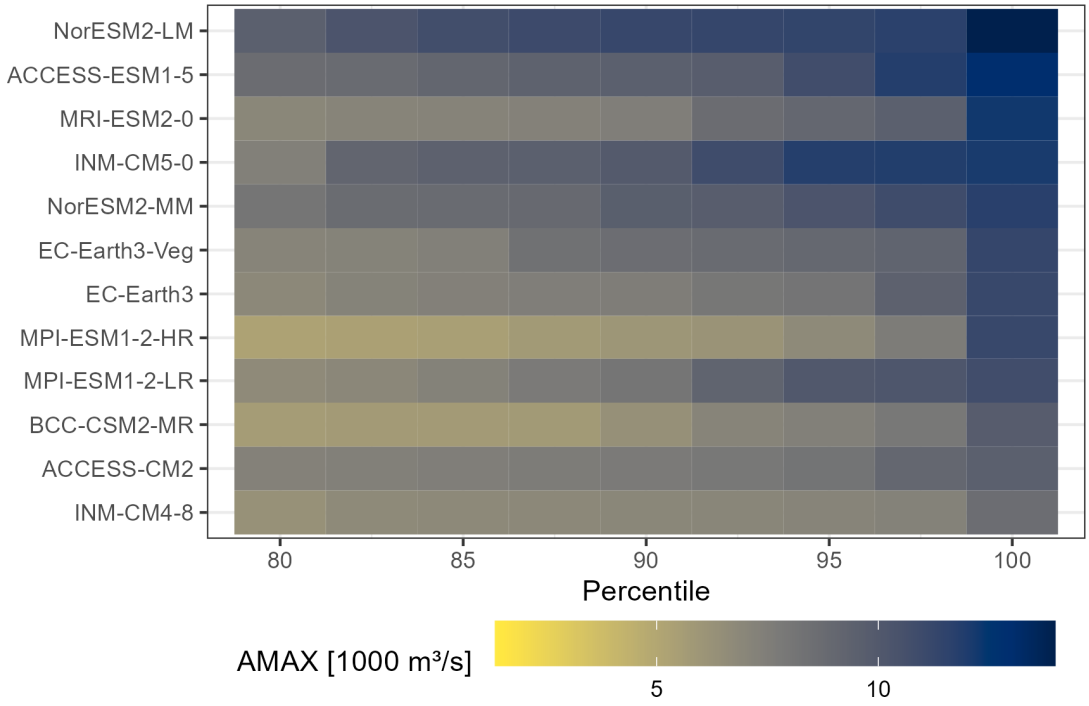


Figure 5.21: AMAX Percentiles of the baseline ensemble members for the period 1975 - 2014. The members are ordered according to the predicted flows at the 99th percentile. The AMAX percentile is the median discharge aggregated by the rank of the flood event.

This section presented the results of the baseline simulations and compared them with the observations and the historical simulations to assess the validity of the CMIP6 simulations. In summary, the baseline simulations reproduce the observed flow seasonality with high flows during the monsoon and low flows during the rest of the year. However, the flow seasonality is less pronounced than in the historical simulations and the observations because the non-monsoon season flow is slightly overpredicted and the monsoon flow underpredicted. The composition of the runoff follows the expected seasonality with snowmelt dominating flows in the pre-monsoon season, rainfall-runoff in the monsoon season, and baseflow in the post-monsoon season and the winter. This reflects our understanding of the hydrological system and is also consistent with the composition of the historical simulations. The main difference between these two simulations is the lower rainfall-runoff and the higher glacier melt in the baseline simulations. Glacier melt is predicted throughout the year which is unreasonable because glaciers grow during the winter months. However, the effect of the overpredicted glacier melt reduces with increasing discharge, and for the flood events, the contribution is below average. The glacier melt contribution to the rare flood event is marginal and hence this glacier melt bias does not hamper the prediction of the future flood flows for the projected climates.

5.4.2 The hydrology of the Karnali River for the projected climates

This section compares the simulated hydrology of the CMIP6 baseline scenario and the projected medium-emission scenario SSP245 and the high-emission scenario SSP585. The model ensemble of each scenario consists of 768 members (64 hydrological models x 12 CMIP6 members). The baseline scenario represents the period 1975 – 2014, and the projected scenarios are divided into the near future (2020 – 2059) and the far future (2060 – 2099). This length of 40 years was chosen over the 30-year climate reference period to increase the size of the flood record and, thus, decrease the sampling error of the following flood frequency analysis.

The following section compares the annual discharge of the scenarios to gain insights into the temporal long-term trends. Afterwards, the monthly discharge predictions are compared to assess changes in flow seasonality. Both sections present the runoff composition to understand the drivers of the projected changes, but also to assess the validity of the parameter transfer between different climatic patterns which provides the fundament for the quantification of changes in the flood hazards in Section 5.4.3.

5.4.2.1 Annual scale

The annual discharge is projected to increase with time and emissions whereas the scenario differences increase in the far future. The ensemble median annual discharge is projected to increase from 37.8 Billion m^3 in the baseline to 43.3 Billion m^3 (SSP245) and 43.6 Billion m^3/a (SSP585) in the near future (Figure 5.22). This change equals an increase of 15% and hence the higher emissions do not, initially, result in climatic conditions which facilitate notably higher discharges than the lower emissions. The predicted annual discharge is projected to increase further in the far future whereas the trends of both scenarios decouple. The trend slows down for the medium-emission scenario SSP245 in which the annual discharge is projected by a further 10% to 47.0 Billion m^3. On the contrary, the annual discharge projected for the high-emission scenario SSP585 increases strongly by another 30% to 54.9 Billion m^3.

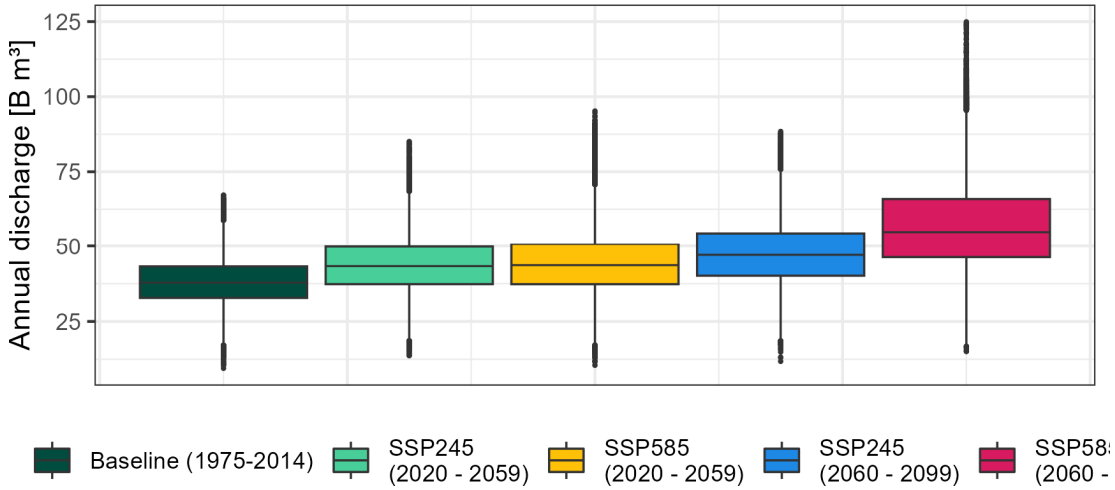


Figure 5.22: The annual discharge predicted for the baseline and projected CMIP6 scenarios. Each record consists of 30,720 samples (12 CMIP6-members x 64 hydrological models x 40 years).

The projected change of the annual discharge increases with the percentile of the ensemble predictions. The annual discharge of the 95th percentile increases with 25% (SSP245) and 61% (SSP585) in the far future stronger than the median. On the contrary, the discharge at the 5th percentile is projected to increase by 22% (SSP245) and 39% (SSP585). Therefore, extreme wet years are projected to become more extreme. However, the discharge of the dry years is projected to increase as well and hence the extreme dry years become wetter and thus less extreme.

The projected increase in the annual discharge is driven by increased rainfall-runoff and, to a lower degree, increased baseflow and glacier melt contributions. These increases compensate for the projected reduction of snowmelt runoff (Table 5.3). Similar trends are predicted for both scenarios, and the scenario differences are low in the near future. The mean annual rainfall-runoff increases from 15.5 Billion m^3 in the baseline to 18.6 Billion m^3 (+20%) and 18.8 Billion m^3 (+21%) for the medium-emission and high-emission scenarios, respectively. It is followed by baseflow which increases by 1.9 and 2.4 Billion m^3 (+18 and +21%). The trends of the temperature-related melt components are more pronounced in the SSP585 scenario where the glacier melt increases by 21% (+1.6 Billion m^3) and the snowmelt decreases by 19% (-0.8 Billion m^3/a) compared to a 17% increase (+1.3 Billion m^3/a) in glacier melt and a 12% decrease (-0.5 Billion m^3/a) in snowmelt in SSP245 (near future).

Table 5.3: Mean composition and the standard deviation of the annual discharge. Note that the mean is used for the composition because the median composition does not add up to 100%.

	1975 - 2014	**2020 - 2059**		**2060 - 2099**	
	Baseline	SSP245	SSP585	SSP245	SSP585
Total runoff [B m^3]	38.3 ±8.3	44 ±9.8	44.6 ±10.9	47.5 ±10.7	56.9 ±15.3
Rainfall-runoff [B m^3]	15.5 ±5.6	18.6 ±6.8	18.8 ±7.6	20.3 ±7.7	25.7 ±11.5
Baseflow [B m^3]	10.8 ±3.1	12.7 ±3.6	13.1 ±3.8	14.2 ±4	16.9 ±5
Snowmelt [B m^3]	4.3 ±2.5	3.8 ±2.3	3.5 ±2.3	3.1 ±1.9	2.5 ±1.7
Glacier melt [B m^3]	7.6 ±2.7	8.9 ±3.1	9.2 ±3.3	10 ±3.5	11.8 ±4.3
Rainfall-runoff [%]	40.1 ±10	41.8 ±10	41.5 ±10.8	42.1 ±10.7	43.9 ±11.5
Baseflow [%]	28.4 ±6.6	29.1 ±6.6	29.5 ±6.7	30 ±6.7	30.1 ±7
Snowmelt [%]	11.4 ±6.2	8.8 ±5.2	8.1 ±4.9	6.7 ±4.2	4.8 ±3.4
Glacier melt [%]	20 ±6.5	20.4 ±6.5	20.9 ±6.9	21.2 ±6.8	21.2 ±6.9

The trends of increasing rainfall-runoff, baseflow and glacier melt, and decreasing snowmelt continue in the far future but are more pronounced in the SSP585 scenario (Table 5.3). The biggest scenario difference is predicted for the rainfall-runoff component which increases by 10 Billion m^3 (compared to the baseline) to 25.7 Billion m^3 for SSP585 which is almost double the projected increase for the SSP245 scenario (+5.4 Billion m^3). The predicted baseflow and glacier melt increases are stronger for the SSP585 scenario in absolute quantities, but the relative contribution is similar for both scenarios. The snowmelt contribution decreases from 11.4% in the baseline to 6.7% for SSP245 and 4.8% for SSP585. The rainfall-runoff gains the most importance in the annual discharge budget in the far-future. Its contribution increases by 2% (SSP245) and 4% (SSP585), whereas the baseflow and glacier melt increase each by 1% in both scenarios and the snowmelt contribution decreases by 4% and 6% respectively. The annual variability is largest for the rainfall-runoff contribution and lowest for the snowmelt contribution (Table 5.3).

All runoff components follow a distinct temporal trend but this trend is different for the medium-emission scenario SSP245 with linear projected increases and the high-emission scenario SSP585 with exponential increases (Figure 5.23). The glacier melt increases consistently with time in both scenarios with small internal variations. However, in the SSP245 scenario, it reaches a plateau around the year 2080 whereas in the SSP585 scenario, the contribution accelerates around the year 2060. The baseflow contribution increases steadily with time in SSP245, and in SSP585 the rate of change increases around the mid-century.

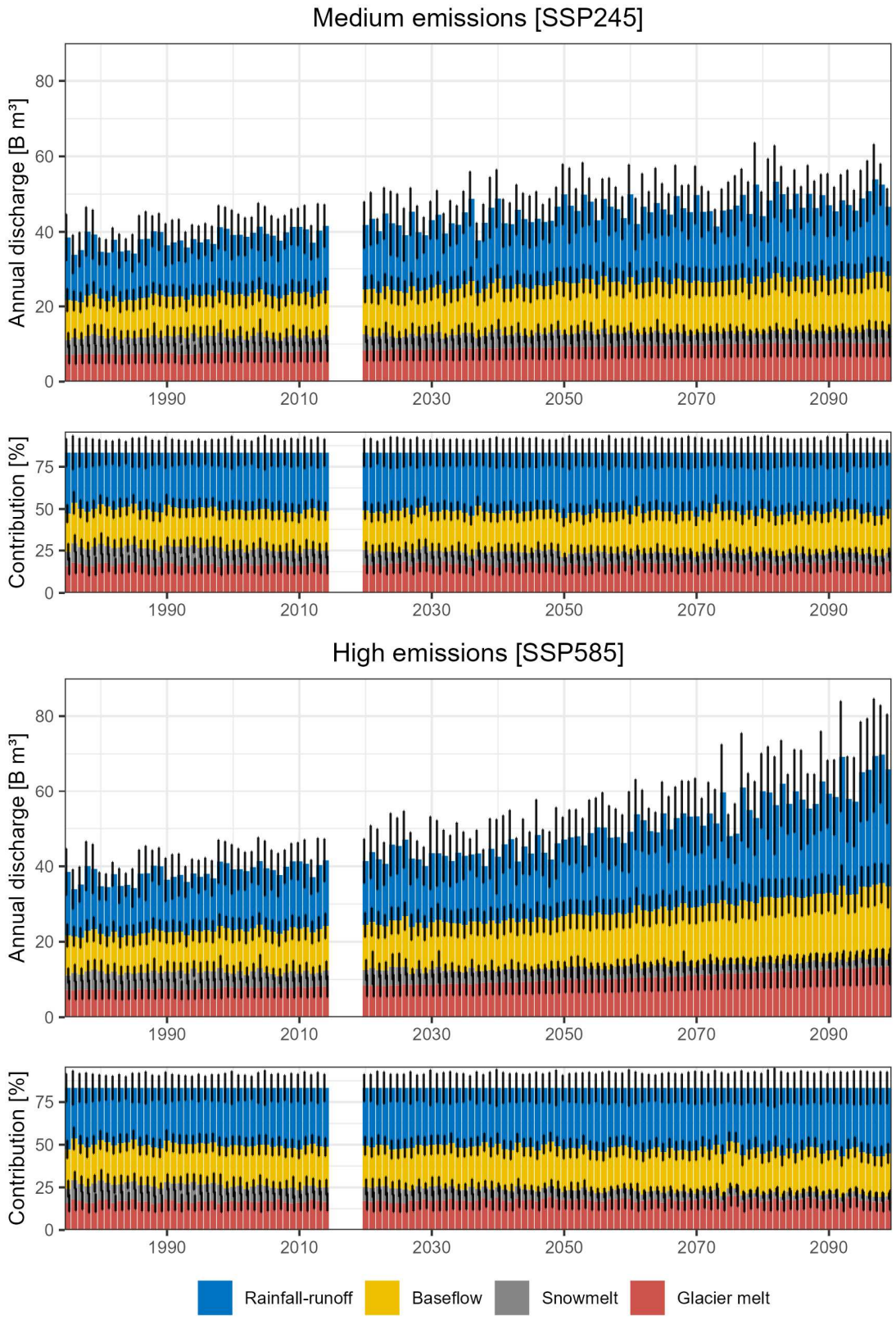


Figure 5.23: Mean annual discharge composition for the baseline (≤ 2014) and both projected scenarios (≥ 2020). The bars indicate the ensemble mean contribution and the error bars their respective standard deviation. The x-axis shows the year.

The different temporal trends of the runoff components and the climate scenarios are highlighted by regression modelling of the annual discharge of the components and the year of the simulation (Figure 5.24). The lower coefficient of determination (R^2) of the rainfall-runoff and the snowmelt components in both scenarios indicate the higher annual variability of these components. The baseflow and glacier melt contributions have low annual variability and increase steadily in both scenarios. However, the SSP245 scenarios project a linear increase with time and the glacier melt contribution remains on a similar level after the year 2080. The rainfall-runoff projections are different for both scenarios. For the SSP245 scenarios, these annual variation increases in time but overall the projected increase is somewhat linear. For the SSP585 scenario, these increases are higher and follow an exponential trend with higher increases around the year 2040 and around the 2075. After the year 2080, all annual contributions exceed the ones predicted for the baseline.

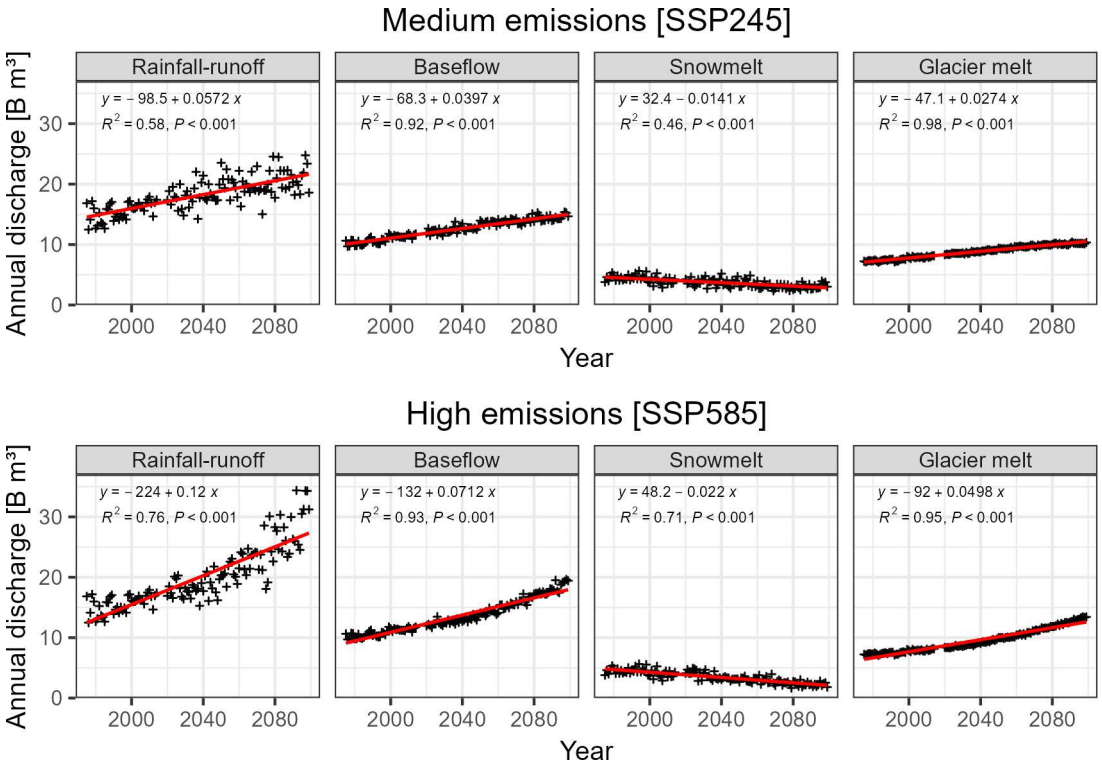


Figure 5.24: The temporal trends of the ensemble mean annual discharge for each runoff component. The red line indicates the predictions of the linear regression models.

5.4.2.2 Monthly scale

The characteristics of the monthly discharge follow the annual trends and the projected changes correlate with time and emissions. The seasonality with high flows during the monsoon season and low flows during the non-monsoon season does not change but the seasonality is projected to be more pronounced because the monsoon season discharge increases more strongly than the non-monsoon season discharge. However, the discharge is projected throughout the year and the driving processes vary for the seasons (Figure 5.25).

The largest changes are predicted for the monsoon season (Jun – Sep) which mean runoff increases from 24.3 Billion m^3 for the baseline to 30.6 Billion m^3 for SSP245 and 37.0 Billion m^3 for SSP585 in the far future (Figure 5.25). Hence, around two-thirds of the projected increase in the annual discharge occurs during the monsoon season. The share of the monsoon season discharge is projected to increase from 63.5% (baseline) to 64.4% (SSP245) and 65% (SSP585) in the far future.

The main source of the monsoon season discharge increase is the rainfall-runoff which increases by 4.7 Billion m^3 (SSP245) and 9.6 Billion m^3 (SSP585) in the far future (Figure 5.25). Consequently, the rainfall-runoff contribution increases from 51% in the baseline to 56% (SSP245) and 60% (SSP585). The baseflow and glacier melt components increase by 1.3 - 1.4 B m^3 for SSP245 and 2.2 – 2.4 B m^3 for SSP585. Their relative contributions remain stable (±0.5%) except for the glacier melt for SSP585. This contribution decreases by 1% the higher melt rates as a result of the comparatively larger increase of rainfall-runoff. The monsoon season snowmelt contribution is projected to decrease from 9% (baseline) to 4% (SSP245) and 2% (SSP585). It is further worth noting that the standard deviation is highest during the monsoon season, particularly for the rainfall-runoff and hence the variation in the annual dischrge is dominated by the variation of the monsoon rainfall (mean predictions).

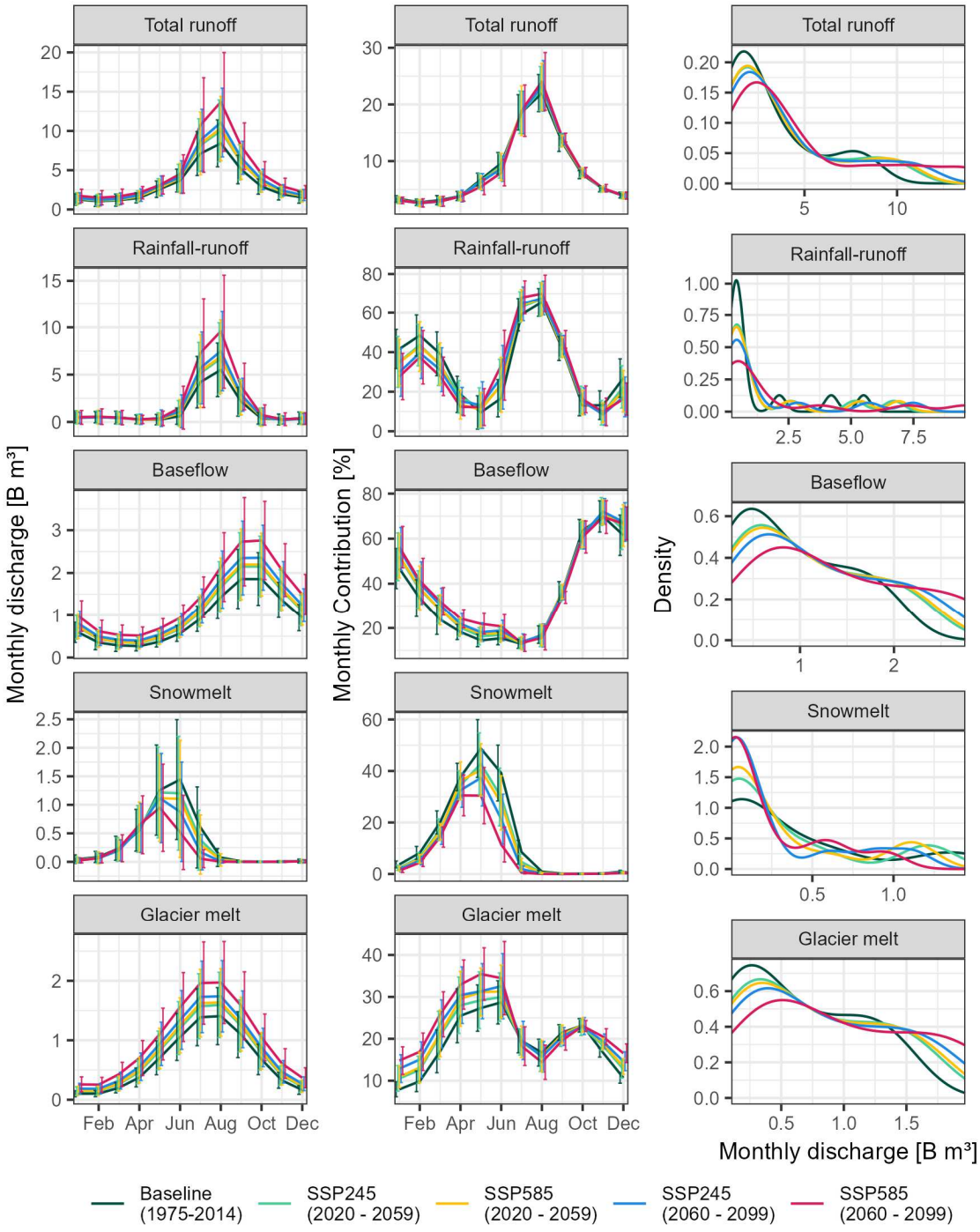


Figure 5.25: Comparison of the monthly runoff composition of the baseline climate and the projected climates. The left column shows the absolute monthly discharge, the centre shows the relative contribution and the right column shows the density of predicted monthly discharge rates. The relative contribution of the total runoff (centre column of the top row) shows the monthly contribution to the annual discharge (%). The lines indicate the ensemble mean and the error bars their standard deviation which arises from the climate variability and the ensemble variation.

A fraction of the monsoon precipitation drains as delayed baseflow during the transition from the high to low flows in the post-monsoon season (Oct – Nov). Hence, the projected increase in the monsoon precipitation translates into increasing post-monsoon season discharge (Figure 5.25). However, these increases are with 1.3 Billion m^3 (SSP245) and 2.6 Billion m^3 (SSP585) considerably lower than the projected increase of the monsoon season runoff (far future). In relative terms, the increase is 25% (SSP245) and 52% (SSP585) and, thus, of similar magnitude as the monsoon season increase. The projected baseflow contribution is 65 – 66% and the main contributor to the Post-monsoon season discharge. The baseflow is also the main factor for the projected increases (SSP245: +0.9 Billion m^3, SSP585: +1.6 Billion m^3). However, also the rainfall-runoff is projected to increase in October which indicates that either, the frequency or intensity of late-monsoon season rainfall is projected to increase. However, these increases are with 0.2 Billion m^3 (SSP245) and 0.4 Billion m^3 (SSP585) too small to provide strong evidence for a projected prolonging of the monsoon season (mean predictions).

The lowest changes are projected for the dry winter season (Dec – Feb). The discharge is projected to increase by 0.7 Billion m^3 (SSP245) and 1.7 Billion m^3 (SSP585) in the far future (Figure 5.25). However, the projected relative changes are with +18% and +44% large. Most of the projected increase originates from higher baseflow. The rainfall-runoff remains constant which indicates that the projected precipitation increase of the climate ensemble concentrates during the monsoon season. The simulations predict increases in glacier melt by 0.2 Billion m^3 (SSP245) and 0.5 Billion m^3. However, while these increases are comparatively small, the predicted winter glacier melt contributions increase from 11% in the baseline to 13% for SSP245 and 16% for SSP585 (mean predictions).

The changes predicted for the pre-monsoon season (Mar – May) are similar to the ones projected for the winter. The discharge increases by 1 Billion m^3 (SSP245) and 1.7 Billion m^3 (SSP585) in the far-future (Figure 2.52). This increase is driven by higher baseflow and glacier melt which both increase roughly at the same rate of 0.5 Billion m^3 (SSP245) and 0.8 – 0.9 Billion m^3 (SSP585). The rainfall-runoff remains stable but loses importance due to the increase in the other components. The snowmelt contributions are projected to decline by 0.2 Billion m^3 (SSP245) and 0.3 Billion m^3 whereas most of the decline is predicted for the late pre-monsoon season in May. The decrease of the pre-monsoon season snowmelt is lower than for the monsoon season. This is explained by changes in the melting season which shifts towards earlier dates with time and emissions.

The flow seasonality remains similar for the projected climates (Figure 5.25). The timing of the rising limb remains in June and the monsoon season ends in September as indicated by the low changes of the October rainfall-runoff. The seasonality of the rainfall-runoff is more pronounced with stable contributions throughout the non-monsoon season but higher contributions during the monsoon season. The baseflow and glacier melt seasonality remains similar but on a higher level throughout the year. However, the snowmelt contributions change in magnitude and timing. The peak snowmelt is predicted for June in the baseline but this shifts to May in both scenarios in the far future.

The comparison of the baseline and projected scenarios indicates an increasing discharge with time and emissions. This increase is driven by the projected increase in monsoon precipitation. However, the baseflow delays part of the catchment response until the pre-monsoon season in the following year. Consequently, the discharge increases throughout the year which is amplified by the projected increase of glacier melt, and reduced by lower snowmelt runoff. However, the discharge is projected to increase across the full range of low and high flows as also indicated by the runoff density curves (Figure 5.25).

The projected trends correlate with time and emissions, particularly for baseflow and glacier melt. The annual variation is highest for the rainfall-runoff. The projected total discharge is similar for both scenarios in the near future (2020 – 2059). However, the higher glacier melt contributions projected in the near future by SSP585 indicate that the temperatures are projected to increase stronger with emissions in the near future but this has low effects on the hydrological regime. In the far future, the projections of both scenarios decouple which is, to a large degree, caused by higher rainfall-runoff which increases around the years 2040 and 2080 in the high-emission scenario. Further difference is predicted for the glacier melt and snowmelt which stabilise around the year 2080 for the medium-emission scenario SSP245 but not for the high-emission scenario SSP585.

5.4.3 The projected flood discharge

The previous two sections established that the hydrological model predicts similar catchment behaviour for the historical and baseline simulations despite the parameter transfer from the observed calibration data to the simulated CMIP6 climate data (Section 5.4.1), and this behaviour does not change for the projected climates which indicates that the parameter transfer to different climatic conditions does not limit the predictive capabilities of the model (Section 5.4.2).

The main differences between the historical and baseline simulations are the lower baseline discharge caused by reduced rainfall-runoff and the increased glacier melt contributions. However, these differences are lower for flood events because the glacier melt contribution reduces with increasing discharge and the predicted flood discharge of the historical and baseline simulations are similar for the extreme flows ≥ AMAX P_{95}. This indicates the good predictive capabilities of the modelling ensemble for the flood flows.

This section presents the flood discharge predictions for the projected climates, focusing on the flood discharge magnitudes, their composition and timing. These results are the foundation of the quantification of the changes in the flood hazard frequencies and magnitudes in Section 5.4.4.

5.4.3.1 The CMIP6 ensemble projections

The projected flood discharge of the CMIP6-ensemble increases with time and emissions (Figure 5.26 A). The median AMAX remains stable between 4,000 – 6000 m^3/s throughout the baseline period. It increases in the projected scenarios, whereas it is similar for the SSP245 and SSP585 scenarios until the year 2050. The median AMAX of both scenarios decouple in the second half of the century, and a stronger increase is projected for SSP585. This trend of increasing flood discharge intensifies after the year 2080 for SSP585. For SSP245, the median AMAX remains stable in the second half of the century, although on a higher level than before the year 2050. The increase in flood discharge is more pronounced for the higher percentiles of the ensemble predictions, and an increase in flood discharge towards the end of the century is also predicted at the Q_{95} (the 95^{th} percentile of the ensemble predictions) for SSP245. However, the strong variation between the years and the CMIP6-ensemble members hampers the identification of less pronounced trends at the lower percentiles, especially for the SSP245 scenario.

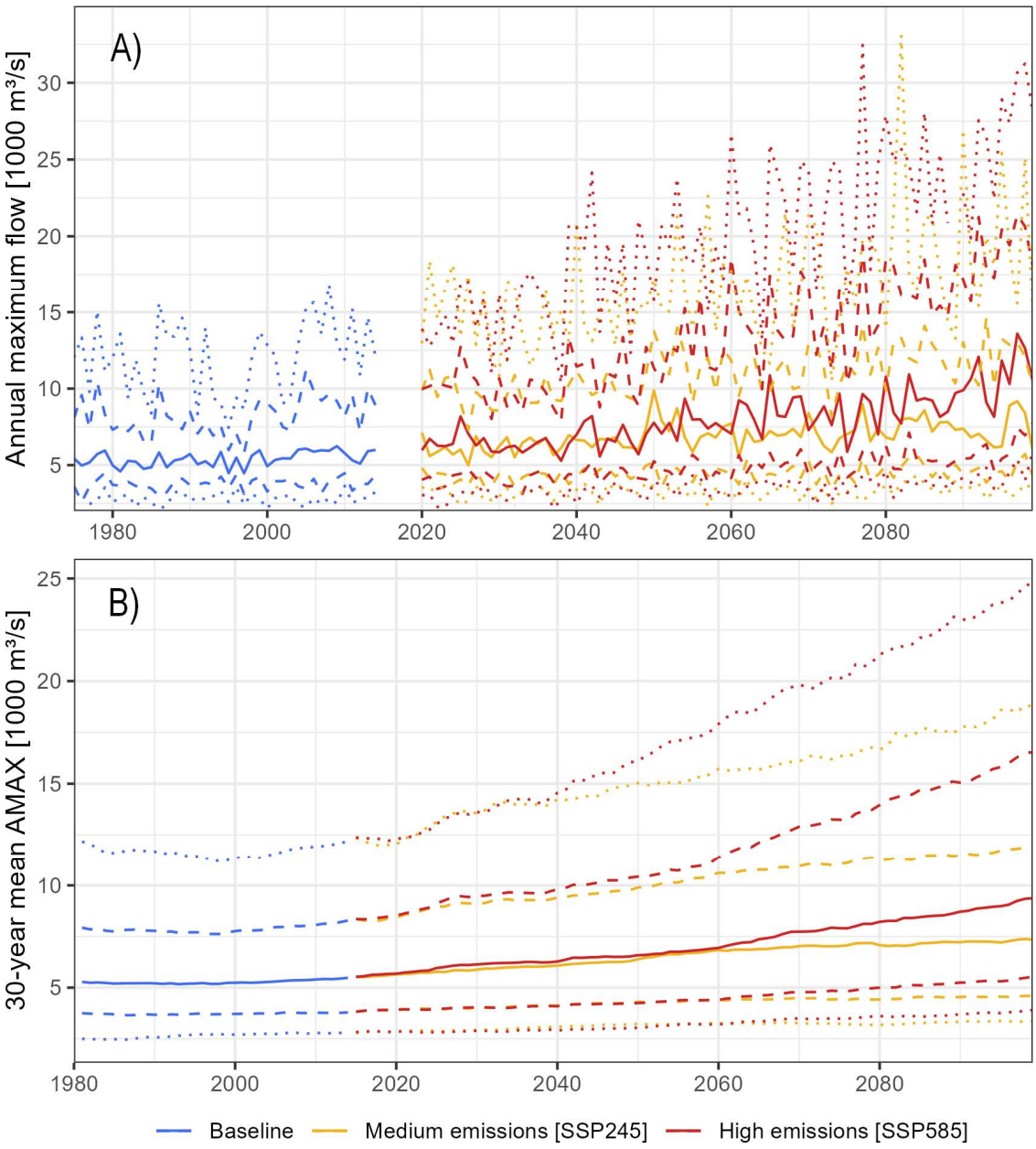


Figure 5.26: A) Time-series of the predicted flood discharge (classified as AMAX). B) Time-series of the 30-year mean flood flows (note that the projected flows before the year 2044 contain predictions of the baseline simulations). The solid line is the median, the dashed lines are the 32nd and 68th percentiles, and the dotted lines are the 2.5th and 97.5th percentiles of the ensemble (n = 768).

The long-term (30-year) mean AMAX negates the annual variation and emphasises the long-term trends. This long-term mean decreases by 2% from 1981 to 1997 after which it increases by 6% to 6,000 m^3/s at the end of the baseline period in 2014 (Figure 5.26 B). This increasing trend is projected into the future for both scenarios, whereas the differences between both increase with time (Figure 5.26 B). The long-term mean AMAX is similar in both scenarios until 2025 because the later baseline AMAX values are used for the calculation at the beginning of the projected period. After the year 2025, the running mean is slightly higher in the SSP585 scenario. However, these scenario differences are small and by the end of the near future (the year 2059) the 30-year mean AMAX increases by 24% and 26% to 6,800 m^3/s and 6,900 m^3/s in the SSP245 and SSP585 scenarios, respectively (Table 5.4). The differences between both scenarios increase with time, particularly after the year 2060. After this year, the long-term AMAX increases moderately to 7,300 m^3/s in 2099 (+34% compared to 2014) for the medium-emission scenario. For the high-emission scenario, however, the trend intensifies and the long-term AMAX increases to 9,400 m^3/s in 2099 which is an increase of 71% compared to 2014 (median predictions).

The increase of the flood discharge scales with the quantile of the ensemble predictions. In 2099, the $Q_{2.5}$ of the long-term AMAX increases by 20% for SSP245 and 40% for SSP585. The $Q_{97.5}$ increases by 54% for SSP245 and 103% for SSP585 compared to 2014. Consequently, the ensemble mean long-term AMAX increases more than the ensemble median, and the standard deviation increases from 43% in the baseline scenario to 54% (SSP245) and 60% (SSP585) in the far future (Table 5.4). This indicates that the increasing uncertainty range is driven by the higher percentiles of the ensemble. However, while the long-term mean emphasises the trends it also attenuates the extremes, and the response of the trends is delayed because it contains information about the previous 29 years.

Table 5.4: Statistics of the long-term AMAX (30-year mean AMAX discharge) of the CMIP6-ensemble. The SD is the 30-year mean standard deviation in percent of the median. Q_n indicates the discharge predicted by the n^{th} percentile of the modelling ensemble (e.g. the Q_{50} is the median).

		30-year mean AMAX [1000 m^3/s]					**Change to Baseline [%]**			
	Year	Mean	Q_{50}	$Q_{2.5}$	$Q_{97.5}$	SD [%]	Mean	Q_{50}	$Q_{2.5}$	$Q_{97.5}$
Baseline	2014	6.0	5.5	2.8	12.2	43	-	-	-	-
SSP245	2059	7.4	6.8	3.3	15.5	47	24	24	16	27
SSP585	2059	7.7	6.9	3.2	17.6	53	29	26	14	44
SSP245	2099	8.2	7.3	3.4	18.8	54	38	34	20	54
SSP585	2099	10.8	9.4	3.9	24.8	60	80	71	40	103

The flood events are projected to increase for all AMAX percentiles (see box 1), whereas this increase scales with time and emissions (Figure 5.27). The projected AMAX discharge is similar for both scenarios in the near future (2020 – 2059). The discharge increases from 5,400 m^3/s in the baseline to 6,500 m^3/s (SSP245) and 6,800 m^3/s (SSP585) for the AMAX P_{50}, which is statistically exceeded once every two years (Table 5.5). For the most extreme floods of the P_{100}, the flow is projected to increase from 11,300 m^3/s in the baseline to 13,700 m^3/s and 13,800 m^3/s for SSP245 and SSP585 respectively. The predicted increases in the near future range between 20 – 28% for both scenarios and the projected increases do not scale with the AMAX percentile (median predictions).

In the far future (2060 – 2099), the projected flood discharge of both scenarios decouples (Figure 5.27). The AMAX P_{50} increase from 5,400 m^3/s in the baseline to 7,200 m^3/s for SSP245 and to 9,300 m3/s for SSP585. The most extreme flows (AMAX P_{100}) increase from 11,300 m^3/s in the baseline to 16,700 m^3/s (SSP245) and 20,900 m^3/s (SSP585). For both scenarios the projected changes increase with the AMAX percentile from +33% (SSP245) and +72% (SSP585) for the AMAX P_{50} to 48% (SSP245) and 85% (SSP585) for the AMAX P_{100} (median predictions).

However, the flood time series (Figure 5.26) indicates that the AMAX flows are not stationary but increases with time, particularly for the SSP585 scenario. This non-stationarity is not reflected in the flow frequency curves of the AMAX percentiles (Figure 5.27) which are based on 40-year records and hence the expected increase for the SSP585 far future is even higher at the end of the century than indicated by this analysis.

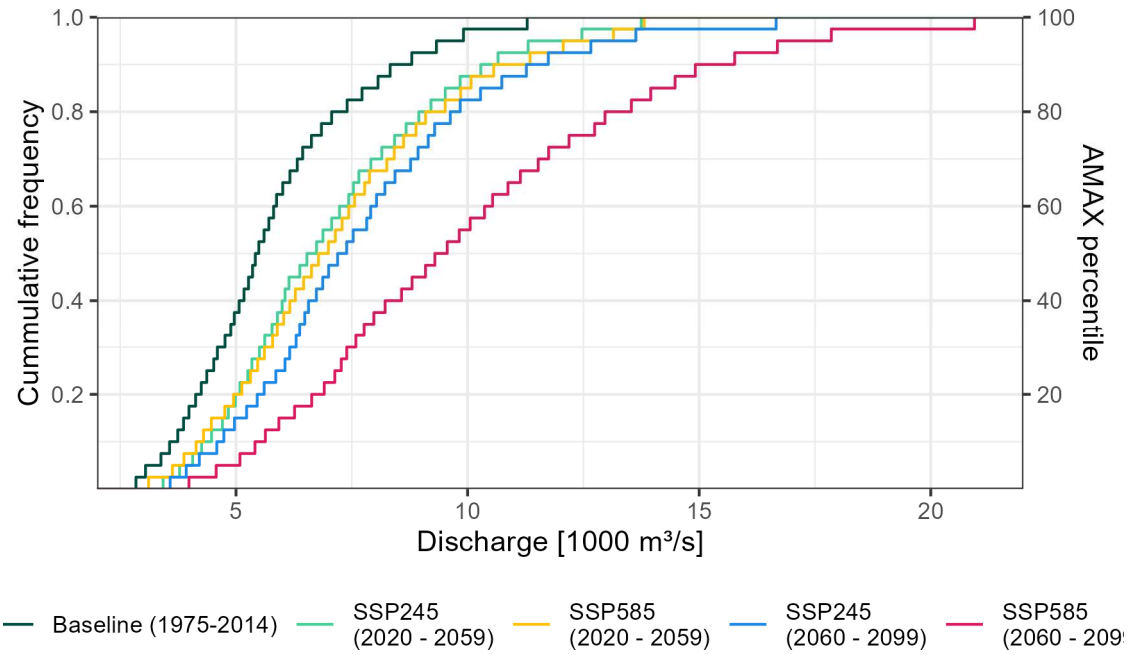


Figure 5.27: Flow duration curves of the CMIP6-ensemble predictions. The ensemble is aggregated by the median.

Table 5.5: Flood flows of the CMIP6-ensemble for different AMAX percentiles (ensemble median). The P_{50} represents the ensemble median AMAX discharge which is exceeded by 50% of the events and the P_{100} is the ensemble median of the highest predicted AMAX event.

	AMAX Percentile [1000 m³/s]					**Change to Baseline [%]**		
	P_{50}	P_{80}	P_{90}	P_{95}	P_{100}	P_{50}	P_{90}	P_{100}
Baseline (1975 - 2014)	5.4	7.1	8.3	9.3	11.3	-	-	-
SSP245 (2020 - 2059)	6.5	8.9	10.3	11.3	13.7	20.4	24.1	21.2
SSP585 (2020 - 2059)	6.8	9.1	10.6	12.1	13.8	25.9	27.7	22.1
SSP245 (2060 - 2099)	7.2	9.6	11.3	12.7	16.7	33.3	36.1	47.8
SSP585 (2060 - 2099)	9.3	13	14.9	16.7	20.9	72.2	79.5	85.0

5.4.3.2 The projections of the CMIP6 members

The large prediction range of the flood discharges of the ensemble (Figure 5.26) indicated the great variation between the CMIP6-ensemble members. This section presents the predictions of the individual ensemble members which differ greatly in their projected magnitudes of the projected flood discharge and change rate.

The ensemble members can be classified into three groups with low, moderate, and high projected changes (the year 2099 compared to the year 2014) according to the long-term 30-year mean AMAX (Figure 5.28 and Table 5.6). The high change group comprises the members EC-Earth3, EC-Earth3-Veg, and ACCESS-ESM1-5, which project increases of 131 – 142% (SSP585) and 55 – 72% (SSP245), and MRI-ESM2-0, NorESM2-LM, NorESM2-MM, and ACCESS-CM2 which project increases of 70 – 90% (SSP585) and 40 – 46% (SSP245). The group projecting moderate changes of 52 – 58% (SSP585) and 21 – 28% (SSP245) includes the INM-CM4-8, INM-CM5-0, and BCC-CSM2-MR members. While these changes are still high, they are considerably smaller than for the previous group. The low change group comprises the MPI-ESM1-2-HR and MPI-ESM1-2-LR members. The long-term AMAX of MPI-ESM1-2-LR increases by 3% for SSP245 and 5% for SSP585. For MPI-ESM1-2-HR, it increases by 6% in the SSP245 scenario. However, for SSP585, the projected increase of 35% is considerably larger than for MPI-ESM1-2-LR (median predictions).

The difference in the projected increase between SSP245 and SSP585 is lowest for the ensemble member with the lowest simulated increase (MPI-ESM1-2-LR: 2%), and highest for the three members with the highest simulated increase (EC-Earth3, EC-Earth3-Veg, ACCESS-ESM1-5: 67 – 87%) (Table 5.6). For the remaining ensemble members, these differences are 25 – 44% (median predictions).

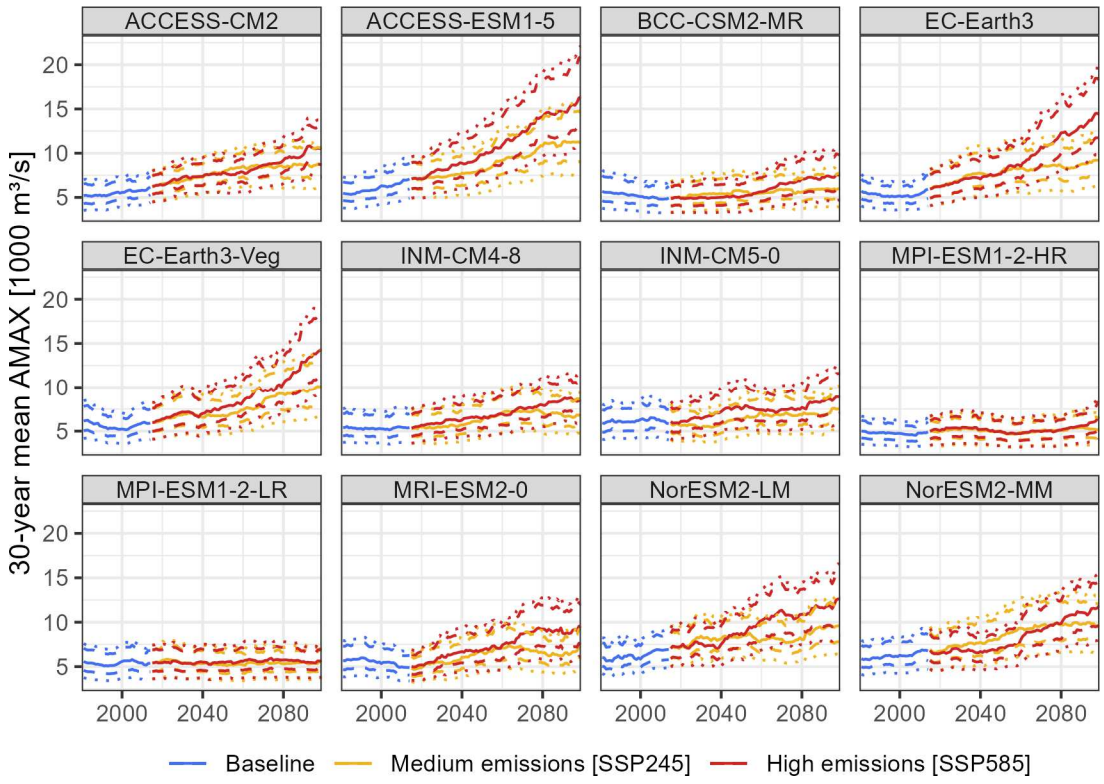


Figure 5.28: Long-term (30-year) mean AMAX predicted for each CMIP6 ensemble member (note that the projected flows before the year 2044 contain predictions of the baseline simulations). The solid line is the median, the dashed lines are the 32nd and 68th percentiles, and the dotted lines are the 2.5th and 97.5th percentiles of the hydrological ensemble (n = 64).

Table 5.6: Statistics of the 30-year mean AMAX in the year 2099 (SSP245 and SSP585) and the year 2014 (baseline). These statistics are the median predictions of the hydrological ensemble. The ensemble members are sorted based on the changes projected for the SSP585 scenario. The predictions at the 2.5th and 97.5th percentiles are presented in Appendix 5.4.

	30-year mean AMAX [1,000 m³/s]			**Change to baseline [%]**	
	Baseline	SSP245	SSP585	SSP245	SSP585
EC-Earth3	5.9	9.2	14.3	55	142
EC-Earth3-Veg	5.9	10.2	14.3	72	142
ACCESS-ESM1-5	7.1	11.1	16.3	57	131
MRI-ESM2-0	4.9	7.2	9.3	46	90
NorESM2-LM	6.9	9.7	12.6	40	82
NorESM2-MM	6.8	9.5	11.9	41	76
ACCESS-CM2	6.2	8.9	10.5	44	70
INM-CM4-8	5.4	6.8	8.5	25	58
BCC-CSM2-MR	4.9	5.9	7.5	21	52
INM-CM5-0	5.9	7.5	8.9	28	52
MPI-ESM1-2-HR	4.9	5.2	6.6	6	35
MPI-ESM1-2-LR	5.4	5.5	5.6	3	5

The ensemble members vary further in the magnitude of the predicted flood discharge (Figure 5.29 and Table 5.7). The AMAX P_{50}, which is statistically exceeded once every two years, ranges from 4,500 m^3/s (MPI-ESM1-2-HR) to 6,700 m^3/s (NorESM2-LM) in the baseline. The highest AMAX P_{50} for SSP245 is predicted for ACCESS-ESM1-5 with 10,100 m^3/s in the far future. The EC-Earth-3-Veg, EC-Earth3, NorESM2-LM, and NorESM2-MM models predict an increase to 8,900 – 9,100 m^3/s. The MPI-ESM1-2-LR and MPI-ESM1-2-HR models predict a very small increase, and the flow remains between 4,600 – 4,800 m^3/s. A similar behaviour is predicted for the SSP585 scenario for which ACCESS-ESM1-5 predicts with 15,600 m^3/s the highest AMAX P_{50}. However, the difference increases between the EC-Earth-3-Veg and EC-Earth3 models (13,200 – 14,300 m^3/s), and the NorESM2-LM, and NorESM2-MM models (11,100 – 11,200 m^3/s). The lowest flow magnitude is predicted for the MPI-ESM1-2-LR and MPI-ESM1-2-HR models (4,900 – 5,300 m^3/s) (median predictions).

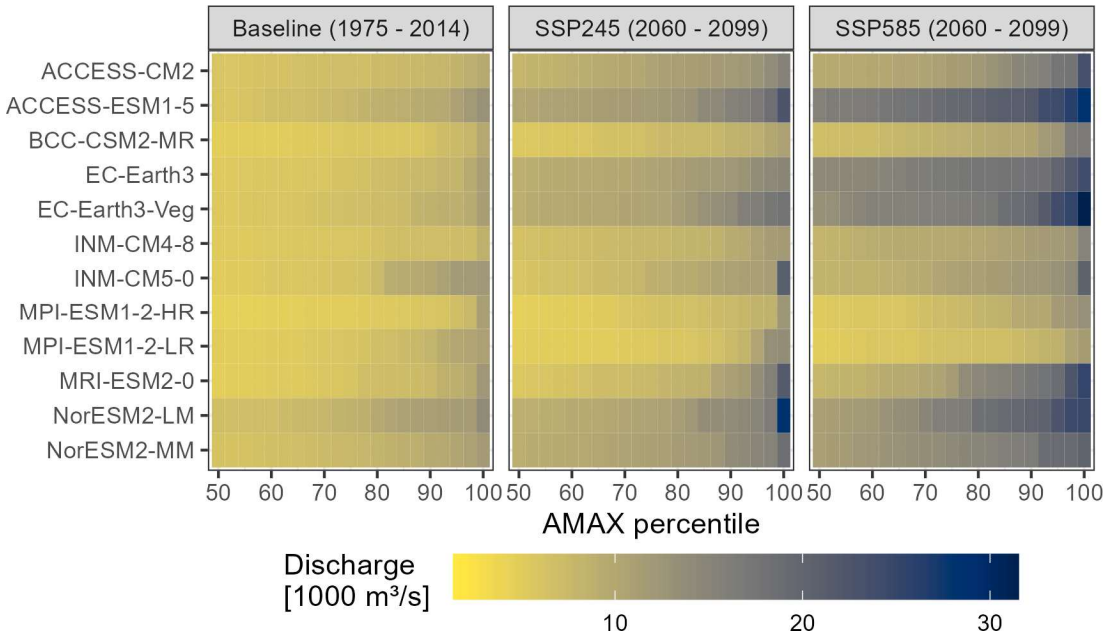


Figure 5.29: The median (hydrological ensemble) discharge of the AMAX percentiles for the CMIP6 ensemble members.

Table 5.7: The predicted discharge of selected AMAX percentiles of the CMIP6-ensemble members for the baseline (1975 – 2014) and the projected scenarios (2060 – 2099). The flood discharge and the projected changes present the median predictions of the hydrological ensemble (n = 64). The non-monsoon events indicate the fraction of all AMAX events that are predicted during the non-monsoon (Oct – May). The AMAX P_{50} is statistically exceeded once every two years and the AMAX P_{100} is the highest simulated flood event in the 40-year record. The predictions for the near future (2020 – 2059) are summarised in Appendix 5.5.

		AMAX percentile [1000 m³/s]					**Change to baseline [%]**			**Non-monsoon events [%]**
		P_{50}	P_{80}	P_{90}	P_{95}	P_{100}	P_{50}	P_{95}	P_{100}	
ACCESS-CM2	Baseline	6.0	7.3	7.8	8.2	9.5	-	-	-	12
	SSP245	7.9	11.2	11.8	12.6	15.4	32	54	62	7
	SSP585	9.5	12.1	15.1	17.3	23.4	58	111	146	4
ACCESS-ESM1-5	Baseline	5.8	8.7	9.5	10.8	13.0	-	-	-	3
	SSP245	10.1	12.4	15.3	16.6	22.6	74	54	74	0
	SSP585	15.6	20.1	21.7	24.3	28.1	169	125	116	0
BCC-CSM2-MR	Baseline	4.8	5.6	6.4	7.4	9.8	-	-	-	5
	SSP245	5.5	7.5	8.5	8.9	10.9	15	20	11	5
	SSP585	6.8	9.4	10.4	12.4	16.6	42	68	69	7
EC-Earth3	Baseline	5.3	6.9	7.6	8.1	11.2	-	-	-	1
	SSP245	8.9	11.2	12.2	13.7	14.7	68	69	31	0
	SSP585	14.3	17.4	17.9	20.1	23.8	170	148	112	0
EC-Earth3-Veg	Baseline	5.4	7.1	8.6	9.0	11.3	-	-	-	2
	SSP245	8.9	11.3	14.1	15.9	17.7	65	77	57	2
	SSP585	13.2	16.6	19.4	23.9	31.5	144	166	179	0
INM-CM4-8	Baseline	5.2	6.3	6.9	7.1	8.7	-	-	-	7
	SSP245	6.4	8.1	9.3	10.8	11.7	23	52	34	12
	SSP585	8.0	10.1	11.4	11.8	15.1	54	66	74	1
INM-CM5-0	Baseline	4.9	7.4	9.9	11.8	12.1	-	-	-	9
	SSP245	6.0	9.3	10.9	11.1	21.0	22	-6	74	6
	SSP585	8.2	11.1	12.3	13.1	20.4	67	11	69	5
MPI-ESM1-2-HR	Baseline	4.5	5.3	6.0	6.8	11.2	-	-	-	13
	SSP245	4.6	6.5	7.4	8.3	12.5	2	22	12	18
	SSP585	5.3	8.1	9.9	11.7	13.3	18	72	19	29
MPI-ESM1-2-LR	Baseline	5.0	6.7	8.1	10.1	10.8	-	-	-	16
	SSP245	4.8	5.7	7.4	10.3	13.6	-4	2	26	10
	SSP585	4.9	6.7	7.8	8.6	11.4	-2	-15	6	8
MRI-ESM2-0	Baseline	4.7	7.0	7.5	9.0	12.3	-	-	-	15
	SSP245	5.5	8.4	11.1	13.5	21.7	17	50	76	1
	SSP585	7.8	14.6	17.0	18.8	25.2	66	109	105	1
NorESM2-LM	Baseline	6.7	9.5	11.2	11.4	14.0	-	-	-	4
	SSP245	9.0	11.3	13.9	15	28.5	34	32	104	5
	SSP585	11.2	18.2	20.1	22.6	24.8	67	98	77	1
NorESM2-MM	Baseline	6.2	8.1	9.6	10.3	11.7	-	-	-	4
	SSP245	9.1	11.6	13.9	14.9	18.5	47	45	58	2
	SSP585	11.1	15.0	16.4	18.9	20.3	79	84	74	0

The predictions of the most extreme flows AMAX P_{100} differ from the AMAX P_{50} (Figure 5.29 and Table 5.7). These flows range between 8,700 m^3/s (INM-CM4-8) and 14,000 m^3/s (NorESM2-LM) in the baseline. In the far future, the NorESM2-LM model predicts the highest discharge for the SSP245 scenario (28,500 m^3/s). This is considerably higher than the second-highest prediction of 22,600 m^3/s of the ACESS-ESM1-5 member. The lowest discharge is predicted for BCC-CSM2-MR with 10,900 m^3/s. This order changes for the high-emission scenario SSP585 for which the highest discharge is predicted by EC-Earth3-Veg with 31,500 m^3/s, followed by ACCESS-ESM1-5 with 28,100 m^3/s. The NorESM2-LM model predicts with 24,800 m^3/s a lower discharge than for SSP245. The lowest AMAX P_{100} is predicted by MPI-ESM1-2-LR with 11,400 m^3/s.

The members differ in the projected changes and the projected magnitudes. However, the models also show different behaviour regarding the AMAX percentile for which they predict the largest changes (Figure 5.30 and Table 5.7). The EC-Earth3-Veg model predicts higher increases above the AMAX P_{90} (SSP585). A similar behaviour is predicted for ACCESS-CM2, and NorESM2-LM (SSP245). On the contrary, the EC-Earth3 and ACESS-ESM1-5 models predict higher changes below AMAX P_{80}. However, the ACCESS-ESM1-5 model predicts high baseline flood discharge and hence this model still projects high increases for the upper AMAX percentiles. The MPI-ESM1-2-LR model stands out from the ensemble because it predicts low changes and is the only model which predicts a decrease in flood flows except for the most extreme AMAX P_{100}.

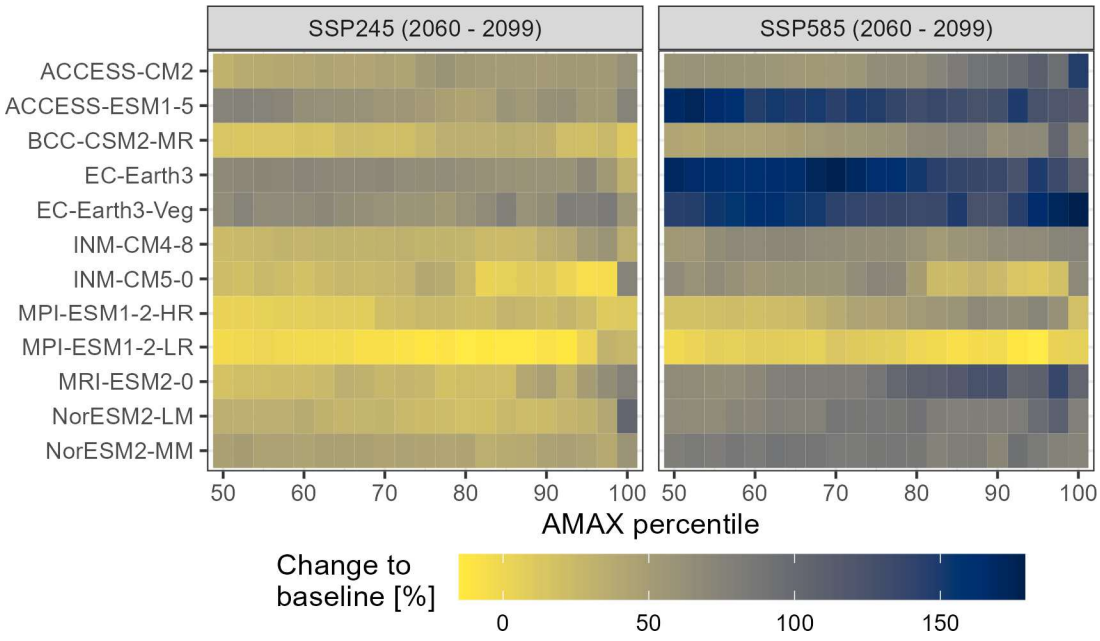


Figure 5.30: The changes of the AMAX percentiles between the baseline (1975 – 2014) and the projected climate scenarios (2060 – 2099) for each CMIP6-ensemble member. These changes are the median of the hydrological ensemble.

5.4.3.3 The composition of the projected flood events

The projected increase of the flood flows is dominated by increasing rainfall-runoff which contributes 90 – 93% to the projected increases in the flood discharge. Rainfall-runoff is the most important source with a mean contribution of 78% in the baseline scenario (Figure 5.31). This share increases to 80% in the near future (2020 – 2059), and to 81% (SSP245) and 83% (SSP585) in the far future (2060 – 2099). The absolution rainfall-runoff contribution increases by 1,953 m^3/s (SSP245) and 4,255 m^3/s (SSP585) in the far future. This projected increase vastly exceeds the projected increase of the baseflow and glacier melt components (100 – 250 m^3/s). The share of the baseflow and glacier melt contributions remain stable at 8 – 10% for SSP245 but decrease by one percent for SSP585. The absolute and relative share of snowmelt decreases and 1.1% (SSP245) and 0.3% (SSP585) of the flood water originates from snowmelt in the far future. This decrease in the snowmelt contribution could either result from decreases in snowfall or an earlier melting season.

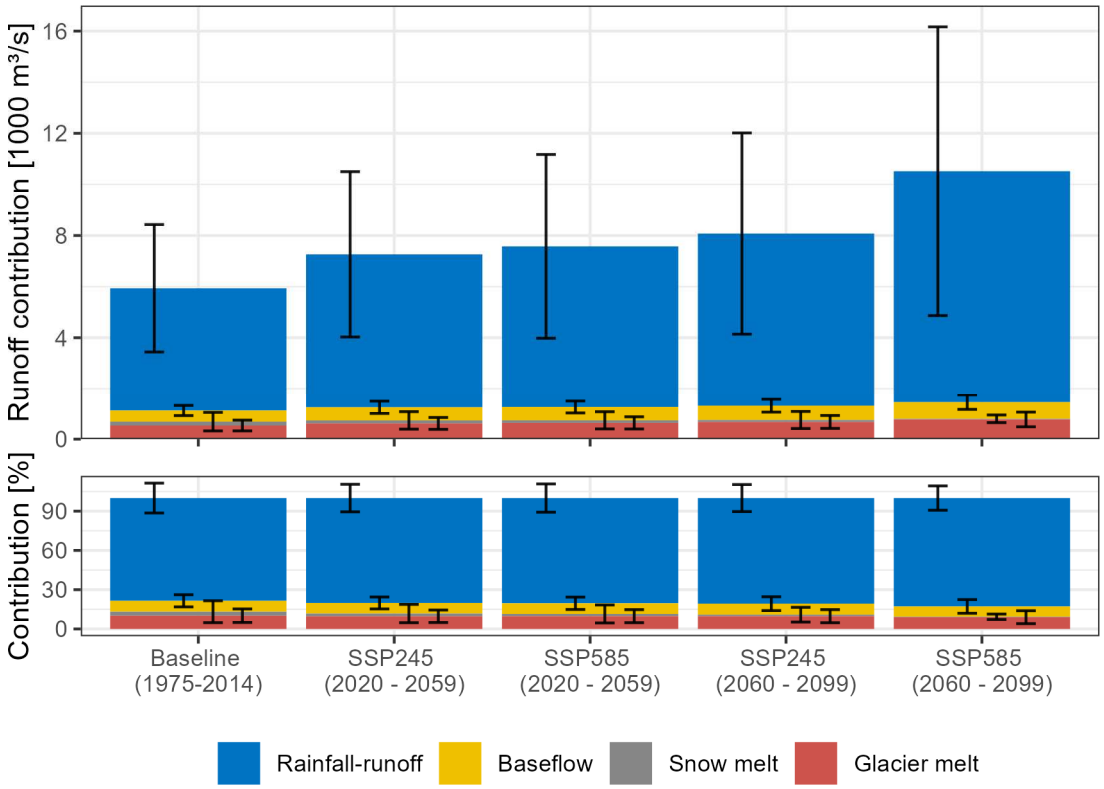


Figure 5.31: mean contribution of the runoff components to the flood discharge for the CMIP6 simulations. The error bars indicate the standard deviation of the composition resulting from the ensemble variation and the event variation.

While rainfall-runoff dominates the periodic mean contribution the other components can contribute a large share to individual events (Figure 5.32). The importance of these components decreases with the flood discharge and the events with the highest magnitudes are composed of ≥ 90% rainfall-runoff. The snowmelt contribution can reach up to 75% for lower flood flows (≤ 5,000 m^3/s), whereas the number of events with high snowmelt contributions decreases with time and emissions. However, the composition of the flood flows shares similar characteristics for all scenarios and the increase in the flow magnitude is more pronounced than the compositional changes.

There are distinct differences in the timing of high contributions between the components. High snowmelt contributions are simulated in the pre-monsoon season, glacier contributions increase during the monsoon season, and the baseflow ones increase towards the post-monsoon season (Figure 5.32). No substantial differences in the timing of large contributions between the scenarios are predicted except for the snowmelt component. High contributions (≥ 50%) are occurring until week 25 in the baseline. The timing of such high snowmelt contributions shifts towards earlier dates and these events occur only until week 22 in the far future of the high-emission scenario SSP585. This coincides with the projected shift of the snowmelt season to earlier months (Figure 5.25).

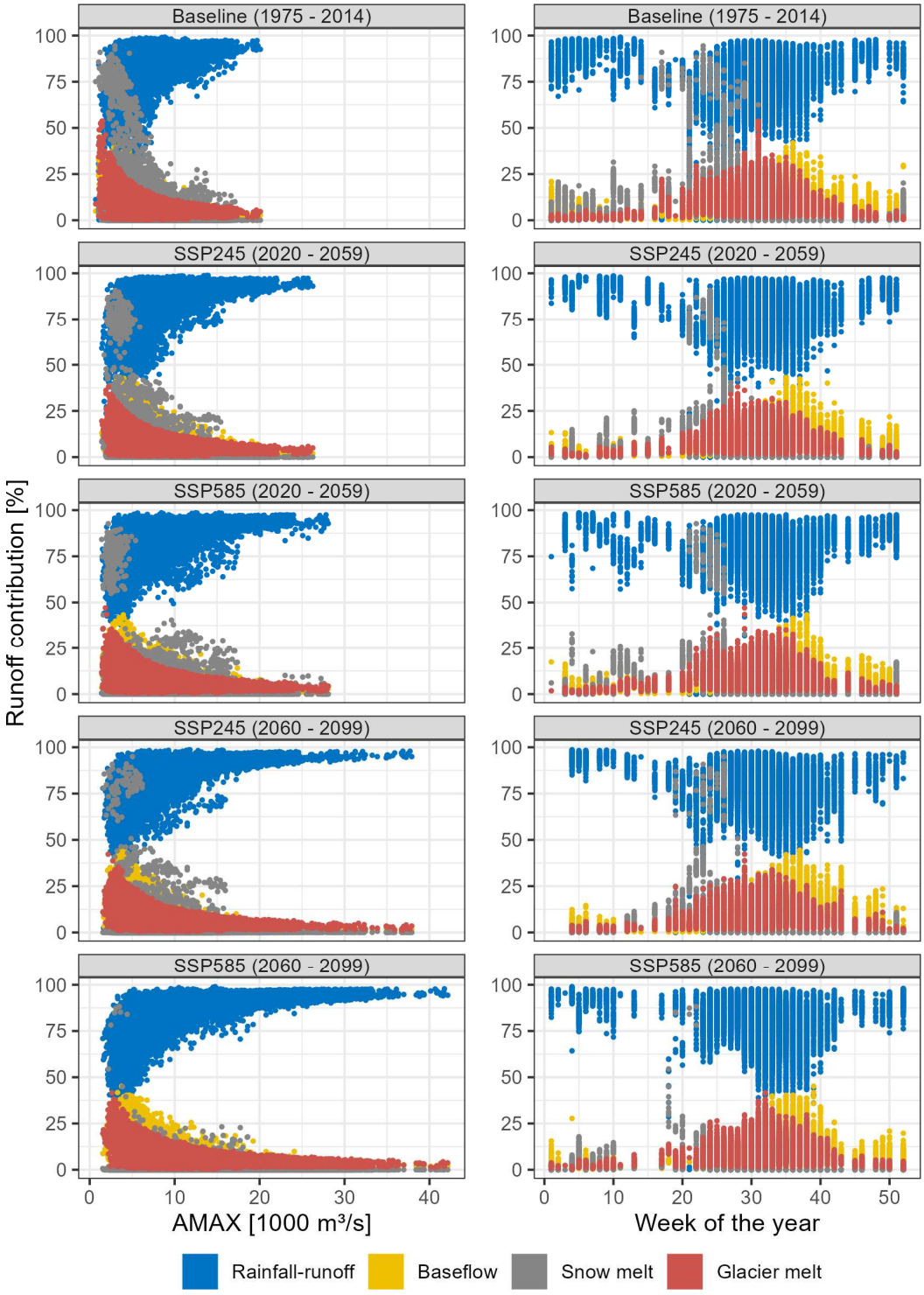


Figure 5.32: The relationship between the runoff contribution and the flood discharge (left) and between the runoff contribution and the timing of the flood event (right) for the CMIP6 simulations. Each plot consists of 30,720 samples (12 CMIP6 models x 64 hydrological models x 40 years).

5.4.3.4 The timing of the projected flood events

The timing of the flood events remains similar in the projected scenarios. The median week of the flood occurrence is week 31 in all scenarios. This week marks the break of July and August which are the months with the highest simulated discharge. The high agreement in the flood timing is also visible in the density curves in Figure 5.33 which indicate no systematic differences between the scenarios.

The number of floods occurring in the non-monsoon season (Oct – May) decreases for the projected scenarios whereas this trend scales with the emission scenario. In the baseline period, 7.5% of the simulated flood flows occur in the non-monsoon season. This decreases to 5.2% (SSP245) and 5.1% (SSP585) in the near future. In the far future, the share of non-monsoon season floods decreases further to 4.7% in the SSP585 scenario, while it increases slightly to 5.6% in the SSP245 scenario. The ensemble members which predict high flood discharges generally predict fewer flood events outside the monsoon season. The models with the highest projected flood flows simulate no such events (e.g. ACCESS-ESM1-5, EC-Earth3) (Table 5.7). The MPI-ESM-1-2-LR and MPI-ESM-1-2-HR models predict the lowest change in the flood discharge and for these models, the largest number of non-monsoon season floods are projected. For MPI-ESM1-2-HR an increase is predicted from 13% in the baseline to 18% (SSP245) and 29% (SSP585) in the far future. Therefore, a contradictory behaviour of more non-monsoon season events for the higher emissions is simulated for this member.

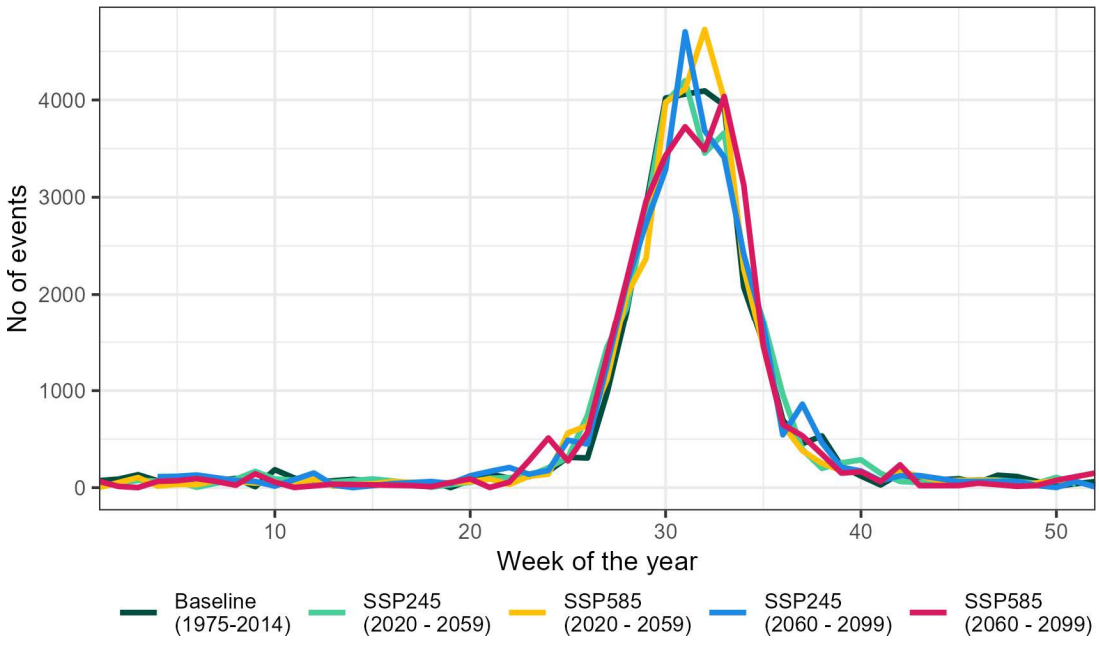


Figure 5.33: The timing of the flood events for the CMIP6 ensemble. Each group consists of 30,720 events (12 CMIP6 members x 64 hydrological models x 40 years).

5.4.3.5 The duration of the projected flood events

Conventional approaches to estimate the flood duration (e.g. flow duration analysis) require the definition of a flow threshold and quantify the duration for which this threshold is exceeded. However, the flood flow threshold is unknown for the Karnali River at the catchment outlet. Therefore, the flow of the rising and recession limbs is scaled relative to the flood discharge to investigate changes in the flood duration between the scenarios. This is only conducted for events ≥ AMAX P_{50} of the flood time series to remove events which are potentially misclassified by the AMAX classification approach.

The scaled flow of the days before and after the flood peak decreases with time and emissions (Figure 5.34). The scaled discharge two days before the flood event is 66% of the flood discharge in the baseline which decreases to 60% in the near future, and to 58 – 60% in the far future. The scaled discharge one day after the flood event reduces from 94% in the baseline to 93% in the far future. This behaviour of reduced relative discharge is predicted consistently for the interval ±10 days to the flood event, and hence the simulations projected a higher increase for the flood discharge than for the rising limb and the flood recession. This suggests that the flood duration is not projected to increase. However, the flood discharge is projected to increase and hence the rising limb and the recession may increase in absolute discharge rates but at a lower rate than the flood discharge. Therefore, the out-of-bank flow might last longer which cannot be investigated without knowing the bankfull flow rate. Nonetheless, the flood projections do indicate that the flood-triggering rainfall events increase in intensity since the flood flow increases, but do not increase in duration which would result in shifts of the rising or recession limbs of the scaled discharge (median predictions).

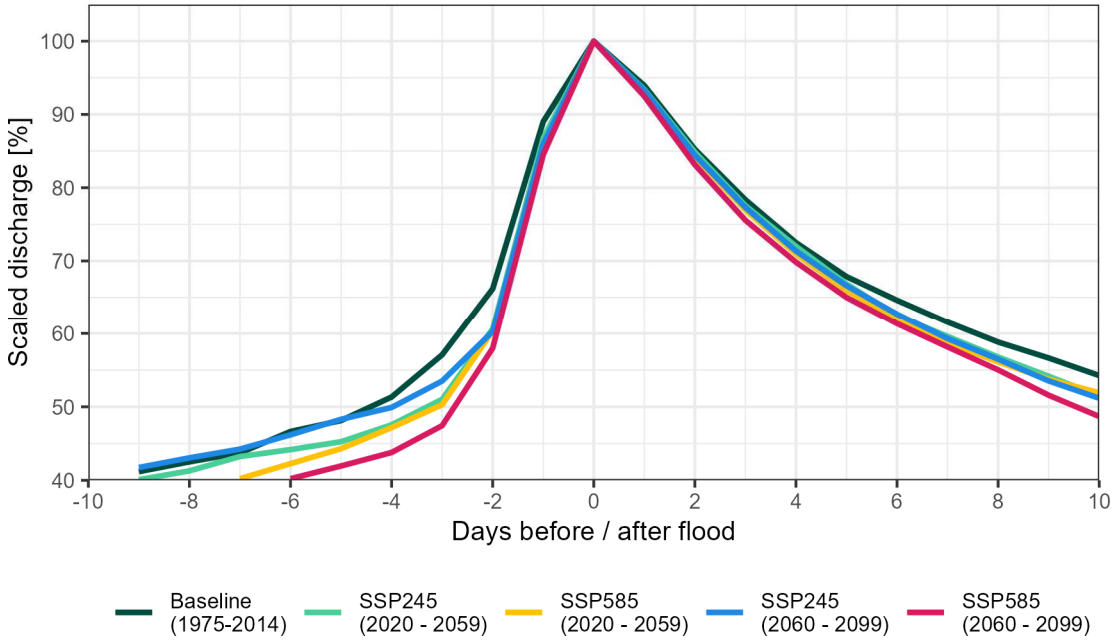


Figure 5.34: The scaled discharge (median) before and after the flood event. The discharge is scaled with the flood peak. Only events ≥ AMAX P_{50} are considered (events which are statistically exceeded once every two years) to remove misclassifications from the AMAX approach.

The comparison of baseline and projected flood discharge indicates an increase with time and emissions. This projected increase is higher for higher percentiles of the modelling ensemble. The increase in the flood discharge is predominantly driven by higher rainfall-runoff which contributes 90 – 93% to this increase. On the contrary, snowmelt loses significance for the flood flows. The comparison of the baseline and projected runoff composition indicates no shift in the model behaviour. Therefore, the parameter transfer to the projected climates does not cause an unreasonable behaviour of the hydrological modelling ensemble. The projected flood discharge indicate no trend in the timing and duration of flood events.

However, while the comparison of baseline and projected flows provides a clear trend of increasing flood discharges, it lacks the quantification of the changes in the flood magnitudes. This is due to the high temporal variability of extreme events which hampers a direct comparison of single events. The mean over a period counteracts this variability but loses information about the largest extremes which are naturally most interesting to flood managers. It is, therefore, necessary to standardise the AMAX events to quantify the changes for past and projected climates. Therefore, a Flood Frequency Analysis is conducted to standardise the flood events and quantify the changes in the projected flood magnitudes.

5.4.4 The flood frequency analysis for the Karnali River

The flood frequencies and magnitudes of the Karnali River are predicted by a Flood Frequency Analysis (FFA) which is conducted for the observed and the simulated floods of the CMIP6-modelling. The predicted flood flows of the CMIP6 simulations comprise three scenarios, the baseline (1975 – 2014) as a reference of the current climatic conditions, the medium-emission scenario SSP245 and the high-emission scenario SSP585 (both 2020 – 2099). The projected scenarios are split into the near future (2020 – 2059) and the far future (2060 – 2099). The floods are classified from the Annual Maximum Flow (AMAX) approach and each record, therefore, consists of 40 flood events.

The FFA is conducted by fitting an Extreme Value Distribution (EVD) to the flood record but it is unknown which EVD best represents the distribution of the flood samples. The first Section 5.4.4.1 presents the EVD selection process and Section 5.4.4.2 compares the FFAs of the observed flood record and the simulated baseline ones to gain insight into the accuracy of the flood frequencies and magnitudes of the CMIP6 simulations. The projected changes in flood frequencies and magnitudes are quantified by comparing the FFA of the baseline and projected flood records in Section 5.4.4.3.

5.4.4.1 Model selection

A set of eight Extreme Value Distributions EVDs were fitted for each simulated AMAX record resulting in 3840 fitted distributions for each EVD (12 CMIP6 x 64 SPHY x 5 periods). The efficiency of each modelled distribution is evaluated by comparing it with the distribution of the AMAX record. The median test statistics are presented in Table 5.8, and the full range of these statistics is visualised in the boxplots of Figure 5.35.

All EVDs have a very high median Coefficient of Determination (R^2) of ≥ 0.96 and low Root Mean Square Errors (RMSE) of ≤ 0.06. The best performance is observed for the Wakeby distribution followed by the Generalised Extreme Value distribution, Generalised logistic distribution, Pearson Type III distribution, Weibull distribution, and the Generalised Pareto distribution. The performances of the distributions are, in general, very similar but the Gumbel distribution and Exponential distribution have the highest number of frequency curves with $R^2 < 0.9$ and RSME > 0.1.

Table 5.8: Median efficiencies for the flood frequency curves fitted for different Extreme Value Distributions.

Extreme value distribution	Anderson-Darling Test statistic	Anderson-Darling Test P-Value	R^2	RMSE
Wakeby	0.205	0.977	0.992	0.028
Gen. extreme value	0.287	0.947	0.989	0.033
Gen. logistic	0.331	0.911	0.988	0.036
Pearson Type III	0.27	0.91	0.988	0.033
Weibull	0.247	0.839	0.988	0.034
Gumbel	0.414	0.833	0.984	0.041
Exponential	0.401	≤ 0.05	0.968	0.056
Gen. pareto	0.277	≤ 0.05	0.984	0.039

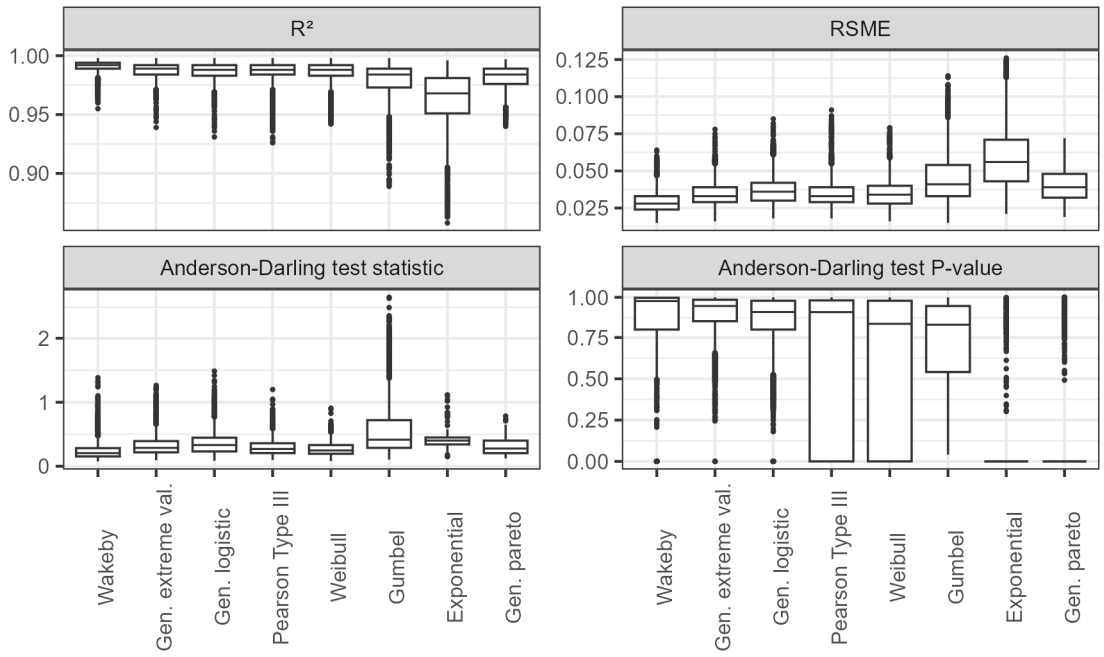


Figure 5.35: The performance of the Extreme Value Distributions fitted to the simulated AMAX records.

A low Anderson-Darling-Test (AD) value indicates a better fit between the modelled distribution and the data. The Wakeby distribution has the lowest AD value but all EVDs except for the Gumbel distribution, the Exponential distribution, and the Generalized logistics distribution are very similar (median AD values 0.20 – 0.29). However, the Pearson Type III distribution, Weibull distribution, the Exponential distribution, and the Generalised Pareto distribution contain higher numbers of EVD for which the null-hypothesis (P-Value ≤ 0.05) cannot be rejected. Therefore, these distributions are excluded. The Gumbel distribution is eliminated for the higher RMSE, and the Generalised logistic distribution is excluded for the higher AD value. The remaining Generalised Extreme Value distribution and Wakeby distribution have high efficiencies and are both suitable candidates. The Wakeby distribution has a slightly better AD test statistic and is therefore selected for the FFA.

5.4.4.2 The comparison of the baseline predictions with observations

The FFA is a statistical data-driven method and, thus, prone to errors in the dataset. The measurement uncertainty which describes errors in the discharge of the flood record is accounted for by the ensemble predictions of 64 hydrological models and the 12 CMIP6 datasets. The sampling error describes the uncertainty arising from deviations between the true flood discharge distribution and the sampled one, which is most commonly exacerbated by the short record length. This sampling uncertainty is estimated from a bootstrap approach (Burn, 2003). The bootstrap size was set to 999 and it was tested that a larger size does not alter the predictions. Consequently, 64,000 frequency curves are for each CMIP6 member (999 synthetic records + 1 simulated record x for each of the 64 hydrological models).

A random sampling approach is used to estimate the measurement uncertainty in the discharge observations. For this, the AMAX events are extracted from the discharge observations. This record is duplicated 64 times (to equal the measurement uncertainty sample size of the hydrological modelling), and each element of these records is multiplied with a random value in the range of 0.6 – 1.4 to incorporate the uncertainty in the discharge measurements which ranges between ± 40% for flood discharges (McMillan *et al.*, 2012 and Figure 4.16). The bootstrap approach (n = 999) is used to estimate the sampling uncertainty which results in 64,000 frequency curves of the observed discharge record. The predictions of these flood frequency curves are hereafter referred to as the observed flood magnitudes. For clarification, these observed flood magnitudes are simulated from discharge observations.

The flood magnitudes (FM) of the baseline record are lower than the observed FM across the full range of investigated flood frequencies (Figure 5.36). The slope of the observed FFA is steeper than the one of the baseline FFA and hence the differences increase with the return period. The baseline FM of the 1-in-10-years event is 8,300 m^3/s and is 43% lower than the observed FM (12,900 m^3/s). This difference increases to 65% for the 1-in-100-years event for which the baseline FM is 11,600 m^3/s and the observed FM 22,600 m^3/s (median predictions).

The differences between the baseline FM and the observed FM decrease with increasing percentile of the ensemble predictions (Figure 5.36). The underestimations are highest for the $P_{2.5}$ and range between 70 – 76% for the 1-in-10-years to 1-in-100-years return periods. For the $P_{97.5}$, these underpredictions range between 20 – 46% for the 1-in-10-years to 1-in-100-years return periods. However, the absolute underpredictions increase with the percentile from 7,400 m^3/s ($P_{2.5}$) to 11,600 m^3/s ($P_{97.5}$) for the 1-in-100-years event.

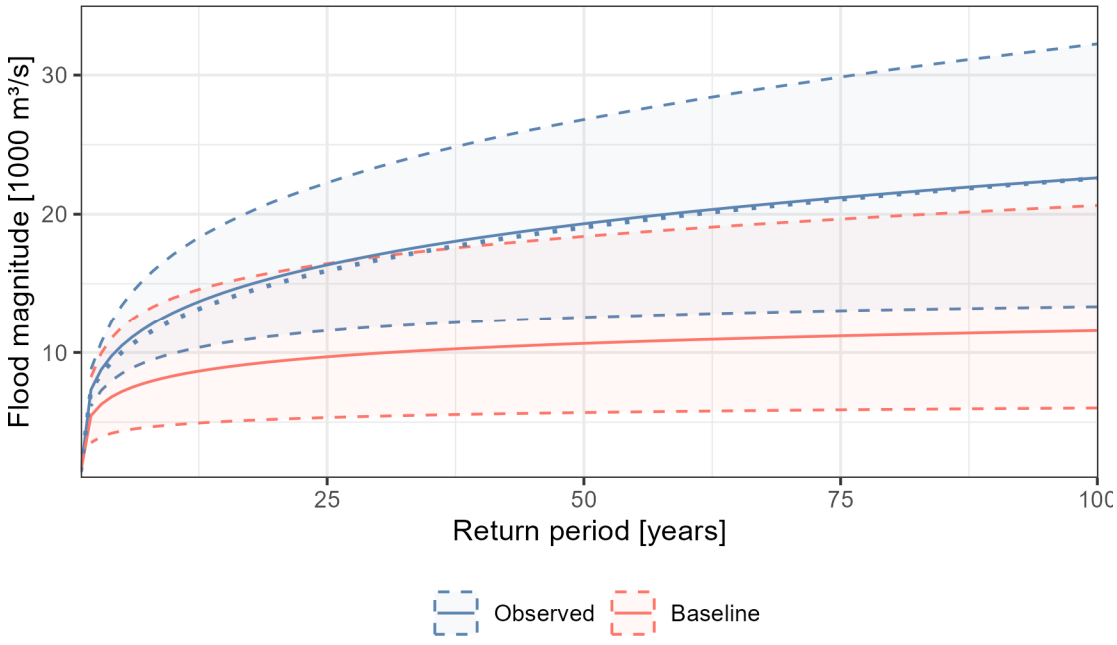


Figure 5.36: Flood frequency curves of the AMAX observations(1975 – 2014), and the one simulated from the CMIP6 baseline predictions (1975 – 2014). The solid line indicates the median and the dashed ones are the $P_{2.5}$ and $P_{97.5}$ of the ensemble. The ensemble sizes are 64,000 for the observed FFA and 768,000 for the baseline FFA.

5.4.4.2 The flood magnitudes of the projected climates

The Flood Magnitudes (FM) and Flood Frequencies (FF) are projected to increase for the projected climate scenarios. This projected FM increase scales with time and emissions (Figure 5.37). The Flood Frequency Analysis (FFA) indicates that the flood magnitudes increase for all percentiles of the ensemble predictions and across the full range of investigated return periods.

The median flood magnitude of the 1-in-10-year event increases from 8,300 m^3/s in the baseline to 10,300 m^3/s (SSP245) and 10,700 m^3/s (SSP585) in the near future, and to 11,300 m^3/s (SSP245) and 15,200 m^3/s (SSP585) in the far future (Table 5.9). Therefore, the projected 1-in-10-years magnitudes are close to or exceed the baseline 1-in-100-year magnitude. This baseline 1-in-100-years magnitude is 11,600 m^3/s, which increases by 23 – 26% to 14,300 – 14,600 m^3/s (SSP245 and SSP585) in the near future, and increases by 40% to 16,200 m^3/s (SSP245) and 79% to 20,700 m^3/s (SSP585) in the far future. The median flood frequency curves of the baseline and projected scenarios differ in magnitude but have a similar shape above the 1-in-10-years return period, and hence similar changes are projected for all return periods between 1-in-10-years and 1-in-100-years (median predictions).

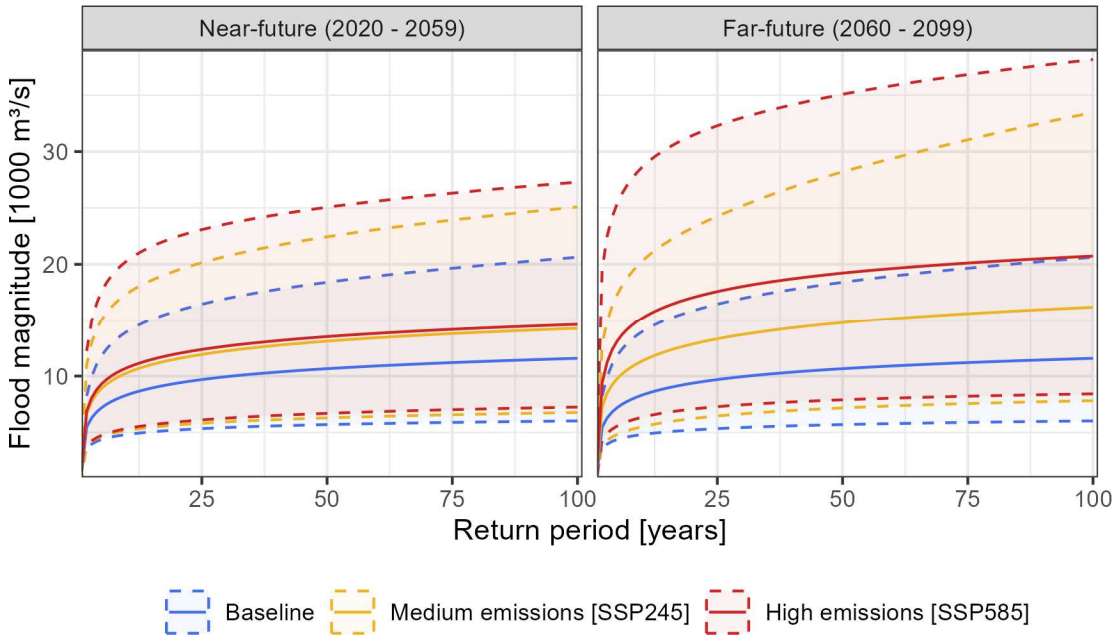


Figure 5.37: The flood frequency analysis for the baseline and projected scenario of the CMIP6 ensemble. The solid lines indicate the median the dashed lines indicate the $P_{2.5}$ and $P_{97.5}$ of the ensemble predictions.

Table 5.9: The median flood magnitudes of the CMIP6 ensemble for selected return periods and the projected changes to the baseline.

Scenario	Flood magnitude [1,000 m^3/s]				Change to baseline [%]			
	10 y	20 y	50 y	100 y	10 y	20 y	50 y	100 y
Baseline (1975 - 2014)	8.3	9.4	10.7	11.6	-	-	-	-
SSP245 (2020 - 2059)	10.3	11.5	13.1	14.3	24	23	23	23
SSP585 (2020 - 2059)	10.7	12	13.5	14.6	29	28	27	26
SSP245 (2060 - 2099)	11.3	12.8	14.8	16.2	37	37	39	40
SSP585 (2060 - 2099)	15.2	17	19.2	20.7	83	82	80	79

The projected changes in the flood magnitudes are similar to the projected flood flow changes (Section 5.4.3.1). These also predict increases with time and emissions as well as growing scenario differences in the far future. Furthermore, the projected changes of the AMAX percentiles are of similar magnitude as for the FFA predictions (Table 5.5 and Table 5.9). The AMAX P_{90} is projected to increase by 24% for SSP245 and 27% for SSP585 in the near future. This rate of change is similar to the projected change of the 1-in-10-year flood magnitude of 24% and 29% for both scenarios respectively. Furthermore, the predicted AMAX percentile and the predicted flood magnitudes are of similar magnitude. The difference between the AMAX P_{90} and the 1-in-10-years magnitude is within ±300 m^3/s for each scenario. The rate of change of the P_{100} is similar (±6%) to one of the 1-in-100-year events and the flow differences are within ±800 m^3/s. The only exception is the SSP245 (far future) for which higher AMAX P_{100} changes are predicted (+48%) than for the 1-in-100-years FM (+40%) (median predictions).

The projected increase of the flood magnitudes scales with the percentile of the ensemble predictions (Figure 5.37 and Table 5.10). For example, the flood magnitudes of the 1-in-100-years event of the far-future SSP585 are projected to increase by 40%, and 85% at the $P_{2.5}$ and $P_{97.5}$ of the FFA ensemble, respectively. This behaviour is consistent for each scenario in both periods for all return periods. Furthermore, the slope of the flood frequency curves is steeper at the higher percentiles which means the prediction range of the ensemble increases with the return period. For example, the prediction range ($P_{2.5}$ – $P_{97.5}$) is 9,200 m^3/s (4,800 – 14,000 m^3/s) for the 1-in-10-years event and increases to 14,600 m^3/s (6,000 - 20,600 m^3/s) for the 1-in-100-years event of the baseline scenario. (Table 5.10).

Table 5.10: The flood magnitude prediction intervals of the CMIP6 ensemble for the 1-in-10-years and 1-in-100-years return periods.

	Flood magnitude [1,000 m^3/s]				**Change to baseline [%]**			
	10 y		100 y		10 y		100 y	
	$P_{2.5}$	$P_{97.5}$	$P_{2.5}$	$P_{97.5}$	$P_{2.5}$	$P_{97.5}$	$P_{2.5}$	$P_{97.5}$
Baseline (1975 - 2014)	4.8	14,0	6,0	20.6	-	-	-	-
SSP245 (2020 - 2059)	5.2	17.4	6.8	25.1	7	24	12	22
SSP585 (2020 - 2059)	5.3	20.3	7.2	27.3	11	45	20	32
SSP245 (2060 - 2099)	5.5	20.3	7.8	33.5	14	45	30	62
SSP585 (2060 - 2099)	6.3	28.6	8.4	38.2	32	105	40	85

While the shape of the frequency curves differs between percentiles of the ensemble predictions, the shape of the respective curves is largely similar for the three scenarios (Figure 5.37). The projected frequency curves at the $P_{2.5}$ are moderately steeper than the baseline one and hence the projected change increases with the return period. For example, the projected change of $P_{2.5}$ increases from 14% to 30% between the 10 – 100 year return period for SSP245 (far future). The shape of the median flood frequency curves is largely similar for the three scenarios. The slope of the baseline $P_{97.5}$ flood frequency curve is steeper than the projected flood frequency curves, and hence the projected change decreases with the return period. For example, the projected changes of SSP585 (far future) decrease from 105% for the 1-in-10-years event to 85% for the 1-in-100-years event (Table 5.10).

The differences between the medium-emission scenario SSP245 and the high-emission scenario SSP585 are stable for the return periods except for the $P_{97.5}$ predictions (Figure 5.38). The $P_{97.5}$ flood frequency curve of the SSP245 scenario (far future) is steeper than the one of SSP585 and hence the difference between both scenarios decreases with the return period. Figure 5.38 shows the cumulative frequency curves of the FFA ensemble predictions for the 1-in-100-years event (far future) and is a cross-section through the flood frequency curves of Figure 5.37 at the 1-in-100-years return period. It shows that the difference in the predicted flood magnitudes between both projected scenarios increases until the 80th percentile of the FFA ensemble. At higher percentiles, this difference decreases, particularly above the 95th percentile. The scenario differences at P_{99} are 3,200 m^3/s (SSP245: 37,600 m^3/s, SSP585: 40,800 m^3/s) and hence very small. Therefore, almost similar results are predicted for the most extreme ensemble predictions of SSP245 and SSP585 for the 1-in-100-years event in the far future.

The flood frequencies are the counterpart of the magnitudes and therefore decrease with time and emissions (Table 5.11). The discharge of the baseline 1-in-100-year flood event (11,600 m^3/s) is projected to be exceeded every 21 years for SSP245 and 16 years for SSP585 in the near future (median predictions). The return periods are projected to decrease further in the far future and the baseline 1-in-100-year magnitude is exceeded every 11 years (SSP245) and 3 years (SSP585) (median predictions). When also including the $P_{2.5}$ and $P_{97.5}$ predictions, the baseline 1-in-100-year magnitude is exceeded every 21 – 34 years (SSP245) and 11 – 23 years (SSP585) in the near future, and every 11- 16 years (SSP245) and 2 – 7 years (SSP585) in the far future.

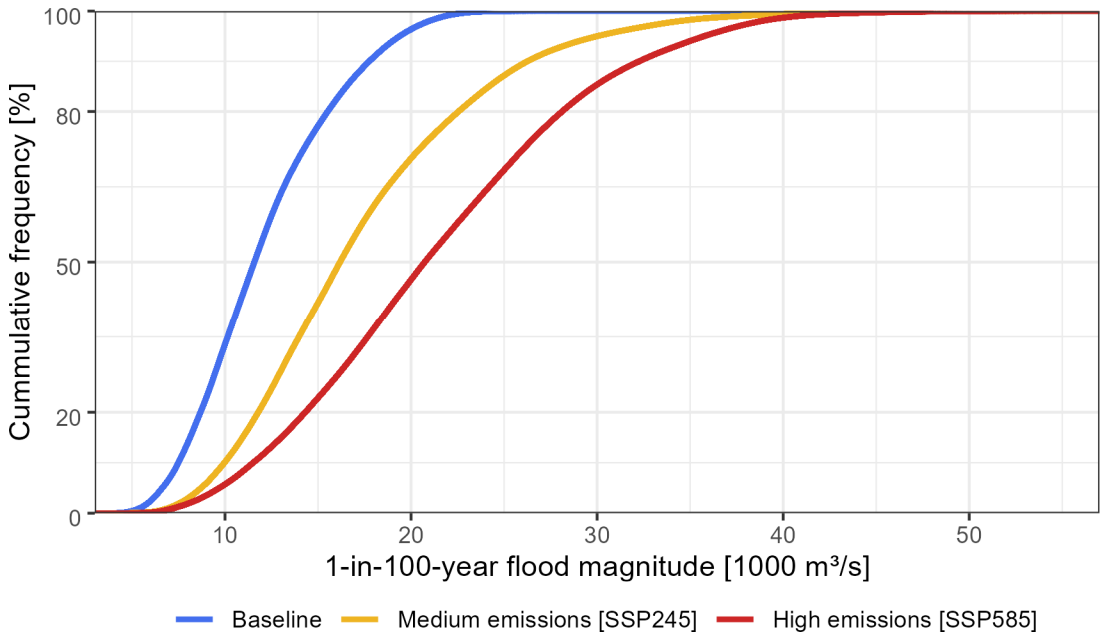


Figure 5.38: Cumulative frequency curves of the flood frequency analysis ensemble predictions for the 1-in-100-years flood event in the far future. This figure shows a cross-section of the flood frequency curves of the 1-in-100-years event shown in Figure 5.37.

Table 5.11: The projected return periods of the baseline 1-in-100-years flood magnitudes for different percentiles of the CMIP6 ensemble predictions.

	$P_{2.5}$	**P_{50}**	**$P_{97.5}$**
Baseline (1975 – 2014)	100	100	100
SSP245 (2020 - 2059)	34	21	29
SSP585 (2020 - 2059)	23	16	11
SSP245 (2060 - 2099)	16	11	11
SSP585 (2060 - 2099)	7	3	2

The FFA of the CMIP6 ensemble predicts increases in the flood magnitudes and frequencies which scale with time and emissions. Furthermore, the steepness of the flood frequency curves increases with the percentile of the ensemble predictions. This leads to an uncertainty interval which increases with time, emissions, and the return period.

5.4.4.3 The flood magnitudes of the CMIP6 ensemble members

The results of the Flood Frequency Analysis (FFA) of the CMIP6-ensemble showed two main characteristics; i) the increase of flood magnitudes with time and emissions and; ii) the increasing uncertainty (indicated by the increasing prediction range) with time, emissions, and return period. Both these characteristics dominate the FFA of the CMIP6-ensemble members.

The CMIP6-ensemble members predict the increase of Flood Magnitudes (FM) with time and emissions. Furthermore, the differences between the ensemble members increase with time and emissions which is one reason for the increasing uncertainty. These increasing differences are illustrated by the flood frequency curves (Figure 5.39) and the predicted flood magnitudes (Tables 5.12 – 5.14) of the CMIP6 ensemble members.

In the baseline scenario, the median 1-in-100-years FM ranges between 8,400 m^3/s (INM-CM4-8) and 15,700 m^3/s (INM-CM5-0). The magnitudes of both members have a considerable distance to the next closest members which are BCC-CMS 2-MR (9,500 m^3/s) and ACCESS-ESM1-5 (14,000 m^3/s) (Table 5.12). This prediction range increases for the projected climates. In the far future of SSP245, the lowest median 1-in-100-years magnitude is 11,100 m^3/s predicted by BCC-CM2-MR, and the highest one is 22,900 m^3/s predicted by NorESM2-LM (Table 5.13). For SSP585, the projected FM ranges between 11,500 m^3/s (MPI-ESM1-2-LR) and 28,100 m^3/s (EC-Earth3-Veg), whereas the MPI-ESM1-2-LR is considerably lower than the second lowest prediction of INM-CM4-8 (14,400 m^3/s) (Table 5.14). Therefore, the difference in the median magnitudes of the CMIP6 members increases from 7,300 m^3/s in the baseline to 11,800 m^3/s (SSP245) and 16,600 m^3/s (SSP585) in the far future. This indicates that the difference between the climate projections of the CMIP6 members increases with time and emissions, which translates into an increasingly uncertain response of the flood hazards to the projected climates and explains the increasing uncertainty interval with time and emissions.

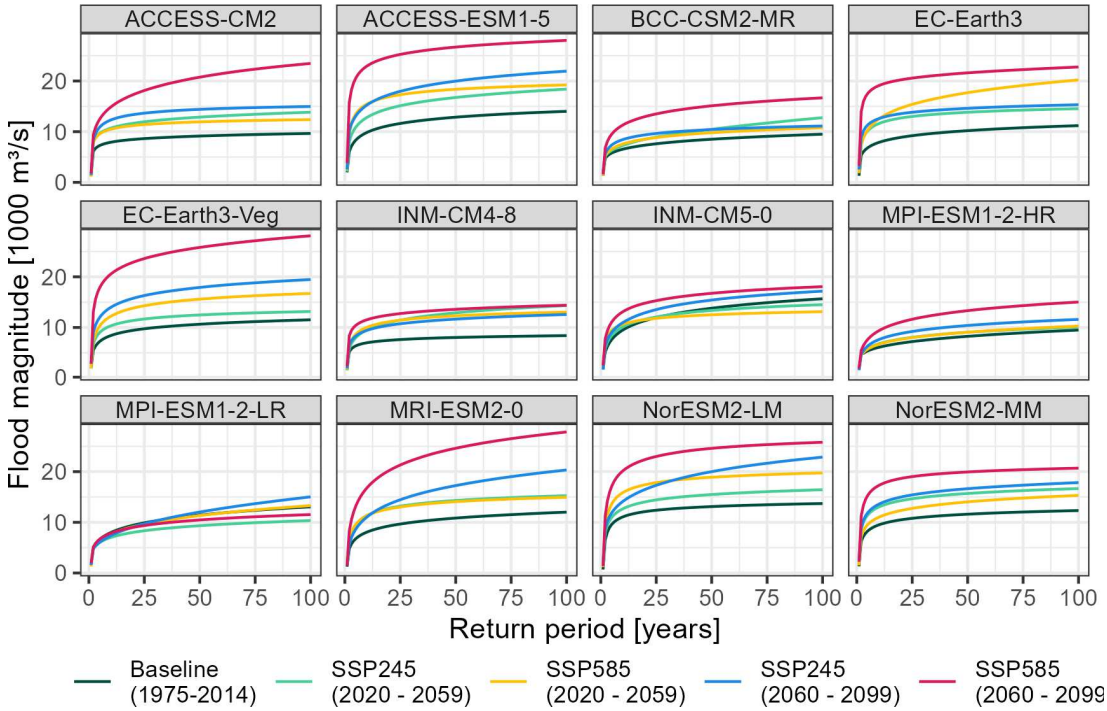


Figure 5.39: Median flood frequency curves of the CMIP6-ensemble members. The flood frequency curves of the 2.5th and 97.5th percentiles are included in the Appendices 5.6 and 5.7.

Table 5.12: Flood magnitudes of the 1-in-100-years event for each CMIP6-ensemble member for the baseline scenario (1975 - 2014).

CMIP6 member	**Flood magnitude [m³/s]**		
	Median	$P_{2.5}$	$P_{97.5}$
ACCESS-CM2	9.7	7.0	12.8
ACCESS-ESM1-5	14.0	8.7	21.5
BCC-CSM2-MR	9.5	5.2	15.2
EC-Earth3	11.2	6.6	18.0
EC-Earth3-Veg	11.5	6.6	18.5
INM-CM4-8	8.4	5.2	12.9
INM-CM5-0	15.7	9.7	23.0
MPI-ESM1-2-HR	9.5	5.1	18.3
MPI-ESM1-2-LR	13.1	7.1	20.9
MRI-ESM2-0	12.0	7.1	20.4
NorESM2-LM	13.7	9.1	21.1
NorESM2-MM	12.3	7.7	18.9

Table 5.13: Flood magnitudes of the 1-in-100-years event for each CMIP6-ensemble member for the SSP245 scenario (2060 - 2099).

CMIP6 member	Flood magnitude [1000 m^3/s]			Change to baseline [%]		
	Median	$P_{2.5}$	$P_{97.5}$	Median	$P_{2.5}$	$P_{97.5}$
ACCESS-CM2	15.0	10.1	22.7	55	44	78
ACCESS-ESM1-5	21.9	12.8	35.3	57	47	64
BCC-CSM2-MR	11.1	6.8	17.4	17	32	14
EC-Earth3	15.3	9.9	23.3	37	51	29
EC-Earth3-Veg	19.5	11.7	28.2	69	78	52
INM-CM4-8	12.6	7.7	17.9	50	46	39
INM-CM5-0	17.2	9.3	32.8	10	-3	43
MPI-ESM1-2-HR	11.6	6.3	22.5	22	23	23
MPI-ESM1-2-LR	15.0	6.9	27.9	15	-3	33
MRI-ESM2-0	20.3	10.1	38.0	69	43	86
NorESM2-LM	22.9	12.1	42.9	67	33	103
NorESM2-MM	17.8	10.5	27.4	45	36	45

Table 5.14: Flood magnitudes of the 1-in-100-years event for each CMIP6-ensemble member for the SSP585 scenario (2060 - 2099).

CMIP6 member	Flood magnitude [1000 m^3/s]			Change to baseline [%]		
	Median	$P_{2.5}$	$P_{97.5}$	Median	$P_{2.5}$	$P_{97.5}$
ACCESS-CM2	23.5	14.1	33.0	143	102	158
ACCESS-ESM1-5	28.0	17.9	38.8	100	106	80
BCC-CSM2-MR	16.7	8.7	27.5	76	67	80
EC-Earth3	22.8	14.5	31.6	103	120	76
EC-Earth3-Veg	28.1	17.0	45.7	144	159	147
INM-CM4-8	14.4	9.7	19.8	71	86	53
INM-CM5-0	18.1	9.8	30.2	15	2	31
MPI-ESM1-2-HR	15.0	6.9	27.4	58	34	50
MPI-ESM1-2-LR	11.5	6.4	20.7	-12	-10	-1
MRI-ESM2-0	27.8	14.8	41.0	132	109	101
NorESM2-LM	25.8	15.0	39.7	88	66	88
NorESM2-MM	20.7	13.5	28.8	68	76	52

The uncertainty of the projected flood magnitudes is, aside from the climate models, caused by the hydrological model and the Flood Frequency Analysis (FFA). The comparison of the prediction intervals of the FM of the CMIP6 ensemble, the FM of the CMIP6 members, and the AMAX percentiles enables the estimation of the impact of the uncertainty sources (Table 5.15). The AMAX percentiles provide information about the uncertainty arising from the ensemble of the hydrological models, the FM of the CMIP6 members provides information about the uncertainty of the FFA, and the CMIP6 ensemble provides information about the uncertainty in the climate projections. The AMAX P_{90} has a similar frequency as the 1-in-10-years event, and the AMAX P_{95} has a similar frequency as the 1-in-20-years event.

The mean (CMIP6-ensemble members) prediction interval of the AMAX percentile increases with the flood discharge. For example, the interval increases from 5,800 m^3/s of the AMAX P_{90} to 8,400 m^3/s of the AMAX P_{100} in the baseline period (Table 5.15). Since the flood discharge increases with time and emissions, the prediction interval increases alongside to 14,300 m^3/s for the AMAX P_{100} projected in the far future of SSP585. Since these are the means of the CMIP6 ensemble members the increasing prediction range indicates that the uncertainty introduced by the hydrological model increases with time and emissions.

The uncertainty introduced by the FFA is added to the hydrological uncertainty. This FFA uncertainty can be estimated by comparing the mean prediction ranges of the AMAX percentiles and the flood magnitudes of respective frequencies. This difference between the 1-in-10-years magnitude and the AMAX P_{90} prediction intervals ranges between 1,100 – 2,000 m^3/s and increases to 1,200 – 2,800 m^3/s for the 1-in-20-years magnitude and the AMAX P_{95}. These differences are low in comparison to the mean prediction range of the AMAX percentiles which indicates that a large fraction of the uncertainty in the FFA originates from the hydrological ensemble. However, the mean prediction interval of the FM consistently increases with the return period, e.g. from 6,900 m^3/s (1-in-10-years event) to 11,400 m^3/s (1-in-100-years event) in the baseline period (Table 5.15). This indicates that the uncertainty of the FFA increases with the return period of the event and hence the prediction of rare events is more uncertain than for more common events.

The differences between the mean FM prediction intervals of the CMIP6 ensemble and the CMIP6 members provide information about the uncertainty that arises from the variation in the climate projections. This difference increases from 2,300 m^3/s (9,200 m^3/s – 6,900 m^3/s) in the baseline to 9,500 m^3/s (22,300 m^3/s – 12,800 m^3/s) in the far future of SSP585 for the 1-in-10-years events (Table 5.15). This indicates that climate uncertainty increases with time and emissions and is one of the main uncertainty sources in the far future.

This comparison of the prediction intervals of the AMAX percentiles, the ensemble members, and the ensemble indicates that the main sources of the uncertainty are the hydrological modelling, and the climate projections. The FFA uncertainty is caused by the short record length mainly affects the extreme 1-in-100-years magnitudes. However, this comparison is an approximation of the uncertainty sources and not an exact quantification.

Table 5.15: The mean prediction range of the flood magnitudes of the CMIP6-ensemble (left), the CMIP6-members (centre), and the mean prediction range of AMAX percentiles of the CMIP6-members. The prediction range is the difference of the predicted values at the 2.5th and 97.5th percentiles of the ensembles. The AMAX percentiles (right) refer to the rank of the AMAX event. The P_{100} is the highest simulated AMAX event, the P_{95} is the 2nd highest event which is statistically exceeded once every 20 years, and the P_{90} is the 4th highest event which is statistically exceeded once every 10 years.

	Prediction range of the flood magnitudes of the CMIP6-ensemble [1000 m^3/s]			**Mean prediction range of the flood magnitudes of the CMIP6-members [1000 m^3/s]**			**Mean prediction range of the AMAX percentiles of the CMIP6-members [1000 m^3/s]**		
	10 years	20 years	100 years	10 years	20 years	100 years	P_{90}	P_{95}	P_{100}
Baseline (1975 - 2014)	9.2	10.6	14.6	6.9	8	11.4	5.8	6.8	8.4
SSP245 (2020 - 2059)	12.2	13.8	18.3	8.5	9.9	14	7.1	8.3	10.2
SSP585 (2020 - 2059)	15	16.5	20	9	10.3	13.8	7.8	9	10.4
SSP245 (2060 - 2099)	14.8	17	25.7	9.8	12	18.5	8.3	9.2	13.5
SSP585 (2060 - 2099)	22.3	24.4	29.7	12.8	14.6	19.6	10.8	12.2	14.3

The uncertainty of the CMIP6 climate projections is indicated by the member variation of the projected changes. The CMIP6 ensemble pronhects and increase with time and emissions and this is, to a varying degree, reflected by the ensemble members but a different behaviour is predicted for several members (Figure 5.40). The median 1-in-100-years flood magnitude is projected to increase by 23% for the CMIP6-ensemble in the near future (2020 – 2059) of the medium emission scenario SSP245. However, the MPI-ESM1-2-LR member predicts a decrease of this magnitude by 21%. This is considerably lower than the second-lowest projection of INM-CM5-0 with an increase of 7%. The largest changes are projected for INM-CM4-8 (+70%) and, with a large distance ACCESS-CM2 (+44%). The ensemble predicts an increase in the median flood magnitude with a 100-year return period of 26% for SSP585 (near future). The INM-CM5-0 contradicts this behaviour and is the only member projecting a decrease of this magnitude (-16%). The largest changes are projected for EC-Earth-3 (+81%) and, with a large distance, INM-CM4-8 (+55%). Several members project a higher increase for SSP245 than for SSP585 in the near future. These models are ACCESS-CM2, INM-CM4-8, BCC-CSM2-MR, NorESM2-MM, and INM-CM5-0 (Figure 5.40).

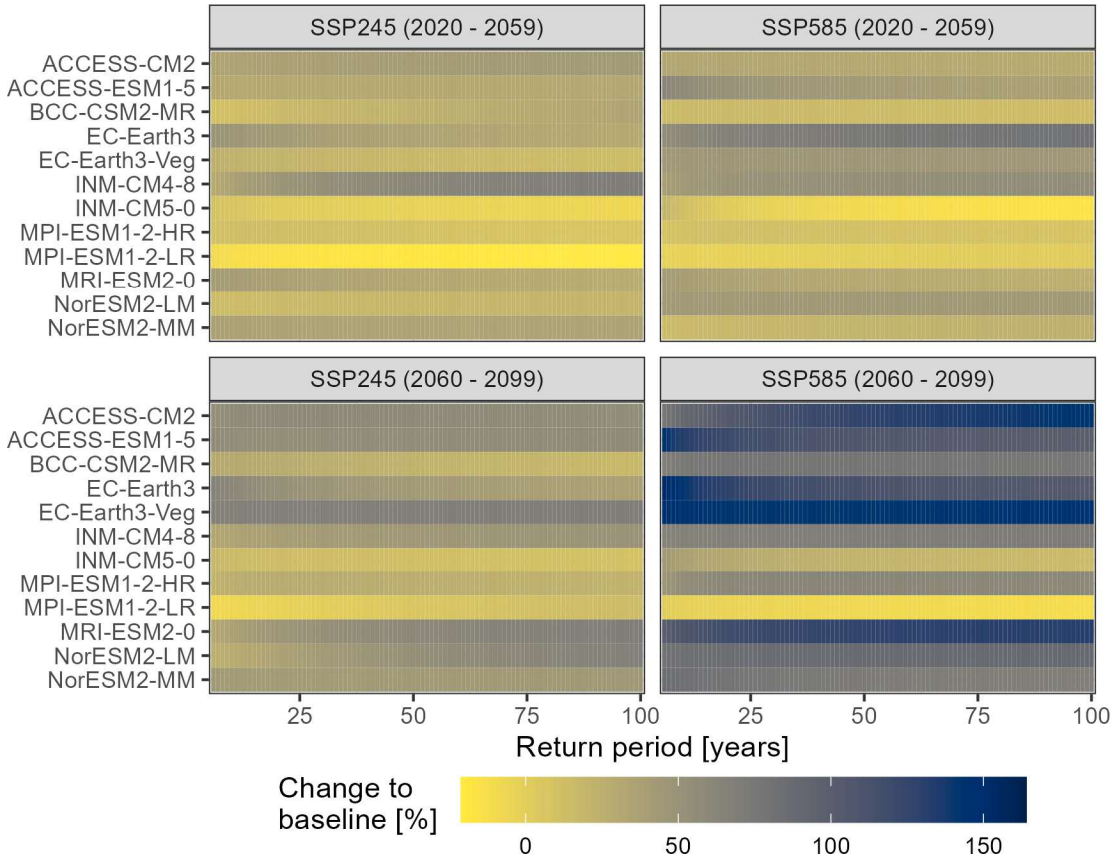


Figure 5.40: Change of the projected flood magnitude to the baseline of each CMIP6-ensemble member (median predictions).

In the far future (2060 – 2099), only the MPI-ESM1-5-LR model predicts a higher increase for SSP245 (+15%) than for SSP585 (-21%) (Figure 5.40). The lowest change for the SSP245 scenario is projected by the INM-CM5-0 member (+10%), and the highest one for the EC-Earth3-Veg and MRI-ESM2-0 members (+69%). For the SSP585 scenario, there is a large difference between the projections of the two members predicting the lowest change which are MPI-ESM1-5-LR (-12%) and INM-CM5-0 (+15%). The largest increases are projected for EC-Earth3-Veg and ACCESS-CM2 (+144 and +143%) (median predictions).

The MPI-ESM-1-5-LR member stands out from the ensemble by consistently predicting low changes and, thus a small response of flood hazards to climate change. This is, to a lower degree also valid for the INM-CM5-0 model which predicts the highest flood magnitudes in the baseline period. However, the members which predict high increases in flood discharge (Section 5.4.3.2) predict high increases in flood magnitudes.

5.5 Discussion of the projected flood hazards

The prediction of future flood hazards was conducted by applying the hydrological model with an ensemble of climate projections from 12 downscaled and bias-corrected CMIP6 models within a parameter uncertainty estimation framework. The projected flood discharge was then analysed with a Flood Frequency Analysis (FFA) to quantify changes in the flood magnitudes and frequencies. Each stage of the modelling process introduces uncertainty which propagates through the modelling cascade, in particular the transfer of hydrological parameters from the observed hydro-meteorological conditions in Chapter 4 to the CMIP6 climate simulations, the transfer of these hydrological parameters between the past and projected climates, the uncertainty of the projections of the future climates, and the statistical analysis of the projected floods.

The first Section 5.5.1 discusses the performance of the hydrological model by comparing the model behaviour of the baseline period with the behaviour of the historical simulations of the observed hydro-meteorological conditions. Section 5.5.2 discusses the projected changes in the hydrological system of the projected climate scenarios. These two sections provide information about the robustness of the model predictions and are the foundation for the discussion of the projected flood flows in 5.5.3.

5.5.1 The effects of the parameter transfer to the climate projections

The hydrological model was calibrated from the observed hydro-meteorological conditions and 64 parameter sets were identified which replicate the observed catchment behaviour (Chapter 4). This hydrological ensemble was applied to the projected climate of an ensemble of 12 CMIP6 models for the baseline scenario, and two projected emission scenarios. This is an accepted and robust modelling strategy for assessing the climate change impact on the catchment hydrology in the region (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014; Nepal *et al.*, 2014; Bhattarai *et al.*, 2018; Chandel and Ghosh, 2021). However, this approach introduces additional uncertainty because parameters are transferred between datasets with individual biases.

The performance of the hydrological simulations regarding the flow seasonality and the runoff composition was discussed for the historical simulations in Chapter 4.5. The behaviour of these simulations reflects our knowledge of Central Himalayan catchment hydrology. Since it is established that the simulations are a behavioural representation of the hydrology of the Karnali River system, this section focuses on the differences

between the baseline and historical simulations. The two main differences are the reduced rainfall-runoff which translates into lower annual discharge predictions due to its dominating role on catchment hydrology, and the increased glacier melt runoff and the prolonged glacier melt period which partly compensates for the reduced rainfall-runoff in the annual budget.

5.5.1.1 The differences in the rainfall-runoff

The median annual rainfall-runoff of the baseline simulations (15 Billion m^3/a) is significantly lower than the historic ones (27 Billion m^3/a) (Figure 5.13). Most of this deficit is generated during the monsoon season, leading to lower predicted monsoon season runoff in the baseline simulations. These large difference in rainfall-runoff between both sets of simulations is systematic and related to the removal of the precipitation correction factor in the baseline simulations. This factor increases the precipitation by 35 – 42% (ensemble median: 40%) in the historical simulations to compensate for a negative bias in the precipitation data. The rainfall-runoff difference between both simulations is 57% and hence it is likely that a large fraction of the lower baseline rainfall-runoff is caused by the removal of the precipitation correction factor. The lower rainfall-runoff is the main cause of a deficit in the annual discharge of the baseline simulations. This suggests that the precipitation correction factor could have transferred from the historical data to the climate scenarios as also conducted by Lutz *et al.* (2014). However, it was not transferred for three reasons: i) the CMIP6 dataset is a regional, downscaled product that is already bias-corrected (Mishra *et al.*, 2020); ii) it is unlikely that each CMIP6 member inherits a similar bias which coincides with the bias of the observed precipitation data and; iii) the climate change impact assessment implies the non-stationarity of the climate which raises the question whether the bias is stationary, and thus whether a stationary precipitation correction factor is applicable (Bérubé *et al.*, 2022; Majone *et al.*, 2022).

The lower rainfall-runoff might further be affected by parameter interactions. While the transfer of the bias-correction factor to the CMIP6 data would have introduced uncertainty, its removal also introduces uncertainty because the other parameters were calibrated in combination with it (Odoni, 2007). The median annual baseflow is 11 Billion m^3 and is similar in the baseline and historical simulations which means, in conjunction with the lower total discharge, that a larger fraction of the rainfall is converted to baseflow in the baseline simulations. The sensitivity analysis of the model parameter in Chapter 4 showed the sensitivity of the rainfall-runoff and baseflow allocation, particularly for the

AlphaInf parameter (fraction of daily rainfall during the hour of most intense rainfall) which simulates the infiltration-excess runoff, and the *kEff* parameter which describes the saturated hydraulic conductivity. These parameters were calibrated in conjunction with the precipitation correction factor and the removal of the correction factor alters their behaviour leading to a larger baseflow contribution. However, these parameters were also calibrated for the historical data and the parameter transfer to the CMIP6 data also changes their behaviour. This type of uncertainty is inherent in the research design.

This different allocation between rainfall-runoff and baseflow raises the question of whether the baseline simulations are less behavioural representations of the hydrological system. This question is difficult to address because the true runoff composition is unknown, and every quantification is based on hydrological modelling. The comparison of the contributions with other studies is hampered by differences in the catchment properties, climate data, study periods and simulation lengths, and model structures. However, the median annual rainfall-runoff composition (40%) is lower than the one simulated for the Upper Ganga Basin (66%) to which the Karnali River contributes (Lutz *et al.*, 2014). On the other hand, the allocation of baseflow and rainfall-runoff is not well understood for Central Himalayan catchments and the simulated combined contribution of rainfall-runoff and baseflow 68% is similar to the 66% estimated for the Dudh Koshi basin (Nepal *et al.*, 2014). The high correlation of the monthly discharge and the monthly runoff contributions between the historical and baseline simulations indicates that the model behaviour has not shifted which would limit the predictive capability of the baseline simulations (Figure 5.15).

5.5.1.2 The differences in the glacier melt runoff

The second difference between the baseline and historic simulations is the increased glacier melt runoff. This component contributes (median) 4% to the annual runoff in the historical simulations but 20% in the baseline ones. While this increase is also caused by the lower rainfall-runoff, the glacier melt also increases strongly in absolute quantities from 1.6 Billion m^3/a in the historical simulations to 7.6 Billion m^3/a in the baseline ones (Figure 5.13). The importance of glacier melt for the annual budget in the region is strongly affected by the glaciated area of the catchment. The Ganga which comprises Central Himalayan catchments but also large parts of the Indo-Gangetic plain is supplied to 4% by glacier melt (Immerzeel *et al.*, 2010). This contribution increases to 11.5% for the Upper Ganga Basin which comprises a lower fraction of the plain (Lutz *et al.*, 2014). The contribution further increases for high mountainous subbasins and in the Dudh

Koshi River and the Langtang River, glacier melt contributes 17% and 50% respectively (Immerzeel *et al.*, 2013; Nepal *et al.*, 2014). The Karnali River is located, in terms of the glaciated area, somewhat between the Langtang River and the Upper Ganga River and hence the 20% contribution is not out of the boundary of predicted contributions in the region.

The glacier melting season increases in the baseline simulations and is longer than for other regional studies. The glacier melt season lasts from June to September in the historical simulations and increases from April to November in the baseline simulations. Other studies predicted melt season lasting from May – October and June – September for the Langtang River (Immerzeel *et al.*, 2012, 2013), May – September for the Dudh Koshi (Nepal *et al.*, 2014), and June – October for the Upper Gahagara River to which the Karnali contributes (Lutz *et al.*, 2014). Furthermore, low glacier melt runoff is also predicted throughout the winter which is not predicted in any other study and which contradicts the seasonality with glacier growth during the winter and glacier melt during the summer. This indicates that the glacier melt runoff is overestimated and that the performance of the glacier module decreases in the baseline simulations.

The glacier melt is temperature dependent and hence the shift in the glacier representation results from temperature differences between the observed temperature data and the CMIP6 baseline temperature data. The glacier processes are controlled by temperature-dependent melt parameters (degree-day factors). Temperature differences between the datasets change the behaviour of these parameters and lead to an increase in the melt rates.

The variation in the predicted melt rates between the 12 CMIP6 members is small (Figure 5.14). This consistent behaviour for 12 different climate models alongside the vast increase from the historical to the baseline simulations provides evidence for a systematic difference between the datasets which may not only result from the parameter transfer. One plausible explanation is that the empirical lapse rate (Chapter 4) which was used for the regionalisation is not transferable to the CMIP6 data and leads to a warm bias in the high mountainous parts of the catchment. However, the uncertainty in the precipitation projections is larger than the uncertainty in the temperature projections. Hence the lower variation in the temperature-driven glacier melt than in the other, precipitation-driven components may also result from the climate model uncertainty (Wu *et al.*, 2022).

5.5.2 The projected changes in the hydrological system

The calibrated parameters are projected onto the future climate. Thus, the modelling strategy is based on the assumption of stationarity because the model is calibrated to climatic conditions that are no longer valid for the projected climates (Jehanzaib *et al.*, 2020; Bérubé *et al.*, 2022). This problem is more pronounced for precipitation than for temperature because the internal variability (the natural variation in precipitation) is larger, the response to anthropogenic forcing is later, and the uncertainty in the climate projections is larger (Giuntoli *et al.*, 2018; Martel *et al.*, 2018; Bérubé *et al.*, 2022; Wu *et al.*, 2022). Thus, the use of global parameters for past and projected climates introduces uncertainty, and this uncertainty might vary between CMIP6-ensemble members due to different degrees of non-stationarity. However, this uncertainty resulting from the non-stationarity is smaller than the structural uncertainty of the hydrological model (Poulin *et al.*, 2011; Bérubé *et al.*, 2022).

The next sections discuss the projected hydrological systems to investigate the drivers of changes in flood hazards and to gain an understanding of the reliability of the simulations.

5.5.2.1 Annual discharge

The annual discharge is projected to increase with time and emissions in comparison with the baseline reference scenario (1975 – 2014). This projected increase is consistent with our knowledge of the climate and hydrological systems and agrees with regional observations. The water vapour content of the atmosphere increases with rising temperatures and hence wet seasons become wetter (Douville *et al.*, 2021). The Himalayas have warmed by 0.2 ˚C per decade since the 1970s and increasing annual precipitation has been observed since the 1990s (Ren *et al.*, 2017; Krishnan *et al.*, 2019). The annual and monsoon precipitation is projected to increase with the temperature in the Himalayas (Krishnan *et al.*, 2019; Douville *et al.*, 2021). It is, thus, reasonable that the annual discharge rates are projected to increase as the monsoon season dominates the hydrological regime of the Karnali River. The projected increases scale with time and emissions and hence this projected evolution correlates with the temperature projections (Figure 5.2).

The differences between the medium and high emission scenarios increase in time, and similar trends are predicted in the near future (2020 – 2059) (Figure 5.22 and Table 5.3). This similar annual discharge can be attributed to similar precipitation projections in the first decade of the period as indicated by the similar rainfall-runoff predictions. Furthermore, the internal variability of precipitation is large and is projected to increase for the South Asian monsoon precipitation (Martel *et al.*, 2018; Douville *et al.*, 2021; Bérubé *et al.*, 2022). Therefore, the precipitation response to climate change, and thus runoff response, is overlayed by the internal variability in the near future and manifests from the 2040s (Massei *et al.*, 2020; Douville *et al.*, 2021). The annual discharge increases by 24% (SSP245) and 56% (SSP585) in the far future (2060 – 2099) and hence the streamflow response to the higher emissions intensifes with time (Figure 5.22 and Table 5.3). This behaviour of increasing discharge, and increasing scenario differences is consistent with global and South Asian projections (Douville *et al.*, 2021) and agrees with assessments for single Central Himalayan watersheds (Immerzeel *et al.*, 2013; Bajracharya *et al.*, 2018; Dahal *et al.*, 2020). Furthermore, the projected change increases with the percentile of the ensemble predictions which is consistent with global streamflow projections and the ones for the South Asian monsoon discharge (Douville *et al.*, 2021).

The simulated changes fall within the range of projected changes in the region, but the differences in catchment characteristics and assessed periods hamper the direct comparison. Previous studies have used the RCP4.5 and RCP8.5 scenarios which are the predecessors of the SSP245 and SSP585 scenarios, respectively. The global mean discharge is projected to increase by 8% (RCP4.5) and 19% (RCP8.5) in the period 2081 – 2100 (Douville *et al.*, 2021). The projected increase of the South Asian monsoon discharge is higher and reaches 11% and 30% for these scenarios, respectively (Douville *et al.*, 2021).

The Himalayas are more sensitive to climate change and hence the higher increase compared to global and regional studies that comprise larger fractions of lowlands is reasonable (Krishnan *et al.*, 2019). The projected response of Central Himalayan Rivers varies between the studies and the projected change of this study falls within the reported range. Bharati *et al.* (2014) project a 2 – 4% increase in the annual discharge until 2050 for the Koshi River. In a climate change impact assessment of the Karnali River, the annual discharge increased by 8% (RCP4.5) and 11% (RCP8.5) in the period 2071 – 2100 (Dahal *et al.*, 2020). The annual discharge of the Upper Ganga River Basin to which the Karnali contributes is projected to increase by 6% for RCP4.5 and 14% for

RCP8.5 (mean), and by 10% for RCP4.5 and 27% for RCP8.5 (maximum) by 2041 – 2050 (Lutz *et al.*, 2014).

Other regional studies project significantly higher annual runoff increases. The discharge of the small and high mountainous Langtang River (160 km^2, 3,800 – 7,300 masl) is projected to increase by up to 88% at the end of the century (Immerzeel *et al.*, 2013), and the annual discharge of the neighbouring, medium-sized Kaligandaki River (12,000 km^2, 200 – 8,100 masl) is projected to increase by ~40% (RCP4.5) and ~60% (RCP8.5) in the 2090s (Bajracharya *et al.*, 2018).

5.5.2.2 Runoff seasonality

The projected increase of the annual runoff is driven by increased rainfall which drains rapidly as rainfall-runoff or drains delayed as baseflow. Consequently, the monsoon and post-monsoon season discharge gains further importance to the annual budget (Figure 5.25). This agrees with our current understanding of the climate change impact on the South Asian monsoon system with projected increases in the monsoon precipitation and projected increases in the runoff seasonality (Douville *et al.*, 2021). The projected increases of rainfall-runoff and baseflow in the Karnali River Basin fall in line with the projections for other Central Himalayan catchments (Immerzeel *et al.*, 2013; Bharati *et al.*, 2014; Bajracharya *et al.*, 2018; Bhattarai *et al.*, 2018) and the large South Asian River systems Ganga and Brahmaputra (Lutz *et al.*, 2014; Scott *et al.*, 2019).

The CMIP6 climate projections provide some evidence for the lengthening of the Indian Summer Monsoon at the end of the century, particularly for the SSP245 scenario (Douville *et al.*, 2021). The increase in the post-monsoon season discharge is dominated by higher baseflow and the rainfall-runoff only contributes a small fraction to the projected increase (Figure 5.25). Therefore, this study does not provide evidence for a prolonging of the monsoon season in the Karnali River catchment. However, it could be that the signal of the lengthening is lost by the monthly aggregation or by averaging the projections over 40 years.

The discharge in the dry winter and pre-monsoon season is projected to increase but to a lower degree than the monsoon and post-monsoon season discharge (Figure 5.25). The rainfall-runoff remains stable and hence there is no evidence of a change in the westerly precipitation during the winter. Instead, the projected changes are caused by higher baseflow, a delayed response to the higher monsoon season precipitation, and changes in snowmelt and glacier melt runoff. The snowmelt runoff is projected to

decrease in the annual budget but the melting season shifts to earlier dates and hence an increase in snowmelt runoff is projected for some months. These projections complement regional climate change assessments which project that the snow cover duration decreases, the snowmelt shifts to an earlier date, and snowmelt contributions decrease (Immerzeel *et al.*, 2013; Bolch *et al.*, 2019; Douville *et al.*, 2021; Chandel and Ghosh, 2021; Khanal *et al.*, 2021).

The implications of these projected snowmelt changes (seasonality and melt rates) for the total discharge are ambiguous. Some studies suggest that the lower snowmelt causes a decrease in the pre-monsoon season discharge of Ganga River tributaries (Bharati *et al.*, 2014; Scott *et al.*, 2019). Other studies suggest that the reduced snowmelt is overcompensated by higher baseflow (Immerzeel *et al.*, 2013), or that the earlier melting season causes higher pre-monsoon discharge (Bajracharya *et al.*, 2018; Dahal *et al.*, 2020; Chandel and Ghosh, 2021).

The glacier melt runoff is projected to increase with time and emissions, but the relative contribution remains stable due to the projected increases in rainfall-runoff (Table 5.3). The large difference between the simulated melt rates of the historical and baseline simulations, and the prediction of glacier melt throughout the winter indicates errors in the representation of glacier processes. Furthermore, the projected linear (SSP245) and exponential (SSP585) increases of glacier melt with time contradicts regional projections. Himalayan glaciers have lost mass as a result of the warming climate with increased loss rates from the 2000s (Bolch *et al.*, 2019; Masson-Delmotte *et al.*, 2021). This trend is projected to continue with glacier volumes decreasing by 33 – 66% (RCP1.26 – RCP8.5) at the end of the century (Bolch *et al.*, 2019). The increased glacier melt results in an increased glacier melt runoff in the near future (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014). Therefore, the projected near future increase in glacier melt is somewhat reasonable.

The projections of the glacier melt in the far future (2060 – 2099) contradict the projections of other regional studies. These studies project decreasing glacier melt runoff in the second half of the century which is caused by the reduction of glacier storage (Immerzeel *et al.*, 2013; Shea and Immerzeel, 2016; Nepal, 2016; Chandel and Ghosh, 2021; Khanal *et al.*, 2021). This is not replicated by the model which projects further increases of the glacier melt runoff. This projected increase is contradictory because the overestimated length of the melt season of the baseline scenario indicates the overestimation of glacier melt and hence the effect of the glacier shrinkage on melt runoff should be more pronounced than in other studies. One reason for the absence of this response could be the overestimation of the glacier volume from the ice thickness and

glacier outlines. The large differences between the historic and baseline simulations and the contradiction with the projected trends indicate that the glacier runoff predictions are uncertain and should be interpreted cautiously. This is particularly the case for the winter and pre-monsoon seasons with low discharge rates and higher glacier melt contributions. The total runoff could decrease during some months in the non-monsoon season which is concealed by overestimated glacier melt rates.

5.5.3 The projected flood hazards

The hydrological simulations and flood frequency analysis project an increase in flood discharge, flood magnitudes, and flood frequencies. The classification of these results is hampered by large biases of hydrological models in the reproduction of extreme flows (Huang *et al.*, 2017; Seneviratne *et al.*, 2021), the uncertainty in the precipitation projections of climate models, particularly over the Himalayas (Scott *et al.*, 2019; IPCC, 2023), and the complex response of rivers to extreme precipitation (Seneviratne *et al.*, 2021; Caretta *et al.*, 2022).

5.5.3.1 The predicted changes in the flood hazards

The flood discharge is projected to increase with time and emission. The ensemble-median long-term flood discharge (30-year mean) increases by 24% (SSP245) and 26% (SSP585) by the year 2059. This projected trend continues in SSP245 to an increase of 34% by the year 2099, whereas the trend accelerates in the high emission scenario SSP585 particularly after the year 2080 to an increase of 71% by 2099 (Figure 5.26 and Table 5.4). The flood magnitudes are projected to increase at slightly higher rates and this increase is consistent across all investigated return periods (1-in-10-years to 1-in-100-years). These increases range between 37 – 40% (SSP245) and 79 – 83% (SSP585) in the far tuture (median predictions) (Figure 5.37 and Table 5.9). Thus, the extreme discharge increases stronger than the annual discharge which increases by 24% and 48% at the end of the century (Figure 5.22).

The higher increase in the extreme flows is consistent with the precipitation over Central Himalayan catchments where the extremes are projected to increase stronger than the averages (Chhetri *et al.*, 2021; Talchabhadel, 2021). However, the increase in flood magnitudes is consistent over the range of frequencies between the 1-in-10-years to 1-in-100-years events, whereas the projected changes in extreme precipitation events are

non-linear and with larger increases for rarer events (Seneviratne *et al.*, 2021). This discrepancy could be related to the complex response of rivers to extreme rainfalls (Seneviratne *et al.*, 2021). It can be further caused by the deficits in the process representation of the hydrological model which causes a delayed drainage of a fraction of the rainfall-runoff (Chapter 4). This delayed response may increase with the intensity of the rainfall and, thus, may lead to an increasing underestimation with increasing flood magnitude.

The projected trends of the floods agree with our current understanding of the climate change impact on flood hazards. Precipitation extremes intensify with increasing temperatures which translates into increasing flood magnitudes and frequencies on the global scale (Hirabayashi *et al.*, 2013, 2021; Seneviratne *et al.*, 2021; Caretta *et al.*, 2022). However, the response of rivers to changing precipitation extremes varies because of the changes in the hydrological system (e.g. melt seasonality), and anthropogenic influences (e.g., land use, water management) and hence the regional climate change impact is more uncertain (Seneviratne *et al.*, 2021; Caretta *et al.*, 2022).

There is evidence that flood hazards in the Central Himalayas and the foreland in the Ganga Plain have intensified in the past decades (Elalem and Pal, 2015; Seneviratne *et al.*, 2021; Tellman *et al.*, 2021). However, these trends are uncertain because of the large internal variability of flood flows in conjunction with the relatively short study period, and the superimposing socio-economic factors (e.g. population growth, urbanisation, land use change). The climate change impact on floods is insufficiently studied in this region but, although the uncertainties remain large, there is a consensus that flood hazards will intensify as a result of the increasing intensity and frequency of heavy rainfall events (Lutz *et al.*, 2014; Pechlivanidis *et al.*, 2017; Wijngaard *et al.*, 2017; Scott *et al.*, 2019; Chhetri *et al.*, 2021; Ranasinghe *et al.*, 2021; Seneviratne *et al.*, 2021; Talchabhadel, 2021; Caretta *et al.*, 2022). The simulated flood hazards fit into this context and are reasonable because they are projected to increase with time and emissions, and this increase is driven to more than 90% by increasing rainfall-runoff (Figures 5.31 and 5.32).

The projected flood magnitudes and frequencies align with regional studies, but the direct comparison is difficult because these studies are applied at different spatial and temporal scales, and use different generations of climate models and different ensemble sizes. Wijngaard *et al.* (2017) project an increase in flood magnitudes (1-in-50-years event for the period 2071 - 2100) in the Upper Ganga Basin using the SPHY model and an ensemble of 8 CMIP5-GCMs. These projected increases are 80% (SSP245) and 108% (SSP585) and are higher than the projected increases predicted for the Karnali

River in this study (+39% for SSP245 and +80% for SSP585) (Figure 5.37 and Table 5.9). Pechlivanidis *et al.* (2017) project an increase of 20 – 100% of the 100-year flood magnitudes (SSP585) for the Ganga River at the end of the century from an ensemble of five hydrological models and 20 CMIP5-GCMs. Hirabayashi *et al.* (2013) project an increase in the 100-year flood magnitude of 20 – 40% (11 CMIP5-GCMs) in the Central Himalayas and the Indo-Gangetic Plain at the end of the century. Dankers *et al.* (2014) project an increase of 20 – 100% for the 30-year flood from an ensemble of 5 CMIP5-GCMs in this region. A higher increase (70 – 120%) is projected for the Brahmaputra River from 12 SRES GCMs for the A1B and A2 scenarios which are located between the SSP245 and SSP585 in terms of global temperature increases (Gain *et al*., 2011). Thus, the projected changes in flood hazard magnitudes for the Karnali River fall within the range of other regional assessments.

The flood frequencies are projected to increase in the Karnali River which agrees with other regional assessments. The baseline 100-year flood event (1975 – 2014) is projected to occur every 11 – 16 years for SSP245 and every 2 – 7 years for SSP585 at the end of the century (Table 5.11). This falls within the range frequency increases of the 1-in-100-years event to 5 – 50 years (SSP245) and 2 – 25 years (SSP585) which is projected from an ensemble of 9 CMIP6-GCMs for the Central Himalayas and the Indo-Gangetic plain at the end of the century (Hirabayashi *et al*., 2021). The vast increase in the occurrence of these rare floods is also projected by Ali *et al.* (2019) which are projected to occur up to 12 times per century if the temperature increases by 4 ºC.

The timing and duration of the flood events are not projected to change. The changes in flood duration are investigated by comparing the fraction of the flood discharge that occurs before/after the event and no shift is detected in these fractions. However, the AMAX flow is projected to increase and hence the discharge before and after the event increases if the relative flow remains stable (Figure 5.34). Therefore, it is likely that the duration that the bankfull flow is exceeded increases. A throughout investigation of flood discharge changes was not conducted because the daily resolution is too coarse to accurately quantify flood durations in the highly responsive Karnali River. The lack of an accurate estimation of the bankfull flow rate hampers the quantification of the event duration and the annual flood duration from the flow magnitude. The number of heavy precipitation events is projected to double in the period 2021 - 2050 (SSP585) (Chhetri *et al*., 2021; Talchabhadel, 2021). It is therefore likely that the number of floods and, thus, the flood days per year increase which cannot be detected from the AMAX flood classification approach.

5.5.3.2 The uncertainty in the flood hazard projections

The uncertainty of the projected flood hazards originates from the uncertainty in the climate projections, the hydrological modelling, and the Flood Frequency Analysis (FFA).

The variation in the climate projections, particularly the precipitation, is a main driver of the uncertainty of the projected flood hazards. This variation increases with time and emissions and hence the prediction of flood hazards becomes increasingly uncertain. This behaviour is reasonable because projections regarding a system tend to become more uncertain with the passage of time and the magnitude of changes that occur within the system. The climatic uncertainty is exaggerated in the Himalayas by the complex topography in conjunction with the coarse modelling resolution, and the data scarcity to identify and correct biases (Nepal *et al.*, 2014; Scott *et al.*, 2019). Precipitation processes are characterised by a more complex spatial and temporal variability and a higher sensitivity to local and regional features than temperature which amplifies the uncertainties for the projection of the rainfall-induced flood events in the Karnali River (Giorgi, 2010; Wu *et al.*, 2022).

The variation of the climate projections translates into a broad range of predicted changes in the flood magnitudes. The projected increases in the median 1-in-100-years flood magnitudes of the CMIP6 members range between +10% and +69% for the medium-emission scenario SSP245, and between -12% and +144% for the high-emission scenario SSP585 in the far future (Tables 5.13 and 5.14). This large prediction range emphasises the uncertainty in climate projections and the necessity to use probabilistic climate projections for climate change impact assessments.

The ensemble-modelling approach assumes that each ensemble member is an equally likely representation of the future climate but the simulations indicate that the members projecting higher changes may be more behavioural representations of the system. The models projecting lower changes, particularly the MPI-ESM1-2-LR and MPI-ESM1-2-HR models predict the largest fraction of non-monsoon season flood events (13 – 16%) and are, thus, inferior in capturing the monsoon seasonality of the Karnali catchment (Table 5.7). While other CMIP6 members predict > 10% non-monsoon season events, this fraction is projected to decrease while the MPI-ESM1-2-LR and MPI-ESM1-2-HR members project increases to up to 29%. Contrarily, the models which predict the highest changes are the ones that reproduce the flood seasonality best indicating that the projected higher increases are more likely than the lower ones. These members also tend to have the highest agreement with the observed distribution of monsoon and non-monsoon season discharge (Figure 5.11). The better performance of high-change

models agrees with comparisons of observed and predicted global temperature changes which are best reflected by the models which predict a higher response (Carvalho *et al.*, 2022).

The conversion of the climate forcing to streamflow using hydrological models adds another layer of uncertainty to the uncertainty in the climate projections (Krysanova *et al.*, 2017; Giuntoli *et al.*, 2018, 2018; Seneviratne *et al.*, 2021). The different structures of hydrological models influence the projected streamflow to a degree that can change the direction of the trends (i.e. decrease/increase) on the local and regional scale (Dankers *et al.*, 2014). The sensitivity of projected streamflow extremes to the choice of hydrological model is illustrated by several studies that use an ensemble of different hydrological models (Dankers *et al.*, 2014; Huang *et al.*, 2017; Giuntoli *et al.*, 2018; Krysanova *et al.*, 2017; Pechlivanidis *et al.*, 2017). The hydrological model uncertainty is high for low flows and for extremely high flows (1-in-50-years and more) and can exceed the climate uncertainty, particularly if snowmelt and glacier melt processes are important sources (Krysanova *et al.*, 2017; Huang *et al.*, 2017; Giuntoli *et al.*, 2018, 2021). The prediction range of the hydrological simulations of the Karnali catchment increases with the discharge and is, together with the climate uncertainty, the main source of uncertainty for the projection of flood flows (Table 5.15).

The comparison of the simulated and observed flood discharge indicates the underestimation of flood flows for the baseline and historical simulations. This hydrological uncertainty is caused by the model's inability to reproduce the fast runoff generation and routing of the Karnali River system to the full extent. The observed flood flows, estimated from stage observations, likely overestimate the true flood flows and hence the actual flood flow deficit is likely to be smaller than indicated by the comparison of simulated and predicted floods. This uncertainty was discussed in Chapter 4.

The hydrological model is calibrated for the historical hydro-meteorological conditions and these parameters are transferred to the CMIP6 climate data. A similar hydrological system is predicted for the historical and baseline simulations with some deviations. One difference is the lower runoff of the baseline simulations which is caused by the lower rainfall in the CMIP6 data. However, the predicted flood flow of the historical and baseline simulation is similar for the most extreme flows which are exceeded statistically once in five years or less (5.17). Therefore, the lower annual and seasonal discharge does not translate into the prediction of lower extreme flows.

The removal of the precipitation correction factor introduces uncertainty due to parameter interactions which particularly affect the distribution of rainfall into rainfall-

runoff and baseflow. These differences mostly affect the below-flood discharge because the composition of the flood flows is similar in the historical and baseflow simulations (Figures 5.19 and 5.20). The CMIP6 simulations predict non-behavioural glacier melt rates which are higher than in the historical simulations. However, the importance of the glacier melt decreases with increasing runoff and therefore, the uncertainty introduced by the parameter transfer is low.

The projected changes are reasonable and agree with other regional climate change impact assessments except for the glacier melt projections. However, the glacier melt contributions to the flood discharge are low and, particularly for the highest floods this contribution is marginal. This suggests that the uncertainty introduced by the parameter transfer is larger for the simulation of low to medium flows than for the high to extreme flows which have the largest effect on the determination of the flood magnitudes.

The Flood Frequency Analysis (FFA) introduces uncertainty to the prediction of flood magnitudes and frequencies. In particular, the difference between the true and sampled (in this case simulated) flood discharge distributions, and the extrapolation from the 40-year records to the 1-in-100-year event (Apel *et al.*, 2008; Kjeldsen *et al.*, 2014; Parkes, 2015). This uncertainty manifests in the larger prediction range of the 1-in-100-year events compared to the more frequent events (Table 5.15). However, the comparatively larger prediction range of the CMIP6-ensemble and the hydrological ensemble, and the high agreement between the simulated flood flows and the predicted flood magnitudes indicates that the uncertainty introduced by the FFA is lower than the ones introduced by the climate projections and the hydrological modelling (Table 5.15).

It is worth mentioning that this assessment focused solely on the climate change impact on flood flows. However, other anthropogenic factors alter flood hazards. For example, land use changes alter the hydrological cycle and can aggravate or reduce flood hazards (Douville *et al.*, 2021; Pearson *et al.*, 2022). This effect could be implemented in the modelling framework by simulating the climate projections with different land cover maps.

5.6 Conclusions

The potential climate change impact on the hydrology of the Karnali River catchment was assessed by forcing the calibrated hydrological modelling ensemble with climate projections of an ensemble of 12 CMIP6 climate models for the medium-emission scenario SSP245, and the high-emission scenario SSP585. It is, to the best knowledge of the author, the first climate change impact assessment of Central Himalayan catchments that uses the latest generation of CMIP6 models. Previous studies on the catchment scale used climate projections of CMIP5 models (Immerzeel *et al.*, 2013; Lutz *et al.*, 2014; Bajracharya *et al.*, 2018; Dahal *et al.*, 2020) or Hadley Centre Coupled Model HadCM3 (Nepal, 2016; Bharati *et al.*, 2014; Bhattarai *et al.*, 2018). These studies use single climate models or a subset of extreme projections (i.e., dry-cold, dry-warm, wet-cold, wet-warm) to quantify the extreme bounds of climate impact on catchment hydrology. Several studies assessed the climate change impact on flood hazards on the global or regional scale which may not capture the small-scale variability of the climatic conditions in Central Himalayan catchments (Gain *et al.*, 2011; Hirabayashi *et al.*, 2013, 2021; Dankers *et al.*, 2014; Lutz *et al.*, 2014; Pechlivanidis *et al.*, 2017; Wijngaard *et al.*, 2017).

This study complements previous studies by using probabilistic climate projections (12 models) of the latest generation of downscaled and bias-corrected climate models (CMIP6) in a hydrological model within an uncertainty framework to assess the climate change impact on flood hazards on the catchment scale. The hydrological ensemble was calibrated and validated for the observed historical catchment behaviour (Chapter 4) and the behavioural parameterisations were transferred to the climate projections. The CMIP6 baseline simulations were compared with the historical simulations (Chapter 4) to investigate whether this parameter transfer altered the predicted catchment behaviour.

The comparison of the baseline and historical simulations indicates that the hydrological predictions reproduce the hydrological system of the Karnali River well and are behavioural. However, the parameter transfer between the simulations and the removal of the precipitation correction factor in conjunction with parameter interactions adds uncertainty to the hydrological simulations. These uncertainties can potentially be reduced by calibrating the model from the climate data of the baseline ensemble. Such an approach could aim to optimise the observed and simulated flood frequency curves, to maximise the probability that flow extremes belong to a similar statistical distribution, or to weight the GCM against the discharge observations (Gain *et al.*, 2011; Majone *et*

al., 2022). The high agreement of the flow seasonality and the runoff composition, particularly for the flood flows, between both simulations indicates that this uncertainty is low compared to the uncertainty of the hydrological model and the climate projections.

The consistent underestimations of the annual, seasonal, and flood discharges of the baseline scenario indicate that the precipitation bias of climate simulations is one main challenge for climate-change impact assessments. This bias could be, potentially, determined by calibrating a precipitation correction factor for the baseline simulations. However, the model uses stationary (in time and space) parameters and assumes a constant bias. Furthermore, it is questionable that this bias is transferable to the climate projections. The precipitation in Central Himalayan Catchments is elevation-dependent and hence the spatial patterns could change with increasing temperatures (Winiger *et al.*, 2005; Lutz and Immerzeel, 2016). Therefore, the bias of precipitation projections remains a challenge for the hydrological climate change impact assessment.

The parameter transfer to the CMIP6 climate data altered the predictions of the glacier dynamics. The model overestimates the glacier melt as indicated by the simulated melt during winter and the large differences between baseline and historical predictions. Furthermore, the glacier melt increases continuously until the end of the century which contradicts a majority of studies that project a decrease of glacier melt in the second half of the century. Therefore, the glacier melt processes are not depicted well by the models which is arguably the biggest weakness of the model. However, the importance of glacier melt decreases with the flow rate and hence the implications for the projections of flood flows are marginal and do not affect the meaningfulness of the projected changes in flood hazard magnitudes and frequencies.

The hydrological simulations and the Flood Frequency Analysis (FFA) provide strong evidence for the intensification of flood hazards with time and emissions which is driven by an increase in the frequency and intensity of heavy rainfall events. The 30-year mean flood flow is projected to increase by 34% in the medium-emission scenario SSP245, and by 71% in the high-emission scenario SSP585 until the end of the century compared to the year 2014 (median projections). The flood magnitudes are projected to increase at slightly higher rates (SSP245: 37 – 40%; SSP585: 79 – 83%) for all investigated return periods (1-in-10-years to 1-in-100-years). The magnitudes of the 1-in-100-years event are projected to increase by 40% (SSP245) and 79% (SSP585) in the far future (2060 – 2099) (median projections). However, a strong increase in flood discharge is projected around the year 2080 for SSP585 which is attenuated in the 40-year record used for the FFA. Hence the flood magnitudes will have intensified more as quantified in this study.

The flood frequencies are projected to increase with the flood magnitudes. The baseline 1-in-100-year flood magnitude is projected to occur every 11-16 years (SSP245) and every 2 – 7 (SSP585) in the far future and hence flows which are currently extremely rare are projected to occur frequently in the future.

Two CMIP6 ensemble members project decreases or low increases in the flood magnitudes. These models do not capture the flood seasonality of the Karnali River and may be less likely projections. The vast majority of the members predict increases between 37 – 69% (SSP245) (8/12 CMIP6 models) and 68 – 144% (SSP585) (9/12 CMIP6 models).

This chapter has updated our knowledge about the climate change impact on the hydrology of Central Himalayan river systems to the latest generation of climate models. It has improved our understanding of how climatic changes alter the magnitude and timing of hydrological processes that characterise the hydrological regime and govern flood water generation. This has improved our understanding of the uncertainty in climatic projections and its implications for projecting future flood hazards. Furthermore, this chapter has highlighted key challenges for the application of hydrological models with probabilistic climate projections.

The high performance of the hydrological modelling ensemble of the historical simulations (Chapter 4), the high agreement between the historical and baseline simulations, and the similar model behaviour of the baseline and projected simulations indicate that the flood hazard projections are behavioural. Furthermore, the projected changes fall within the range of projections of other studies in this region which further increases the trust in the simulations. The projected flood discharge (the AMAX discharge) is used in the next chapter to simulate the morphological evolution of the Karnali fan. These projected fan topographies are combined with the projected flood magnitudes to predict the potential future flood hazards in Chapter 7.

6 The prediction of the morphological evolution of the Karnali Fan

This chapter presents Stage 3 of the modelling framework to predict the morphological evolution of the Karnali fan downstream of the mountain outlet of the Karnali catchment to address O4. This stage uses the projected flood discharge (O2) (Stage 2, Chapter 5) of the modelling framework to predict potential future fan topographies from morphodynamic modelling. These topographies will be used in Stage 4 (Chapter 7) to map the potential flood hazard characteristics from hydrodynamic modelling.

6.1 Introduction

Alluvial fans form at the intersection of the mountainous catchments and the downstream plains. These fans aggrade as they store the sediments that are discharged from the mountains but can also degrade and deliver previously stored sediments downstream. Hence, alluvial fans are dynamic landscape features which act as sediment sinks and sources (Coulthard *et al.*, 2002; Harvey, 2018). Large alluvial megafans have formed at the mountain outlets of the Central Himalayan Rivers and there is increasing evidence that the flood risk is connected to the morphodynamic behaviour of these fans because it alters the channel capacity and flow pathways (Sinha, 2008; Dixit, 2009; Sinha *et al.*, 2014; Dingle *et al.*, 2016, 2020a).

The fan topography and flood hazards are interlinked because floods are the primary driver of sediment influx to the fan and are essential for the redistribution and downstream discharge of sediments (Kleinhans *et al.*, 2013; Harvey, 2018; Dingle *et al.*, 2020b). On the other hand, the flow pathways and, thus, the flood hazard are controlled by the fan topography, the sediment characteristics, and the fan behaviour (i.e. aggradation vs degradation). The channel degradation (also referred to as channel incision) increases the channel capacity and, thus, reduces the overbank flow. It further decreases the risk of channel avulsions because it increases the elevation difference between channel and floodplain (Dingle *et al.*, 2016). Contrarily, the channel aggradation (the deposition of sediments in the riverbed) decreases the channel capacity and, thus, increases the overbank flow. It decreases the elevation difference between the channel and the floodplain which increases the risk of channel avulsions (Lane *et al.*, 2007; Sinha, 2008; Dixit, 2009; Sinha *et al.*, 2014; Dingle *et al.*, 2016). Furthermore, lateral processes (outer bank erosion and inner bank deposition) may

cause the migration of channels and, thus, alter the pathways of flood water, and potentially alter the channel capacity (e.g. channel widening).

The morphological evolution of alluvial fans is controlled by endogenic and exogenic processes. The tectonically-induced basin subsidence (endogenic) is a primary control for the evolution of fans in the region (Dingle *et al.*, 2016). This is superimposed by exogenic processes which control the sediment production (e.g. landslides, rockfalls, glacial erosion) in the upstream catchment and the delivery to the fan by the flood flows (Dingle *et al.*, 2016, 2017; Harvey, 2018; Quick *et al.*, 2019). Changes between degrading and aggrading states can be caused by changes in the base level, land use, sediment production, and flood magnitudes and frequencies (Coulthard *et al.*, 2002; Lane *et al.*, 2007; Harvey, 2018; Quick *et al.*, 2019). Furthermore, the degradation and aggradation are controlled by the fan characteristics whereas sediments are relocated from steeper to flatter sections (Coulthard *et al.*, 2002; Dingle *et al.*, 2020a). Consequently, alluvial fans are a complex system whereas processes of aggradation alternate in time and space.

Climate change potentially impacts the morphological evolution of the Karnali fan by altering external and internal processes. The sediment production is projected to increase as a result of an increase in the frequency of rainfall-induced landslides (Crozier, 2010; Gariano and Guzzetti, 2016; Wijaya *et al.*, 2023). Furthermore, permafrost melting in the high mountains decreases the slope stability and increases the frequency of rockfalls (Savi *et al.*, 2021). The projected increase in the flood discharge at the fan inlet (which is the mountain outlet) (Chapter 5) increases the sediment transport capacity and, thus, may increase the sediment delivery to the fan. This projected increase in the flood flows may also change the internal fan behaviour (degradation and aggradation) by altering the redistribution of fan sediments. However, we currently lack an understanding of the climate change impact on the sediment cascade including the sediment production, the sediment delivery, and the morphological evolution in Central Himalayan River systems.

This research applies a morphodynamic model in the Karnali fan to better understand the controls of the topographic change under consideration of the projected climate change impact on flood flows. The study area is a 545 km^2 large subsection of the complete Karnali fan that covers the gravel reaches between the mountain outlet in the north and the Gravel-Sand Transition (GST) 30 – 35 km south of the outlet (Figure 6.1). The terrace in the east of the Karnali River is excluded because it is superelevated over the channel and not at risk of channel avulsion, channel migration or flooding. The study area has been described in Section 2.3 and is hereafter referred to as the Karnali fan.

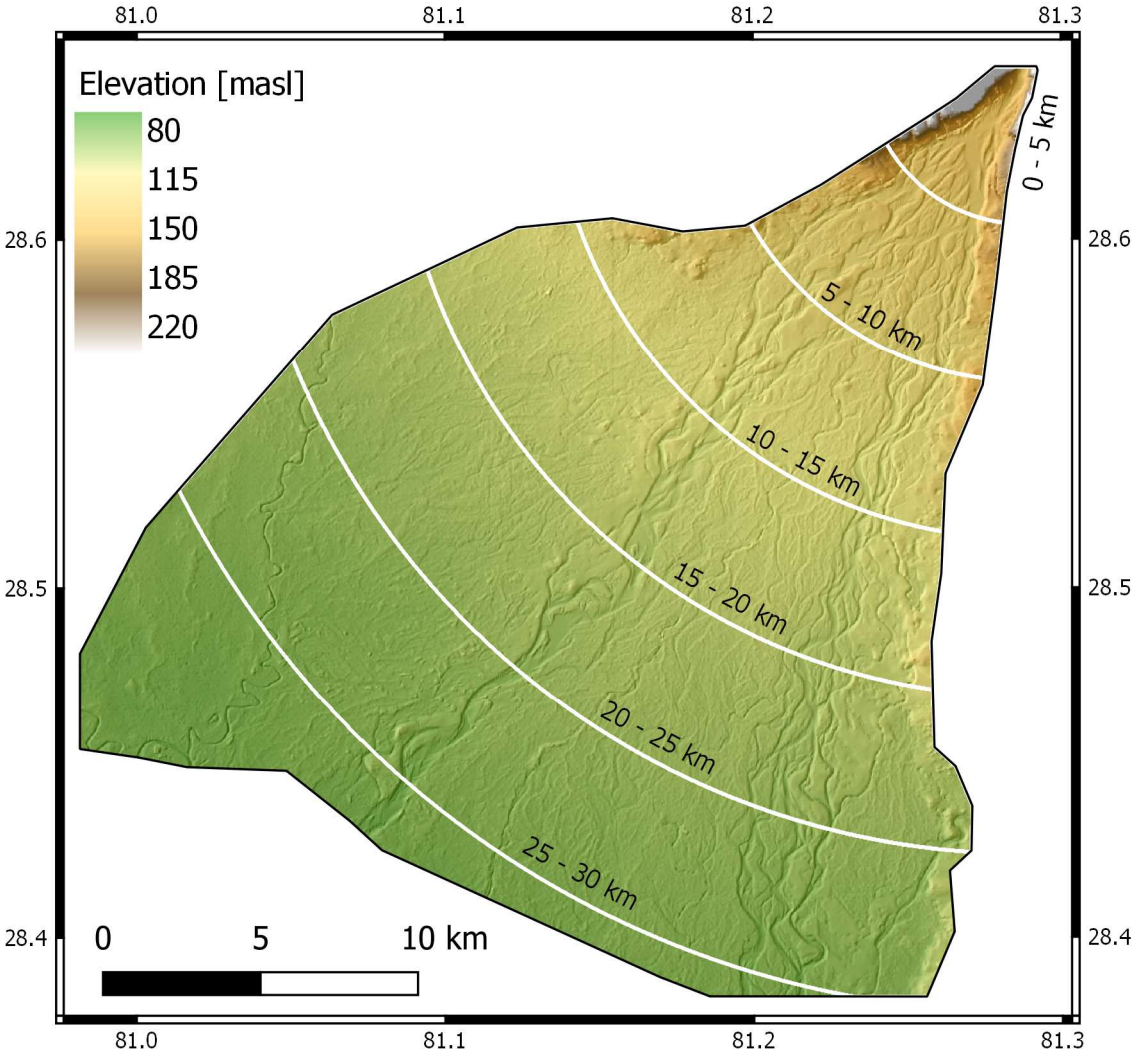


Figure 6.1: The study area of the morphodynamic modelling in the Karnali fan. The white lines show the distance from the fan inlet. The topography is the DEM used in the morphodynamic modelling which is described in Section 6.3.1.

The morphodynamic simulations are designed as an experiment to explore the sensitivity of the morphological evolution to the climate-change-induced flood flow changes until the end of the 21st century at the mesoscale (5 – 50 km). Our current understanding of the morphological system with multiple processes, which vary in their temporal and spatial scales and interact nonlinearly, is limited and the models with the best representation of the physical processes are too complex for the application at the required spatial and temporal scales (Section 3.5).

All morphodynamic models are sensitive to the boundary condition which is problematic as there is limited data on the morphological characteristics of the Karnali fan. The lack of knowledge about key variables (e.g., bedload sediment transport rates) forces the use of assumptions (e.g. sediment transport at capacity). Furthermore, there is a lack of data

which could be used for model calibration (e.g., consecutive topographic measurements). Therefore, there are many unknown variables and hence this study is designed as a conceptual experiment to provide an exploratory investigation of the climate change impact on the morphological evolution of the Karnali fan and better understand the factors which control the fan evolution.

The rest of this chapter is structured as follows, Section 6.2 presents the methodology, Section 6.3 presents the data which define the boundary conditions, followed by the results (Section 6.4), the discussion (Section 6.5), and the conclusions (Section 6.6).

6.2 Methods

The model selection is presented in Section 6.2.1, followed by a description of the morphodynamic model (Section 6.2.2), the model parameterisation (Section 6.2.3), the experimental design (Section 6.2.4), and the methodology to prepare the topographic data of the modelling domain (Section 6.2.5).

6.2.Model selection

The morphodynamic model must be able to simulate the governing processes which control the morphological evolution of the braided river shaping the Karnali fan. The dominant sediment transport mode for such rivers is the bedload transport of the sand and gravel fractions (Murray and Paola, 1994; Williams *et al.*, 2016; Harvey, 2018; Dingle *et al.*, 2020b). Braided river systems are characterised by the splitting, switching and migration of channels, and the formation and migration of bars (Murray and Paola, 1994; Kleinhans *et al.*, 2013). This evolution of channels and bars is attributed to mechanisms which are affected by multi-directional and lateral flow such as central bar development, transverse bar conversion, lateral bar development, chute cutoff of point bars, bar edge trimming, bank erosion, channel degradation and aggradation, and overbank deposition (Murray and Paola, 1994; Kleinhans *et al.*, 2013; Williams *et al.*, 2016). The simulation of the morphological evolution in this system, therefore, requires a model which incorporates bedload sediment transport models. Furthermore, the model must simulate flows in the longitudinal and lateral directions to predict the evolution of braided channels.

The main criteria for the model selection are the process description and the computational efficiency. The morphodynamic simulation experiment with ensemble simulations (12 climate projections each comprising three scenarios and three flow percentiles) of a 545 km^2 large area over 40 years (baseline) to 80 years (projected scenarios) is computationally expensive and excludes 3D models which do not exist at the scale of enquiry. The computationally efficient 1-D models do not predict lateral movement and are not suitable for the application in the Karnali River. Therefore, a 2-D model is selected.

These 2-D models are classified into lower-complexity cellular models and higher-complexity physics-based models. The benefit of the cellular models is the computational efficiency which enables the simulation of an ensemble over larger spatial and temporal scales (Nicholas *et al.*, 2006; Nicholas, 2013a). However, physics-based models are superior in capturing the morphological processes in braided river systems (Williams *et al.*, 2016). The inferior performance of cellular models arises from the simple flow routing algorithms which neglect the momentum conservation, and the use of local bed slope to calculate sediment transport rates (Nicholas, 2013a). The selection of a 2-D model needs a balance between the process representation and the computational efficiency.

The CAESAR-LISFLOOD model (Bates *et al.*, 2010; Coulthard *et al.*, 2013) is a hybrid cellular model which overcomes the strongest critic of the poor flow algorithms by incorporating a more complex hydrodynamic flow model (Coulthard *et al.*, 2013; Ziliani *et al.*, 2020). This hydrodynamic model is an approximation of the Shallow Water Equations (SWE) which is computationally efficient while providing good estimates of flow velocity and depth and significantly improves the flow representation in braided rivers compared to the original CAESAR flow model (Neal *et al.*, 2012; Ziliani *et al.*, 2020). This model has a good balance between the process representation for simulating braided river systems and computational efficiency for ensemble simulations.

6.2.2 The CAESAR-LISFLOOD model

The CAESAR-LISFLOOD (C-L) model combines the cellular morphodynamic CAESAR model (Coulthard *et al.*, 2002; Van De Wiel *et al.*, 2007) with the hydrodynamic LISFLOOD-FP model (Bates and De Roo, 2000; Bates *et al.*, 2010). The initial CAESAR model (Coulthard *et al.*, 2002) is a catchment scale model to simulate the landscape evolution on large temporal scales (up to millennia). Van De Wiel *et al.* (2007) modified this model to improve the representation of hydraulic and geomorphic processes in alluvial environments. The included processes improve the representation of overbank flow, sediment erosion and deposition, suspended and bed load transport, lateral erosion and bank failure which enables the simulation of channel degradation and aggradation, channel migration, bank erosion, point bar formation, and terrace formation (Van De Wiel *et al.*, 2007). This model was then combined with the physics-based LISFLOOD-FP model to improve the representation of flow processes (Coulthard *et al.*, 2013).

The C-L model can be applied either in the catchment mode to simulate the rainfall-runoff conversion and erosion within a catchment, or in the reach mode to simulate the evolution of a channel and floodplain for predefined water and sediment influxes. The reach model comprises a hydrodynamic flow model and a sediment transport model.

6.2.2.1 The LISFLOOD-FP flow model

The flow model is a 2D storage cell inundation model (Bates *et al.*, 2010) which approximates the full Shallow-Water-Equations SWE and integrates an inertia term. These equations simulate 1D flow but are solved in the x and y directions to simulate 2D flow over a raster grid. The discharge Q is simulated from the equation (Bates *et al.*, 2010; Coulthard *et al.*, 2013):

$$Q = \frac{q - gh_{flow}\,\Delta t\,\frac{\Delta(h+z)}{\Delta x}}{(1 + gh_{flow}\,\Delta t n^2\,|q|/\,h_{flow}^{10/3}}\,\Delta x \qquad \text{Eq. 6.1}$$

The water depth of a cell is calculated from the equation:

$$\frac{\Delta h^{i,j}}{\Delta t} = \frac{Q_x^{i-1,j} - Q_x^{i,j} - Q_y^{i,j-1} - Q_y^{i,j}}{\Delta x^2} \qquad \text{Eq. 6.2}$$

Where t is time (s), q is the flow of the previous time step ($m^2\ s^{-1}$), g is the acceleration due to gravity ($m\ s^{-1}$), n is Mannings roughness coefficient ($m^{1/3}\ s^{-1}$), h is the depth (m),

z is the elevation (m), hflow is the maximum depth of flow between the cells, x is the grid cell width (m), and i and j are the cell coordinates.

The modelling time step is controlled by the Courant-Friedrichs-Levy condition and hence the time step scales linearly with the grid resolution unlike diffusive models for which it scales exponentially with the grid resolution (Bates *et al.*, 2010):

$$\Delta t_{max} = \propto \frac{\Delta x}{\sqrt{gh}} \qquad \text{Eq. 6.3}$$

where $\propto$ is a coefficient which enhances the stability of shallow floodplain flows and is typically within the range of 0.2 – 0.7 (Bates *et al.*, 2010).

This SWE approximation has lower computational costs due to the larger time steps imposed while the predicted flow velocity and depth are not inferior to other approximations and fall within 10% interval of full SWE codes (Neal *et al.*, 2012). However, this model should only be used in sub-critical flow conditions and with steady or gradually-varying flows and cannot predict shallow flows over steep terrain (Bates *et al.*, 2010; Neal *et al.*, 2012). Furthermore, the model simulates flow in the four cardinal directions and cannot predict flow to the four diagonal cells which can cause inaccuracies if the grid resolution exceeds the channel width (Coulthard *et al.*, 2013).

6.2.2.2 The sediment transport model

The predicted flow depths and velocities are used to calculate the sediment transport processes (erosion, transport, deposition). C-L allows the separation of the sediments into nine grain size fractions and the sediment transport is calculated for each fraction separately (Coulthard *et al.*, 2013). The finest fraction can be transported as suspended load which is simulated based on the sediment concentration and the settling velocity of the water column (Coulthard *et al.*, 2013). The bedload can be simulated from three equations (Meyer-Peter and Müller, 1948; Einstein, 1950; Wilcock and Crowe, 2003). The model of Wilcock and Crowe (2003) is an empirical model which is chosen because it was developed for the sand and gravel fractions, performs well for braided river systems, does not require information on subsurface grain sizes, and requires no calibration (Wilcock and Crowe, 2003; Ziliani *et al.*, 2020).

The bedload sediment is transported and deposited in the downstream cells with lower bed elevation whereas the proportion is based on the local bed slope between the source cell and the target cells. The deposited material can be re-eroded in the next time step. The suspended load is distributed to all neighbouring cells for which the bed elevation is

lower than the water level of the source cell (Van De Wiel *et al.*, 2007). The sediment transport of bedload and suspended load is both, capacity-limited whereas this limit is defined by the sediment transport equation, and supply-limited by the volume of the material in the active layer of the riverbed (Van De Wiel *et al.*, 2007). The model includes a threshold for the maximum erosion rate which limits the time step of the sediment transport calculations. However, since the flow model also includes a time step calculation, the lower time step of both models is chosen at every iteration (Coulthard *et al.*, 2013).

This active layer is the upper layer of the bed sediments. The thickness of the active layer is flexible as it increases when sediment is deposited and decreases when sediment erodes. However, it is converted to a strata layer when the thickness exceeds 150% of the strata thickness which is a predefined model parameter. The strata layers (up to 20 layers) are sediment storages which represent the alluvial sediments. The first strata layer underneath the active layer can be added to the active layer when the active layer thickness is smaller than 25% of the strata thickness. The bedrock layer is the lowest layer representing the bedrock which can erode by a predefined fixed erosion rate (Van De Wiel *et al.*, 2007; Coulthard *et al.*, 2013).

The model uses a simple rule-based algorithm to predict bank erosion rates and is explained in detail in Van de Wiel *et al.* (2007). In short, this algorithm predicts bank erosion as a function of near-bank flow velocity and water height to the radius of the curvature of the river bank. The eroded bank sediment is deposited within the channel whereas it is distributed laterally based on the cross-channel gradient of the curvature. This representation of lateral transport lacks physical justification and simulates the symptoms rather than the causes of bank erosion (Van De Wiel *et al.*, 2007).

The model further includes routines to simulate in-channel lateral erosion. The in-channel lateral erosion is integrated to prevent unreasonable narrowing of channels to single cells caused by the positive feedback between flow concentration and higher erosion rates. This lateral in-channel erosion is controlled by a parameter which represents the cohesion of the modelled substrate. A high cohesion results in narrower channels and a low cohesion in the formation of flatter and wider channels (Coulthard et al., 2013).

6.2.3 Model parameterisation

The model parameterisation is based on values provided in the literature, particularly the C-L user manual (C-L Manual, 2023), and the calibration of the CAESAR and C-L models for two braided rivers in the Italian Alps (Ziliani *et al.*, 2013, 2020). The parameter descriptions, the used values, and their justification are presented in Table 6.1. The reader is referred to the user manual (C-L Manual, 2023) for further details about these parameters. Most of the parameter values can be determined either from the grid resolution (25m in this study – see Section 6.2.4) or the modelled river system.

Three parameter values are obtained from measurements. The first of these parameters is the number of passes for the edge smoothing of the lateral erosion algorithm. This value should be set to the distance between river bends which was measured in Google Earth for several well-defined bends in the Eastern and Western river branches. The second parameter is the input/output difference which determines the selection of the time step of the sediment transport and flow models. Usually, C-L uses the lower of both time steps for both models, but a threshold can be defined to speed up the simulations during low flow conditions. This threshold should describe the low flow discharge and was obtained from the DHM discharge observations. The third parameter describes the slope of the edge cells at the boundaries of the modelling domain. This slope was measured from the DEM (Section 6.2.4).

Table 6.1: The description and definition of the CAESAR-LISFLOOD parameters

Module	*Parameter*	*Unit*	*Value*	*Description*	*Justification*	*Reference*
Sediment transport	Max. velocity	-	5	Maximum velocity for sediment transport calculations	Default. Only required for very steep rivers	C-L Manual, 2023
	Max. erode limit	m	0.02	Limits the amount of erosion / deposition per cell and, thus, controls the time step	Cell size dependent. 0.01 for 10 m, higher for coarser grids	C-L Manual, 2023
	Active layer thickness	m	0.1	The thickness of the active layer	0.1 – 0.2. Minimum 4x max. erode limit	C-L Manual, 2023
	in-channel lateral erosion rate	N m^{-2}	25	Represents the cohesion of the sediments	Typically between 10 – 20. Tested in initial runs	C-L Manual, 2023 Tested
	Lateral erosion rate		0.002	Parameter for the bank erosion calculation.	0.001 – 0.01 for braided rivers	Ziliani *et al*., 2013; C-L Manual, 2023
	Number of passes for edge smoothing	grid cells	60	Describes the smoothing of the curvature for the lateral (bank) erosion calculation	Distance between river bends in no. of cells.	(C-L Manual, 2023; Measured in Google Earth
	Downstream shift of lateral erosion	grid cells	6	Number of cells to shift lateral (bank) erosion downstream	10% of edge smoothing passes	C-L Manual, 2023
	Max difference in cross-channel smoothing	-	0.00001	Affects the lateral sediment transport of the sediments eroded from river banks	0.00001 for channels wider than 10 grid cells	C-L Manual, 2023
Flow module	input/output difference	m^3/s	360	Affects runtime as it enables a fast model during low flows	Low flow or mean annual flow	DHM Discharge record
	Water depth threshold	m	0.01	Water depth above which erosion is calculated	0.01 for DEM > 5m and < 50m	C-L Manual, 2023
	Slope for edge cells	m m^{-1}	0.001	Defines water depth for outflow cells. Affects deposition / scouring at outflow cells.	Slope at downstream boundary	Measured from DEM
	Courant number	-	0.7	Controls the numerical stability and speed of the calculations	0.7 for 20 – 50 m grids	C-L Manual, 2023
	hflow threshold	m	0.00001	Prevents water flow between cells if the water surface gradient is very small	Default	C-L Manual, 2023
	Froude number	-	0.8	Ratio of inertial and gravitational forces	Default; 0.8 causes sub-critical flow	C-L Manual, 2023
	Manning's n	-	0.04	Roughness coefficient	Surface roughness	Section 6.2.4.3

The lateral erosion rate is a sensitive parameter and, thus, should be calibrated for the study area (Ziliani *et al.*, 2013). However, a calibration requires information about the morphological change which is not available for the Karnali fan. Furthermore, morphological models are sensitive to the boundary condition and hence even a behavioural model might produce simulations which deviate from the observations (Nicholas *et al.*, 2006; Papanicolaou *et al.*, 2008; Hardy, 2013). This lateral erosion rate should fall within 0.01 – 0.001 for braided rivers (C-L Manual, 2023). In the absence of field observations, the value calibrated to the braided Tagliamento River (0.002) was tested and transferred to the Karnali River (Ziliani *et al.*, 2013).

The in-channel lateral erosion rate affects the lateral transfer of bank erosion material. This parameter represents the cohesion of the sediments and the values range between 10 – 20 for most river systems (C-L Manual, 2023). However, braided rivers and alluvial fans are characterised by a low sediment cohesion and hence a lower cohesion (higher value) is used (Murray and Paola, 1994; Kleinhans *et al.*, 2013). A high cohesion produces narrow channels, and a low cohesion produces shallower and wider channels. The value of 25 was determined from initial test runs aiming to predict reasonable channel widths. However, the channel widths are wider at the inlet and in the Western branch than in the Eastern ones and hence this global value cannot represent the heterogeneity of the study area. However, the value of 25 neither simulated neither unreasonable wide nor unreasonable narrow channels and is deemed as a sufficient approximation of the cohesion of the Karnali fan sediments.

6.2.4 Experimental design

The C-L model is used to simulate the morphological evolution for the projected hydrology of the CMIP6 scenarios. However, several assumptions are made to account for the data scarcity and the available computational resources. These assumptions are described in the following Sections 6.2.4.1 – 6.2.4.3. The model setup is then described in Section 6.2.4.4.

6.2.4.1 Sediment influx

The dominant sediment transport mode of braided rivers on alluvial fans is bedload transport for which no records are available for the Karnali River (Murray and Paola, 1994; Kleinhans *et al.*, 2013; Harvey, 2018). The bedload transport dominance in the Karnali river system is also emphasised by the dominance of gravel (and the lack of fine sand, silt and clay) in the grain size distributions of the bed and bank material (Quick *et al.*, 2019; Dingle *et al.*, 2020b).

Several studies estimate the sediment influx into the fan from either suspended sediment records or detrital Cosmogenic Radionuclide (CRN) analysis of the fan sediments both of which cannot be used for this study. The first method estimates the average erosion rates of Himalayan catchments and these studies suggest that the sediment transport is supply-limited (Andermann *et al.*, 2012a; Morin *et al.*, 2018; Sinha *et al.*, 2019) However, suspended sediments bypass the fan and deposit in the Ganges-Brahmaputra delta in Bangladesh (Lupker *et al.*, 2011; Dingle *et al.*, 2020b). The CRS Analysis also considers coarser sediments which are transported as bedload but the obtained estimates are averaged over long periods (250 – 600 years) and cannot be linked to single flow events (Lupker *et al.*, 2012). Hence, the estimates of both methods cannot be used to determine the sediment influx into the Karnali fan.

The sediment which is released into the fan and makes up the riverbed and banks is sourced in a small fraction of the catchment located 10 – 30 km upstream of the fan (Quick *et al.*, 2019). Most of the gravel which is supplied to the rivers in the upstream catchment abrades to sand and bypasses the fan (Dingle *et al.*, 2017). Instead, most of the coarse sediments (up to 100%) which make up the fan originate from the loose and poorly consolidated Upper Siwalik conglomerates (Upreti, 1999; Quick *et al.*, 2019). The high density of gullies and landslides demonstrates the high sediment production and these sediments are efficiently transported in the river network (Hurtrez *et al.*, 1999;

Upreti, 2001; Ghimire *et al.*, 2013). Furthermore, extensive gravel bars are exposed in the Karnali and Bheri rivers upstream of the mountain gauge which indicates the presence of gravel storage along the rivers that can be mobilised during high flow events. In these sections, the river flows through wide valleys and is, unlike other rivers upstream of the mountain gauge (e.g. the Koshi River), not laterally confined which further suggests the presence of vast gravel deposits. The high sediment production, the efficient sediment transport and the vast gravel storage in proximity to the mountain outlet do not indicate a limited supply of coarse sediments.

It is, therefore, assumed that the sediment influx is limited by the transport capacity of the river. This capacity is calculated from the Bedload Assessment for Gravel-bed streams (BAGS) model (Pitlick *et al.*, 2009). This model calculates transport capacities at the cross-section resolution (1-D) and is established at the inlet of the C-L modelling domain. The bedload transport is calculated from the Wilcock and Crowe (2003) bedload equation for the same nine grain size classes used in the C-L model. The grain size distribution of the Siwalik conglomerates is used as surface grain size distribution (Quick *et al.*, 2019). Manning's roughness coefficient is set to 0.04 s $m^{-1/3}$ (Section 6.2.4.3). The cross-section topography and the longitudinal slope (0.0013 m m^{-1}) are derived from the DEM used in the C-L model (Section 6.3.1).

6.2.4.2 Temporal subset

The morphological evolution of alluvial fans is dominated by flood flows which deliver most sediments to the fan and redistribute the fan sediments (Kleinhans *et al.*, 2013; Harvey, 2018). The largest sediment pulses are delivered by extreme floods and these sediments are redistributed during smaller floods (Kleinhans *et al.*, 2013). Hence, it is assumed that flood flows are the dominant control on the morphological evolution of the Karnali fan.

The importance of flood flows for the morphological evolution varies in the Karnali fan because the grain sizes decrease with distance to the mountain outlet. The coarse bed material in the fan head is only transported during extreme and rare floods (≥ 1-in-500-years flood) but the grain sizes decrease in the lateral and longitudinal direction and are transported at lower and more frequent flows (Dingle *et al.*, 2020b). Therefore, the impact of this assumption on the predictions varies in space and is higher for the southern channel sections. However, this assumption that the morphological evolution is controlled by the flood flows is required to maintain the feasibility of the simulations. An initial model run of 80 flood events predicted for one climate scenario (SSP585) by

one CMIP6 member (ACCESS-CM2) required 281 hours of simulation. The simulation of longer periods (e.g., the monsoon season) is, due to the complex computations and the required runtimes, beyond the scope of this research.

The model runtimes further restrict the simulated period of each flood event to three days. The two consecutive days following the flood are simulated to account for the impact of the flood recession on the morphological evolution and allow for the redistribution and sorting of delivered sediments. However, initial tests with variable periods (flood event ±10 days) indicated that it is not feasible to simulate periods longer than two days after the flood.

6.2.4.3 Surface roughness

The surface roughness describes the effect of the surface friction of the channel and floodplain on the flow. The flow models are sensitive towards this roughness coefficient which is commonly calibrated for the channel and floodplain separately. This calibration is commonly conducted from the comparison of the simulated and observed inundation extents or flood wave travel times (Horritt and Bates, 2002; Horritt, 2006; Di Baldassarre *et al.*, 2010; Komi *et al.*, 2017). Furthermore, it can also be estimated from high-resolution topographic data and flow measurements (Pearson, 2020).

The roughness coefficient is not calibrated but obtained from the literature in this study. Several factors hamper the calibration of the roughness coefficient which is represented by Manning's n friction coefficient in the C-L model. The main reason is the lack of accurate datasets of flood inundations of past events. The calibration would be challenging even with such datasets available due to the high uncertainty in the discharge data during flood events, and the temporal discrepancy between the daily averaged discharge data and the inundation maps which either describe the flood patterns at a specific point in time or the maximum inundation. This hampers the identification of the site-specific roughness coefficients, and hence literature values are transferred to the Karnali River instead.

The Manning's n roughness coefficient of mountain streams with no in-channel vegetation and gravel and cobbles in the riverbed is taken from Chow (1959). The Manning's n of such streams ranges between 0.03 – 0.05. The normal roughness of these streams is 0.04 which is used in this study for the channel and floodplain. This selection is supported by the roughness obtained from field measurements of the Central Himalayan East Rapti River (Pearson, 2020; Pearson *et al.*, 2022), sensitivity analyses

of the LISFLOOD-FP model (Di Baldassarre *et al.*, 2010; Komi *et al.*, 2017), and literature values used in studies of comparable river systems (Yan *et al.*, 2014; Feeney *et al.*, 2020).

This study assumes a uniform roughness coefficient in the channel and floodplain whereas most studies use a higher roughness in the floodplain than in the channel (e.g., Yan *et al.*, 2014; Komi *et al.*, 2017; Pearson, 2020). This assumption is made because the locations of the channel and floodplain change with the morphological evolution of the fan which would require the classification of channel and floodplain and roughness mapping after each flood event.

6.2.4.4 Modelling setup

The C-L model is applied for the flood discharge projected by the 12 CMIP6 models (Chapter 5). These projections comprise the baseline (1975 – 2014) which is applied as a reference to simulate the morphological evolution under the current conditions, the medium-emission scenario SSP245, and the high-emission scenario SSP585 (both 2020 – 2099) which are applied to simulate the morphological evolution for the projected hydrological conditions.

The flood events are classified using the Annual Maximum flow (AMAX) classification approach and hence one flood event is simulated per year (i.e. 40 events for the baseline scenario, and 80 events for the projected scenarios SSP245 and SSP585). For each flood event, 3.5 days are simulated; the first 12h are simulated with a constant influx of 3,000 m^3/s to wet the channel cells, followed by the flood hydrograph of the flood event and the consecutive two days of the flood recession. However, each flood event consists of 64 predictions of the hydrological modelling ensemble. This ensemble is reduced to the median, the 2.5^{th} percentile ($P_{2.5}$), and the 97.5^{th} percentile ($P_{97.5}$)because it is unfeasible to simulate each ensemble member. Therefore, 36 flood records are simulated for each scenario (12 CIMP6 model X 3 flow percentiles). The model is restarted after each year whereas the output (topography, grain size distribution) of the previous simulation is used as the boundary conditions for the following simulation.

As previously mentioned, the lack of data to determine the boundary conditions, and the available computational resources requires the use of assumptions. These assumptions are:

- Flood events dominate the morphological evolution of the Karnali fan, and hence only the flood events are simulated (Section 6.2.4.1).
- The sediment influx into the fan is limited by the transport capacity and not by the sediment supply (Section 6.2.4.2).
- The surface roughness is uniform in time and space (Section 6.2.4.3).
- Bedload transport is the dominant sediment transport mode and the suspended sediment transport is not simulated. This assumption is supported by our knowledge of braided rivers with gravel beds and alluvial fans (Murray and Paola, 1994; Kleinhans *et al.*, 2013; Ziliani *et al.*, 2013; Williams *et al.*, 2016; Harvey, 2018). It is also supported by our knowledge of the sediment dynamics of the Central Himalayan River systems and the Karnali fan although the sand may be transported in suspension during floods with a high magnitude (Lupker *et al.*, 2011; Dingle *et al.*, 2020b).
- The vegetation does not control the morphological evolution of the fan. Vegetation stabilises islands and bars in braided river systems and retains sediments in the floodplains (Murray and Paola, 2003; Kleinhans, 2010; Kleinhans *et al.*, 2013; Ziliani *et al.*, 2013). However, the interactions between vegetation, flow, and sediment transport are complex as the vegetation can be destroyed by water and sediments (Coulthard *et al.*, 2002; Kleinhans *et al.*, 2013). The effect of sediments on the morphology of the Karnali fan is unclear, and the growth and destruction of vegetation cannot adequately be simulated in this simulation framework which only simulates 3.5 days per year. Therefore, the simulation of the vegetation effects introduces another layer of uncertainty and is, thus, disregarded in this study.
- The basin subsidence does not control the morphological changes for the simulated flows and time scale. The basin subsidence is an important control of the fan evolution on longer time scales and the subsidence rate of the Karnali is 0.4 ± 0.2 mm/yr (Dingle *et al.,* 2016, Harvey, 2018). However, the gravel supply of the mean monsoon season flow (1985 – 2014, excluding AMAX flood events and assuming sediment transport at capacity) exceeds the accommodation space of the basin subsidence. It is therefore argued that it is unreasonable to include the basin subsidence in the current modelling setup which only simulates the flood events.

The simulations of the morphodynamic evolution for the projected climates are designed as an experiment to gain insight into the morphodynamic system of the Karnali fan. This design as an experiment is constraint by:

- the limited understanding of morphological processes (Mosselman, 2012; Hardy, 2013; Ancey, 2020a);
- the sensitivity of morphodynamic models towards the boundary conditions (Papanicolaou *et al.*, 2008; Hardy, 2013; Williams *et al.*, 2016);
- the lack of information to determine the boundary conditions and to calibrate the model for the Karnali fan; and consequently the use of the described assumptions.

These experimental simulations provide information about the climate-change impact on the morphological evolution, the controls of the morphological evolution, and the uncertainty propagation when linking climate models (the CMIP6 projections), hydrological models (Chapter 5), and morphological models. The results are further used to gain a better understanding of the impact of morphological changes on spatial flood hazard patterns (Chapter 7).

6.2.5 The generation of a digital terrain model

A method is developed to generate a Digital Terrain Model (DTM) from stereo satellite images. These images are converted to point clouds which are the basis for the generation of the DTM. The main challenge is the lower density of ground points compared to scans at lower altitudes (i.e. drones or airborne surveys). Due to this low ground-point density, the point cloud can be processed to a Digital Surface Model (DSM) which includes surface objects (e.g., buildings and trees) but the ground elevation cannot be estimated from existing algorithms.

The generation of a DTM is further complicated by the lack of in-situ data of the ground elevation (e.g. in forested areas). This lack prevents the removal of surface objects in the DSM using machine learning techniques (e.g., Hawker *et al.*, 2022). An approach to estimate the terrain elevation without in-situ data was tested but did not provide reliable results in areas with dense object cover (Lindsay, 2018). Therefore, a novel method is developed which uses interpolation techniques and focal filters to remove trees, forests, and buildings, estimate the riverbed elevation, and remove sensor noise.

The basis of the DTM are three panchromatic stereo images from the WorldView-1 and WorldView-2 satellites which were captured during low flow conditions (29/11/2020, 07/01/2021; 03/03/2021). These images were converted to a point cloud using the Agisoft Metashape software. These points were then classified into ground points (e.g. bare ground, roads) or surface object points (e.g. trees, buildings) and interpolated to generate one DTM and one DSM which both include the elevation of objects (hereafter called raw DTM and raw DSM). The raw DTM does not, despite its name, represent the ground surface which is attributed to the lower point density compared to drone/airborne products which results from the larger distance of the satellite sensor.

The generation of the DTM from the raw DTM includes four processing stages (Figure 6.2). The first stage masks areas that are covered by forests and water which have a poor elevation representation in the raw DTM. The second and third stage estimate the ground elevation of these inundated and forested areas. The final stage applies filters to remove noise to generate the final DTM.

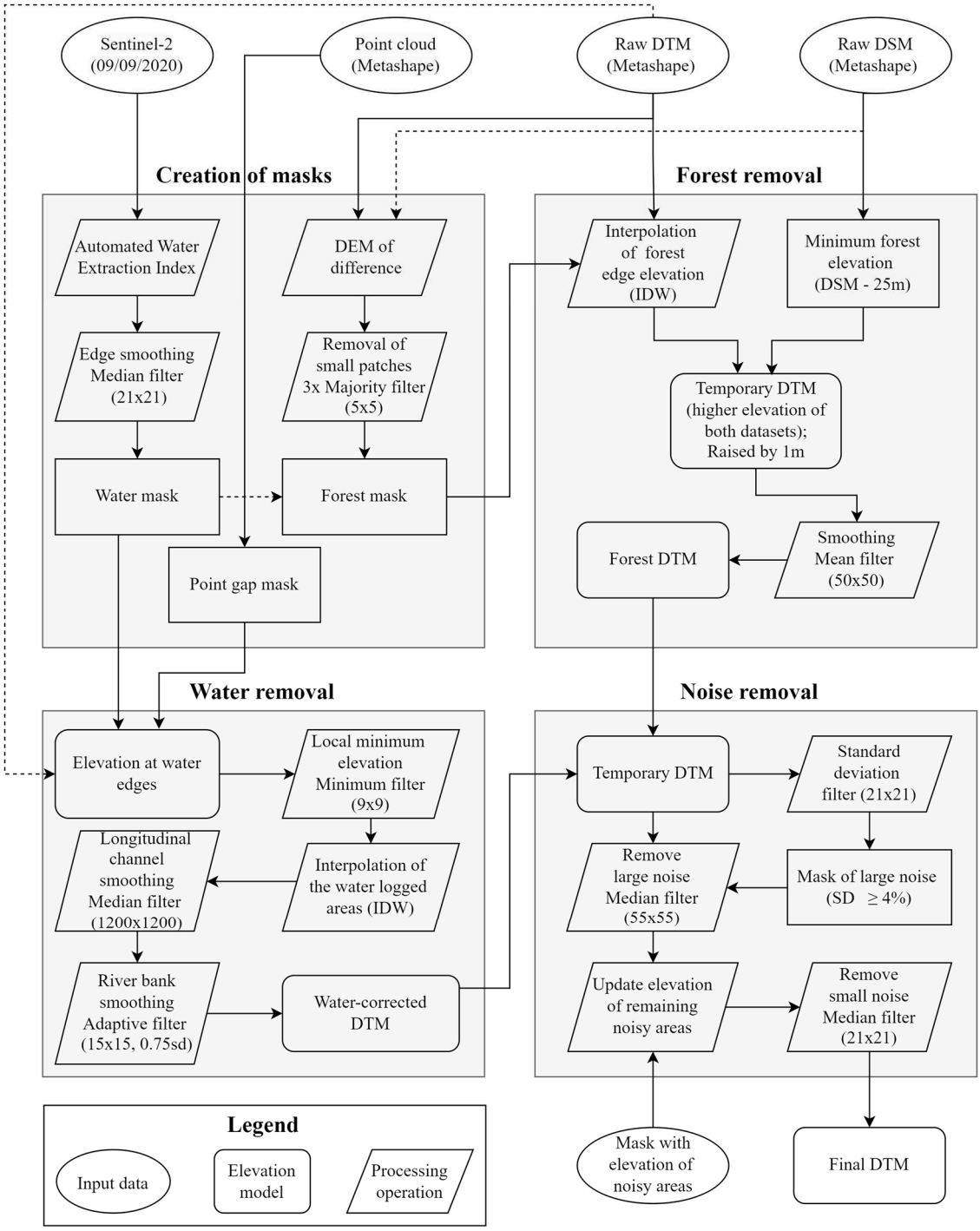


Figure 6.2: Processing workflow of the Digital Terrain Model (DTM) generation. The DTM is generated from a Digital Surface Model (DSM) by estimating the ground elevation of forests and inundated areas. The raw DTM does contain, despite its name, the surface elevation in areas with dense object cover. The different arrow styles (dashed, solid) are used to enhance visibility and do not represent conceptual differences.

6.2.5.1 Masking

The initial step is the identification of areas covered by forests and water. The water areas are classified from Sentinel-2 imagery using the Automated Water Extraction Index (AWEI) (Feyisa *et al.*, 2014). This water mask is resampled from the Sentinel-2 resolution (10 m) to the resolution of the raw DTM (1.8m) and a median filter (21x21 pixels) is used to smooth the riverbanks.

The forests are masked from a DEM of Difference (DoD) of the raw DTM and the raw DSM. It is found that this DOD is a good descriptor of forested areas. However, these differences occurred also in some channel sections and hence areas classified as water are removed from the forest mask. Small patches of a few pixels are removed using a majority filter (5x5 pixels) three consecutive times because these small patches may be misclassifications and will be removed in the final noise-removal stage.

A point gap mask Is created which includes pixels which lack any points in the point cloud. These gaps are most frequent in the channel. This mask is used to prevent the use of such pixels in the later interpolation process.

6.2.5.2 Ground elevation in forests

A statistical approach is used to estimate the ground elevation of forests. It was tested to remove forests from two global Canopy Height Models (CHM) (Potapov *et al.*, 2021; Lang *et al.*, 2022). However, these products overestimate the tree height which causes large pits and depressions in the DTM. Therefore, the forest is removed from the interpolation of edge elevations.

The forest ground elevation is estimated by projecting the elevation at the forest outline onto the forested areas using Inverse Distance Weighting (IDW) interpolation (Philip and Watson, 1982; Watson, 1985). The dominating tree species in the area is the Sal tree with an average height of 23.8 m (DFRS, 2014). A maximum threshold of 25 m is used to prevent the prediction of unreasonably low ground elevations. The ground elevation is replaced by the raw DSM height minus 25 m for the pixels that exceed this 25 m threshold. Afterwards, the ground elevation is smoothed using a mean filter (50x50 pixels) to remove local peaks and pits.

In the final step, the ground elevation of forested areas is raised by 1 m. This follows the rationale that forests develop in less frequently flooded areas and, therefore, are elevated higher than their surrounding which is not accounted for when interpolating the

edges. This 1 m value was obtained from iterative testing and provides elevations which appear reasonable but cannot be validated due to the lack of in-situ data. The resulting temporary DTM is referred to as the forest DTM.

6.2.5.3 Riverbed elevation

The riverbed elevation is poorly represented in the raw DTM because the water surface changes between the satellite images. The areas of the water mask and point gap mask are removed and a local minimum filter (9x9 pixels) is applied to identify the lowest local elevation. This elevation is interpolated using IDW interpolation (Philip and Watson, 1982; Watson, 1985) and is then smoothed with a large median filter (1,200x1,200 pixels) to obtain a downstream gradient. The predicted channel elevation is then lowered by 1 m to cut the channel into the DTM because the water-covered pixels must be lower than their neighbours. This 1 m value was tested iteratively and provided reasonable cross-section topographies but cannot represent the variation of water depths across the fan. However, this operation is conducted on the inundated channel section and hence the effect of the under- or overestimation of the riverbed elevation decreases with increasing flow rate.

The obtained channel DTM is then merged with the forest DTM and an adaptive filter (3x3 pixels, standard deviation threshold: 0.75) to smooth the channels in the updated DTM in the lateral direction while maintaining the steep riverbanks observed in the Eastern branch. This filter is only applied within a 20-pixel buffer of the channel outline.

6.2.5.4 The removal of small features and noise

The updated DTM still contains smaller objects such as buildings and tree patches, and noise from the generation of the raw DTM. Larger noise and surface objects (e.g. trees and buildings) are masked using a standard deviation filter (21x21 pixels). Pixels for which this standard deviation ≥ 4% are buffered and replaced by the elevation of a smoothed DTM (median filter: 55x55 pixels). This approach was tested by comparing it with satellite images in Google Earth and efficiently classified and removed surface objects of lower density (e.g. tree patches, buildings).

The resulting DTM is checked for remaining larger areas which are noisy and are updated with manually derived elevations. This DTM is then smoothed using a median filter (21x21 pixels) to remove small noise relics. This final DTM is shown in Figure 6.1.

6.3 Boundary conditions

The C-L model requires data about the surface topography, the bedrock layer, the grain size distribution across the modelling domain, and the influx of water and sediments.

6.3.1 Topography

The DTM is generated following the method described in section 6.2.5. The spatial resolution of the modelling is a trade-off between the model runtimes and the preservation of topographic details. Cross-section profiles showed that the topography of smaller channels in the floodplain is maintained at the 25 m resolution. Therefore, the DTM is then resampled from 1.8 m to 25 m (bilinear interpolation) which reduces the number of grid cells from ~170M to ~1M. This reduction of the spatial resolution decreases the model runtime and enables the simulation of the ensemble.

An artificial drainage ring is added to the boundaries of the study area. This ensures that water and sediments can leave the modelling domain at any cell of the model boundary and are not artificially trapped within the fan. The slope at the downstream border is set to 0.001 m m^{-1} which represents the river slope at the downstream edge measured from the DTM and agrees with Dingle *et al.* (2020a).

The C-L model requires the definition of the bedrock elevation. However, the study area does not contain a bedrock layer but is an alluvial fan located on the Indo-Gangetic plain which comprises up to 1,500 m thick alluvium (Upreti, 1999). Therefore, a theoretic bedrock surface is set 40 m below the DTM elevation. This layer is not activated in the modelling because the erosion does not reach 40 m in any cell.

6.3.2 Grain sizes

The Grain Size Distribution (GSD) is estimated from the measured GSD of seven gravel bars along a transect of the Karnali River in the fan (Figure 6.3 A). C-L simulates sediment transport for up to nine grain sizes. The taxonomic grain size classification (coarse sand, fine gravel, medium gravel, etc.) is not feasible due to the large variations in the GSD with coarse sediments at the fan apex (D_{50}: 238 mm) and much finer sediments at the downstream border of the study area close to the Gravel Sand Transition GST (D_{50}: 53 mm). Ziliani *et al.* (2013) determined the grain size classes from the measured GSD and not from the grain size taxonomy. A similar approach is used in this study where the class boundaries are chosen to reflect the mean grain size distribution of the seven samples (Figure 6.3 B).

The C-L model uses a single grain size for each class (and not the grain size range of each class) and these grain sizes are defined as the grain size with the mean density of each group (Table 6.2). The only exception is the finest class for which the size is set to 2 mm to have a sand fraction in the system. Sand gains importance towards the downstream border but this is not reflected by the mean GSD of the seven samples (Quick *et al.*, 2019; Dingle *et al.*, 2020b).

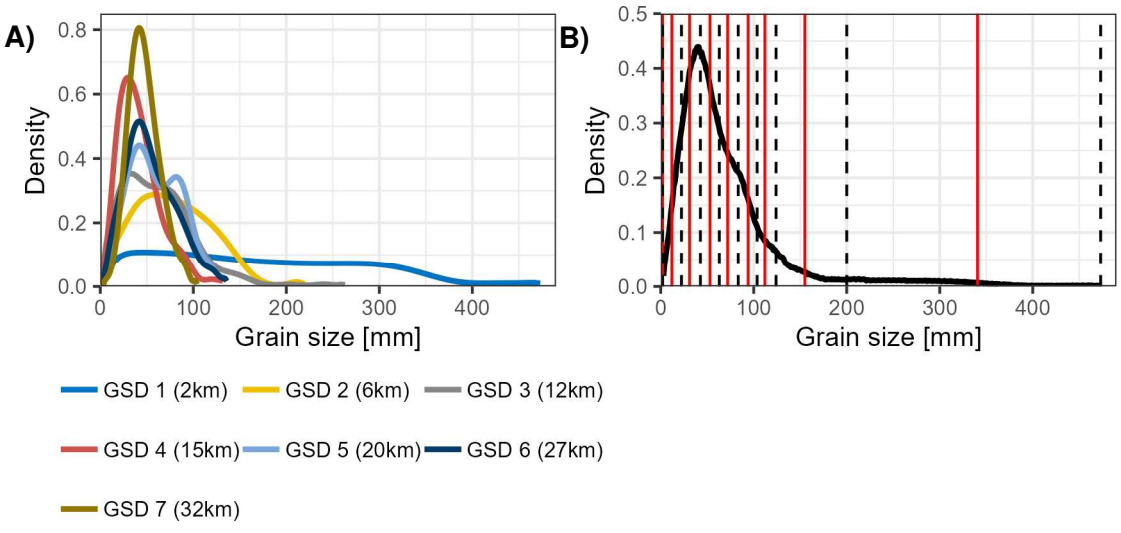


Figure 6.3: The grain sizes of the Karnali fan. Figure A shows the grain size distributions of gravel bars measured at seven locations along the Karnali River in the fan. The distance upstream inlet (the mountain outlet) is included in the brackets. The grain size distribution is obtained from Quick et al. *(2019). Figure B shows the mean grain size distribution of the seven samples (solid black line), the boundaries of the nine grain size classes used in the C-L modelling (dashed black lines), and the grain sizes used in the C-L model (red lines). See also Table 6.2 for the grain sizes of the classes.*

Table 6.2: The grain size classes used in the C-L modelling. The boundaries are the lower and upper boundaries of each class and the grain size is the grain size used in the C-L model.

Grain size class	Lower boundary [mm]	Upper boundary [mm]	Grain size [mm]
1	2	2	2
2	2	22	12
3	22	43	31
4	43	63	53
5	63	83	72
6	83	104	94
7	104	124	112
8	124	200	155
9	200	630	341

The fraction of each grain size class on the total composition along the longitudinal river profile is then interpolated from statistical modelling (Figure 6.4). The choice of the model is based on the longitudinal distribution. An exponential model is used for the smaller grain size classes which increase in the downstream direction (classes 1-4; 2 mm – 63 mm). The medium grain sizes classes that have the highest fraction in the central areas (10 – 25 km downstream of the inlet) are predicted from General Additive Models (GAM) (Class 5 – 7; 63 mm – 124 mm), and an exponential decay function is used for the largest two classes 8 and 9 (124 mm – 475 mm) for which the fraction decreases in the downstream direction. This approach produces a map with grain size fining in the longitudinal direction but without fining in the lateral direction and hence the gravel bar and bed are assumed to have a similar GSD.

The information about the presence of sand (≤ 2 mm) in the fan is contradictory. The grain size measurements of Quick *et al.* (2019) indicate that sand is absent in the study area. Grain size measurements shared by Dingle (personal comment) indicate that the riverbank at the bifurcation (5-6 km downstream of the inlet) comprises 3% sand. Furthermore, the downstream border is located close to the gravel-sand transition (GST) which marks a sudden shift from gravel to sand dominated material (Dingle *et al.*, 2020b). The presence of sand in the steeper upstream parts and the location of the GST indicate that it is unreasonable to simulate this system without any sand. Therefore, it is assumed that the sand fraction increases moderately from 3% at the inlet to 6% at the outlet. This moderate increase is somewhat subjective but the decision is a trade-off between observed downstream grain size fining and the measured sand at the bifurcation on the one hand (Dingle *et al.*, 2020b), and the sand-free downstream samples of Quick *et al.* (2019).

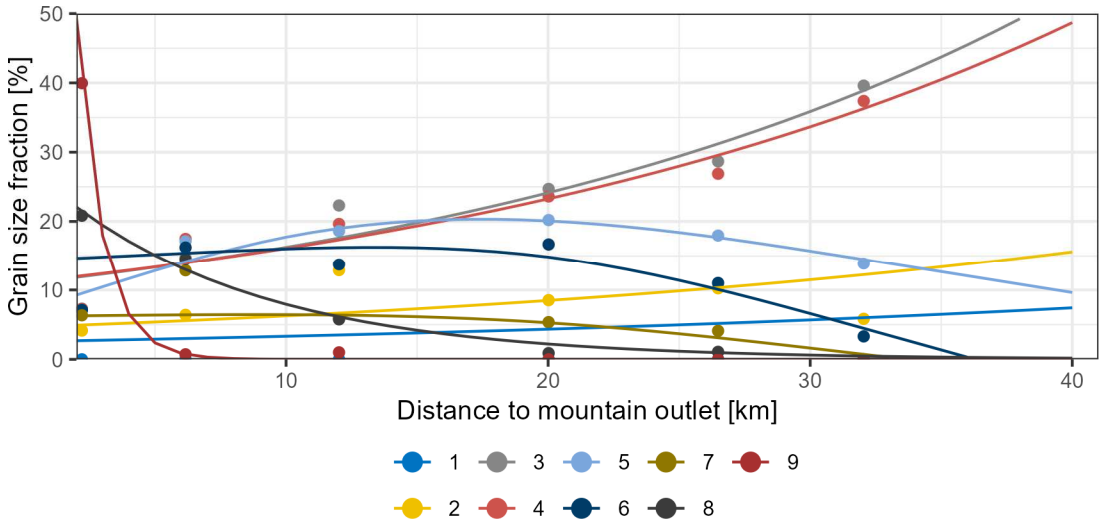


Figure 6.4: The interpolated grain size class fractions used to generate the grain size layer for the C-L modelling. The points indicate the fractions of each class measured on the riverbank at seven locations (Quick et al.*, 2019). The lines show the statistical models used for the interpolation of the grain size fractions to obtain the spatially distributed grain size layer. The grain sizes of the classes are presented in Table 6.2.*

6.3.3 The influx of water and sediment

The influx of water is determined from the hydrological modelling of the CMIP6 ensemble (Chapter 5). This water influx comprises the hydrograph of the AMAX event and the two consecutive days as described in section 6.2.4.2.

The sediment influx is the transport capacity of the respective flow as described in Section 6.2.4.2. This transport capacity is calculated for all projected flows for the nine grain size classes using the bedload equation of Wilcock and Crowe (2003). The sediment transport capacity is converted from kg/min to m^3/day assuming a rock density of 2,650 as also used by Dingle *et al.* (2017) for the Karnali fan. The predicted transport capacity for the range of simulated flood flows used to determine the sediment influx is presented in Figure 6.5.

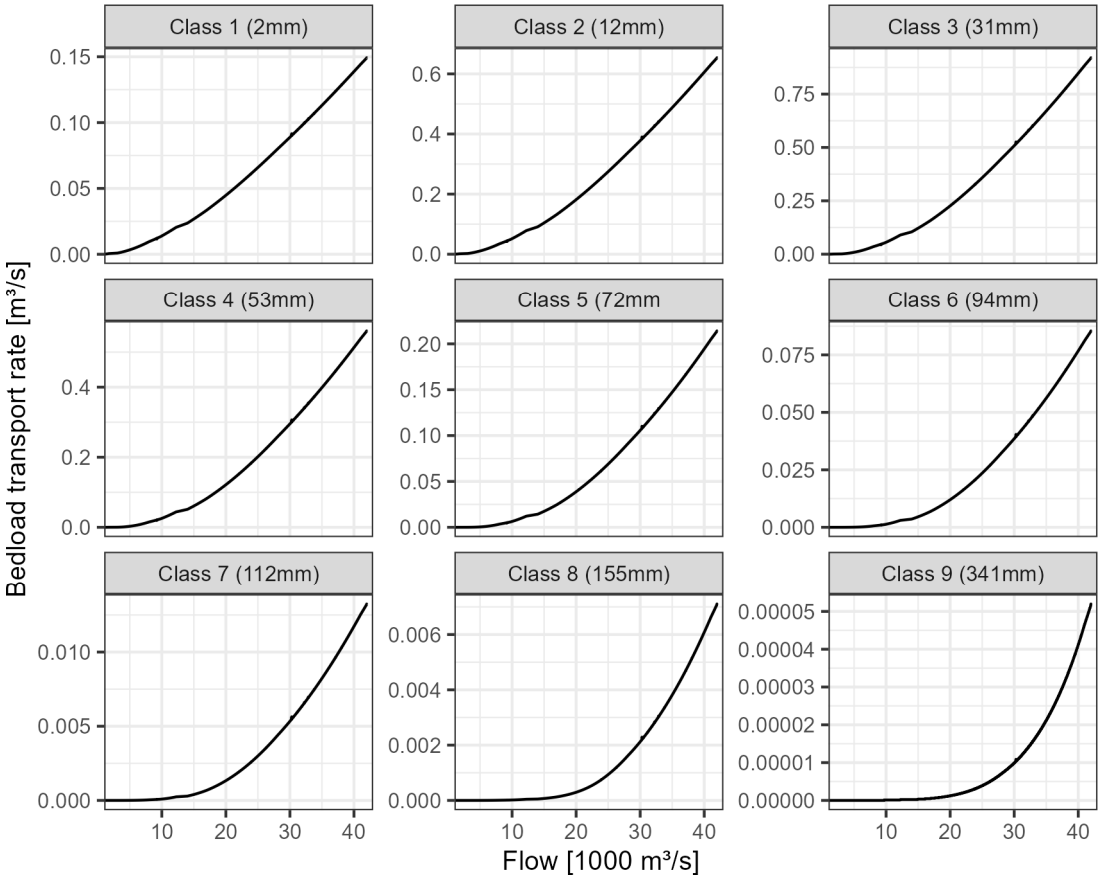


Figure 6.5: Bedload sediment transport capacity predicted by the BAGS model which is used to determine the sediment influx into the Karnali fan.

6.4 Results: The predicted evolution of the Karnali fan

This section presents the analysis of the morphodynamic simulations of the baseline scenario (1975 – 2014), the medium-emission scenario SSP245 (2020 – 2099), and the high-emission scenario SSP585 (2020 – 2099). The first section 6.4.1 describes the results of the initial model evaluation to establish the internal validity (mass balance) of the simulations and identify the warm-up period of the model predictions. This provides the basis for the analysis of the morphological evolution (Section 6.4.2), the sensitivity of the simulations to the sediment influx (Section 6.4.3) and the evolution of the flow pathways (Section 6.4.4).

6.4.1 The initial model evaluation

The mass balance of sediment and water provides information about the model validity as no mass should be gained or lost except at the model boundaries. The difference between inflow, outflow, and the mass stored in the modelling domain is 2% for the flow model and hence 2% of water is lost in the calculations (Figure 6.6 A). No sediments are lost as the balance is even for the sediment transport model (Figure 6.6 B).

The model initialises the topography and grain sizes during the first five years of the simulations. All ensemble members simulate a decrease in the fan elevation in the first five years (Figure 6.7). This predicted change does not appear to be driven by the flow rate as the behaviour is similar for every ensemble member and all three scenarios. Rather, this is a period in which the interpolated elevation and grain sizes are initialised by flushing out fine sediments and degrading (incising) channels into the predefined flat (in the lateral direction) interpolated channel sections of the input DTM. The highest change is simulated in the first year, and the initialisation is finished after five years. Therefore, the first five years are disregarded in the further analysis.

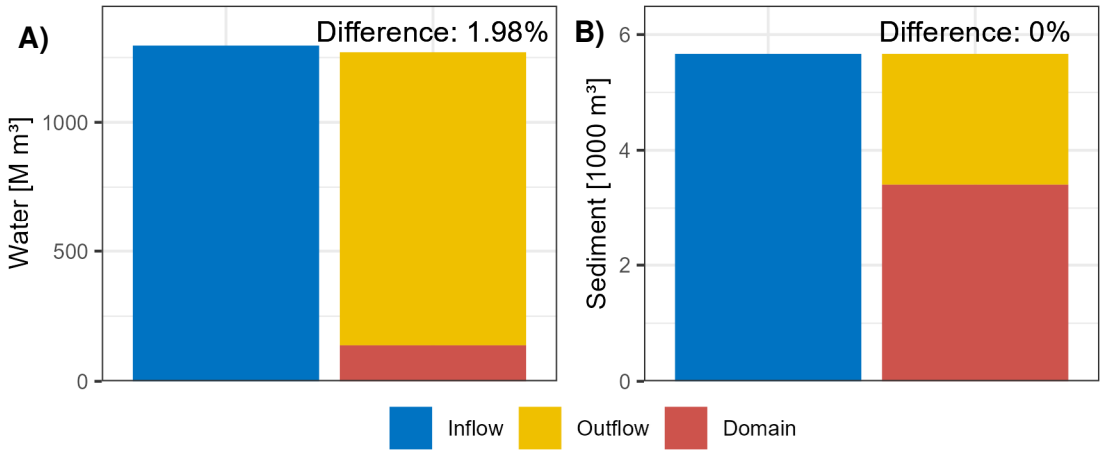


Figure 6.6: The mass balance of the C-L simulations. Figure A presents the water balance and Figure B presents the sediment mass balance. The domain refers to the mass which is stored within the modelling domain at the end of the simulation. The mass balance is calculated from a model run with an influx of 5,000 m^3/s and sediment influx at transport capacity for 72 hours.

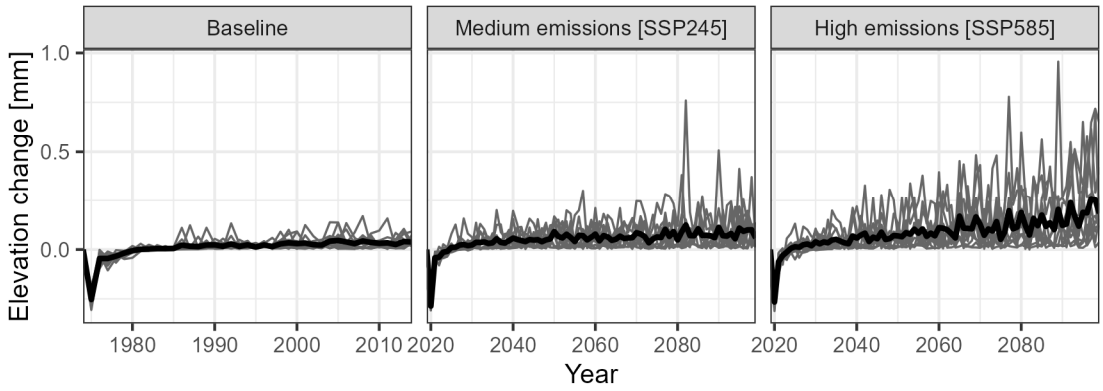


Figure 6.7: The mean topographic change of the fan to the previous year. The black line indicates the ensemble mean and the grey lines present the evolution of the ensemble members for the median flood discharge.

6.4.2 The predicted morphological evolution

The simulations predict morphological changes at different spatial scales from a few hundred metres to a few kilometres. The fan is divided into three zones with distinct characteristics to analyse the large-scale changes. These three zones are:

- the fan head which covers the area between 0 – 10 km downstream of the fan inlet (the fan inlet is the mountain outlet). This is the steepest fan section with a channel slope of ~ 0.002 m m^{-1} (Dingle *et al.*, 2020);
- the fan centre which covers the area between 10 – 20 km downstream of the fan inlet. The Karnali flows in two branches (the Eastern and the Western branches) and the slopes decrease from ~ 0.002 m m^{-1} to ~ 0.0015 m m^{-1} (Dingle *et al.*, 2020b);
- the lower fan which covers the area between 20 – 30 km downstream of the fan inlet. The channel slopes decrease further to ~ 0.001 m m^{-1} in this zone (Dingle *et al.*, 2020b).

6.4.2.1 Longitudinal evolution

The highest morphological activity is predicted for the fan head in the upstream 10 km of the fan (Figure 6.8). The single-threaded channel degrades at the inlet and aggradation is predicted after the first kilometre where the Karnali River widens as it enters the fan. The largest aggradation is predicted at the transition from the single-threaded channel to the braided channels 2-3 km downstream of the inlet. A second smaller aggradation peak is predicted at the bifurcation into the Eastern and Western branches 7 km downstream of the inlet.

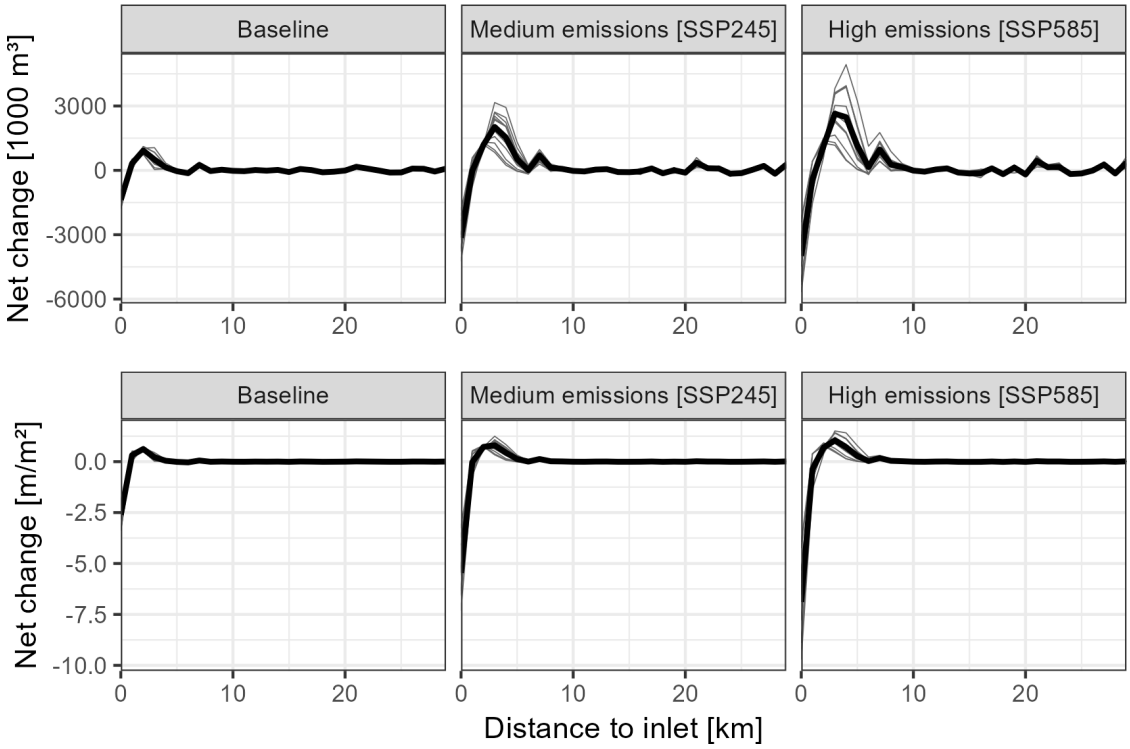


Figure 6.8: The longitudinal fan evolution of the CIMP6 scenarios at 1 km resolution predicted for the median flood discharge. The top row shows the total predicted topographic change and the bottom row indicates the mean topographic change of the active fan area (grid cells with simulated elevation change). The solid line indicates the ensemble mean and the grey lines the individual CMIP6 members.

The geomorphic activity decreases in the fan centre (10 – 20 km downstream of the inlet) and the lower fan (20 – 30 distance to the inlet). The predicted change alternates between degradation and aggradation in the fan centre but this change is low and the channel remains comparatively stable (Figure 6.8). The morphological change increases in the lower fan whereas alternating sections of degradation and aggradation are also predicted in this zone. However, the predicted changes are much lower than predicted for the fan head. This difference in upstream and downstream activity is particularly pronounced when considering the mean predicted change in the active fan area (the area for which changes are predicted). The ensemble mean change for the baseline scenario ranges between -2.6 m/m^2 to 0.6 m/m^2 in the fan head, between -0.008 m/m^2 and 0.007 m/m^2 in the fan centre, and between -0.006 m/m^2 and 0.01 m/m^2 in the lower fan (Figure 6.8).

The predicted spatial patterns of degradation and aggradation are consistent for the scenarios, but the magnitude increases with the emissions (Figure 6.8). The degradation at the inlet increases from -2.6 m/m^2 in the baseline to -5.5 m/m^2 and -6.9 m/m^2 in the medium- and high-emission scenarios, respectively. Consequently, more sediment deposits at the transition from the single-threaded to the braided channel (2-3 km

distance to the inlet) and at the bifurcation (7 km downstream of the inlet) and hence these depositional areas expand. The magnitudes of degradation and aggradation also increase further downstream in the fan centre and lower fan.

Similar trends are predicted for the projected scenarios but some sections switch the mode (degradation and aggradation) between the baseline and the projected scenarios. The ratio of the morphological change between the baseline and projected scenarios is shown in Figure 6.9. A negative ratio indicates a shift between degradation and aggradation. Such shifts occur at the transition between degradation and aggradation zones and are caused by a spatial extension of the aggradation zone (e.g. at 5, 8, and 23 km distance to the inlet) or an extension of the degradation zone (e.g. at 1, 13 -16 km distance to the inlet). The ratios of the medium- and high-emission scenarios follow a similar trend whereas it is more pronounced for the high-emission scenario. However, the ratios of both scenarios are similar in some sections (e.g. at 15 – 17, 24 – 26 km distance to the inlet). This might indicate that the differences between the baseline and projected scenarios are caused by the longer simulation length (baseline: 35 years, projected scenarios 75 years) rather than the projected changes in the flood discharge.

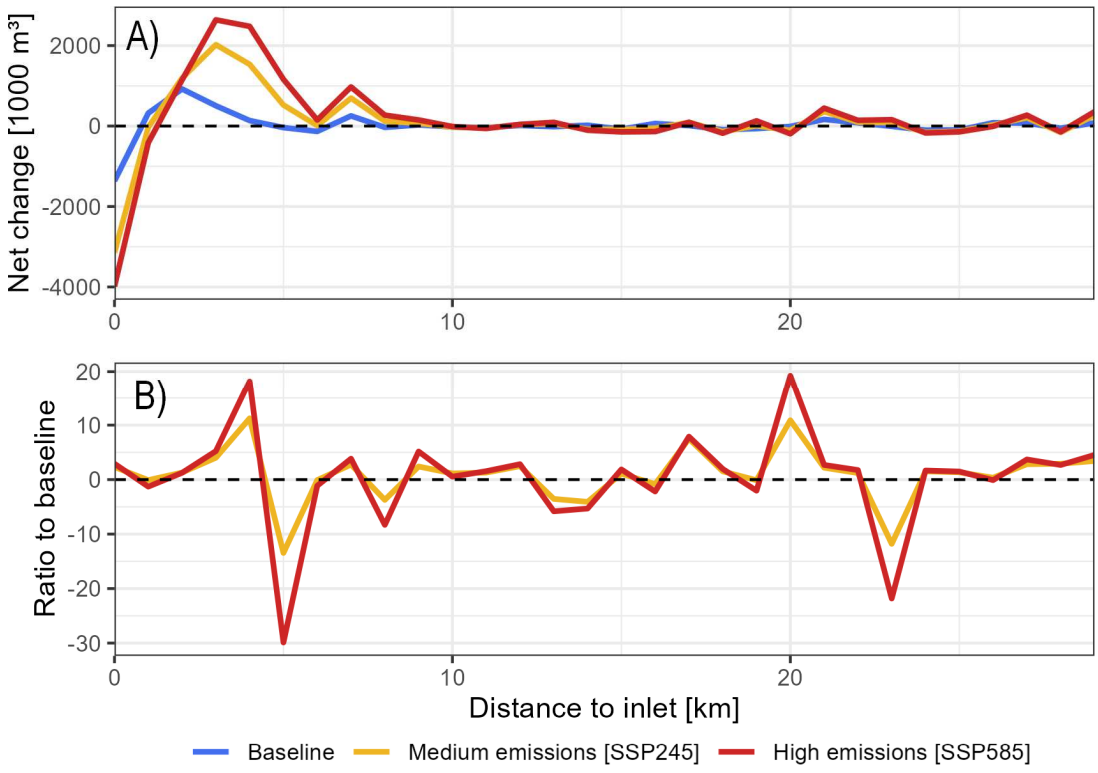


Figure 6.9: A) The mean (CMIP6 ensemble) topographic change at the end of the simulation period for the median flood discharge. B) The ratio of the topographic change between the ensemble means of the baseline and projected scenarios. A negative ratio indicates a shift between degradation and aggradation. The topographic change is the net change at 1 km resolution and is predicted for the median flood discharge.

A net aggradation is predicted for the fan head and the lower fan while a net degradation is predicted for the fan centre (Figure 6.10). The zonal sediment budget indicates that the fan head is the most active zone and acts as a sink for the sediments discharged into the fan. The sediment discharge increases with the flood discharge (assumption of sediment transport at capacity) and hence the amount of stored sediments increases from 0.6 Million m^3 in the baseline to 3 Million m^3 and 4.6 Million m^3 in the medium-emission and high-emission scenarios whereas a part of this increase is also facilitated by the longer simulation length of the projected scenarios. A net loss of 0.2 – 0.3 Million m^3 is predicted for the fan centre for the three scenarios. Hence, this area is less sensitive to the flood discharge changes compared to the other zones. The lower fan stores 0.2 Million m^3 in the baseline which equals the net loss of the fan centre. However, for the projected scenarios, the stored sediment increases to 0.5 Million m^3 and 0.7 Million m^3 and exceeds the sediment loss of the fan centre. This means that some sediments originating from the upstream catchment and/or from the fan head must bypass the fan centre and deposit in the lower fan.

The aggradation of the fan head (0 – 10 km distance to the inlet) increases with the simulation time for each of the three scenarios whereas this trend is more pronounced for the projected scenarios (Figure 6.11). The mean sediment aggradation is 0.6 Million m^3 after 35 years in the baseline scenario, and 1.1 Million m^3 and 1.3 Million m^3 in the medium-emission and high-emission scenarios, respectively. The projected change increases linearly after 35 years in the medium-emission scenario (after the year 2060) while it intensifies in the high-emission scenario, particularly after the year 2080.

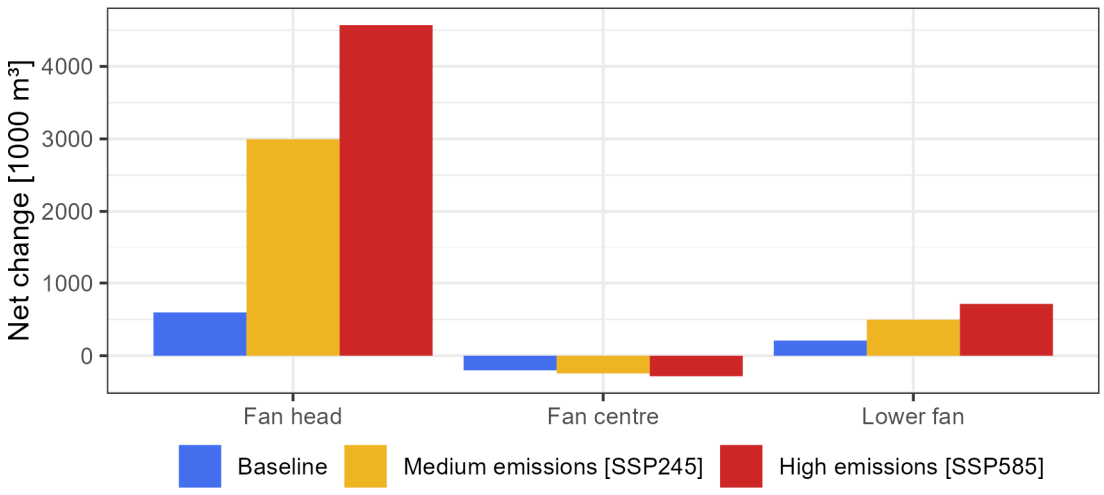


Figure 6.10: Ensemble mean topographic change for the fan head, the fan centre, and the lower fan for the median flood discharge. These areas are classified by the distance to the inlet whereas the fan head the area at 0 – 10 km distance, the fan centre at 10 – 20 km, and the lower fan at 20 – 30 km.

The model predicts the net degradation for the fan centre (10 – 20 km distance to the inlet) in all three scenarios. The degradation rate is highest in the first 10-15 years of the simulations after which it decreases. The fan remains stable with low predicted changes after the year 2000 in the baseline scenario. However, in the projected scenarios, the fan shifts from the net degradation to the net aggradation around the period 2060 – 2070. The lower fan (20 – 30 km distance to the inlet) is projected to aggrade continuously even after the predicted shift from degradation to aggradation in the fan centre. However, the aggradation rate decreases at the time of this upstream shift for the medium-emission scenario SSP245, while the aggradation rate do not change for the high-emission scenario SSP585 (Figure 6.11).

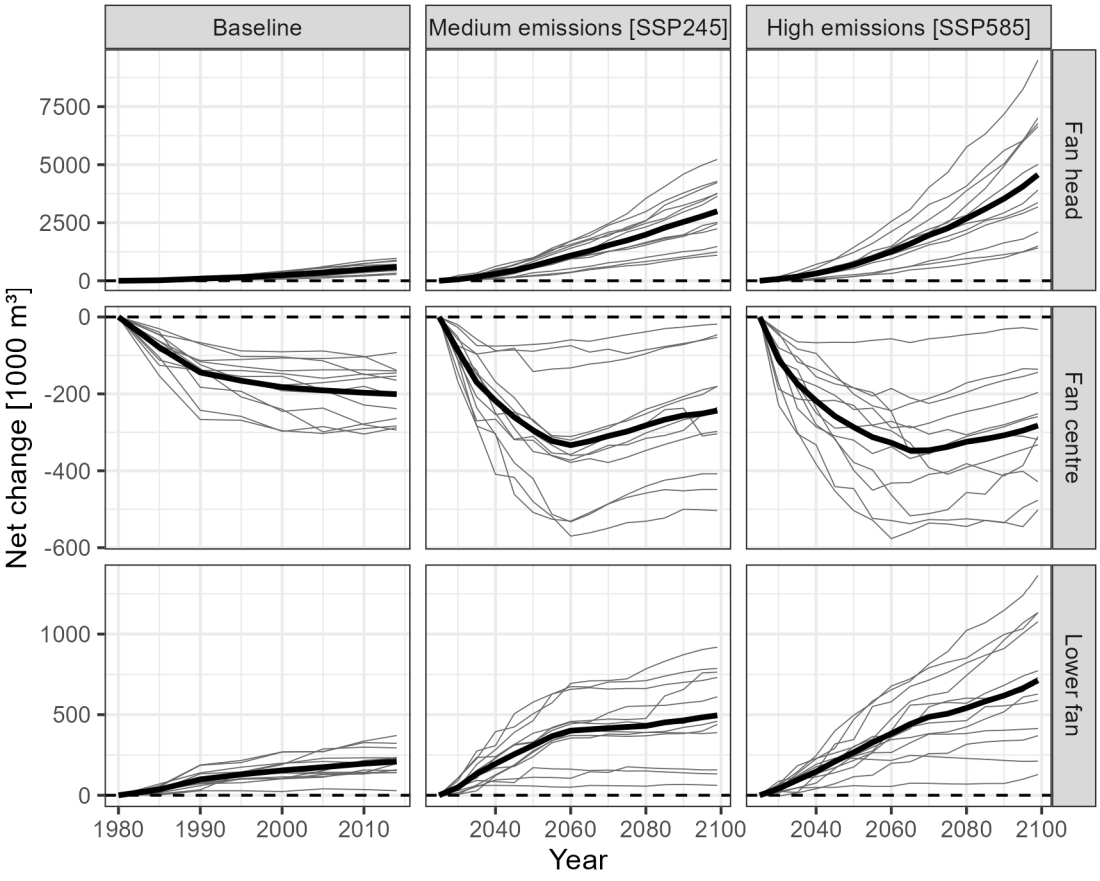


Figure 6.11: The predicted topographic change in time for the fan head (0 – 10 km distance to the inlet), the fan centre (10 – 20 km), and the lower fan (20 – 30 km). The solid line indicates the ensemble mean and the grey lines indicate the ensemble members. These changes are predicted for the median flood discharge.

The longitudinal evolution is characterised by alternating sections of degradation and aggradation on different spatial scales (Figure 6.12). The extent of these sections ranges between several hundred metres to a few kilometres. The largest sections are predicted at the fan head, particularly at the inlet (degradation), the conversion from the single-threaded to the braided channel (2 – 3 km distance to the inlet), and the bifurcation (7 km distance to the inlet) (both aggradation). Further downstream, extended sections are predicted (e.g. 24 – 26 km distance to the inlet) but these are intersected by sections with low changes. In general, the spatial scale of degradation and aggradation is smaller in the fan centre and lower fan. Here, the evolution can change multiple times between degradation and aggradation within a few kilometres (e.g. 12 – 15 km distance to the inlet).

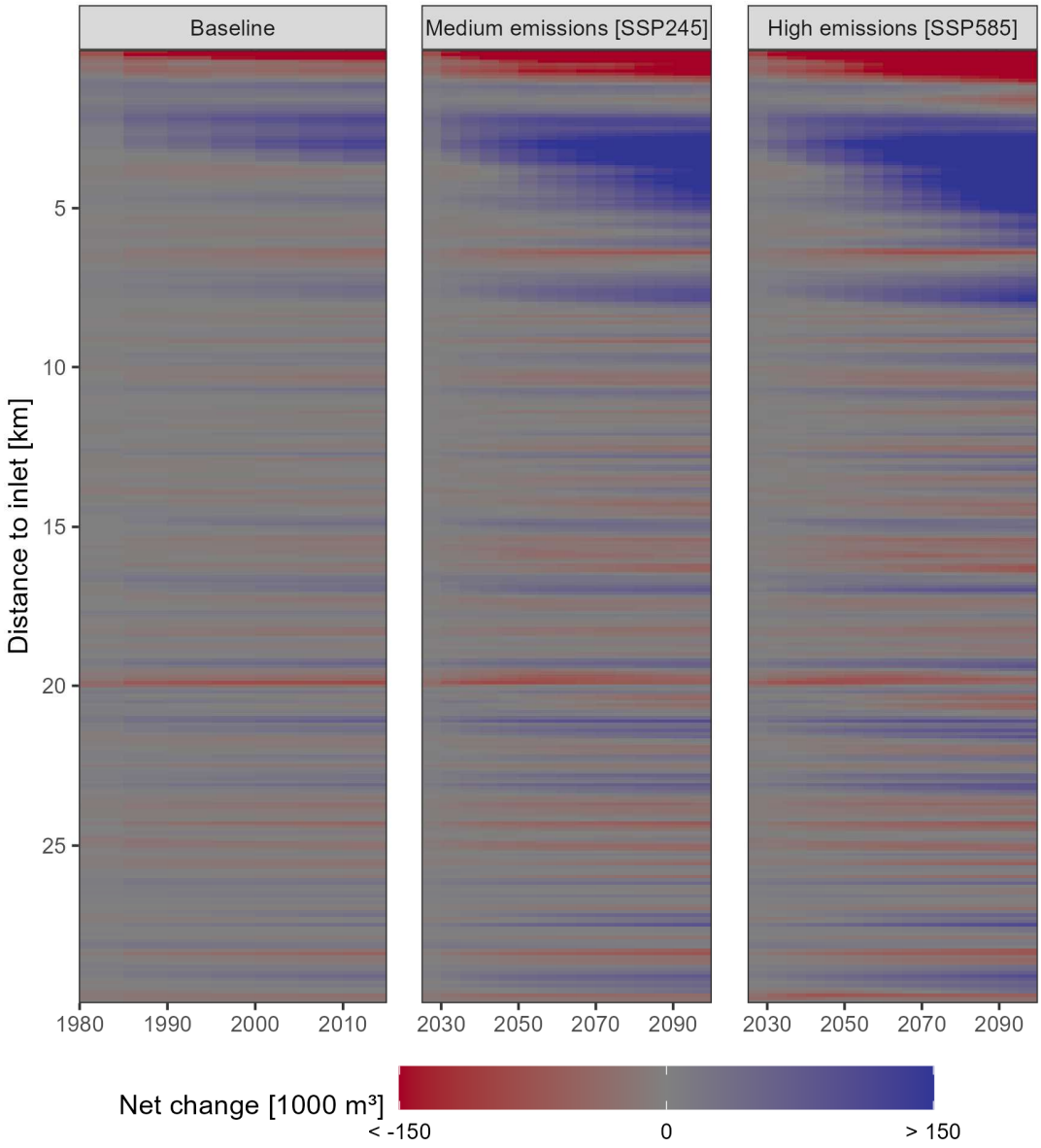


Figure 6.12: The ensemble mean topographic change at 100 m resolution for the simulation period. The a-axis shows the year of the simulation. The changes are predicted for the median flood discharge.

The predicted longitudinal evolution is consistent in time and space (Figure 6.12). Similar trends are predicted for each scenario at the same location. In other words, the sections with high degradation or aggradation are similar in the baseline and projected scenarios. However, the magnitude of the projected change increases in the projected scenarios, and the spatial extent of these zones increases which is most prominent in the fan head. In the fan centre and the lower fan, sections of degradation and aggradation extend further and the transitional zones with low predicted change decrease. The predicted trend intensifies with time in most sections. In a few sections however, the net change decreases at a certain point in time, indicating a shift from degradation to aggradation (e.g. 20 km distance to the inlet) or from aggradation to degradation (e.g. 15 km distance to the inlet).

6.4.2.2 Lateral evolution

The lateral evolution comprises in-channel, riverbank, and floodplain processes such as channel degradation and aggradation, bank erosion and development, and floodplain deposition. The predicted lateral evolution has two distinct characteristics; i) the geomorphic activity is highest in the channel and decreases with the distance and; ii) the evolution intensifies with emissions whereas the predicted spatial patterns remain similar (Figure 6.13). Furthermore, the predicted evolution differs in the fan head, fan centre, and lower fan.

The fan head is the zone with the highest predicted lateral changes but the evolution is constrained to the channel (Figure 6.13). The mean aggradation is 17 mm/m^2 in the baseline scenario which increases to 79 mm/m^2 and 124 mm/m^2 in the medium-emission scenario SSP245 and the high-emission scenario SSP585, respectively. The higher aggradation rates of the projected scenarios are, in part, caused by the longer simulation length (75 years compared to 35 years of the baseline). However, the difference between both projected scenarios indicates that the evolution is influenced by the increase in projected flows and the higher sediment influx. The geomorphic activity decreases with the channel distance and no changes are predicted at ≥ 500 m distance. This is because the channels cover most of the fan head.

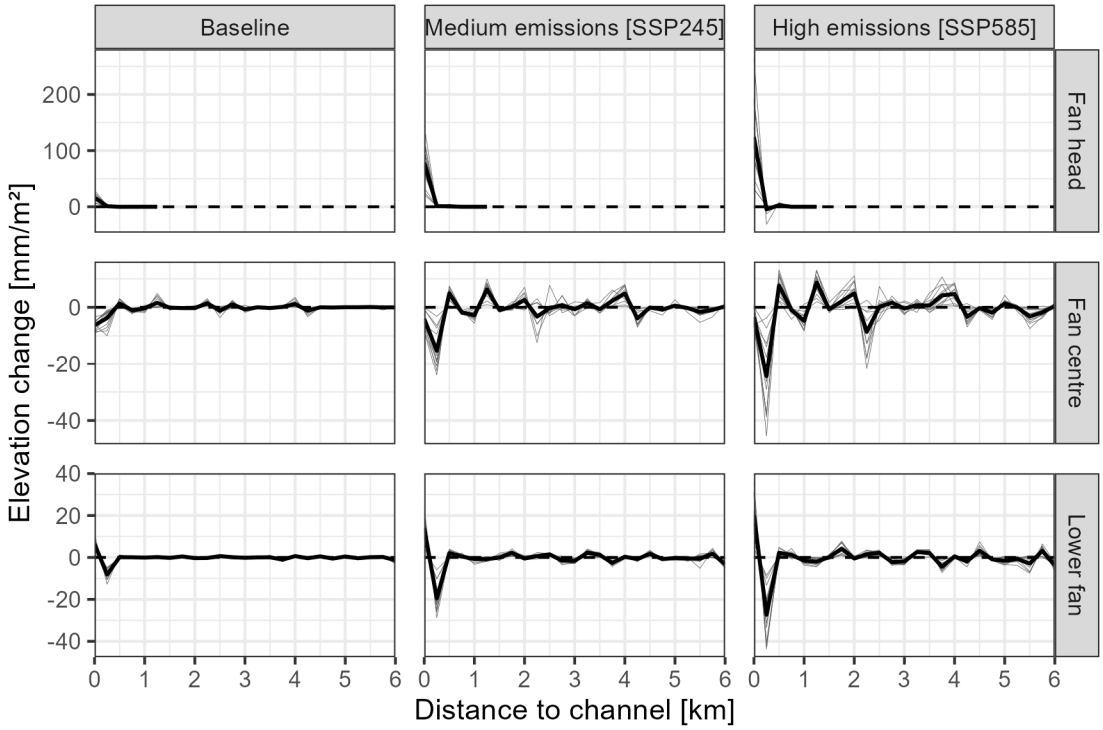


Figure 6.13: Mean topographic change at the end of the simulation periods with distance to the channel for the fan head (0 – 10 km), fan centre (10 – 20 km), and the lower fan (20 – 30 km). The solid line indicates the ensemble mean change and the grey lines the individual ensemble members. The channel is defined as the inundated area during medium flow (3,000 m^3/s) for the initial topography after the initialisation period. The changes are predicted for the median flood discharge and are the net change predicted for each zone.

The fan centre is characterised by the degradation of the initial channels. Alternating patterns of degradation and aggradation are predicted for the floodplain with a channel distance ≥ 500 m whereas the areas of degradation are smaller channels which drain the floodplain. However, the mean elevation change of the floodplain (> 500 m distance to the channel) is < 1 mm/m^2 in each scenario, which indicates that the floodplain is reworked but does not act as a significant storage for the upstream sediments. The channel (≤ 250 m distance to the channel) degrades but this degradation decreases with the emissions from 6 mm/m^2 in the baseline to 4 mm/m^2 and 3 mm/m^2 in the medium-emission and high-emission scenarios, respectively. However, the degradation of the transition zones (250 – 500 m distance to the channel) increases from 4 mm/m^2 (baseline) to 15 mm/m^2 (SSP245) and 24 mm/m^2 (SSP585).

The channels in the lower fan aggrade by 6 mm/m^2 (baseline) which increases to 14 mm/m^2 (SSP245) and 20 mm/m^2 (SSP585). The transitional zones (250 – 500 m distance to the channel) degrade by 8 mm/m^2 (baseline), 19 mm/m^2 (SSP245) and 27 mm/m^2 (SSP585). The mean change in the floodplain (> 500 m distance to the channel) is < 1mm/m^2. Therefore, the predicted lateral evolution is similar to the one predicted for

the fan centre except for the channels which aggrade in the lower fan and degrade in the fan centre.

6.4.2.3 The evolution of the Eastern and Western branches of the Karnali River

The Karnali River splits after the bifurcation (7 km downstream of the inlet) into the Eastern and Western branches. The Eastern branch is characterised by braided channels throughout the fan, while the channels of the Western branch are generally wider and confluence to a single channel in several sections. These different channel characteristics manifest in different predicted evolutions. Higher change rates are predicted for the Eastern branch at 10 – 15 km distance to the inlet, particularly in the baseline scenario (Figure 6.14). Further downstream, higher change rates are predicted for the Western branch in each scenario (15 – 27 km distance to the inlet). Similar change rates are predicted at the downstream end of the modelling domain (≥ 27 km distance to the inlet). In the lower fan, aggradation dominates the evolution of the braided Eastern branch. In the Western branch, aggradation sections are intersected by sections with high degradation in the lower fan (e.g. 24 – 25 distance to the inlet) (Figures 6.14 and 6.15).

The model predicts a different lateral evolution of both branches (Figure 6.15). Generally, the predicted changes are larger in the Western branch, both in the channel (lateral distance < 250 m), and the floodplain (lateral distance ≥ 250 m). One distinct difference is the predicted change at 250 – 500 m channel distance. In the Western branch, this section is dominated by degradation, and aggradation is only predicted in two areas (10 and 23 km distance to the inlet). The predicted rates are particularly large between 17 – 30 km distance to the inlet. In the Eastern branch, the highest degradation rates are predicted between 10 – 15 km distance to the inlet, and ≥ 20 km distance to the inlet the evolution is dominated by aggradation. Alternating patterns of degradation and aggradation are predicted for both branches in areas with ≥ 500 m channel distance, whereas these changes are larger in the Western branch. However, in both branches, the predicted changes increase with the emissions, but the spatial patterns remain similar.

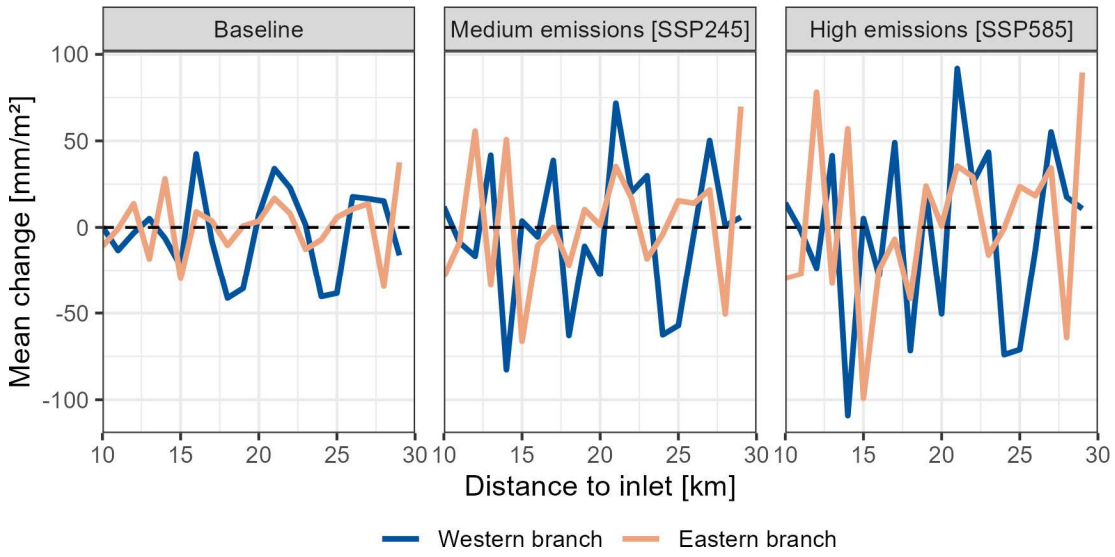


Figure 6.14: Ensemble mean longitudinal change at the end of the simulation period for the Eastern and Western branches of the Karnali River. Areas with a greater distance to the initial channel of 500 m are excluded from the calculation of the mean. These changes are predicted for the median flood discharge.

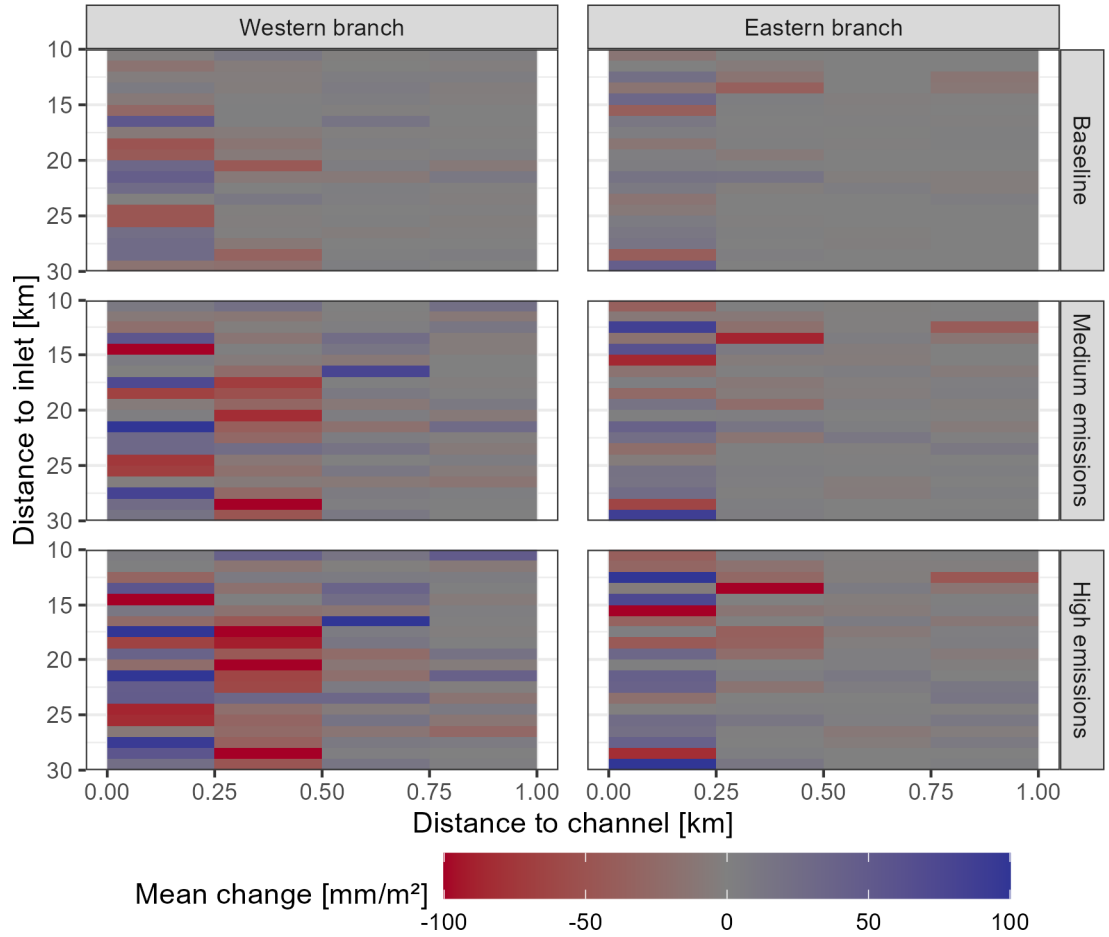


Figure 6.15: The ensemble mean longitudinal and lateral elevation change at the end of the simulations of the Eastern and Western branches. The distance to the channel refers to the initial channel after the initialisation periods. These changes are predicted for the median flood discharge.

6.4.2.4 The predicted changes at modelling resolution

The predicted changes at the modelling resolution (25 m) of the three scenarios are presented as DEMs of Difference (DoD) in Figure 6.16. These DoDs indicate that the areas of degradation and aggradation are in proximity. Furthermore, the predicted changes are in the magnitude of metres which is considerably larger than the scale of the mean changes in the longitudinal and lateral directions (mm – cm). This emphasizes that a significant contribution of the predicted changes is the redistribution of fan sediments on the local scale. The model predicts two different modes of sediment redistribution:

- The redistribution of bank material: The river banks erode and this material deposits within tens to hundreds of metres in the channels. This process dominates the evolution of the channels throughout the fan (e.g. the inset of the Eastern branch in Figure 6.16).
- The redistribution of channel material: The simulations predict the channel degradation and the downstream channel aggradation. This process is predicted in in channel sections where the river flows in a single channel (e.g. at the fan inlet), or in sections in which the channel width decreases (for example the inset of the Western branch in Figure 6.16). This eroded channel material is then deposited further downstream where the channel width increases. This redistribution occurs on a scales of up to a few kilometres.

Furthermore, the model predicts the channel aggradation downstream of channel bifurcations. This is most distinctively predicted in the channel head at the conversion from the single-threaded to the braided channel. However, this behaviour is predicted throughout the fan but is superimposed by the lateral redistribution.

The locations of degradation and aggradation are consistent in the different scenarios. However, the magnitude and the spatial extent of these degrading or aggrading areas increase with the emissions. One difference between the scenarios is the evolution of the floodplains. Few areas along small channels in the floodplains change in the baseline scenario but this area expands in the projected scenarios (Figure 6.16). However, the change concentrates on the Western floodplain and the Island between both branches, while the areas located east of the Eastern branch remain unchanged in each scenario.

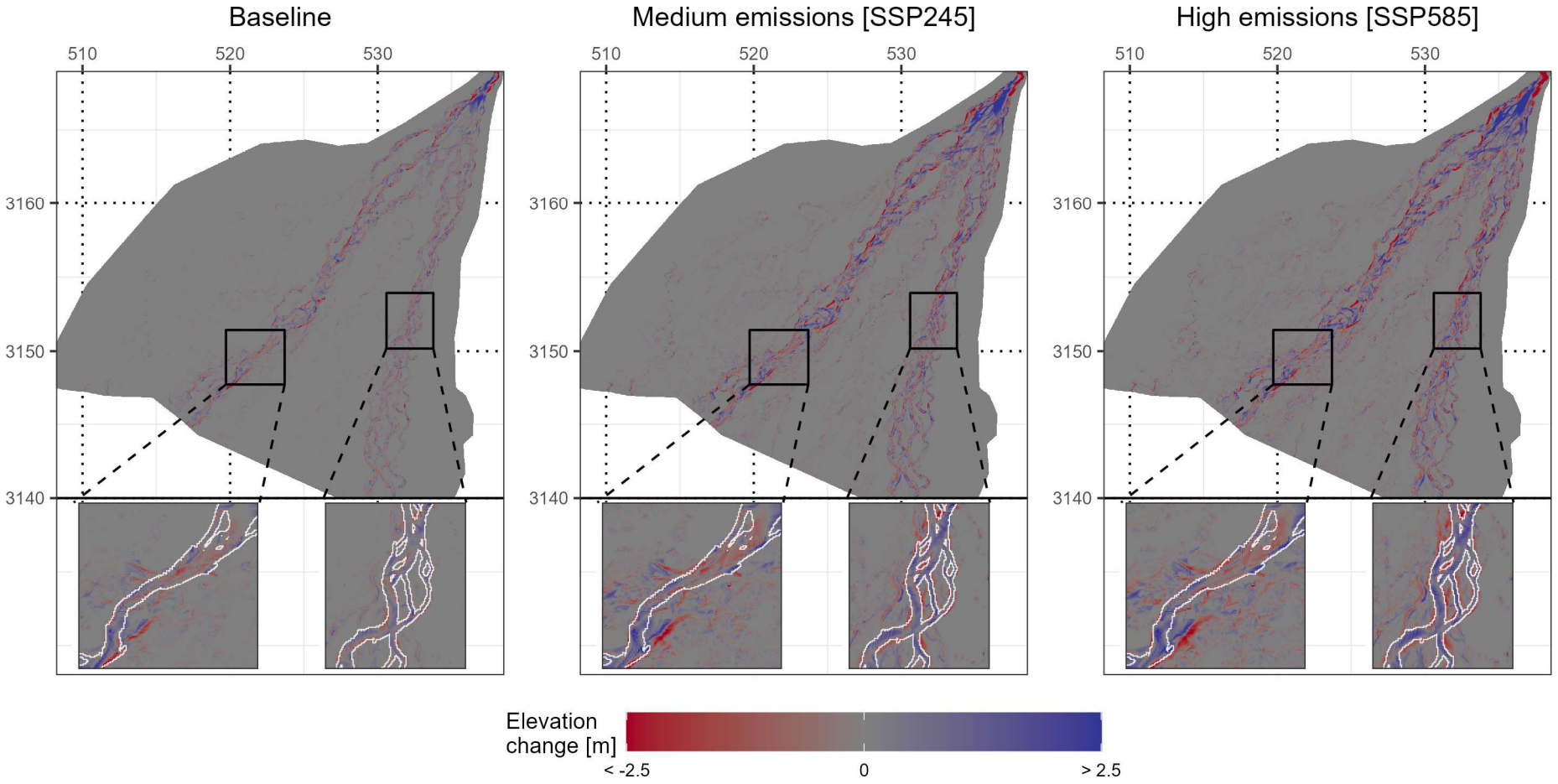


Figure 6.16: DEM of Difference (DoD) of the three CMIP6 scenarios. These DODs show the ensemble mean elevation difference between the topographies at the beginning of the modelling period (after the model initialisation; baseline: 1980; projected scenarios: 2025) and the end of the modelling period (baseline: 2014; SSP245 and SSP585: 2099). Positive values indicate predicted aggradation and negative values indicate predicted degradation. The changes are predicted for the median flood discharge. The white lines in the insets present the outline of the initial channel at 3,000 m^3/s. The coordinate system is WGS84 UTM Zone 44 N in kilometres.

6.4.2.5 The predicted evolution of the ensemble members

The individual ensemble members predict similar spatial patterns of degradation and aggradation whereas the manifestation of these patterns increases with the emission scenario, the inflow water and sediments, and the inflow percentile (the percentiles of the hydrological ensemble predictions) (Figure 6.17). This behaviour is most pronounced in the fan head (0 – 10 km distance to the inlet) where all models predict the degradation at the inlet, the aggradation at the conversion from the single-threaded to the braided channel (2 - 3 km distance to the inlet), and the aggradation at the bifurcation (7 km distance to the inlet).

The spatial extent and the magnitude of degradation and aggradation increase with the emissions and the inflow percentile. The initial degradation expands, the 1st aggradation zone expands and shifts downstream, and the 2nd aggradation zone expands with emissions and the inflow percentile (Figure 6.17). These trends are also predicted for the ensemble members (note that the ensemble members are sorted in decreasing order according to the mean (time) inflow rate of the SSP585 scenario). This is particularly evident when comparing the position and magnitude of change in these zones of the MPI-ESM1-2-HR (lowest flood flows) and the ACESS-ESM1-5 (highest flood flows) members for the projected scenarios.

The ensemble members predict similar patterns of degradation and aggradation for the fan centre (10 – 20 km distance to the inlet) and lower fan (20 – 30 km distance to the inlet). Degradation and aggradation are predicted at similar locations for each ensemble member (Figure 6.17). The predicted changes increase with the emissions and inflow percentile, and the area increases by a few 100 m for simulations with higher inflows. However, no downstream shift of the centre of the degradation and aggradation zones is predicted unlike for the fan head. The predicted changes scale with the flood discharge, and the predicted change is generally lower for models with lower flood flows, as also predicted for the fan head.

The predicted morphological change scales with the flood discharge. Figure 6.18 presents the relationship between the mean (time) flood discharge and the predicted morphological change at the end of the simulations. A strong linear relationship is observed for the fan head and the lower fan. The predicted change of the baseline period is lower compared to the projected scenarios which is also affected by the shorter simulation period (35 years compared to 75 years).

The relationship between the inflow rate and the predicted morphological change is weaker for the fan centre, particularly for the high-emission scenario SSP585. The predicted degradation rates decrease above mean inflow rates ≥ 12,000 m^3/s and hence this relationship appears non-linear but rather v-shaped. This different relationship might be caused by the switch from degradation to aggradation after ~40 years of simulation (Figure 6.11). However, the predicted change is low compared to the fan head and the lower fan and hence the model predicts a strong linear relationship between inflow rate and morphological change.

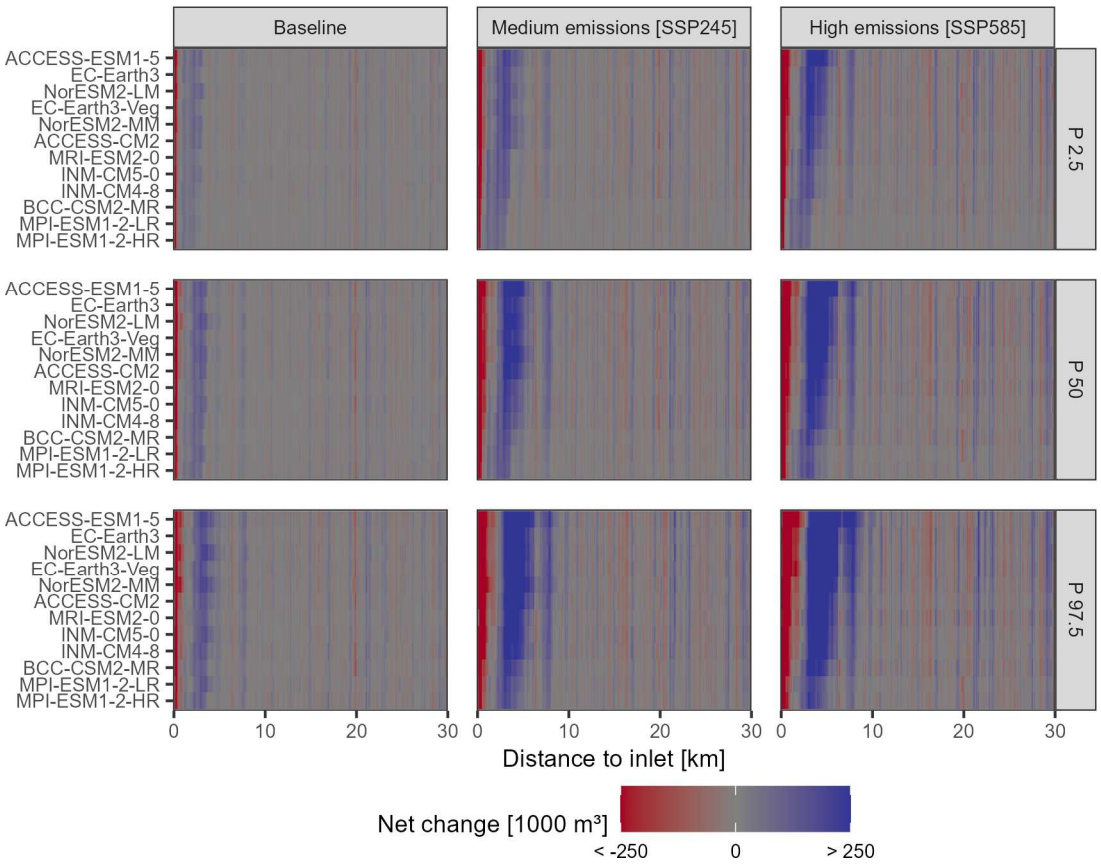


Figure 6.17: The longitudinal evolution of the individual ensemble members at 100 m resolution. The evolution is presented for the 2.5th, 50th and 97.5th percentiles of the flood discharge predictions. The models are sorted in decreasing order according to the mean (time) median (ensemble) flood discharge of the high-emission scenario SSP585.

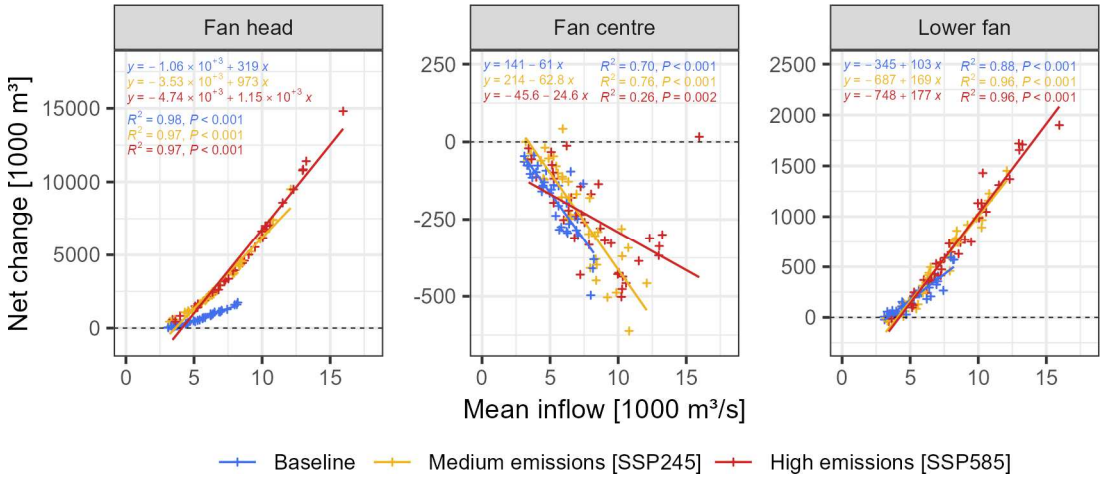


Figure 6.18: The relationship between the mean flood discharge (time) and the predicted topographic change at the end of the simulation period for the fan head (0 - 10 km distance to the inlet), the fan centre (10 - 20 km distance to the inlet), and the lower fan (20 - 30 km distance to the inlet). This plot combines the simulations of the $P_{2.5}$, P_{50}, and $P_{97.5}$ flood discharge percentiles.

6.4.3 The sensitivity of the morphological change to the sediment influx

The NorESM2-LM SSP585 scenario was simulated with different sediment influx rates to investigate the sensitivity of the morphological simulations to the assumption of sediment influx at the transport capacity. This model was run with the sediment influx at 0%, 50%, and 75% of the sediment transport capacity. Furthermore, one model was run with a random influx of 50 – 100% of the sediment transport capacity to investigate how varying sediment pulses affect the fan evolution. The NorESM2-LM SSP585 scenario was selected because it is one of the models with the highest predicted flood flows, and thus, sediment influxes.

The sediment influx strongly affects the evolution of the fan head (Figures 6.19 and 6.20). The aggradation peaks at the shift from the single-threaded channel to the braided channel (3 km distance to the inlet) and the bifurcation (7 km distance to the inlet) develop also for the simulation without any sediment influx. This indicates that the material of the single-threaded channel is deposited at the top of the fan. However, degradation is simulated between both peaks which indicates that the second peak at the bifurcation is, at least to some degree, a product of the reworking of fan material.

The differences between the sediment scenarios are largest at the transition of single-threaded to the braided river, whereas the predicted aggradation rates increase with the sediment influx. Therefore, most of the sediment influx is stored within a zone of 3 – 5 km distance to the inlet. Hence, the predicted change scales with the sediment influx which is also indicated by the sediment budget (Figure 6.20). The fan head degrades if no sediment enters the modelling domain and it aggrades increasingly with higher sediment influx.

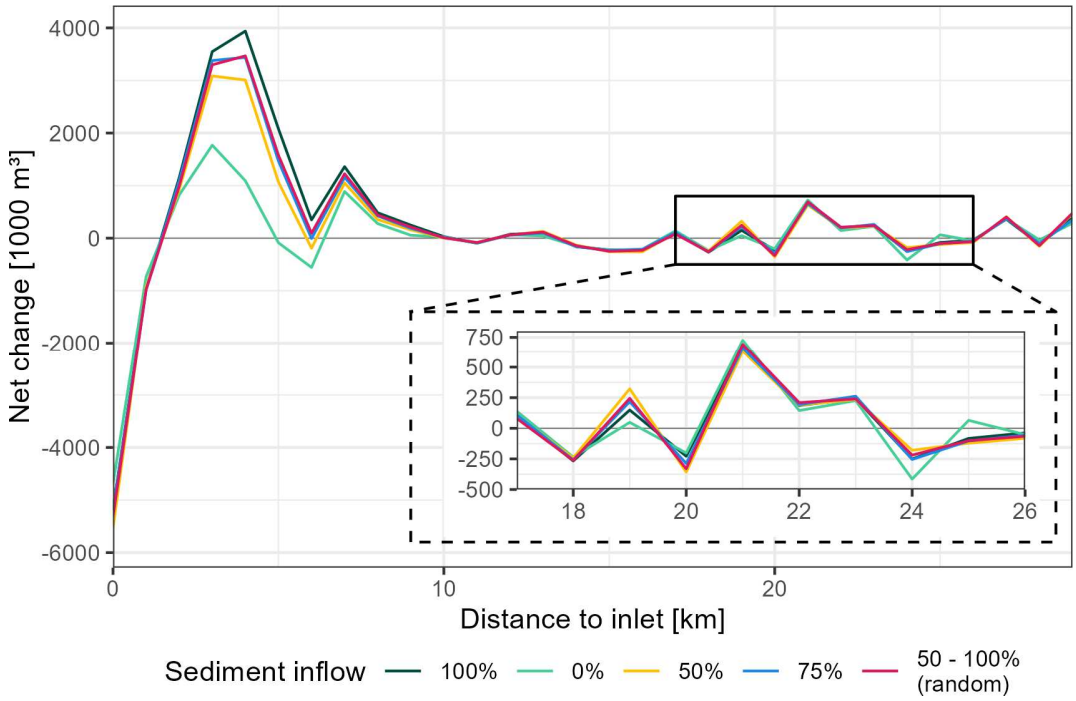


Figure 6.19: Topographic change of the NorESM2-LM SSP585 (median flood discharge) scenario for different sediment scenarios. The topographic change is the net change at 1 km resolution. The sediment influx is scaled with the transport capacity (0%: no sediment influx; 100%: sediment influx at transport capacity). The influx of the 50 – 100% sediment scenario varies randomly between the years and is within 50 – 100% of the transport capacity.

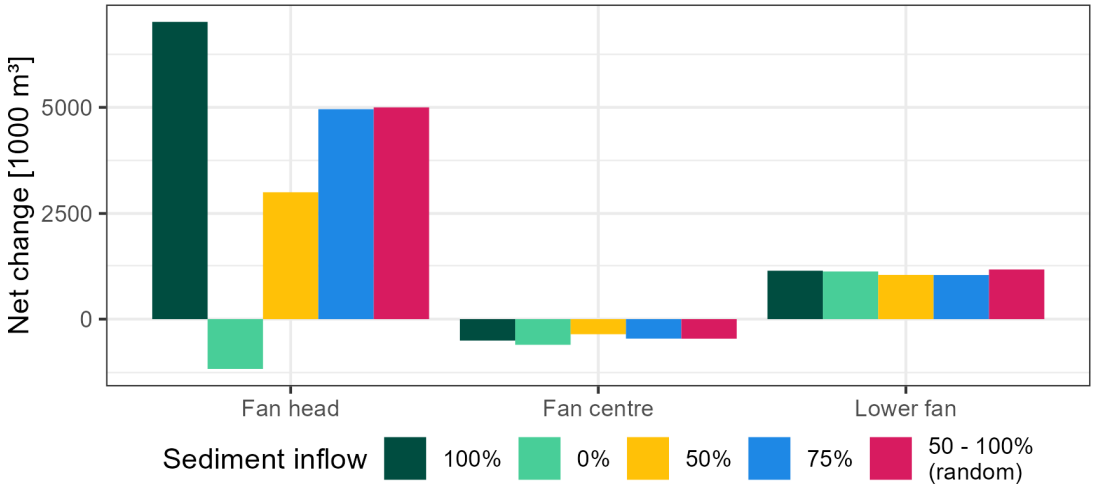


Figure 6.20: The sediment budget of the fan head (0 – 10 km distance to the inlet), the fan centre (10 – 20 km), and the lower fan (20 – 30 km) for the different sediment influxes in percentage of the transport capacity.

The longitudinal evolution of the fan centre and the lower fan are insensitive to the sediment influx (Figures 6.19 and 6.20). The spatial patterns of degradation and aggradation remain similar regardless of the sediment influx, and for most parts, the scenario differences are barely noticeable. However, scenario differences manifest in two regions, 19 – 20 km and 25 – 25 km downstream of the inlet. The simulation without sediment influx has a lower aggradation rate at 19 km downstream of the inlet and a lower degradation rate at 20 km downstream of the inlet than the simulations with sediment influx. Further downstream, at 24 km distance to the inlet, the simulation without sediment influx degrades stronger than the scenarios with sediment influx. The section at 25 km distance to the inlet aggrades in the 0% sediment influx scenario, whereas this section degrades for the remaining sediment scenarios. Interestingly, the scenario differences do not scale linearly in both sections. The differences between the simulation without sediment influx and 50% sediment influx are larger than the ones between no sediment influx and 100% sediment influx (Figure 6.19).

The fan centre degrades in each sediment influx scenario whereas this degradation does not scale linearly with the sediment influx as the highest degradation is predicted for the 0% sediment influx scenarios and the lowest one for the 50% sediment influx (Figure 6.20). This non-linearity is also predicted for the lower fan whereas the scenario differences are very low. This indicates that the aggradation of the lower fan is caused by sediments originating from the fan head and not from the sediment influx. Furthermore, the degradation of the fan head and fan centre of the 0% sediment scenario (1,8 Million m^3) exceeds the aggradation of the lower fan (1,1 Million m^3) which indicates that some fan sediments are discharged from the modelling domain.

The Western branch is more sensitive to the sediment influx than the Eastern branch (Figure 6.21). The sediment scenario differences in the fan centre and the lower fan (Figures 6.19 and 6.20) are mainly caused by differences predicted in the Western branch. The Eastern branch is insensitive towards the sediment influx except for the first few kilometres after the bifurcation. The Western branch is more sensitive, particularly in the lower fan where the channel is wider and less braiding occurs. However, the fan head is the by far most sensitive towards the sediment influx.

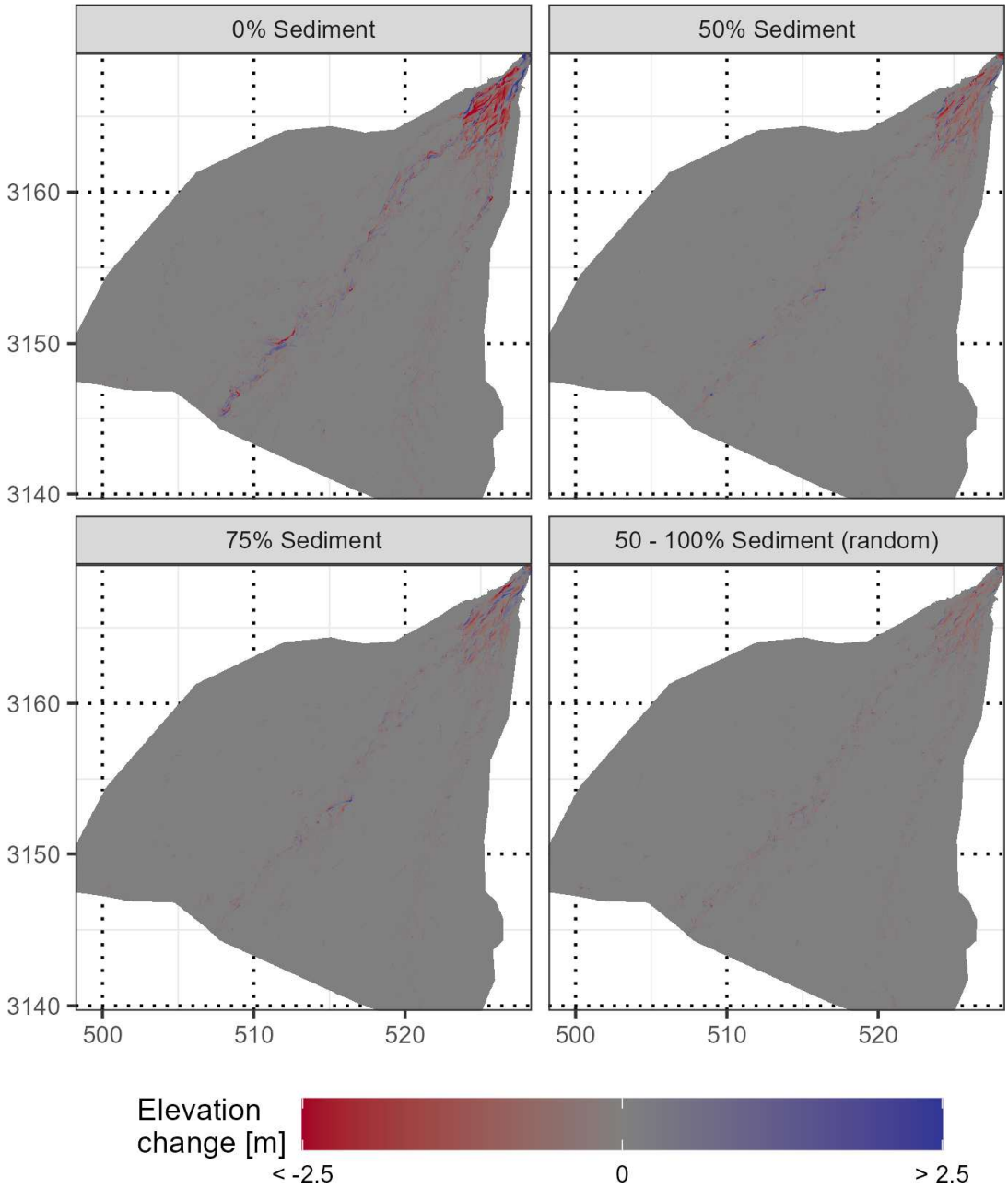


Figure 6.21: DEMs of Difference (DoD) of the NorESM-LM SSP585 sediment scenarios (median flood discharge). These maps show the elevation difference between the predictions of the sediment scenarios (influx at 0%, 50%, 75%, 50 – 100% of the transport capacity) and the predictions with sediment influx at the transport capacity (DEM sediment scenario – DEM transport capacity). The coordinate system is WGS84 UTM Zone 44 N in kilometres.

6.4.4 The evolution of flow pathways

The flow pathways are controlled by the topography and may change with the morphological evolution. The observed changes in the channel location provide information about topographic changes which can be used to evaluate the model performance (Section 6.4.4.1). The largest topographic changes are predicted in the fan head which is a crucial area because it controls the distribution of flow to the downstream fan. The impact of these predicted changes on the flow pathways is presented in Section 6.4.4.2.

6.4.4.1 Observed and simulated changes in the channel locations

The observed changes in the channels concentrate on the fan head and the Eastern branch. Figure 6.22 shows the observed channels in the period 1990 - 2021 classified from Landsat images (Landsat 5 – Landsat 8) using the Automated Water Extraction Index (AWEI) (Feyisa *et al.*, 2014). The channels are characterized by stable sections which have been active throughout the past 30 years and unstable sections in which the channels migrate and are only active during a few years. While stable and unstable sections are observed across the whole fan, the stable sections dominate in the Western branch. The channels in the fan head and the Eastern branch are mostly unstable except for the first 2-3 km at the upstream model boundary and the last 2-3 km at the downstream model boundary. The change of the location of channel bars, bends and banks occurs on the scale of 50 – 500m in unstable sections and these changes occur between the acquisition of consecutive images (~1 year).

The flood event in 2009 altered the topography at the bifurcation and shifted the allocation of water in favour of the Western branch (Figure 6.23). This event indicates the potential of flood flows to alter the topography and the flow pathways. However, the two largest flood events in the past 30 years (2013 and 2014) have not caused a significant change in the location of the channels of the Eastern branch (Figure 6.23). This indicates that the channel mobility decreased in the Eastern branch after the bifurcation shift in 2009. Conversely, these floods have caused the reshifting of the location of multi-threaded channels in the Western branch which indicates that the channel stability in this branch decreased after the flood event. However, these changes occur within the bankfull channel which might be attributed to the ongoing construction of embankments since the year 2011.

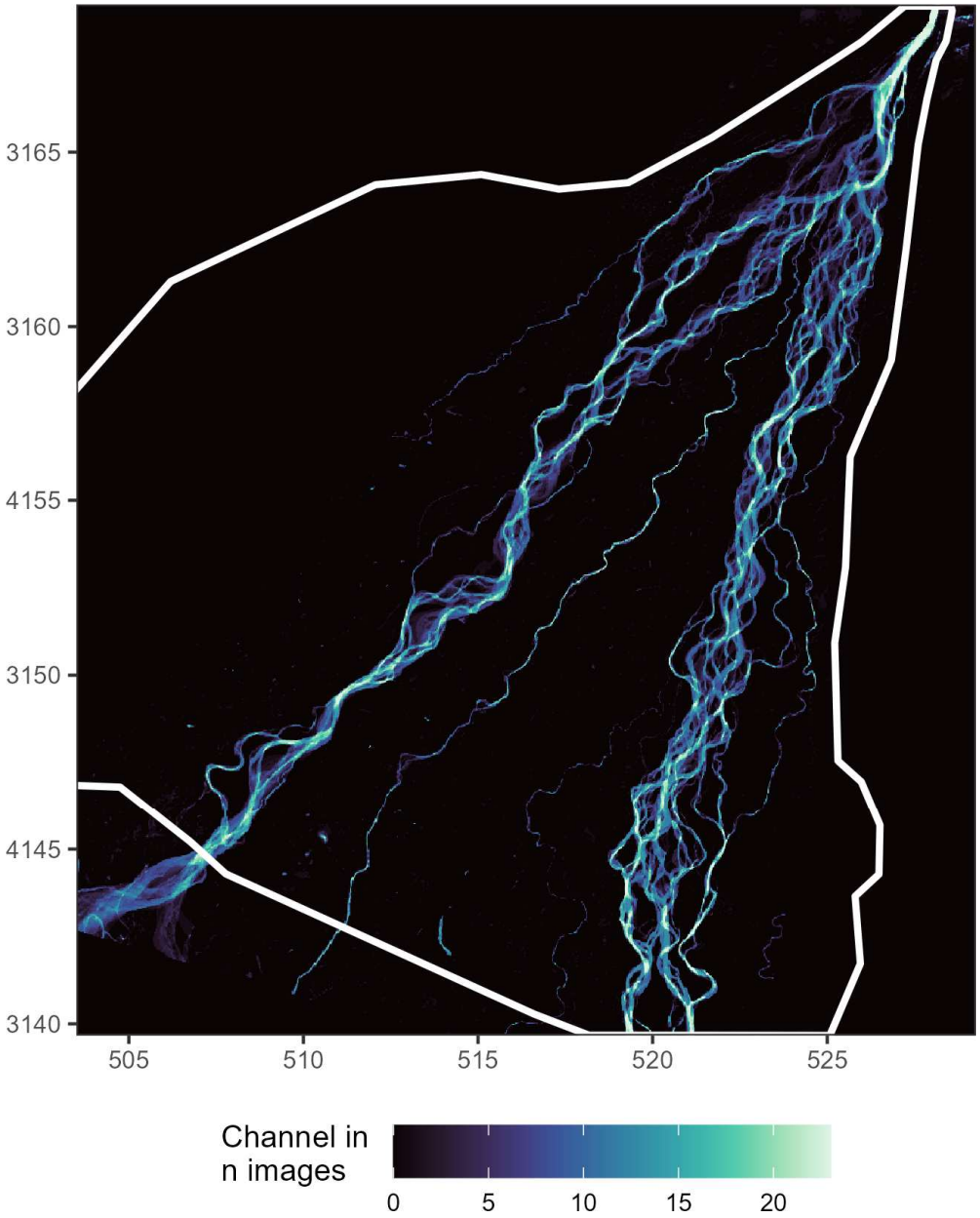


Figure 6.22: Observed channel composite generated from 23 Landsat-5 to Landsat-8 images covering the period 1989 - 2021. The images are taken during the post-monsoon season or, if post-monsoon season images are unavailable, from the pre-monsoon season. Images captured between 2003 – 2012 have data gaps due to an error of the Landsat-7 sensor and are not used for the composite image. The water is classified by the Automated Water Extraction Index (AWEI). The observed discharge at the fan inlet (DHM) of the acquisition dates ranges between 300 m^3/s and 1470 m^3/s (median: 1,000 m^3/s). The discharge of the images taken after 2016 is unknown. The white line indicates the boundary of the modelling domain. The coordinate system is WGS84 UTM Zone 44 N in kilometres.

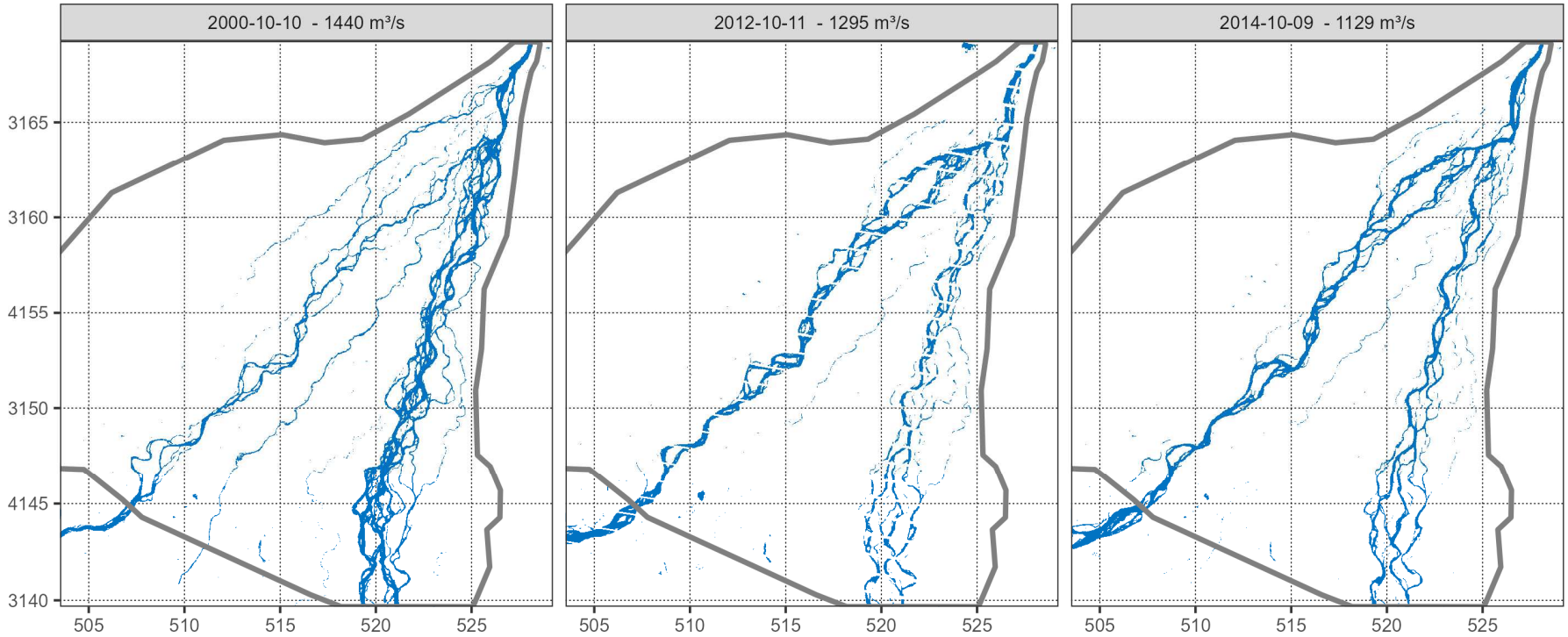


Figure 6.23: The channels observed in individual Landsat-7 and Landsat-8 images. The images are captured before (left) and after (centre) the bifurcation change in 2009. The two largest flood events since 1990 (17,000 – 18,000 m^3/s) occurred between the acquisition of the images in 2012 (centre) and 2014 (right). The striped data gaps in the 2012 image are caused by the failure of the Landsat-7 scan line corrector. The water is classified by the Automated Water Extraction Index (AWEI). The grey line indicates the boundary of the modelling domain. The observed discharge at the fan inlet (DHM) is presented at the top of each image. The coordinate system is WGS84 UTM Zone 44 N in kilometres.

The model predicts the reoccupation and abandonment of channels rather than the channel migration for the baseline simulations. Figure 6.24 presents the number of years during which a channel is active during post-monsoon season flow conditions (1,160 m^3/s) for the three ensemble members with the lowest, medium and highest predicted morphological changes.

The location of the channels remains stable for the member with the lowest predicted morphological change, whereas a few channels in the upper third of the Eastern branch are abandoned and one channel of the Western branch is reoccupied (Figure 6.24). Both, the abandonment of channels in the Eastern branch, and the reoccupation of channels in the Western branch increase for the medium- and high-change members.

The member with the highest predicted morphological change predicts the abandonment of most channels in the Eastern branch which indicates that the bifurcation changes and more water is routed through the Western branch. Therefore, this model predicts a behaviour that is also observed in the Landsat images (Figures 6.22 and Figures 6.23). However, the location of the channels remains similar and hence the model does not predict considerable migration of channels and bars.

The ensemble variation of the channel locations at the end of the simulation periods increases with the emissions (Figure 6.25). This variation is low in the baseline scenario for which some members predict channels in the Island between both branches but, overall, the ensemble variation is low. This variation increases for the projected scenarios, particularly in the Island between the branches. The divergence of channels into the Island occurs at the fan head (~ 10 km distance to the inlet), at the Eastern branch (~20 km distance to the inlet) and at the lower Western branch (~25 km distance to the inlet). The Eastern and Western floodplains remain free of channels except for one member which predicts a channel in the Western floodplain in the medium-emission scenario SSP245.

Generally, the location of the main branches remains stable between the three scenarios which indicates that the ensemble variation is caused by the abandonment and reoccupation of channels rather than the channel migration (Figure 6.25). One exception is the fan head, particularly in the transition from the single-threaded to the braided channel (~2 – 3 km downstream of the inlet). This transition migrates upstream whereas this shift is larger for the high-emission scenario SSP585. In the braided channels of the Eastern branch, more members predict the inundation of more channels with the emissions which indicates an increase of the braiding intensity with higher flows.

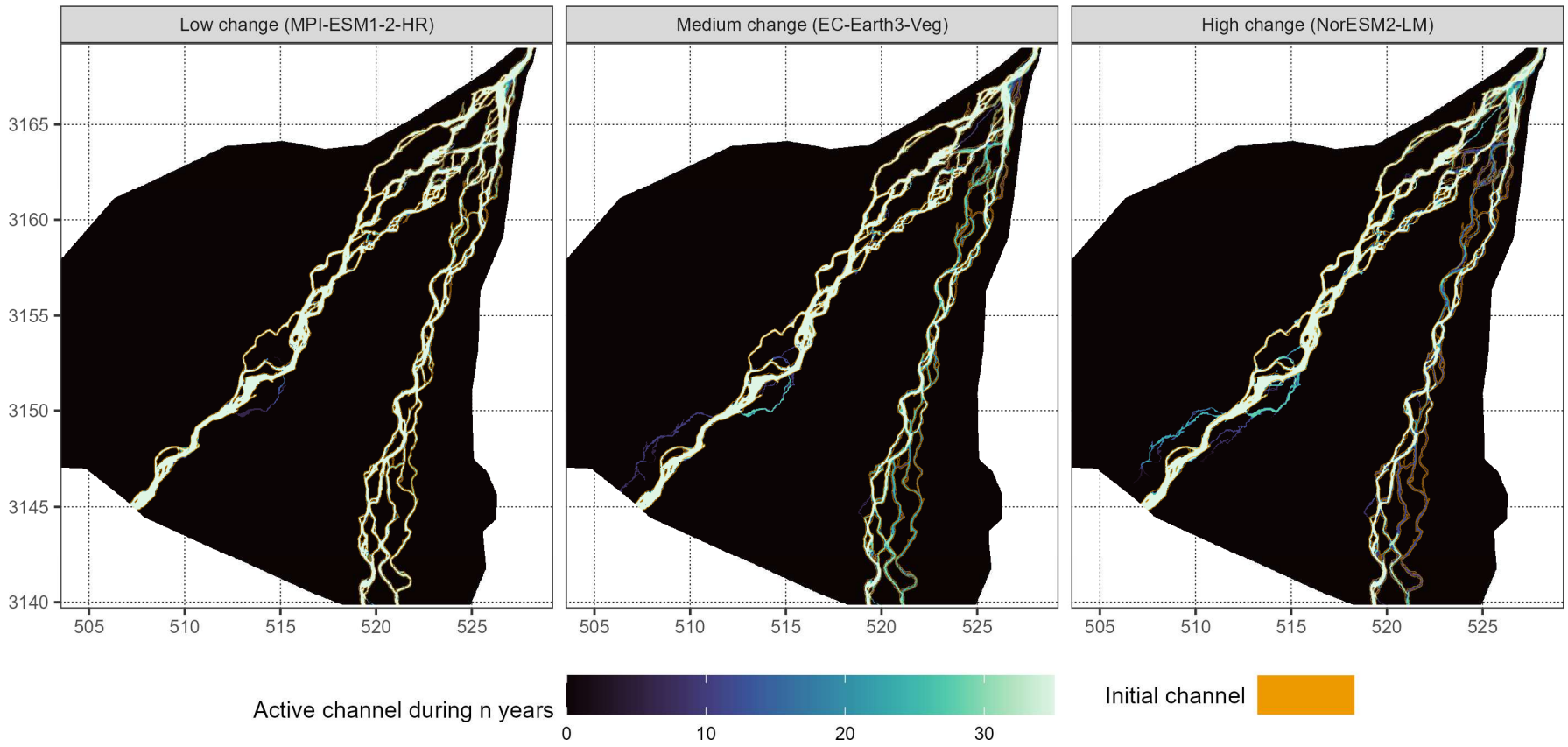


Figure 6.24: The predicted evolution of the channel locations for the baseline period for three ensemble members. The gradient colour indicates for how many years in the period 1980 – 2014 a channel is active and the orange outline indicates the location of the channel at the beginning of the simulation period in 1980. The ensemble members are chosen by their predicted elevation change (baseline scenario) and represent the full range of ensemble predictions from the lowest change (left), medium change (centre), and highest change (right). The channels are classified as pixels which are inundated at 1,160 m^3/s inflow to the fan. This inflow is the median discharge observed by DHM during October (1990 – 2016) and was chosen to ensure the comparability with the observed channels in Figures 6.22 and 6.23. The coordinate system is WGS84 UTM Zone 44 N in kilometres.

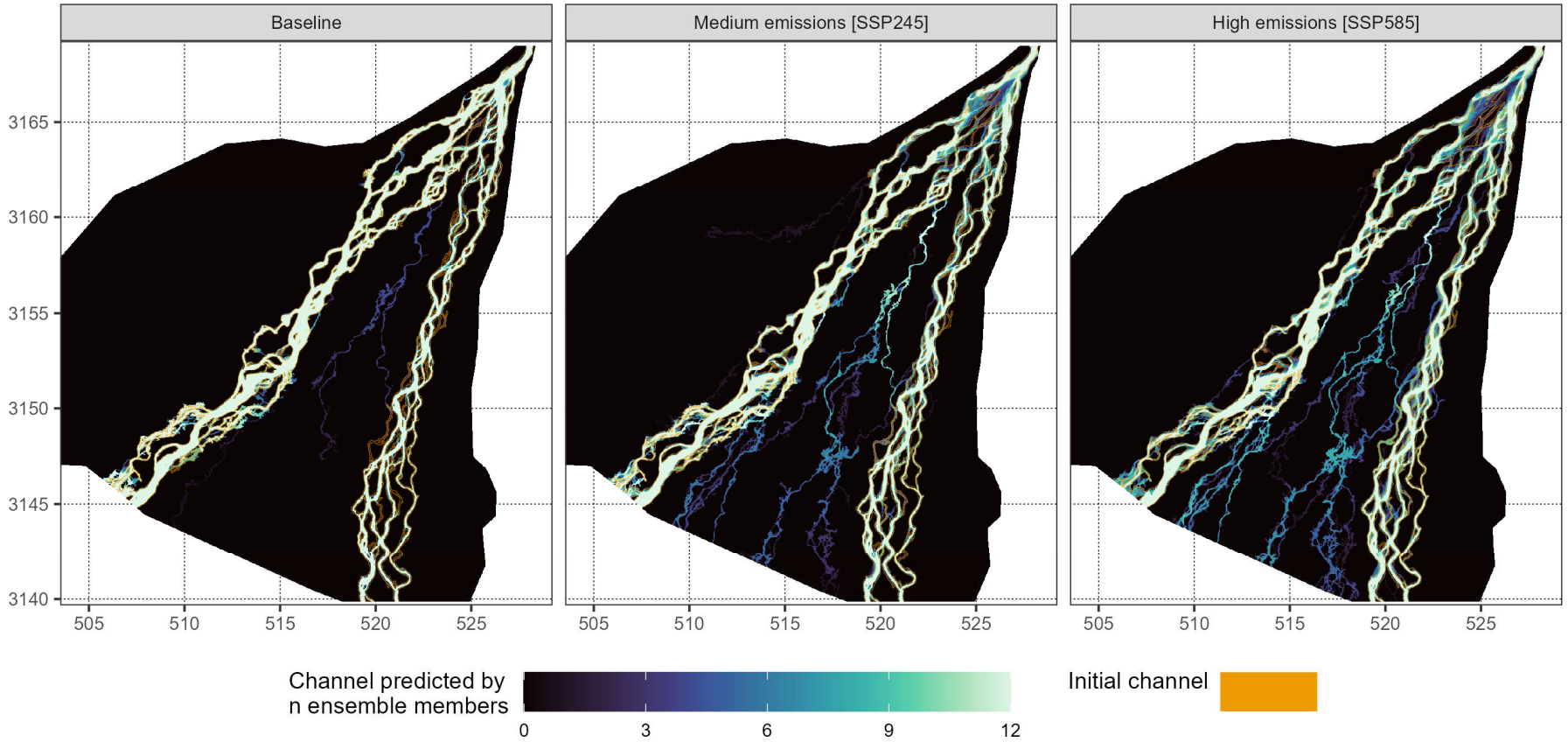


Figure 6.25: The location of the channels at the end of the morphodynamic simulations for the ensemble members. The gradient colour indicates how many of the 12 CMIP6 members predict a channel at a particular location and the orange outline indicates the location of the channel at the beginning of the simulation period. The channels are classified as pixels which are inundated at 3,000 m^3/s inflow to the fan. Note that this inflow rate is higher than the one in Figure 6.24. The higher inflow rate was chosen because it delineates a larger fraction of the channel without inundating the floodplain. The coordinate system is WGS84 UTM Zone 44 N in kilometres.

6.4.4.2 The prediction of bifurcation changes

The changes in the observed and simulated channels indicate the sensitivity of the flow pathways to the topographic change at the fan head. The bifurcation controls the allocation of the flow between the Eastern and Western branches and hence controls the hydrodynamic and morphodynamic processes in the downstream fan. The projected changes in this flow distribution are analysed by applying a flow model for the topographies projected for the high-emission scenario SSP585 for three members which represent the prediction range of the ensemble. The selection of the ensemble members is based on the total predicted change across the whole fan and the members with the lowest, highest, and median predicted change are selected.

The impact of the bifurcation changes on the flow allocation to the branches varies with the flow rate, and the ensemble differences decrease with increasing inflow (Figure 6.26). For the initial conditions, two-thirds of the inflow drains through the Western branch and one-third through the Eastern branch at lower flows (3,000 m^3/s). With increasing inflow, the fraction draining through the Western branch decreases and 40% of the inflow drains through the Eastern branch (< 7,000 m^3/s). For higher inflows, the fraction draining through the Eastern branch remains stable at ~40%, while the fraction decreases for the Western branch and increases for the Island. For the highest investigated inflow (17,500 m^3/s), 48% drain through the Western branch, 37% through the Eastern branch, and 15% through the Island channels.

The projected flow allocation varies between the ensemble members, particularly for inflows < 10,000 m^3/s. For the lower-change member, the Western branch gains importance and 85% drains through it at 3,000 m^3/s inflow. The contrary change is predicted for the high-change member for which the fraction draining through the Western decreases to 36%. The differences between the initial and projected flow allocations decrease with the inflow rate. However, the Western branch drains the largest fraction of water for all investigated flows (≤ 17,500 m^3/s) for the low-change projections, while the Eastern branch drains the largest fraction of water for the high-change projections. The allocation of water for the median-change projections deviates least from the one of the initial topography. Hence, the larger topographic change of this member compared to the low-change member does not translate into higher changes in the flow allocation between the branches.

The channels of the Island between the two branches become more connected to the channel system for all ensemble members. For the initial topography, the Island channels start to receive water at 7,500 m^3/s (Figure 6.26). This threshold decreases to 5,000 m^3/s for the low- and medium-change members, and for the high-change member, the channels in the Island carry flow throughout the investigated inflow range.

The morphological changes at the bifurcation control the allocation of water between the branches. The predictions of this allocation vary between the ensemble members. However, this variation mainly concerns smaller flood events. The projected flow allocation approaches the one predicted for the initial topography for inflows ≥ 10,000 m^3/s (Figure 6.26).

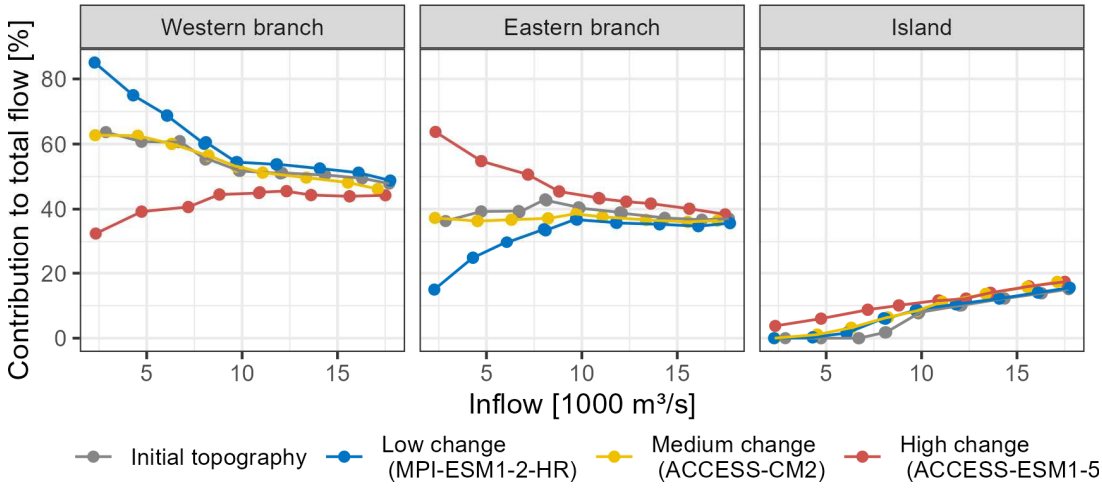


Figure 6.26: The allocation of water between the Western and Eastern branches and the Island channels. The allocation of the initial topography is compared with the topographies predicted for the high-emission scenario SSP585 (2099). Three ensemble members are selected based on the projected changes of the fan topography and cover the full range of predictions for the SSP585 scenario from the smallest projected change (MPI-ESM1-2-HR) to the highest projected change (ACCESS-ESM1-5). The ACCESS-CM2 model represents the median projected change.

6.5 Discussion of the morphological simulations

The morphodynamic simulations predict an increase in the morphological changes which scales with the flood discharge and, thus, with the emissions. These changes concentrate in the fan head at the transition from a single-threaded to the braided river and the bifurcation into the Eastern and Western branches. The fan head (0 – 10 km downstream of the inlet) acts as the main sediment sink and hence most of the sediments which are discharged from the Himalayas deposit within the first few kilometres downstream of the mountain range. The evolution of the fan centre (10 – 20 km downstream of the inlet) and the lower fan (20 – 30 km downstream of the inlet) is mostly caused by the redistribution of fan sediments, and hence these areas are insensitive to the magnitude of sediment inflow for the investigated period. Furthermore, the allocation of water between the river branches is sensitive to the fan head topography.

This section discusses the predicted morphological evolution of the fan and their controls (Section 6.5.1), the changes in the flow pathways (Section 6.5.2), the predicted sediment dynamics (Section 6.5.3), and the uncertainty of the simulations in the final Section 6.5.4.

6.5.1 The predicted morphological evolution

The predicted morphodynamic evolution varies in different zones of the fan. Generally, the model predicts alternating patterns of degradation and aggradation in both, the longitudinal and the lateral directions. The spatial scales of these patterns vary from a few hundred metres in the braided river sections of the fan centre and lower fan to kilometres at the fan head and in sections where the Western branch flows in a single wide channel. The predicted magnitude of change scales with the inflow of water and sediments but the ensemble members predict degradation and aggradation at similar locations. Hence, increasing flood flows intensify the morphological evolution but do not cause a shift in the behaviour (e.g., from degradation to aggradation).

6.5.1.1 Fan head

The spatial variation of the dominant processes (degradation and aggradation) in the fan head (0 – 10 km distance to the inlet) reflects our understanding of the system. This area is the most active zone in which most of the predicted change occurs, and in which areas of degradation and aggradation have the largest extent (Figure 6.8). The single-threaded channel at the inlet degrades in the first 1 – 2 km after which the channels start aggrading. Both, the predicted degradation and aggradation appear reasonable from our understanding of the evolution of alluvial fans.

These landscape features develop at the intersection between the parts of a catchment which deliver sediments and the parts which store these sediments (Coulthard *et al.*, 2002). The degradation is predicted for the mountain gauge in which the river flows in a narrow and well-confined channel which is still part of the delivery (degrading) system and hence the channel degradation is reasonable (Kleinhans *et al.*, 2013). The change to an aggrading state occurs after 1 – 2 km where the river widens which results in the decrease of the stream power and, thus, in the sediment transport capacity. The highest aggradation rates are predicted at the transition from the single-threaded to the braided river which is reasonable because the water is distributed between different channels resulting in the decrease of the sediment transport capacity (Coulthard *et al.*, 2002). A second, albeit smaller, aggradation peak is predicted at the bifurcation where the transport capacity reduces further as the number of channels increases. Therefore, the fan head aggradation is a reasonable prediction for the modelled system which is further complemented by flume experiments of alluvial fans which show similar patterns of aggradation just downstream of the feeder channel for high sediment supply rates (Leenman and Eaton, 2022). This predicted aggradation of the fan head increases the risk of an avulsion because it increases the topographic gradient between the channel and the western floodplain (Kleinhans *et al.*, 2013; Sinha *et al.*, 2014).

The fan head evolution is controlled by the sediment influx and the channel slope. This area can switch between aggrading and degrading phases whereas the direction of change is commonly controlled by the sediment influx (Schumm *et al.*, 1987; Kleinhans *et al.*, 2013; Harvey, 2018; Leenman Eaton, 2022). The model predictions indicate that the fan head is currently in an aggrading phase. Furthermore, the model replicates the sensitivity towards the sediment influx because the net degradation of the fan head is predicted when no sediment enters the fan (Figure 6.20). However, the predicted phase must switch to degradation at some point in time even with continuous sediment supply because the aggradation increases the slope between the fan head and the downstream

fan which facilitates higher sediment transport from the head to the downstream fan areas (Dingle *et al.*, 2020a). Therefore, the morphological predictions cannot be extrapolated beyond the simulated period (the year 2099).

The fan sections at the transition from the single-threaded to the braided river and the bifurcation aggrade even in the scenario without sediment influx. These sediments originate from the degradation of the channel at the mountain gauge. In this section, the river flows in a confined channel and no changes in either width or location have been observed in the satellite images of the past 30 years (Figure 6.22). The predicted channel widths increase from 250 m up to 600 m in the high-emission scenario SSP585 (Figure 6.16). Thus, while the degradation of the bed appears reasonable, the bank erosion is overestimated (the reasons are discussed in Section 6.5.4). Consequently, the aggradation is overestimated because these zones act as the sink of the bank sediments. This aggradation affects the downstream allocation of water (i.e. the Eastern and Western branches) and, thus, controls the evolution of the downstream fan.

6.5.1.2 Fan centre and lower fan

The predicted evolution of the fan centre (10 – 20 km distance to the inlet) and the lower fan (20 – 30 km distance to the inlet) is characterised by:

- the lower predicted change rates compared to the fan head, e.g. the predicted net sediment change is ~ 4.6 Million m^3 in the fan head and ~ 0.75 Million m^3 in the lower fan at the end of the high-emission scenario SSP585 (Figure 6.10);
- the alternating patterns of degradation and aggradation in the longitudinal direction which occur on a smaller scale of a few hundred metres (Figures 6.12 and 6.17);
- the insensitivity of most river sections (particularly the Eastern branch) towards the sediment influx into the fan (Figures 6.19 and 6.20).

The morphological evolution of the fan centre and the lower fan is mainly caused by the redistribution of fan sediments in both, the longitudinal and lateral directions as indicated by the insensitivity of the morphological evolution to the sediment influx. This redistribution occurs in most parts on a scale of less than one kilometre. However, net degradation is predicted for the fan centre and net aggradation is predicted for the lower fan. The lower fan aggradation exceeds the fan centre degradation and is of similar magnitude for each of the sediment influx scenarios. Therefore, fine sediments are eroded in the fan head and these sediments bypass the fan centre and deposit in part in the lower fan. This predicted behaviour is reasonable because the channel slope is

steeper in the fan centre and decreases in the lower fan, and the sediment transport capacity decreases in the lower fan (Dingle *et al.*, 2020b).

The evolution of the fan centre and the lower fan differ in the temporal characteristics and the drivers of the predicted change. The fan centre degrades in the first ~35 years of the simulations whereas the degradation rate decreases with time and after these 35 years it shifts to an aggradation (Figure 6.11). The differences between the medium-emission scenario SSP245 and the high-emission scenario SSP585 are small and hence this shift from degradation to aggradation is not driven by the increasing flow rates. This is also supported by the variation between the ensemble members which differ greatly in the projected flood flows, but all predict the initial degradation and the shift to aggradation. Therefore, the shift from degradation to aggradation is not caused by the climate change-induced increase in flood flows but may occur in the baseline scenario on a longer time scale. However, the rate of the initial degradation scales with the emissions and is considerably lower in the baseline scenario. Therefore, the evolution of the fan centre is, to some degree affected by the flood flow changes but this sensitivity is small compared to the fan head and the lower fan.

Contrary, the evolution of the lower fan is driven by the projected increase in flood flows. This is particularly evident in the temporal evolution because the aggradation rates decelerate around the year 2060 in the medium-emission scenario SSP245 but continue at similar rates for the high-emission scenario SSP585 (Figure 6.11). This is the time at which the projected flood flows of both scenarios decouple which indicates that the aggradation is affected by the flood flow (Figure 5.26). However, it is not controlled by the sediment influx into the fan because similar aggradation rates are predicted for the different sediment inflow scenarios (Figures 6.19 and 6.20). This means that the increasing flood flows facilitate the higher redistribution of fan head sediments to the lower fan.

6.5.1.3 The controls of the morphological evolution

The heterogeneity of the fan enables the identification of the controls of the fan evolution. Three characteristics can be identified that control the predicted evolution which are the number of channels, the channel width, and the channel slope.

The number of channels controls the predicted rate of change in the longitudinal and lateral directions. The predicted change in both longitudinal and lateral directions decreases with the increasing number of channels. This predicted behaviour is reasonable as the flow rate and, thus, stream power in the individual braid-channels decreases as the water is distributed over more channels. A good example of this behaviour is the lower predicted change of the braided Eastern branch in the lower fan (high number of channels) compared to the Western branch (low number of channels) with larger channel changes and bank erosion (Figure 6.15). The confluence of channels increases the stream power and may cause downstream channel degradation. However, this process is overlayed by other processes, such as the lateral redistribution of sediments, and the degradation is not predicted downstream of each confluence.

The channel width is another important control of the morphological evolution. An increase in the width facilitates the aggradation of the channel bed. The most prominent example is the channel at the inlet which degrades initially but starts aggrading as the channel width increases (Figure 6.16). The converse effect is predicted in the Western branch 24 – 26 km downstream of the inlet and at the outlet where degradation is predicted as the channel width decreases (Figure 6.16). These predictions are reasonable from the physical perspective because a decrease in the channel width increases the velocity and, thus, stream power which increases the sediment transport capacity (Ritter, 2003).

The reduction of the channel slope favours the aggradation. The channel slopes have not been systematically mapped due to technical challenges in accurately determining the flow length of the multi-threaded channels for the ensemble of predicted topographies. However, the fan slope decreases from ~0.002 m m^{-1} at the mountain gauge to 0.001 m m^{-1} at the downstream model boundary (Dingle *et al.*, 2020b). The reduction of the slope reduces the stream power and consequently, the sediment transport capacity in the downstream direction. The variation of the slope explains the large-scale redistribution of fan material from the fan head to the lower fan (Figures 6.10 and 6.20).

6.5.2 Channel characteristics

The quantitative analysis of the predicted channel characteristics such as the channel width, bank erosion, channel migration, and bar development and migration is not trivial for ensemble simulations. The reasons for these difficulties are:

- the determination of the bankfull flow rate: the bankfull flow might vary in time, space, and between the ensemble members. The bankfull flow is, however, important to determine the channel width which then can be used to determine the channel capacity;
- the sensitivity to bifurcation changes: The distribution of water between the branches varies with the inflow rate (Figure 6.26). For example, it was tested to determine changes in the channel width by counting the inundated cells for a given flow rate. However, the calculated channel widths were the product of changes in the water allocation rather than projected channel changes. This sensitivity also hampers the determination of the braiding intensity which depends on the number of channels but varies with the flow rate (Kleinhans *et al.*, 2013).
- The small-scale variation of degradation and aggradation: This variation hampers the analysis of cross-section profiles which could be used to determine the channel widths and capacity but are very sensitive to the selection of the location.
- The ensemble size: The quantitative analysis requires automated approaches to determine the channel characteristics because manual approaches are unfeasible for larger ensembles. However, such approaches are either computationally expensive (e.g. the flow routing through the modelling domain to delineate the channels for different flow rates) or need to be developed for braided river systems.

Due to these challenges, the channel characteristics were analysed qualitatively by comparing the observed and predicted variations of the channel locations. This analysis of the channel characteristics indicates that the model predicts the reoccupation and abandonment of channels but does not replicate the dynamics at the bar scale.

The model predicts a stable evolution in time and space (i.e. the location of degrading and agrading channel sections remains stable) which indicates that channels do not migrate. Instead, the changes in the course of water are dominated by the abandonment and the reoccupation of previously abandoned channels (Figure 6.24). Bars grow in some sections and the most prominent example is the bar at the transition from the single-threaded to the braided channel at the inlet (Figure 6.25). In other sections, in particular, in the braided Eastern branch, bars are eroded and/or submerged during medium flow (3,000 m^3/s). However, the observed channels vary stronger than the

model predictions which indicates that the model underpredicts the reworking of the bars and banks and, consequently, the channel migration (Figures 6.22 and 6.25). This indicates that the model has limited predictive capabilities to replicate the morphodynamic processes at the bar scale which agrees with simulations of a braided river in the Italian Alps (Ziliani *et al.*, 2020).

The simulations predict shallower and wider rivers which is caused by the deposition of bank material on the riverbed. The DEMs of Difference (DoD) (Figure 6.16) indicate the model's tendency to erode the river banks and deposit the material in the channel in many sections throughout the fan. The aggrading sections are intersected by sections of degradation where channels confluence or the channel width reduces. This pattern of bank erosion and in-channel aggradation suggests that the channel width increases and the channel depth decreases. This predicted evolution to generate wider and shallower channels has been observed in natural and experimental rivers in braided river systems characterized by the low cohesion of bank sediments (Xu, 2004; Kleinhans, 2010). This suggests that the predictions are reasonable. However, the magnitude of this phenomenon may be underestimated as suggested by the lower predicted than observed channel migration (Figures 6.22 and 6.24). The algorithms controlling the bank erosion and channel aggradation have a weak physical basis and are controlled by parameters that were transferred from a different river system to the Karnali fan (Van De Wiel *et al.*, 2007; Ziliani *et al.*, 2013). However, this bias cannot be quantified without information about the past topographic change.

The sediment delivery and the flow rate are the primary control of the evolution of braided rivers (Leopold and Wolman, 1957; Ferguson, 1987; Kleinhans, 2010; Kleinhans *et al.*, 2013). However, the similar predicted morphological evolution for the sediment scenarios downstream of the fan head (0 – 10 km downstream of the inlet) indicates that the sediment delivery is not a main control of the channel evolution of the fan centre and the lower fan (Figures 6.19 and 6.20). The scenario differences are larger for the Western branch which could imply that this branch is more sensitive to changes in the sediment delivery (Figure 6.21). However, these scenario differences may also be caused by the different fan head evolution which alters the allocation of water between the branches.

The braiding intensity should increase with the projected increase in flood flows (Egozi and Ashmore, 2008). The model predicts a decrease in the number of channels in the Eastern branch and an increase of these in the Western branch in the baseline scenario (Figure 6.24). This shift is caused by bifurcation changes and the increased routing of water through the Western branch. For the projected scenarios, the number of channels

increases throughout the fan which results from the increased flood flows (Figure 6.25). Therefore, the model replicates the increased braiding intensity with the flow volume caused by morphological changes (i.e. changing flow due to alterations of the bifurcation) and upstream hydrological changes (i.e. the increased flood flows of the projected climates).

The fan head is the most active area and impacts the morphological changes in the downstream fan because it controls the distribution of water between the branches. The temporal evolution of the channels indicates that this distribution changes in the baseline which is consistent with the observed changes indicating an unstable bifurcation (Figures 6.22 – 6.24). For the current conditions, more water drains through the Western branch but the allocation assimilates with increasing flow. The higher drainage through the Western branch for lower flows and the assimilation with increasing flow is consistent with field measurements and hydrodynamic simulations (Dingle *et al.*, 2020a). The bifurcation projections range from a continuing shift in favour of the Western branch to the increased routing through the Eastern branch. Furthermore, the changes in the allocations do not scale linearly with flow (Figure 6.26). Therefore, this evolution is most difficult to predict but it has implications for the water and flood risk management of the downstream areas. However, the bifurcation changes have higher implications for smaller and more frequent floods (< 10,000 m^3/s). This suggests that the morphological evolution may be more relevant for water management and nature conservation because it affects the water availability throughout the year.

The braided channels of the Eastern branch are stable in the baseline but become increasingly variable in the projected scenarios (Figure 6.25). These channels are characterized by the migration and reworking of the channel as indicated by the variation in the observed course of water whereas the activity has decreased after the bifurcation change in 2009 (Figure 6.23). Therefore, the low geomorphic activity in the baseline simulations is reasonable since the initial DEM marks the topography after the year 2009. For the projected simulations, the flood flows increase and the projected changes in the fan head increase the routing through the Eastern branch resulting in an increasing channel activity which was observed before the year 2009 (Figure 6.22 and Figure 6.25).

The Karnali River reoccupies the Island in the projected scenarios whereas this reoccupation scales with the emission scenario (Figure 6.25 and Figure 6.26). The water inflow into the Island is controlled by irrigation gates which hampers the comparison of simulations and observations. One divergence is predicted at the lower Western branch (25 km distance to the inlet). In this area, embankments are constructed to prevent the flooding of the Island. Satellite images indicate high bank erosion before the construction

of the embankment in 2011. Furthermore, a 250 m long section of the embankment was washed away between 2021 and 2023 emphasising the force of the river in this section (Figure 2.2). Therefore, the location of the predicted divergence is reasonable. Furthermore, water enters the Island at the fan head and a bifurcation of the Eastern branch. These two locations are controlled by irrigation gates (which are represented as a embankments in the DEM). The erosion of these gates appears to be a reasonable location for the reoccupation of the Island. In general, it appears reasonable that the river reoccupies the Island when it is not controlled by human interventions.

6.5.3 The fate of Himalayan sediments

The evolution of the sand-bed rivers downstream of the simulated fan is a product of the redistribution of floodplain material through bank erosion (Dingle *et al.*, 2020b). This area is insensitive to the sediment discharge of the Himalayas because the coarser gravel sediments are deposited in the fan and finer sediments bypass this area and deposit in the Ganga-Bramaputhra basin (Lupker *et al.*, 2011; Dingle *et al.*, 2020b). The predictions suggest that the fan evolution (≥ 10 km distance to the inlet) is the product of the redistribution of fan material rather than of the discharge of Himalayan sediments. These upstream sediments deposit on the fan head as indicated by the similar sediment budgets of the different sediment scenarios in the fan centre and the lower fan (Figure Figure 6.20). This redistribution occurs in the lateral direction as bank erosion and in-channel deposition, and in the longitudinal direction by the degradation and aggradation of channels which is e.g. caused by changes in the channel width or the confluence of channels (Figure 6.16).

The fan evolution is, however, also controlled by the sediment supply from the mountain range, whereas this control is active on larger time scales than the simulations. The aggradation of the lower fan indicates that sediments originating from the fan head (and the fan centre) deposit in the lower fan (Figures 6.10 and 6.11). This aggradation of the lower fan is independent of the sediment supply as indicated by the sediment scenarios and is, hence, caused by the redistribution of fan material (Figure 6.20). However, this fan material originates from the Himalayas. Hence, the Himalayan sediments are deposited throughout the fan but this incorporation to the downstream fan occurs on longer time scales than the simulated 75 years. Therefore, the lag time of the geomorphic response to sediment pulses is considerably larger in the lower fan than in the fan head.

The predicted evolution is a snapshot in time that will shift in the future. The predicted fan head aggradation leads to an increase of the channel slopes to the downstream fan which then facilitates the fan head degradation and an increased sediment transport to the fan centre and the lower fan. The predicted shift from degradation to aggradation in the fan centre might be the beginning of such a change in fan behaviour. Thus, the predictions provide information about the fan development for the current fan state but should not be extrapolated in time (centuries – millennia) or space.

6.5.4. Uncertainty sources

The morphological simulations were designed as an experiment to understand the controls of the fan evolution but were constrained by data scarcity. This section discusses the main sources of uncertainty which can be attributed to the determination of the boundary conditions, the model parameterisation, and the assumptions made in the experimental design which may help to improve the quality of future studies.

The predicted evolution of the fan head illustrates the challenges of determining the boundary conditions and parameterising the model in a heterogeneous mountain environment. The evolution of the fan head controls the flow of water and, thus, the morphological evolution of the downstream fan which demonstrates how inaccuracies in the predictions propagate from the upstream model boundary to the downstream boundary. The overestimated lateral erosion at the inlet, which is indicated by the predicted but not observed channel widening (Figures 6.16 and 6.22), and the subsequent overestimation of the downstream aggradation at the transition from the single-threaded to the braided river and the bifurcation is affected by the model parameterisation using global parameter in a heterogeneous environment.

This overestimation of the lateral erosion is caused by the different cohesion of the material at the mountain gauge and the downstream fan which cannot be represented with a single global parameter. The channel at the inlet is confined by the Siwalik mountains and these properties are very different to the fan characteristics for which the model is parameterised (Upreti, 1999). The cohesion is controlled by the in-channel lateral erosion rate parameter which is parameterised for the fan and not for the mountain gauge. This leads to the underestimation of the cohesion and the overestimation of the bank erosion at the inlet.

The overestimation of the channel bed degradation at the inlet is affected by the determination of the boundary conditions. The channel degradation may reasonable from a physical perspective because the channel at the inlet is still part of the feeder system (Kleinhans *et al.*, 2013). However, this channel should degrade in the modelling setup because the sediment influx is at transport capacity and hence the river should not have the energy to mobilize any more bed sediments. A combination of factors can influence this degradation:

- transport capacity: There can be differences between the transport capacity calculations of the C-L model and the BAGS model used to calculate the sediment influx. Furthermore, the cross-section of the BAGS calculation is static while the cross-section of the C-L model is constantly changing;
- grain sizes: The grain sizes of the bed sediments are underestimated and are, thus, mobilised at lower flows. This is illustrated by grain size measurements at the bifurcation where the D_{50} is 65 mm on the bar and 231 mm in the bed (Dingle *et al.*, 2020b). This is particularly relevant at the inlet where the river flows in a single narrow channel and has the highest stream power to mobilise sediments;
- model initialisation: The flow predictions are noisy at the upstream boundary because the model requires some distance to fully develop the flow patterns (Ingham & Ma, 2005). This noise can be reduced by moving the boundary further upstream and increasing the flow length. However, such an extension might increase the sediment delivery to the fan because the channel section of the poorly parameterised mountain channel would be extended.

The model parameterisation is further hampered by the lack of calibration data. One example is the underprediction of the bar-scale dynamics. These dynamics are affected by the lateral erosion rate which is one of the most sensitive parameters (Ziliani *et al.*, 2013). The used rate (0.002) is within the range of recommended rates for braided rivers (0.001 – 0.01) (C-L Manual, 2023). However, it is difficult to determine this rate from the satellite observations of rivers alone because water classification from the images is sensitive to the flow rate during the acquisition date, contains misclassifications, and contains no information about quantitative sediment changes. Furthermore, the model is sensitive towards the thickness of the active layer (Mosselman, 2012; Ziliani *et al.*, 2013). The lack of data about topographic changes hampers a quantitative calibration using a combination of performance measures. The model calibration is further complicated by the sensitivity to the boundary conditions and parameter interactions (model equifinality) (Hardy, 2013).

The model equifinality can be illustrated by the lateral erosion rate used by Ziliani *et al.* (2020). The authors used a very high rate of 30 (low bank stability and high erosion) in combination with a vegetation module which stabilises the banks. This rate was tested for the Karnali fan in the initial model testing without the vegetation module which resulted in the massive overestimation of the bank erosion and the dissolution of the channels. One solution to account for the model equifinality is the application of a GLUE approach as used for the hydrological simulations (Chapters 4 and 5). However, the simulation of the 36 ensemble members (12 CMIP6 models X 3 flow percentiles) required > 15,000 h of computations for a single parameter set. This illustrates that a GLUE approach is unfeasible for ensemble modelling even using a reduced-complexity approach.

The model predictions are further affected by the assumptions, namely:

- The simulation of flood discharge: The modelling framework simulates only one flood event per year. While this is the flow with the highest stream power and, thus, the highest potential to alter the fan topography, the fan is also reworked by lower flows (Kleinhans *et al.*, 2013). The simulation of more days per year may increase the sediment delivery from the fan head to the downstream areas. Furthermore, it potentially increases the geomorphic activity of the finer-grained lower fan because these sediments are mobilised at lower flows. Furthermore, the frequency of floods increases with time and emissions, and hence the AMAX classification approach underestimates the number of flood events increasingly with time and emissions.
- The sediment transport capacity: More fan sediments can be mobilised if the sediment delivery from the Himalayas reduces. One sediment scenario included a random sediment influx which did not result in systematic differences compared to the other scenarios. This indicates, alongside the low sediment scenario variation in the fan centre and lower fan, that the sediment transport capacity assumption has no significant impact on the predicted fan evolution. However, it is assumed that the grain size distribution (GSD) of the sediment supply does not change over time. This is an important aspect because the area of deposition is determined by this GSD (finer sediments should deposit further downstream). It is further worth noting that the definition of the grain size classes in the model might affect the modelling results and that this uncertainty was not investigated in this study.
- The focus on bedload sediment transport: Sand is the finest simulated sediment fraction and may be transported as bedload or suspended load depending on the flow magnitude. The focus on the bedload sediment potentially causes an underestimation of the sand transfer from the channel to the floodplain particularly

during high-magnitude flows that transport sand in suspension and inundate larger areas of the floodplain. This suspended sediment would deposit in the floodplain and raise the floodplain relative to the channel, but this process cannot be predicted from bedload sediment transport.

- The role of vegetation: The vegetation stabilises river banks and affects the flow field (Murray and Paola, 2003; Kleinhans *et al.*, 2008, 2013; Kleinhans, 2010; Ziliani *et al.*, 2013). The implementation of the vegetation cover is somewhat difficult for long-term predictions because it depends on the land use and hence ideally the climate projections are combined with land use scenarios which would increase the ensemble size. However, the role of vegetation for the morphological evolution needs to be investigated in future studies to increase the quality of the predictions.
- The basin subsidence: The basin subsidence is an important control of the fan evolution (Dingle *et al.*, 2016; Harvey, 2018). The accommodation space generated by basin subsidence is lower than the gravel supply during the average monsoon season (excluding the flood event) which is why the basin subsidence was excluded in the experimental design. However, it may be required to incorporate basin subsidence if the simulation period is extended (e.g. to monsoon seasons).

Further sources of uncertainty are errors in the topography, particularly in forested areas which have been interpolated without knowledge of the surface elevations, and smaller channels where the banks are vegetated by trees. Furthermore, uncertainty is inherent in the models due to our limited understanding of the sediment transport processes (Mosselman, 2012; Hardy, 2013; Ziliani *et al.*, 2013, 2020; Ancey, 2020a).

6.6 Conclusions

The experimental simulations of the fan evolution provide robust evidence that the geomorphic activity increases with the flood discharge and, thus, with the emissions. Consequently, the prediction of the evolution becomes increasingly uncertain because the uncertainty of the climate projections increases with time and emissions (Chapter 5).

The fan head evolution is sensitive to the changes occurring in the fan head which control the flow of water downstream. The sediment delivery from the Himalayas deposit, at the current development stage of the river and for the investigated period (until the year 2099) in the fan head. This leads to an increased risk of channel avulsions because the topographic gradient between the channel and the Western floodplain increases (Kleinhans *et al.*, 2013).

The evolution of the fan centre and lower fan (≥10 km distance to the inlet) is insensitive to the sediment delivery from the Himalayas in the investigated period. Instead, the evolution is driven by the redistribution of fan sediments on different spatial and temporal scales. The response time of these areas to sediment pulses (e.g. induced by earthquakes) is longer than for the fan head.

The evolution is characterised by a sequence of degradation and aggradation in the longitudinal and lateral directions. The channels degrade in sections where the channel width decreases which leads to the subsequent downstream aggradation. This has important consequences for flood management because embankments constrain the channel width fostering the local degradation and downstream aggradation which potentially aggravates the downstream flood hazard. The projected increase in the flood discharge increases this problem because in natural systems the river would respond by increasing the width in low-cohesion systems.

The predictions provide further evidence that the bifurcation is unstable and that this variation particularly affects lower flood flows (< 10,000 m^3/s). Therefore, the potential bifurcation changes affect the water availability for ecosystems, agriculture, industrial and household use.

These findings illustrate the necessity to consider the morphodynamic evolution in water resource management and flood risk management and to evaluate the impact of structural engineering (e.g. the construction of embankments) on the morphological evolution. Such assessments require the consideration of the climate change impact on the discharge since such changes alter the geomorphic evolution.

The climate change impact assessment poses a dilemma which inevitably increases the uncertainty in the predictions. This dilemma results from the incomplete understanding of the processes and their interactions, and the knowledge about the simulated river system whereas this knowledge decreases with time (i.e. we can measure the current state but need to estimate the future states). Therefore, we need to apply ensemble modelling to consider a range of future pathways (i.e. climate ensembles and scenarios). This ensemble modelling can be achieved by reducing the complexity of the physical process representation which in turn increases the uncertainty in the morphodynamic predictions. Therefore, it is beyond the scope of most applications to account for the morphological and climatic uncertainties.

Alluvial fans are systems with heterogeneous conditions which complicates the application of morphodynamic models. This heterogeneity challenges the determination of the boundary conditions because it requires denser sampling (e.g. of the grain sizes). It increases the parameter uncertainty because global parameters cannot represent the full range of channel and floodplain characteristics. Furthermore, the determination of the sediment delivery to the fan is a challenge. A common approach for rivers with unknown sediment influx is to reroute the sediment discharge (or a fraction of it) at the downstream model boundary back into the upstream model boundary (Ziliani *et al.*, 2013, 2020). This is not possible for fans due to the high rates of downstream fining which would result in the lack of coarse sediment influx when rerouting the sediments (Quick *et al.*, 2019; Dingle *et al.*, 2020b).

The mountain gauge is a crucial area for the simulations of alluvial fans because it affects the simulations until the downstream model boundary and it needs to be investigated how this area is best simulated. There are two conflicting aspects for the simulation of the mountain gauge. Firstly, the different characteristics of the gauge cannot be adequately parameterised when the modelling focus is on the fan. Secondly, the flow model requires some space to develop the flow patterns and hence the predictions at the upstream boundary are noisy. The parameterisation issue suggests moving the model boundary downstream to exhibit the mountain gauge from the model while the flow initialisation suggests to extent the model boundary upstream which would move the noise predictions further away from the fan. One solution could be to reduce the erodible layer at the upstream boundary to prevent degradation. However, this might move the noise downstream because it introduces a sudden jump in the erodibility. Therefore, it needs to be investigated how to incorporate the mountain gauges best into a morphodynamic model.

The accuracy of the predictions is constrained by our knowledge of the Karnali fan. The following studies have the potential to improve future predictions:

- Topography: The generation of multiple DTMs to quantify the morphodynamic change which can be used to calibrate and validate the model. Furthermore, ground control points of forested areas are valuable to improve the quality of the DTM in forested areas. Cross-section surveys would further improve the channel representation and enable a more precise representation of the channel capacity.
- Grain size distributions of the bed material: The lateral variation of the GSD could not be considered in the model which leads to the mobilisation of the bed at lower flows. The sampling in the bed along the fan would fill this data gap and enable a better representation of the fan heterogeneity.
- Sediment delivery: The fan head evolution is controlled by the sediment supply from the mountains. A better understanding of the sediment production in the Siwaliks and the transport to the fan is crucial for a better understanding of the long-term fan evolution and the avulsion risk.
- The effect of vegetation on bank stability: Vegetation is an important factor of the morphological evolution and should be considered in future studies. This might provide useful information about the potential of revegetating riverbanks for sediment and flood management.
- Algorithms for the morphological characteristics: There is a lack of open-source algorithms to determine the channel characteristics (e.g. channel width, slope) for high-resolution datasets where the channel is wider than a grid cell and for braided rivers with divergences and confluences. The development of such algorithms would benefit the analysis of ensemble predictions which depend on automated approaches for the analysis.

This chapter presented, to the best knowledge of the author, the first application of a morphodynamic model to predict the potential evolution of large alluvial fans in the Terai for climate projections. It, therefore, extends our knowledge of the climate change impact on Himalayan river systems by assessing its impact on geomorphological processes. The simulations improved our understanding of the fan behaviour by identifying processes that control the morphological evolution and contributed to our knowledge about the sediment cascade. Furthermore, this chapter identified key challenges to simulate these large alluvial fans which may provide useful guidelines for future modelling applications.

7 The prediction of the potential flood hazards on the floodplain scale

7.1 Introduction

This chapter investigates the spatial patterns of the current and projected flood hazards in the Karnali fan in the Terai to locate those areas that are potentially affected by the projected changes in the flood magnitudes and fan topography. The flood magnitude of the 1-in-100-years event is projected to increase by 40% (medium-emission scenario SSP245) and 79% (high-emission scenario SSP585) at the end of the century (median predictions) (Table 5.9). The morphological simulations indicate that the geomorphic changes increase with the projected flood discharge (Chapter 6). This chapter evaluates how the projected changes of the flood magnitudes and fan topographies alter the spatial flood hazard characteristics (i.e. inundation extent and depth) in the Karnali fan.

The Karnali fan is located in the flood-prone and densely populated Terai (Figure 7.1). The Terai occupies 13% of Nepal but is home to 50% of the population and hence floods in this region affect many people and cause high damage to properties (Khanal *et al.*, 2007; MoHA, 2009; CBS, 2017). Furthermore, the Terai accommodates most of the fertile land in Nepal and is important for the food supply of the Nepalese population (Perera *et al.*, 2015; CBS, 2017). Floods destroy crops and irrigation infrastructure and kill livestock which threatens the livelihood of the local population and the food security of the country (Dixit *et al.*, 2007; MoHA, 2009; Perera *et al.*, 2015; NPC, 2017; Okura *et al.*, 2020).

The Karnali fan is densely populated and intensively used for agriculture. Approximately 240,000 people live in the 545 km^2 study area (Figure. 7.1 and Table 7.1). The population density of 441 persons/km^2 is more than twice the national average of 198 persons/km^2 (NSO, 2022). The population is concentrated on the Western floodplain (142,000 people) which is located below the elevation of the river banks, and on the Island between both Karnali branches (74,000 people). The outskirts of Tikapur city, the only city in the study area located in Tikapur municipality, are located within two kilometres of the Western branch. Most of the fan is used either for housing or for agriculture, except for the extended forest in the Bardiya National Park which stretches along the Eastern edge of the fan. It is therefore important to understand how changes in the flood hazard

manifest spatially because the potential increases in flood extents have potential implications for the local population and the food security of the country.

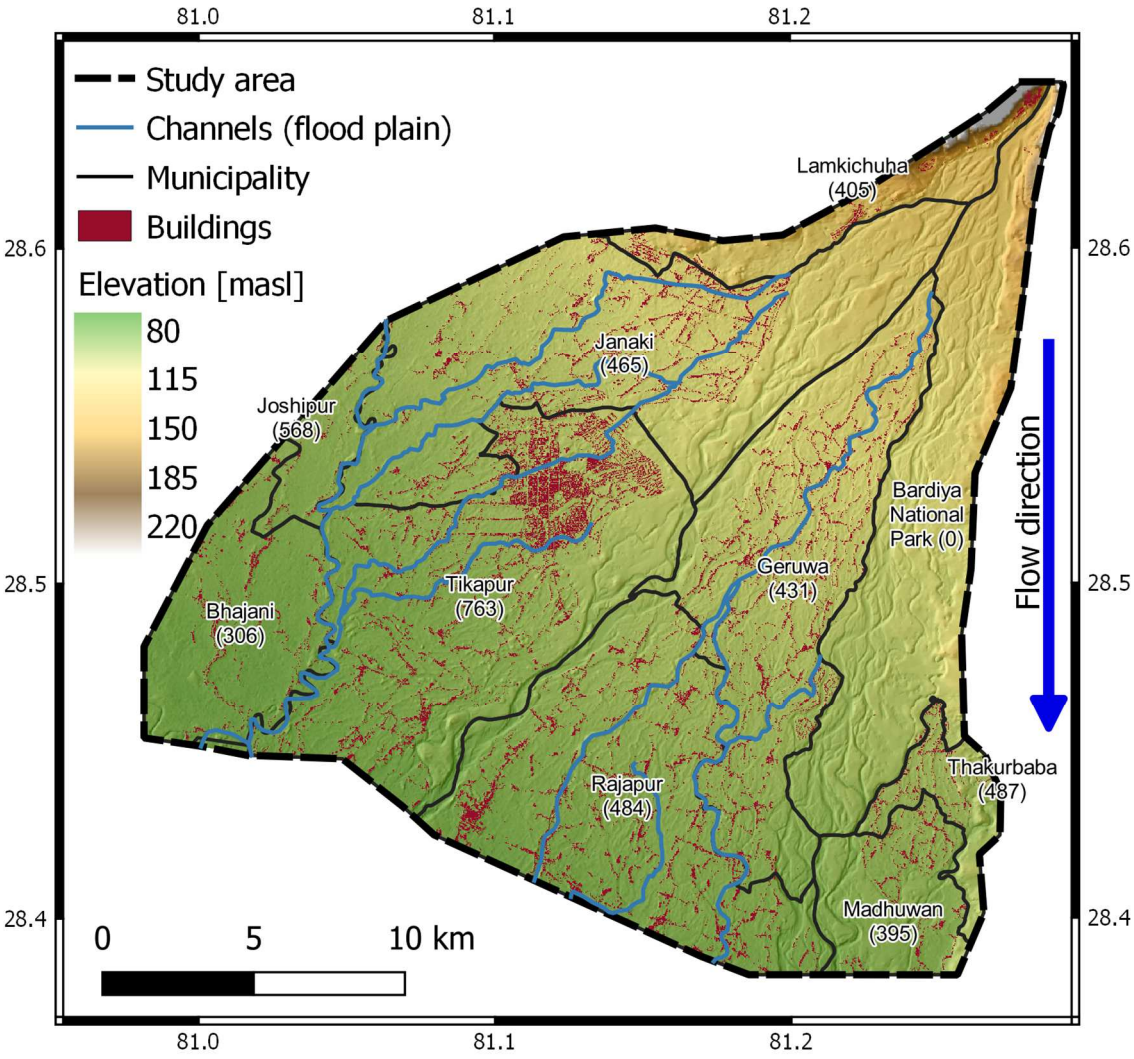


Figure 7.1: The hydrodynamic modelling domain in the Karnali fan. The topography is derived from the DTM described in Section 6.3.1. The population density of the municipalities (persons / km^2) is shown in the brackets (CBS, 2022). The buildings are obtained from OSM (2023). The channels are manually delineated based on Google Earth images.

Table 7.1: The population statistics in the study area. The statistics are provided for the total municipality (including areas outside the study area) and the fraction of each municipality located within the study area. The population of the study area is estimated from the population density of the municipality. The population is obtained from the Nepal census 2021 (CBS, 2022), and the number of buildings from OSM (2023).

Region	Municipality			Karnali fan		
	Area [km^2]	Population	Population density [km^2]	Area [km^2]	Population	No. of buildings
Karnali fan	-	-	441	544	239,778	76,870
Bardiya N.P.	896	0	0	73	0	134
Bhajani	176	53,795	306	39	11,962	3,786
Geruwa	78	33,742	431	78	33,742	14,033
Janaki	107	49,835	465	104	48,221	10,557
Joshipur	65	37,167	568	11	6,493	728
Lamkichuha	225	90,941	405	20	8,169	2,228
Madhuwan	130	51,173	395	24	9,357	2,906
Rajapur	127	61,431	484	83	40,187	17,692
Thakurbaba	104	49,360	487	12	5,986	1,457
Tikapur	118	90,115	763	99	75,661	23,349

The objective of this chapter is the mapping of the spatial flood hazard patterns for the projected flood magnitudes and topographies (Objective 5). This objective is addressed by applying a hydrodynamic model with the current and projected flood magnitudes of the 10-year, 50-year, and 100-year return periods, and with the current and projected topographies. The hydrodynamic simulations provide maps of the flood inundation depths to identify the areas which are most prone to floods, and which are projected to be most affected by the changes in the flood magnitude and/or the morphology.

Furthermore, it is assessed which models contribute to the uncertainty in the flood hazard predictions. This research stage builds upon the hydrological simulations and the flood frequency analysis for the quantification of the flood magnitude (Chapter 5), and the morphodynamic simulations for the prediction of the topography (Chapter 6). These simulations contain uncertainty that propagates to the hydrodynamic predictions, particularly the uncertainty in the climate projections, the hydrological model, the Flood Frequency Analysis (FFA), and the morphodynamic simulations. The quantification of the sources of uncertainty is needed to improve the predictions in future studies by targeting the largest sources of uncertainty, e.g. by focusing the field surveys on specific elements of the modelling cascade.

Section 7.2 introduces the hydrodynamic model, the ensemble design, and the used datasets. Section 7.3 presents the results of the hydrodynamic simulations, which are then discussed in section 7.4. The last section 7.5 concludes this chapter.

7.2 Methods

7.2.1 The hydrodynamic model

The 2-D LISFLOOD-FP inertia flow model (Bates *et al.*, 2010) is applied in the Karnali fan to conduct the mapping of the spatial flood hazard characteristics for the current and projected flood magnitudes and topographies. This model simulates the flow of water in the two cardinal directions based on derivatives of the 1-D shallow water equations which are solved in the longitudinal and lateral directions (Bates *et al.*, 2010). It is the same model that was used to simulate the water flow in the morphodynamic simulations of Chapter 6 and was described in Section 6.2.2.1. The model was chosen because it is computationally efficient while maintaining good accuracy for gradually varied, subcritical flow conditions that prevail in the Karnali fan (Bates *et al.*, 2010; Neal *et al.*, 2012). This fast computation is beneficial because the research design requires the simulation of ensembles which will be described in the next section.

The model is applied within the CAESAR-LISFLOOD C-L code (Bates *et al.*, 2010; Coulthard *et al.*, 2013) which was used for the morphodynamic simulations in Chapter 6. The parameterisation is maintained from the morphodynamic simulations (see Table 6.1) but the sediment transport model is deactivated and only the flow of water is simulated. The parametric uncertainty is not estimated for the hydrodynamic models to maintain a feasible ensemble size for the simulations.

7.2.2 Ensemble design

The simulation ensemble is, conceptually, divided into two groups based on their objective. The first group is the ensemble to map the current and future flood hazards to quantify the changes (O5) and the second group is simulated to understand the uncertainty in the simulations . These sources of uncertainty are:

- The climatic uncertainty: An ensemble of 12 CMIP6 climate models was used to predict the flood discharge at the fan inlet for the projected climates (Chapter 5). These 12 models represent climate uncertainty.
- The hydrological uncertainty: The flood magnitudes were obtained from an ensemble of 64 hydrological models and 1,000 Flood Frequencies Analyses resulting in 64,000 flood magnitudes for each climate model and scenario (Chapter 5).

- The topography: The topographies were simulated for three flood discharge percentiles ($P_{2.5}$, P_{50}, $P_{97.5}$) of the hydrological ensemble (64 models) for each climate model (Chapter 6). The parameter uncertainty of the morphological model was not assessed. However, the impact of the morphological evolution can be quantified by comparing simulations with the initial and the projected topographies.

The first ensemble is applied for the baseline scenario (1975 – 2014) to represent the current climatic conditions, and the medium- and high-emission scenarios SSP245 and SSP585 (2060 – 2099) which represent the potential future climatic conditions. This ensemble comprises three members to consider the lower, median, and upper flood mangitudes predictions. These are represented by the climate model CM with the lowest, median, and highest predicted flood magnitudes at the 2.5^{th} (CM_{lower}), 50^{th} (CM_{median}), and 97.5^{th} (CM_{upper}) percentile of the FFA of each climate scenario (Table 7.2). The topography is represented by the topography simulated for the respective flood discharge interval, climate model, and climate scenario.

The second ensemble to assess the uncertainty sources consists of 18 models for each climate scenario. Each ensemble member of the first ensemble (Table 7.2) is simulated with the $P_{2.5}$, P_{50}, and $P_{97.5}$ percentiles of the FFA (the ensemble of flood frequency curves of the respective CMIP6 member). Hence, nine ensemble members are simulated for each climate scenario and return period. These nine members are simulated for both, the initial and projected topographies resulting in 18 ensemble members.

The number of climate models is uneven which hampers the selection of the median model which would be the mean of the models with the sixth and seventh highest predictions. However, each ensemble member has a corresponding projected topography and calculating the mean topography of two members might distort the channel patterns and hence both members would need to be simulated. Therefore the climate model with the sixth highest flood magnitude is considered the median member to reduce the required simulations.

The flood hazard is simulated for the flood magnitudes generated from the observed discharge record for the $P_{2.5}$, P_{50}, and $P_{97.5}$ percentiles (Figure 5.36). These simulations are conducted with the initial topography and used for comparison with the baseline simulations to investigate how the underestimation of the baseline magnitudes translate into the spatial hazard characteristics.

Table 7.2: The ensemble members for the quantification of the climate change impact on the spatial flood hazard characteristics. The descriptor is the synonym used in the text to refer to the ensemble member. The CMIP6 interval represents the lowest, median (6th highest) and highest predicted flood magnitudes of the CMIP6 ensemble, and the hydrological interval refers to the percentile of the Flood Frequency Analysis predictions used for the ensemble member.

Return period [years]	Climate Scenario	Des-criptor	CMIP6 interval	Hydro. Interval	CMIP6-member	Flood mag. [m^3/s]
10	Baseline	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-HR	4,087
	SSP245	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-HR	4,552
	SSP585	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-HR	5,153
	Baseline	CM_{median}	Median	P_{50}	MRI-ESM2-0	8,079
	SSP245	CM_{median}	Median	P_{50}	ACCESS-CM2	12,327
	SSP585	CM_{median}	Median	P_{50}	MRI-ESM2-0	16,814
	Baseline	CM_{upper}	Highest	$P_{97.5}$	NorESM2-LM	15,779
	SSP245	CM_{upper}	Highest	$P_{97.5}$	ACCESS-ESM1-5	23,400
	SSP585	CM_{upper}	Highest	$P_{97.5}$	ACCESS-ESM1-5	31,760
50	Baseline	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-HR	4,818
	SSP245	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-HR	5,732
	SSP585	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-LR	6,177
	Baseline	CM_{median}	Median	P_{50}	MRI-ESM2-0	10,837
	SSP245	CM_{median}	Median	P_{50}	INM-CM5-0	15,417
	SSP585	CM_{median}	Median	P_{50}	ACCESS-CM2	20,755
	Baseline	CM_{upper}	Highest	$P_{97.5}$	INM-CM5-0	20,206
	SSP245	CM_{upper}	Highest	$P_{97.5}$	NorESM2-LM	34,802
	SSP585	CM_{upper}	Highest	$P_{97.5}$	EC-Earth3-Veg	40,065
100	Baseline	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-HR	5,104
	SSP245	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-HR	6,266
	SSP585	CM_{lower}	Lowest	$P_{2.5}$	MPI-ESM1-2-LR	6,434
	Baseline	CM_{median}	Median	P_{50}	MRI-ESM2-0	12,009
	SSP245	CM_{median}	Median	P_{50}	INM-CM5-0	17,174
	SSP585	CM_{median}	Median	P_{50}	EC-Earth3	22,763
	Baseline	CM_{upper}	Highest	$P_{97.5}$	INM-CM5-0	22,954
	SSP245	CM_{upper}	Highest	$P_{97.5}$	NorESM2-LM	42,916
	SSP585	CM_{upper}	Highest	$P_{97.5}$	EC-Earth3-Veg	45,705

7.2.3 Boundary conditions

The model requires topography and flood inflow as boundary conditions. The initial and projected topographies of the morphological simulations are used to represent the topographic boundary conditions. However, the initial topography is represented by a DEM after five years of morphological simulations because it has a more realistic representation of the bed topography which has no lateral gradient in the raw DEM. The spatial resolution of the modelling is similar to the morphodynamic modelling in Chapter 6 (25 X 25 m).

The flood flow is represented by the flood magnitudes predicted by the FFA in Chapter 5. These were generated from the daily mean discharge and, thus, represent daily means and are simulated for 24 hours. The model is initiated for 12 hours with a constant discharge of 3,000 m^3/s which is below bankfull flow. The rising and recession limbs have a duration of 12 hours each and are linearly interpolated between 3,000 m^3/s and the flood magnitude.

7.2.4 Modelling setup

The hydrodynamic model is run for 72 hours whereas the flood discharge is inserted into the upstream border between the hours 24 – 48. It takes 12 hours for the water to travel from the upstream border to the downstream border in the channels. This travel time is longer for floodplain flows. However, the water levels remain similar after 24 hours after the peak discharge which indicates that all water that is not trapped in depressions drains within a day of the flood. Therefore, the model is run for three days, the first day is the warmup period, the second day is the flood event, and the third day is the flood recession.

The predicted water levels are written as raster files at hourly resolution. The maximum water level of the 72 rasters is calculated for each grid cell and this maximum water level is used to determine the inundation extent and the inundation depth. The inundation extent is classified as cells with water depths ≥ 0.1 m. The inundation extent refers unless stated otherwise, to the area inundated by depths ≥ 0.1 m.

The depth-damage ratio describes the damage potential of different inundation depths (although the flood duration is an important factor too). The depth-damage ratios vary between different building types, crop types and growth stages, and are unknown for Nepal due to the lack of inundation mapping and damage surveys (ADPC, 2010). Shallow inundations have damage potential (e.g. structural and functional damages to buildings, loss of hosing interiors, etc.) (ADPC, 2010). Therefore, a shallow 0.1 m inundation depth threshold is chosen for the calculation of the inundation extent to prevent the underestimation of the damage potential by overestimating the unknown depth threshold.

The population affected by the flood hazard is estimated from the location of the buildings in the study area (Figure 7.1) (OSM, 2023). For this, a boolean raster is created which describes whether a cell contains a building or not. However, this raster is an approximation of the population at risk and not an exact quantification because one cell may include multiple buildings, one building may occupy multiple cells, and the number of people living in a populated cell is variable.

7.3 Results: The predicted flood hazard characteristics

This section presents the results of the hydrodynamic simulations. The flood hazard characteristics of the three climate scenarios are compared in the first section to investigate the projected changes (O5). Section 7.3.2 compares the flood hazard patterns simulated for the observed flood magnitudes and the simulated flood magnitudes of the baseline scenario which provides insight into the BIAS of the flood magnitude predictions. Section 7.3.3 investigates the uncertainty propagation through the modelling cascade by presenting the predictions of the uncertainty ensemble. The last Section 7.3.4 compares the simulations conducted with the initial topography and with the ones conducted with the projected topographies to assess the sensitivity of the flood hazards to the morphological evolution.

Certain areas are of particular interest for the analysis and discussion because they are at higher flood risk or the predictions are more uncertain. The locations of these areas are presented in Figure 7.2. The area between the Eastern and Western branches is prone to flooding and is referred to as the Island. Furthermore, the flood water accumulates in a topographic depression (hereafter referred to as the depression) in the South-Western corner of the study area. Tikapur city has a high population density and is used as a landmark. Tikapur city is separated from the Western branch by a forest. This forest is relevant because the ground elevation in the DTM is estimated (Chapter 6) and hence the flood predictions in this area have a higher uncertainty. This forest is referred to as the Tikapur forest. The Western floodplain describes the area West of the Western branch.

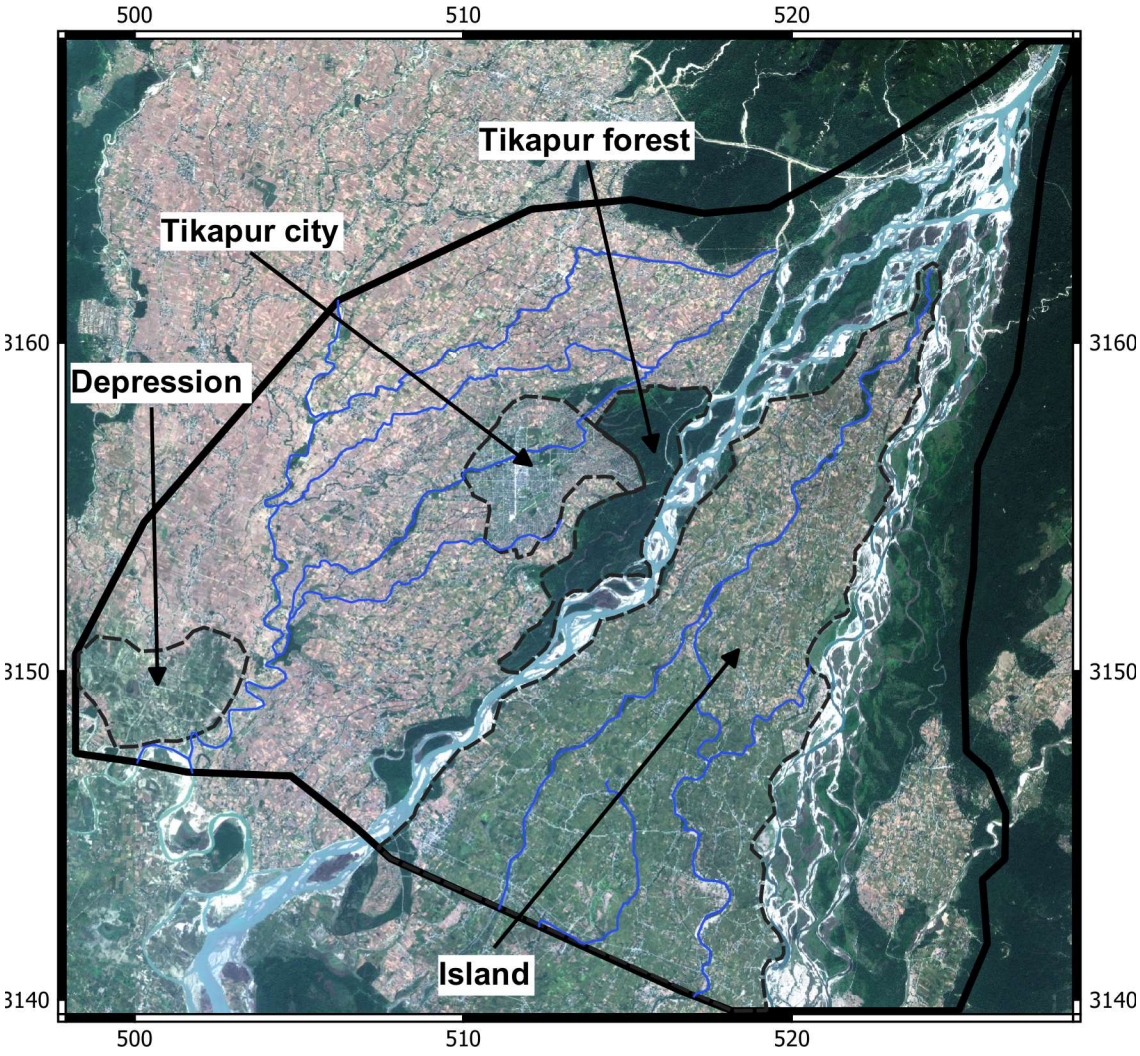


Figure 7.2: The locations of areas of particular interest for the flood inundation simulations. These areas of interest are delineated by the dashed black line. The solid black line demarks the modelling domain and the blue lines indicate smaller channels which drain the floodplain. The background image is a Sentinel-2 composite.

7.3.1 The predictions for the climate scenarios

The comparison of the climate scenarios provides information about how the projected increases in the flood magnitudes translate into the spatial flood hazard characteristics in the Karnali fan. This section compares these characteristics for the prediction interval, whereas the CM_{lower} is the $P_{2.5}$ flood magnitude of the climate model with the lowest respective flood magnitudes, CM_{median} is the median flood magnitude (P_{50}) of the climate model with the 6th highest respective flood magnitude, and CM_{upper} is the $P_{97.5}$ of the climate model with the highest predicted respective flood magnitudes. These simulations are conducted with the projected topographies of the respective climate model and flood discharge percentile of Chapter 6.

The median inundation extent (inundation depth ≥ 0.1 m) increases with the return period and the emission scenario. The 1-in-10-year flood magnitude inundates 141 km^2 in the baseline (Table 7.3). This extent increases by 55% to 217 km^2 for the medium-emission scenario SSP245 and by 94% to 273 km^2 for the high-emission scenario SSP585 in the far future (2060 – 2099). The projected increase is lower for the higher return periods; the 1-in-50-year flood extent increases from 197 km^2 in the baseline to 261 km^2 (+33%) for SSP245 and to 309 km^2 (+57%) for SSP585. The 1-in-100-years flood extent increases from 214 km^2 in the baseline to 276 km^2 (+29%) for SSP245 and to 324 km^2 (+51%) for SSP585. This projected increase in the inundation extent concentrations on the Island and the Western floodplain (excluding the northern parts) while the areas along the eastern and north-western borders remain dry in all scenarios (Figure 7.3). The more hazardeous inundations (inundation depth ≥ 1 m) increase with the emissions particularly along the channels in the island and in the South-Western depression (Figure 7.4). For the 1-in-10-years event, such floodplain inundations are not predicted for the baseline but for both projected scenarios whereas the extent increases with the emissions (median predictions).

Table 7.3: Inundation extents of the Karnali fan (inundation depth ≥ 0.1 m) predicted for the climate scenarios.

		Inundation extent [km^2]			**Change to baseline [%]**	
Return period	**Prediction interval**	**Baseline**	**SSP245**	**SSP585**	**SSP245**	**SSP585**
1-in-10-years	Lower	62	69	74	11	20
	Median	141	217	273	55	94
	Upper	259	331	366	28	41
1-in-50-years	Lower	70	84	89	19	28
	Median	197	261	309	33	57
	Upper	302	375	387	24	28
1-in-100-years	Lower	73	90	93	23	27
	Median	214	276	324	29	51
	Upper	324	392	397	21	23

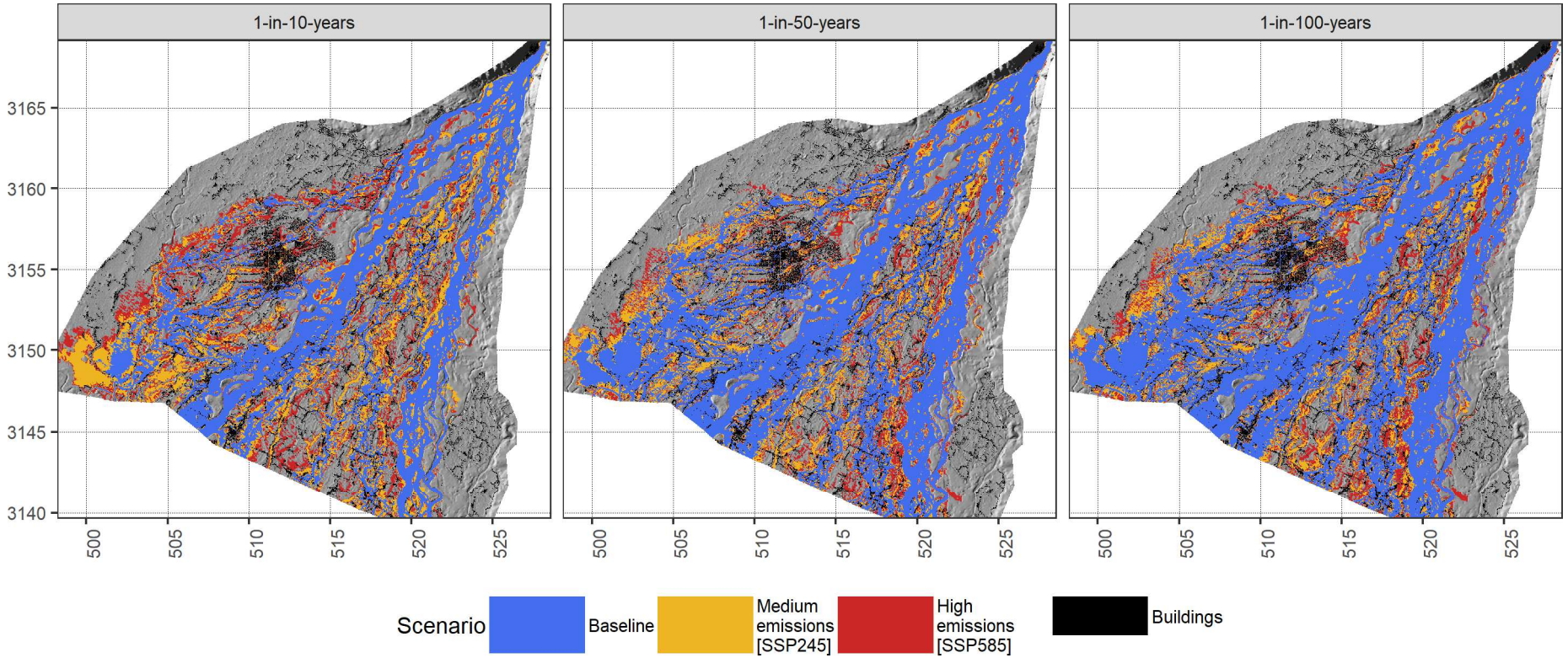


Figure 7.3: The inundation extent (inundation depth ≥ 0.1 m) for the median predictions (CM_{median}) of the three CMIP6 scenarios for the 1-in-10-years, 1-in-50-years and 1-in-100-years flood events. The buildings are obtained from OSM (2023). The coordinate system is UTM Zone 44 N in km. The inundation extents (inundation depth ≥ 0.5 m) are presented in Appendix 7.1.

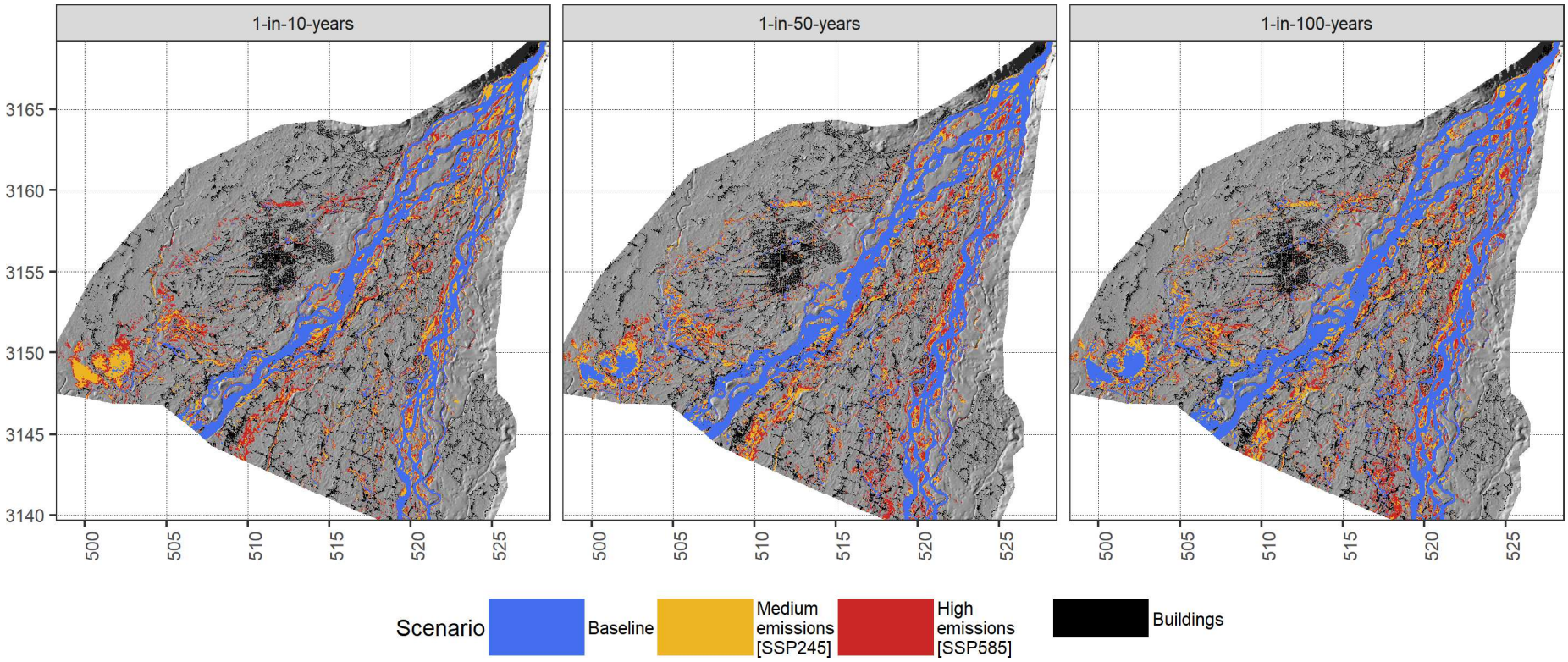


Figure 7.4: The inundation extent (inundation depth ≥ 1.0 m) for the median predictions (CM_{median}) of the three CMIP6 scenarios for the 1-in-10-years, 1-in-50-years and 1-in-100-years flood events. The buildings are obtained from OSM (2023). The coordinate system is UTM Zone 44 N in km. The inundation extents (inundation depth ≥ 0.5 m) are presented in Appendix 7.1.

Climate change increases the frequency of extreme events and hence current extremes are projected to occur frequently in the future. The 1-in-100-years event of the baseline (1975 – 2014) inundates 39% of the fan (Figure 7.5). This extent occurs statistically once every 10 years in the far future (2060 – 2099) of the medium-emission scenario SSP245 (40%). Furthermore, the flood frequencies vary greatly between both projected scenarios. The inundation extent of the 1-in-100-years extent projected for the medium-emission scenario SSP245 (51%) is similar to the extent of the 1-in-10-years event projected for the high-emission scenario SSP585 (50%). Hence, the frequency of flood hazard increases with the emissions and the current extremes are projected to occur frequently in the future (median predictions).

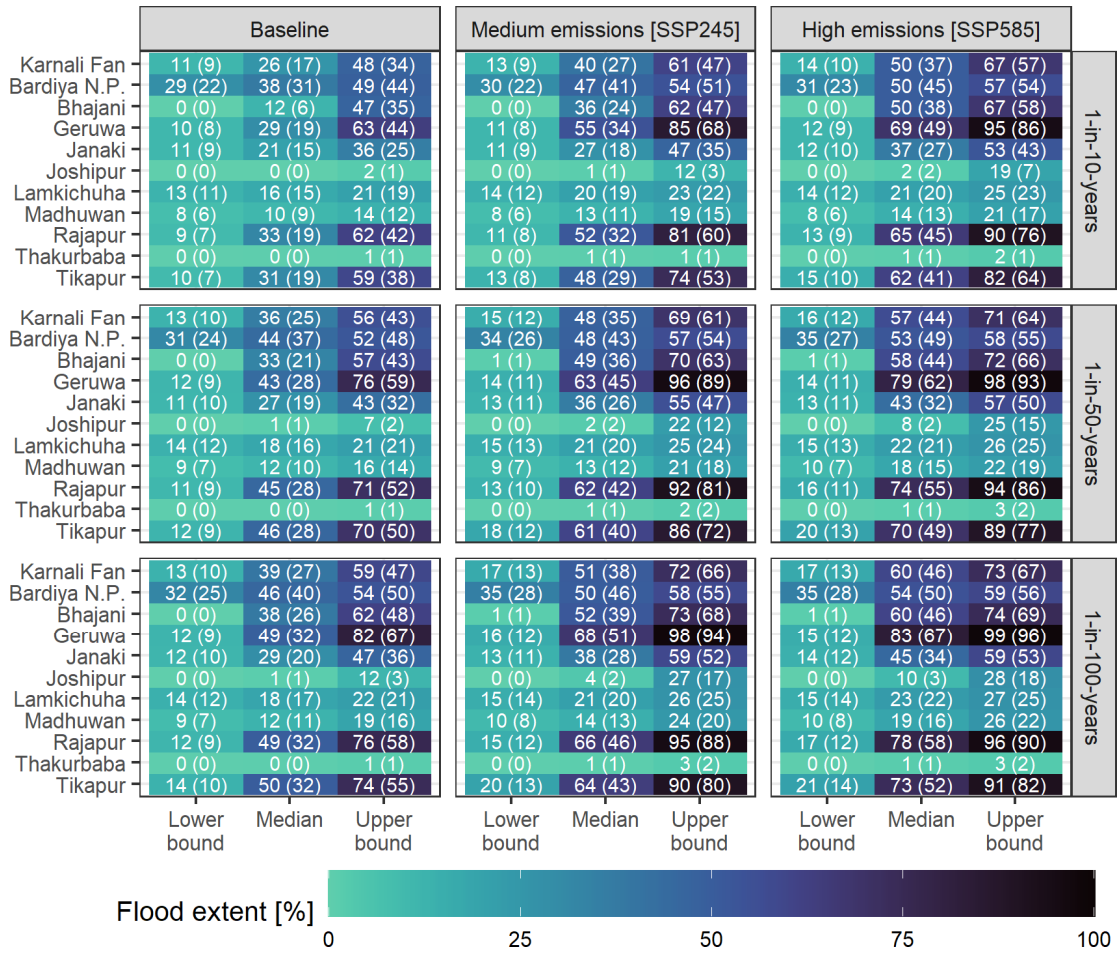


Figure 7.5: The proportion of the study area (Karnali fan) and the municipalities which are inundated for different climate scenarios during flood events of different return periods. The numbers indicate the percentage of the municipality that is inundated deeper ≥ 0.1 m. The numbers in the brackets indicate the inundation extent that is inundated deeper ≥ 0.5 m. The x-axis is the prediction interval of the modelling ensemble. The spatial extent of these predictions is shown in Figure 7.6.

The municipalities most affected by the projected increase in flood magnitudes are the densely populated municipalities Geruwa, Rajapur (both located on the Island) and Tikapur (Western floodplain). These municipalities are inundated to 49 – 50% (inundation depth ≥ 0.1 m) during a 1-in-100-years event in the baseline (Figure 7.5). This fraction increases to 64 – 68% (SSP245) and 73 – 83% (SSP585). Large fractions of these municipalities are projected to be inundated by inundation depths with a higher damage potential (inundation depth ≥ 0.5 m). These higher depths are predicted for 32% in the baseline and increase to 43 – 51% (SSP245) and 52 – 67% (SSP585). Furthermore, the Bhajani municipality and the Bardiya National Park are projected to be inundated (inundation depth ≥ 0.1 m) by ≥ 50% in the projected scenarios. However, the inundation in the Bardiya National Park is predicted for the channels while the floodplain remains dry due to its superelevation over the channel (Figure 7.3) (median predictions).

The variation of the ensemble predictions exceeds the variation of the climate scenarios. The inundation of CM_{lower} is restricted to the channels and a small section in the South of Tikapur city and no floodplain flow is predicted for the 1-in-10-years event (Figure 7.6). The variation between return periods and climate scenarios is low for CM_{lower} and ranges between 11% (baseline 1-in-10-years) to 17% (SSP585 1-in-100-years) of the Karnali fan (Figure 7.5). The variation in the inundation extent is larger for CM_{upper} and ranges between 48% (Baseline 1-in-10-years) to 73% (SSP585 1-in-100-years). The variation in the inundation extents is larger for the prediction interval than for the climate scenarios for all return periods (Table 7.3). For example, the difference between CM_{lower} and CM_{upper} 1-in-10-year extents is 197 km^2 (62 – 259 km^2) in the baseline, which is larger than the climate scenario variation of 132 km^2 between the baseline and SSP585 for CM_{median}.

The inundation extent increases with the emissions for each prediction interval. These increases are lower for the boundaries CM_{lower} and CM_{upper} than for CM_{median} (Table 7.3). The inundation extents of CM_{lower} increases by 11% (SSP245) and 20% (SSP585) for the 1-in-10-years event and by 23% (SSP245) and 27% (SSP585) for the 1-in-100-years event. The inundation extent of CM_{upper} is projected to increase by 28% (SSP245) and 41% (SSP585) for the 1-in-10-years event and by 21% (SSP245) and 23% (SSP585) for the 1-in-100-years event. The highest increases are predicted for CM_{median} in absolute and relative terms for each return period (Table 7.3).

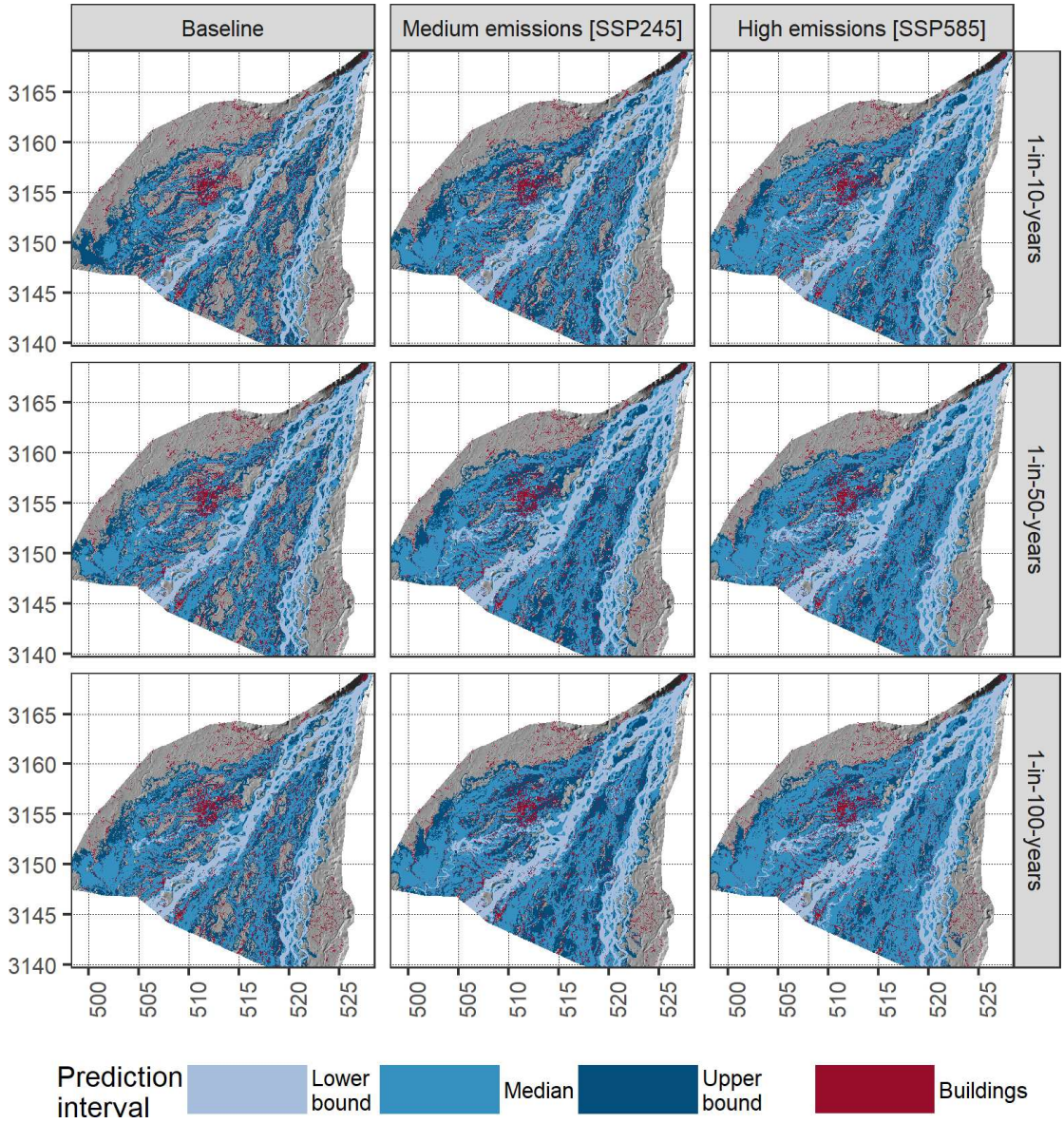


Figure 7.6: The inundation extents (inundation extents ≥ 0.1 m) in the Karnali fan for the lower, median, and upper prediction intervals. The buildings are obtained from OSM (2023). The coordinate system is UTM Zone 44 N in km.

The projected increase of the flood magnitudes leads to deeper inundations in most parts of the Karnali fan. In the baseline, the water level is between 5 – 10 m in the channels of the Eastern and Western branches (Figure 7.7). The floodplain water levels are mostly below 1 m except for the depression in the South-Western part of the fan. This depression fills with floodwater and the inundation depth increases with the return period and the flood magnitude interval to 2 – 3 m, and up to 5 m for the upper prediction boundary (Figure 7.7).

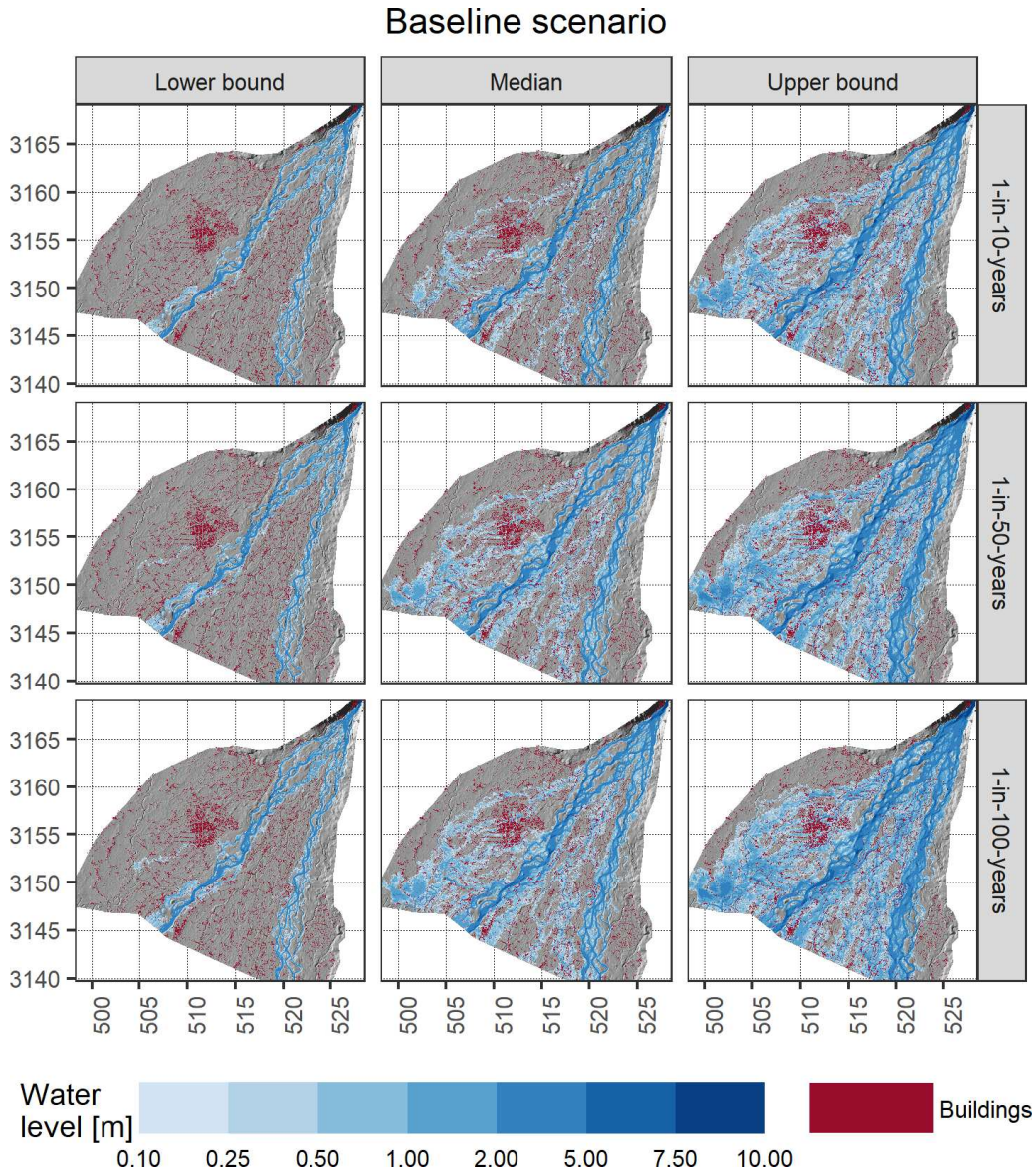


Figure 7.7: The maximum water level (≥ 0.1 m) of the baseline predicted for flood events with different return periods. The left column shows the predictions of the lower boundary of the prediction interval (CM_{lower}), the centre column the ones of the median predictions (CM_{median}), and the right column the ones of the upper boundary of the prediction interval (CM_{upper}). The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km. Note that the scale is non-linear.

The inundation depths increase with the emissions for most areas throughout the floodplain (Figure 7.8). The depths increase by 0.1 – 0.5 m in most parts of the Island and the Western floodplain in the medium-emission scenario SSP245 (median predictions). Larger increases (> 0.5 m) are predicted in the depression and along floodplain channels (see Figure 7.2 for their locations) in the northern half of the Island. These increases of > 0.5 m are predicted for larger areas across the fan, particularly in the Island and the depression, for the SSP585 scenario (median predictions). For CM_{upper}, increases of > 0.5 m dominate throughout the fan for the 1-in-50-years and 1-in-100-years flood events in both scenarios.

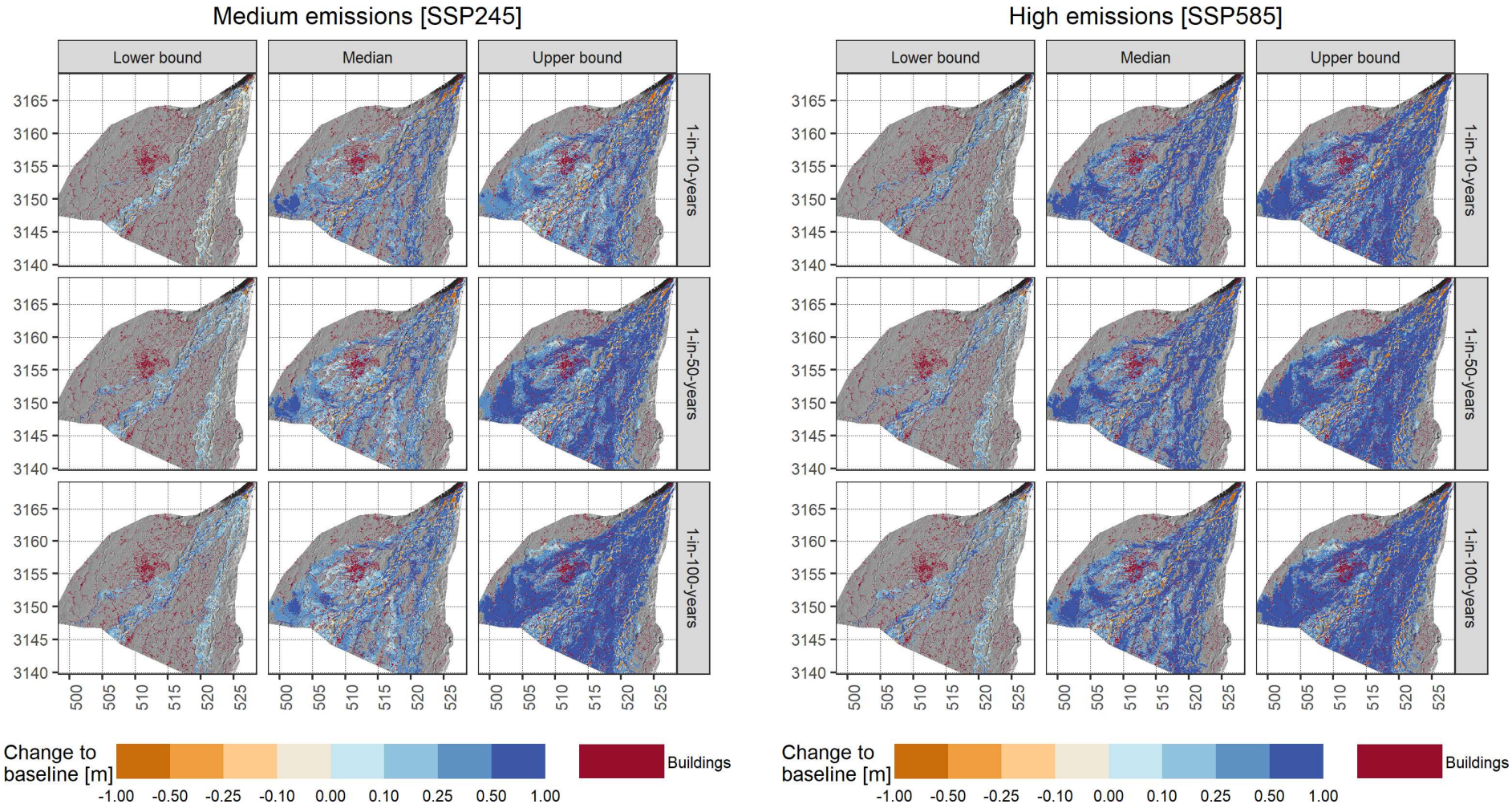


Figure 7.8: The difference between the inundation depths predicted for the projected scenarios (left: medium-emission scenario; right: high-emission scenario) and the baseline for flood events with different return periods. Negative values indicate a decrease of the inundation depth for the projected scenarios and positive values indicate an increase. The columns show the predictions for CM_{lower}, CM_{median} and CM_{upper}. The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km. Note that the scale is not linear and that the changes are capped at -1 m and +1 m and might exceed these values locally. See the Appendices 7.2 and 7.3 for the water levels.

For some areas, a decrease in the inundation depths is predicted despite the projected increase in the flood magnitudes (Figure 7.8). These decreases are generally predicted along the riverbanks of the Karnali branches. Decreasing inundation depths are furthermore predicted locally along floodplain channels around Tikapur city. However, the projected increases dominate the fan and areas with decreasing inundation depths are sparse.

Despite these local inundation depth decreases, the projections predict increases in the for most flood-prone areas which is indicated by the increase in the median inundation depth (Table 7.4). The median depth divides the inundated area in half and hence 50% of the inundated area is inundated by shallower depths and 50% by deeper depths. This median depth increases with the emission scenario for all return periods. The median depth of CM_{median} increases from 0.86 m in the baseline scenario to 0.95 m for SSP245 and 1.05 m for SSP585 for the 1-in-100-years event. The median depth of CM_{upper} increases from 1.09 m in the baseline to 1.6 m (SSP245) and 1.66 m (SSP585). The exception is CM_{lower} for which the median depth decreases with the emissions. For example, the median depth decreases from 1.27 m in the baseline to 1.15 m (SSP245) and 1.11 m (SSP585) for the 1-in-100-years event.

The projected change in the inundation depth varies locally but the projected increases dominate which is also indicated by the distribution of inundation depths for the total fan (Figure 7.9). The area inundated by a certain depth increases with the emissions throughout the range of predicted depths for the CM_{lower} and CM_{median}. The inundated area decreases with increasing depth and hence the largest fraction is inundated by shallow inundations (< 0.5 m). The distribution of CM_{upper} has a different shape, particularly for the 1-in-50-years and 1-in-100-years events. The inundated area increases until inundation depths of 1.0 – 1.5 m after which it declines. This means that larger areas are inundated by deeper flows than by shallower flows. Very deep inundations (≥ 4 m) are predicted within the channel and hence the variation between the prediction intervals is small for the very deep flows.

Table 7.4: Median inundation depths (ID) of flood events with different return periods. The area columns present the spatial extent of the median ID. This area is inundated by depths < the median ID and depths > the median ID and hence the total inundated area is twice this area. Only inundated pixels with depths ≥ 0.1 m are included in the calculations.

		Baseline		Medium-emissions [SSP245]		High-emissions [SSP585]	
		ID [m]	Area [km²]	ID [m]	Area [km²]	ID [m]	Area [km²]
1-in-10-years	Lower	1.36	31	1.2	34	1.23	37
1-in-10-years	Median	0.83	70	0.81	109	0.91	136
1-in-10-years	Upper	0.9	129	1.04	165	1.27	183
1-in-50-years	Lower	1.29	35	1.15	42	1.12	45
1-in-50-years	Median	0.83	99	0.9	131	1.01	154
1-in-50-years	Upper	1.02	151	1.37	187	1.51	193
1-in-100-years	Lower	1.27	37	1.15	45	1.11	46
1-in-100-years	Median	0.86	108	0.95	138	1.05	162
1-in-100-years	Upper	1.09	162	1.6	196	1.66	198

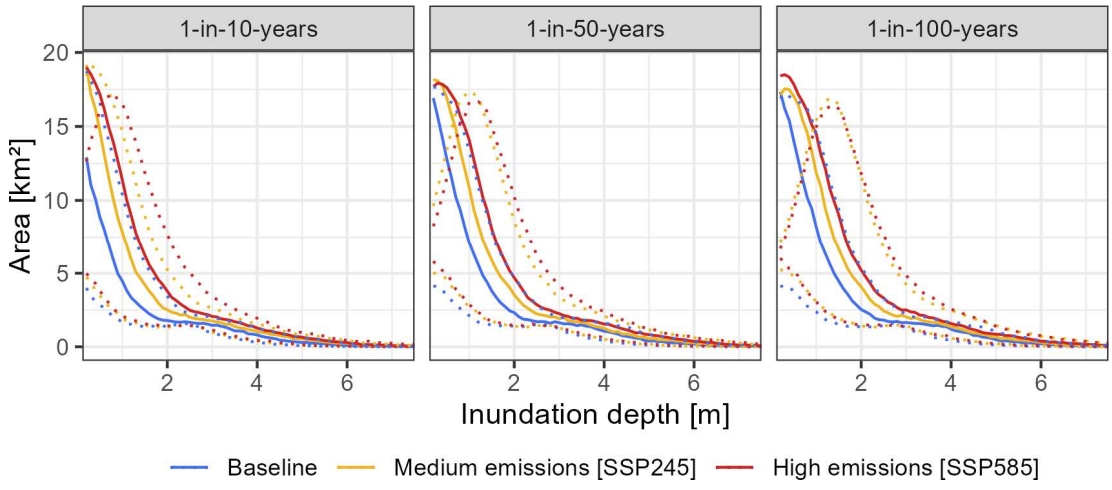


Figure 7.9: The spatial distribution of inundation depths for different return periods. The x-axis shows the depth (< 7.5 m) and the y-axis shows the area which is inundated by the respective depth. The solid lines indicate the distributions of CM_{median} and the dotted lines indicate the distributions of CM_{lower} and CM_{upper}.

The population affected by flooding is projected to increase with emissions for all return periods. The fraction of inundated cells containing buildings (hereafter referred to as populated cells) quadruples from 8% in the baseline to 32% in the high-emission scenario SSP585 during the 1-in-10-years event (Figure 7.10). The projected 1-in-10-years event of SSP245 affects a larger population (20%) than the baseline 1-in-100-years event (19%). The most affected municipalities are the densely populated municipalities of Geruwa and Rajapur in the Island, and Tikapur in the Western floodplain (see Figure 7.1 for the location of the municipalities). In these municipalities, 22 – 27% of the populated pixels are inundated during the 1-in-100-years event in the baseline, which increases to 37 – 49% (SSP245) and 52 – 69% (SSP585) (median predictions).

The predictions of the affected population vary strongly between the prediction intervals (Figure 7.10). CM_{lower} predicts the inundation of 1-3% of the populated cells, and CM_{upper} predicts the inundation of 45 – 72% of the populated cells for the 1-in-100-years event (all scenarios). As for the inundated area, the variation within the prediction interval exceeds the variation of the climate scenarios. For example, the scenario variation of the 1-in-10-years event of CM_{median} is 16% (8 – 32%), while the difference between CM_{lower} and CM_{upper} is 29% (0 – 29%) in the baseline and 59% (1 – 60%) in the high-emission scenario. However, an increase in the inundated populated pixels is predicted for each prediction interval and return period. The most extreme projections (CM_{upper} for SSP245 and SSP585) predict that 80 – 100% of the populated pixels of the densely populated municipalities Geruwa, Rajapur and Tikapur are inundated during the 1-in-100-years flood event.

The inundation depth, and thus the damage potential is projected to increase with the emissions (Table 7.5). The median inundation depth (50% of the populated cells are inundated deeper and 50% are inundated shallower than the median depth) increases from 0.41 m in the baseline to 0.51 m (SSP245) and 0.57 m (SSP585) for the 1-in-100-years event. The scenario differences in the inundation depths increase with the depth percentile, and the depth at the 90th percentile (10% of the populated cells are inundated deeper) increases from 1.05 m in the baseline to 1.15 (SSP245) and 1.35 (SSP585) for the 1-in-100-years event (median predictions). Both these characteristics, the increase of inundation depth with emissions, and the increasing variations with the depth percentile are consistent for all return periods for CM_{median} and CM_{upper}. CM_{lower} projects a decrease in the median depth for some return periods. This is caused by the increase of inundated populated cells which leads to a relative decrease in the inundation depths.

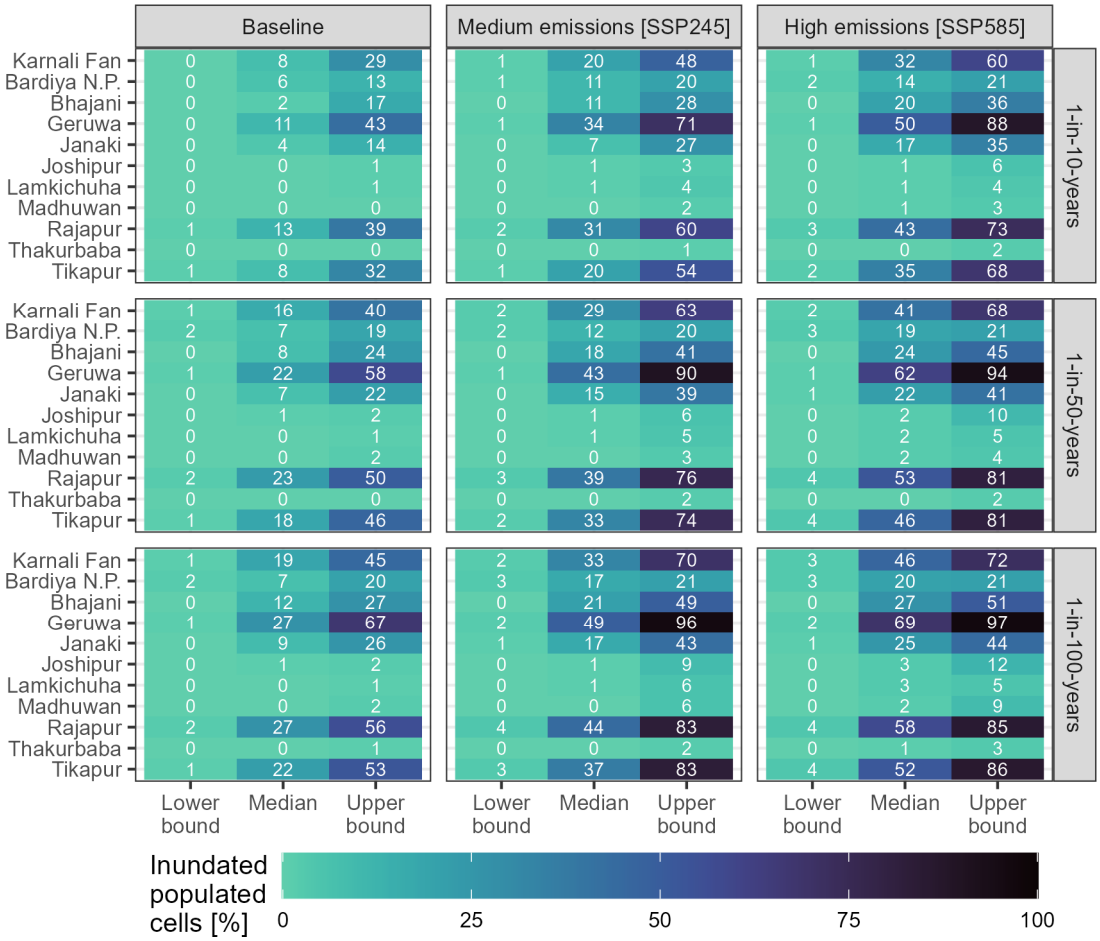


Figure 7.10: The fraction of the populated cells in the total study area (Karnali fan) and the municipalities which are inundated by at least 0.1 m. The populated cells are classified as the cells of the modelling grid which are occupied by at least one building in the OSM (2023) dataset The x-axis plots the prediction interval of the ensemble.

Table 7.5: Statistics of the inundated populated area. The inundated cells present the number of cells which contain at least one building (OSM, 2023) and are inundated by ≥ 0.1 m. The inundation depths present the depths of the median and the 90th percentile.

Return period	Prediction interval	Inundated cells			Median inundation depth [m]			P_{90} inundation depth [m]		
		BL	SSP-245	SSP-585	BL	SSP-245	SSP-585	BL	SSP-245	SSP-585
1-in-10-years	Lower	230	471	646	0.36	0.38	0.41	0.75	0.85	0.95
1-in-10-years	Median	4,100	10,411	16,578	0.36	0.39	0.48	0.95	1.05	1.15
1-in-10-years	Upper	14,779	24,461	30,517	0.45	0.6	0.77	1.15	1.35	1.75
1-in-50-years	Lower	420	902	1,240	0.42	0.42	0.4	0.95	0.95	1.05
1-in-50-years	Median	8,179	14,968	21,165	0.39	0.47	0.55	0.95	1.15	1.25
1-in-50-years	Upper	20,349	32,427	34,889	0.55	0.88	0.97	1.25	1.75	1.95
1-in-100-years	Lower	498	1,186	1,388	0.44	0.42	0.4	0.95	1.05	1.05
1-in-100-years	Median	9,849	16,860	23,583	0.41	0.51	0.57	1.05	1.15	1.35
1-in-100-years	Upper	23,294	35,918	36,782	0.6	1.05	1.1	1.35	1.95	2.15

7.3.2 The comparison of the flood hazard predicted for the baseline and observed flood magnitudes

This section compares the hydrodynamic simulations for the flood magnitudes predicted from the simulated flood flows of the baseline simulations and those predicted for the observed discharge record. This provides information about the bias in the flood magnitudes of the hydrological simulations, which results from the climate and hydrological models, and propagates into the spatial flood hazard predictions. For this, the lower prediction interval CM_{lower} (lowest climate model and 2.5th percentile of the flood magnitude), the median prediction interval CM_{median} (6th highest climate model and median flood magnitude), and the upper prediction interval CM_{upper} (highest climate model and 97.5th percentile of the flood magnitude) are compared with the 2.5th percentile, the median, and the 97.5th percentile of the flood frequency analysis of the observed discharge record (referred to as the observed predictions) (Figure 5.36). These simulations are conducted with the initial topography to compare them with hydrodynamic simulations of other studies and hence the baseline extents and depths are different to the ones in the previous section.

The flood magnitudes of the baseline record are lower than the observed ones which results in the underprediction of the spatial flood hazard characteristics. The baseline simulations underpredict the inundation extent for all return periods. These underpredictions decrease with the prediction interval. The lower prediction interval CM_{lower} predicts an inundation extent (inundation depth ≥ 0.1 m) of 73 km^2 (Baseline) and 224 km^2 (observed) for the 1-in-100-years event (Table 7.6). This equals a percentage difference of 102%. This percentage difference decreases to 43% for the CM_{median} and further to 13% for CM_{upper}. These percentage differences in the inundation extent are consistent for the three investigated return periods.

Table 7.6: The inundation extent and depth predicted for the observed (obs) and baseline (base) flood magnitudes. The floods are predicted for the initial topography and hence the baseline extents are different to the ones in Table 7.3. The Diff-column present the percentage difference (inundation extent) and absolute difference (inundation depth) between the observed and baseline predictions. The prediction interval represents the lower, median, and upper predictions of the baseline ensemble and the 2.5th, median and 97.5th percentiles of the flood magnitudes of the observed record.

Return period	Prediction interval	**Inundation extent**			**Median depth**		
		Obs [km^2]	Base [km^2]	Diff [%]	Obs [m]	Base [m]	Diff [m]
1-in-10-years	Lower	175	58	100	0.71	1.41	0.70
1-in-10-years	Median	219	139	45	0.78	0.68	0.10
1-in-10-years	Upper	271	257	5	0.88	0.84	0.04
1-in-50-years	Lower	215	68	104	0.77	1.23	0.46
1-in-50-years	Median	294	190	43	0.94	0.73	0.21
1-in-50-years	Upper	346	302	14	1.16	0.96	0.20
1-in-100-years	Lower	225	73	102	0.79	1.16	0.37
1-in-100-years	Median	322	207	43	1.03	0.75	0.28
1-in-100-years	Upper	369	324	13	1.32	1.04	0.28

The spatial characteristics of the observed and baseline predictions are similar for the median and upper prediction interval (Figure 7.11). The inundation patterns of the median predictions are, on a large scale, similar to the ones of the median observed predictions. The inundation of the Western floodplain and the Island is predicted for both simulations. However, the inundated areas have a larger lateral extent in the observed predictions which is particularly evident in the South-Western border of the Karnali fan. Therefore, the baseline predictions underestimate the full extent of the inundation but the large-scale patterns are well reproduced. The differences between the baseline and observed inundation extents are small for the upper prediction interval. Particularly for the 1-in-10-years events, differences are marginal and only predicted on the local scale.

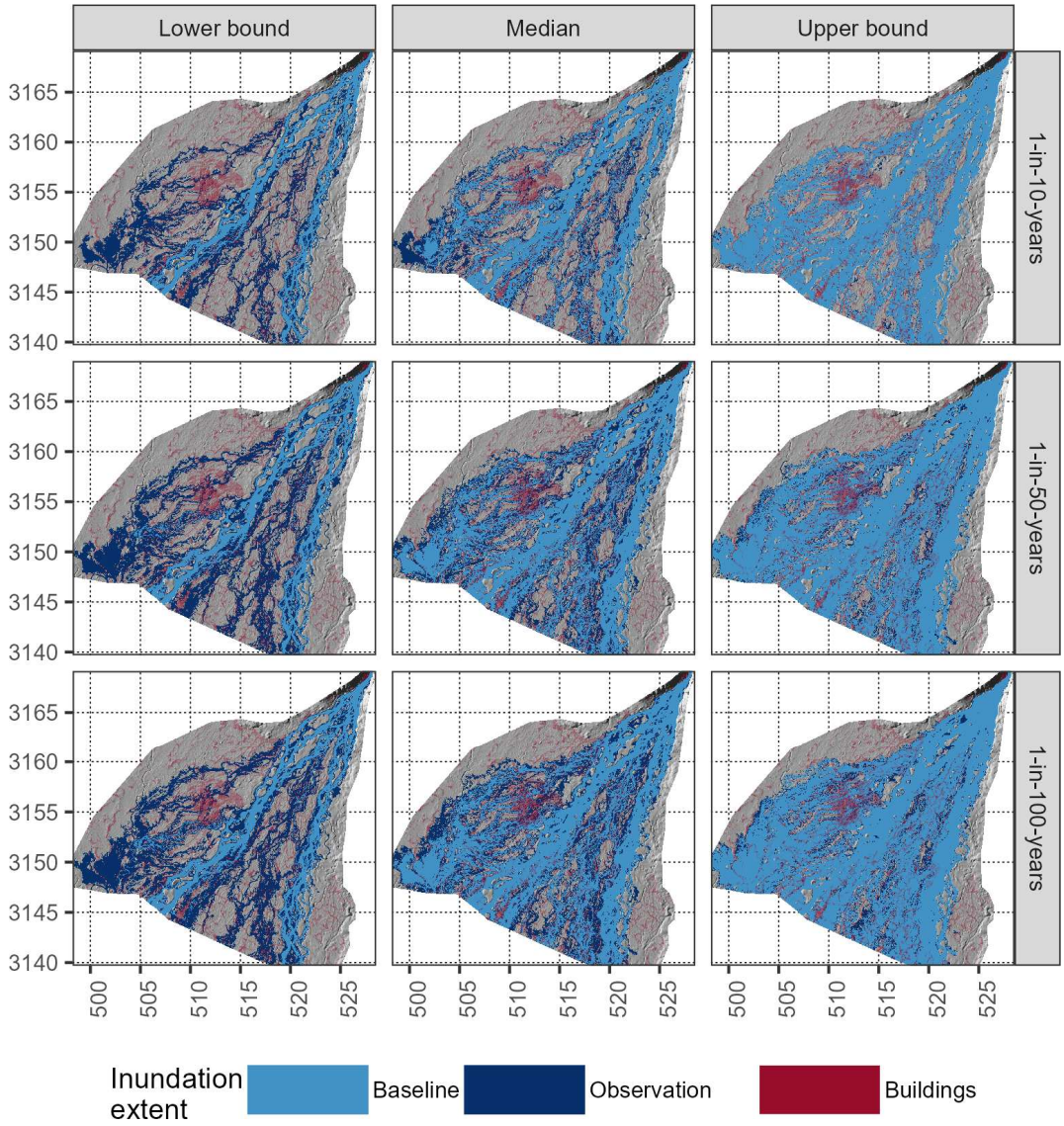


Figure 7.4: The inundation extents (inundation depth ≥ 0.1 m) predicted for the flood magnitudes derived from the observed discharge record and the simulated record of the baseline scenario. The columns show the prediction intervals of the lower, median, and upper predictions of the baseline ensemble, and the 2.5th, median and 97.5th percentiles of the flood frequency analysis ensemble of the observed record. The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km.

The baseline inundation extent of the lower prediction interval vastly underestimates the observed one and does not replicate its spatial characteristics (Figure 7.11). The Western floodplain and the Island are flooded in the observed predictions for all three return periods. In the baseline simulations, the inundation is restricted to the main channels for the 1-in-10-years event. A narrow area is flooded in the South of Tikapur City for the higher return periods but the inundation in the Western floodplain is largely underestimated. The Island remains dry in the baseline simulations for all return periods while it is flooded in the observed predictions.

The baseline simulations underpredict the inundation depth alongside the inundation extent (Table 7.6). The median inundation depth of the lower prediction interval is greater in the baseline predictions than in the observed predictions. This is caused by the larger fraction of channel cells due to the underestimation of floodplain inundation in the baseline predictions. The median depths are higher in the observed predictions for the median and upper prediction interval. This difference increases from 0.1 m for the 1-in-10-years event to 0.28 m for the 1-in-100-years event (median predictions). The difference between the median depths of the upper prediction interval is marginal for the 1-in-10-years event (0.04 m). Hence, the baseline predictions do not only reproduce the spatial inundation extent but also the depth with high accuracy. For the 1-in100-years event, this difference increases to 0.28 m.

The spatial flood hazard characteristics (inundation extent and depth) of the baseline simulation are lower compared to the observed simulations but the differences are small for the upper prediction interval and reasonable for the median predictions. However, the lower prediction interval underestimates the flood hazard characteristics as most of the flood water remains within the channels and hence the flood-prone areas are not identified. The highest agreement between the baseline and observed simulations is predicted for the 1-in-10-years event of the upper prediction interval for which the percentage difference of the inundation extents is 5% and median depth difference is 0.04 m (Table 7.6).

7.3.3 The sources of uncertainty in the flood hazard predictions

Section 7.3.1 has shown that the ensemble variation in the predictions of the inundation extent is great and exceeds the variation of the climate scenarios (Table 7.3). This variation arises from the uncertainty in the flood magnitudes, the climate models, and the projected topographies. This section quantifies the contribution of each of these sources to the total variation. For this, the hydrodynamic simulations have been conducted for each combination of the:

- Climate models (CM): the climate models with the lower, median (6th highest), and highest predicted flood magnitudes CM_{lower}, CM_{median} and CM_{upper};
- Flood magnitudes (FM): the flood magnitudes predicted at the 2.5th, median and 97.5th percentiles of the FFA ensemble of each climate model 2.5th FM, median FM and 97.5th FM;
- Morphology: the initial topography and the projected topography of the morphodynamic simulations (Chapter 6).

This results in 18 simulations with an individual combination of the three CM, the three FM, and the two topographies. The predicted inundation extents of these simulations are presented in Figures 7.12 and 7.13 and summarised in Table 7.7.

The uncertainty in the predictions of the inundation extent is driven by the uncertainty in the Climate Models (CM) and Flood Magnitudes (FM) while the influence of the topography is low. This influence of the CM and FM is illustrated by the variation in the inundation extents of the 1-in-100-years event (Figure 7.13). The extent predicted by CM_{lower} and CM_{upper} varies by 109 km^2 (146 – 255 km^2) for the median FM of the baseline scenario with the initial topography (Figure 7.13). The extent predicted by the 2.5th and 97.5th FM varies by 179 km^2 (116 – 289 km^2) for the CM_{median} (baseline, initial topography).

Table 7.7: The mean prediction range of the inundation extent. This range is calculated for each combination of the remaining variables and then aggregated to the mean. For example, the mean prediction range of the topography is the mean range of the inundation extents predicted for the initial and projected topographies for each combination of climate scenario, climate model, and flood magnitude (27 combinations). The prediction range of the flood magnitudes is the difference between the extent predicted for the 2.5th and the 97.5th percentiles for each combination of climate scenario, climate model, and topography (18 combinations).

	Variable	1-in-10-years	1-in-50-years	1-in-100-years
Mean prediction range [km²]	Climate scenarios	93	87	85
	Climate models	123	120	124
	Flood magnitudes	136	157	163
	Topography	7	6	5

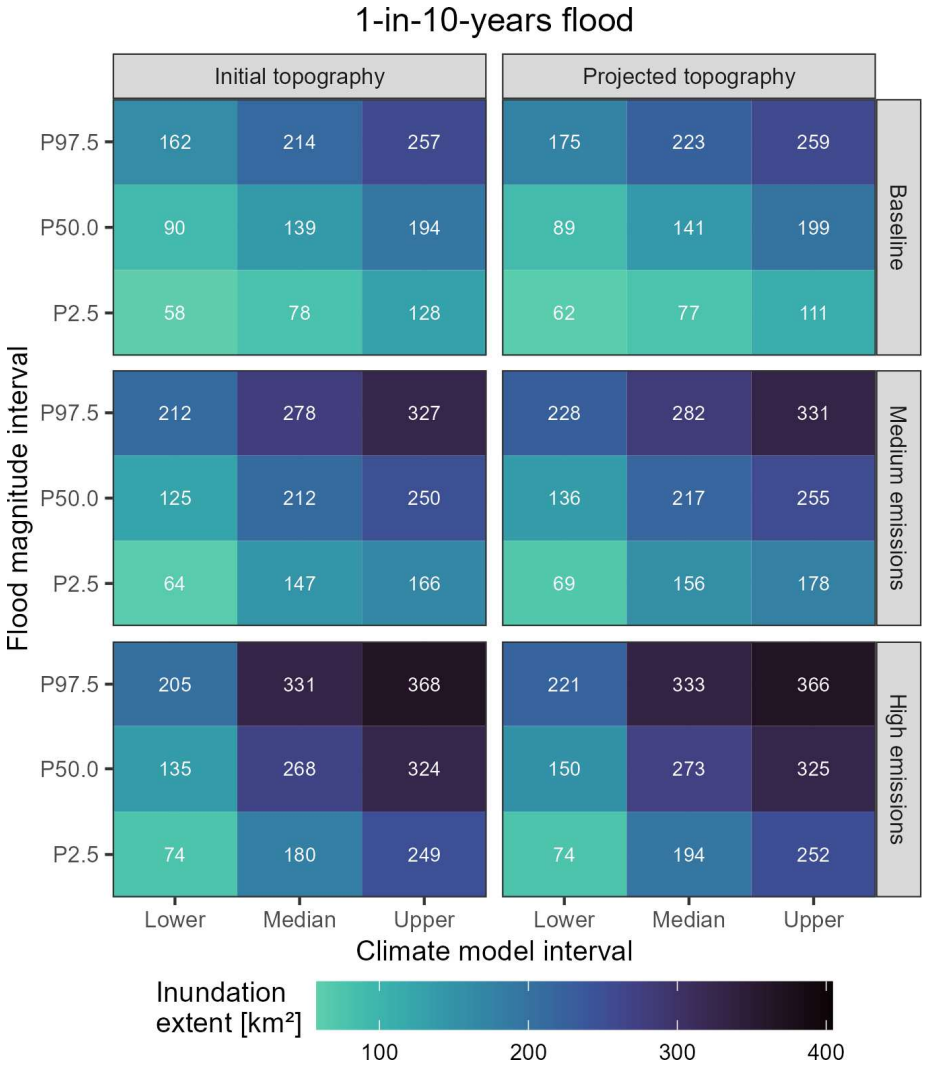


Figure 7.5: The ensemble variation of the inundation extents predicted for the 1-in-10-years flood event. The x-axis shows the prediction interval of the climate models (CM) (lower: CM with the lowest predicted flood magnitude; median: CM with the 6th highest predicted flood magnitude; upper: CM with the highest predicted flood magnitude). The y-axis is the percentile of the flood frequency analysis ensemble of the respective climate model.

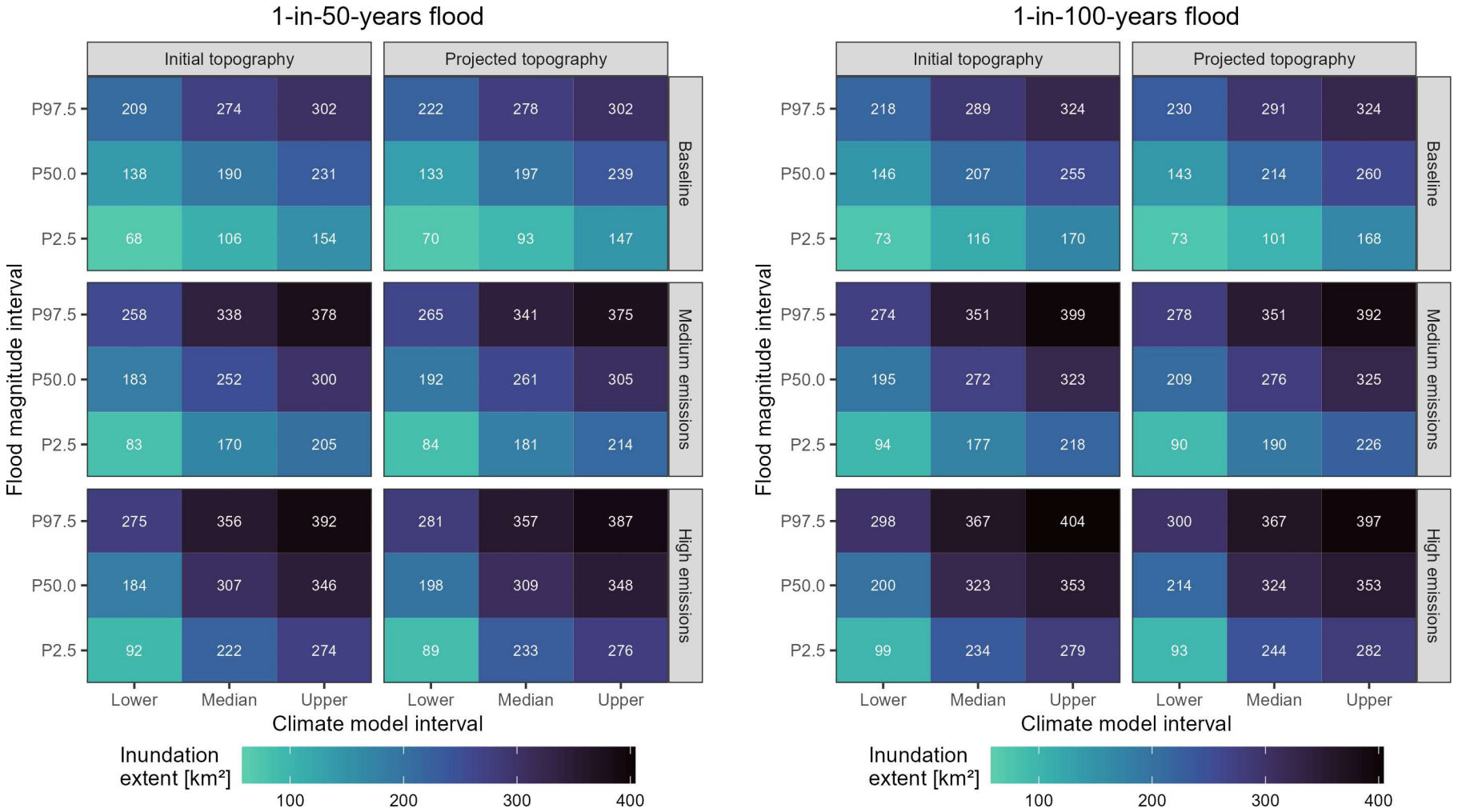


Figure 7.6: The ensemble variation of the inundation extents predicted for the 1-in-50-years flood event (left) and the 1-in-100-years flood event (right). The x-axis shows the prediction interval of the climate models (CM) (lower: CM with the lowest predicted flood magnitude; median: CM with the 6th highest predicted flood magnitude; upper: CM with the highest predicted flood magnitude). The y-axis is the percentile of the flood frequency analysis ensemble of the respective climate model.

Both, the FM and CM are important sources and none dominates the other. This is illustrated by the fact that the inundation extent of one category (CM or FM) is not larger than each inundation extent of a lower category and hence the three highest extents of each scenario and return period are not predicted for either, the 97.5th FM or CM_{upper}. For example, the extent of the median FM of CM_{upper} exceeds the 97.5th FM of CM_{lower}, and the extent of the 97.5th FM of CM_{lower} exceeds the 2.5th FM of CM_{upper} (Figures 7.12 and 7.13).

The dominance of the CM and FM for the uncertainty is also illustrated by the great mean range of the predicted inundation extents (Table 7.7). This range describes the mean prediction range of the inundation extents in Figures 7.12 and 7.13 for the different categories (i.e. climate scenario, climate model, flood magnitude, and topography). The mean prediction interval is greatest for the FM and increases with the return period from 136 km^2 (1-in-10-years) to 163 km^2 (1-in-100-years). The second greatest variation is predicted for the CM but this uncertainty is independent of the return period as indicated by the stable prediction range for the return periods (120 – 124 km^2). The uncertainty of the FM and CM exceeds the variation of the climate scenarios as indicated by the lower mean prediction range of the climate scenarios (85 – 93 km^2). This range decreases with the return periods. The topographic changes only cause a small variation in the predicted inundation extent which is indicated by the low mean prediction range of 5 – 7 km^2 (Table 7.7).

All ensemble members predict an increase in the inundation extent from the baseline scenario to the projected scenarios for each return period (Figure 7.14). This increase ranges between 10 – 103% for SSP245 and between 19 – 151% for SSP585. Higher increases are predicted for SSP585 than for SSP245 except for one member (CM_{lower} 97.5th FM) for the 1-in-10-years event.

The projected changes decrease with increasing flood magnitude interval, particularly for CM_{median} and CM_{upper} (Figure 7.14). The climate models have no clear trend of higher or lower changes with the interval. The highest changes are predicted for the 2.5th FM of CM_{median} and these changes are greater for the projected topographies than for the initial topographies. For this member, the projected topographies reduce the inundation extent in the baseline and increase the inundation extent in the projected scenarios which causes these greater increases (Figures 7.12 and 7.13). The initial topographies generally predict greater increases than the projected topographies for CM_{median} and CM_{upper} except for the 2.5th FM combinations) (Figure 7.14).

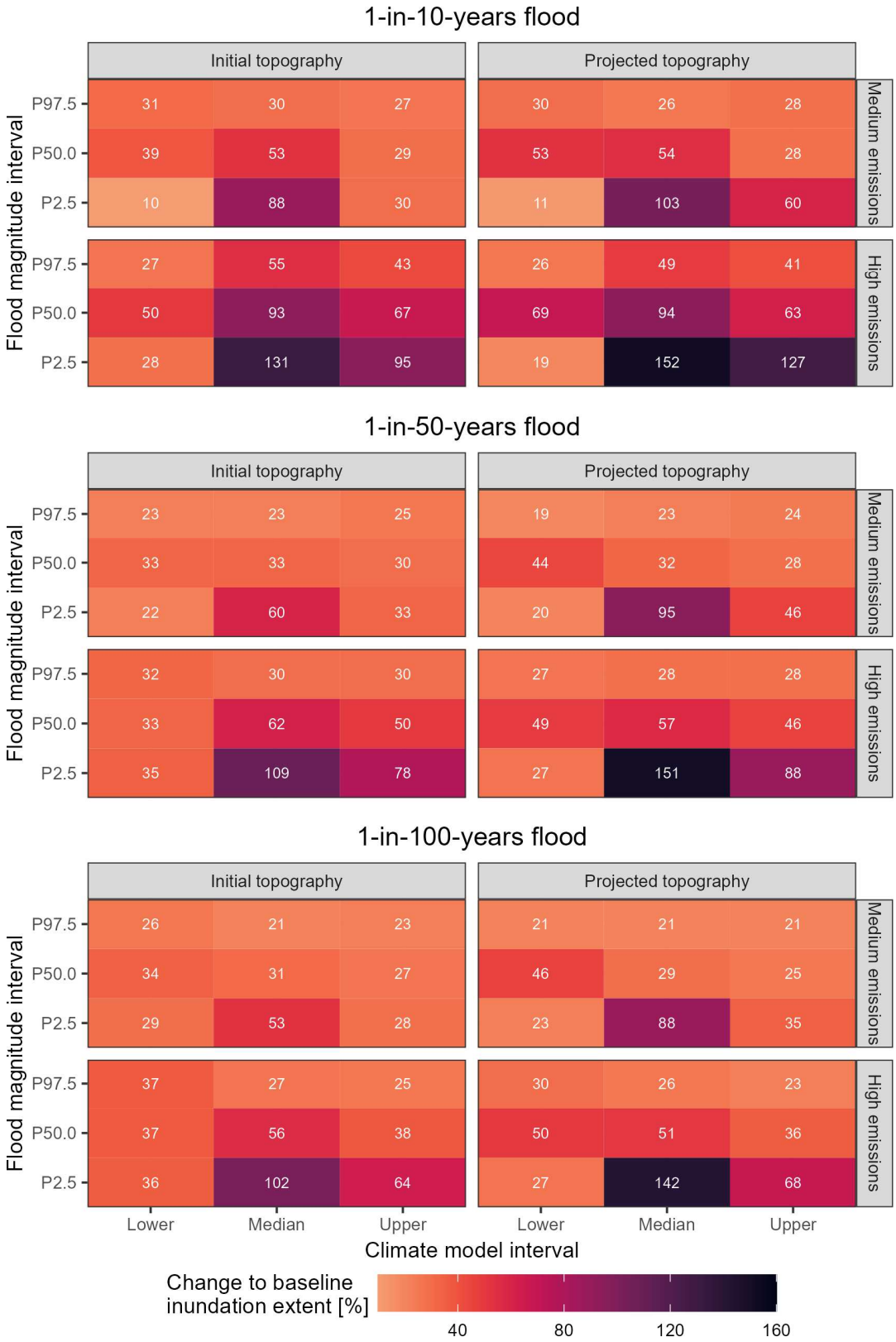


Figure 7.7: The projected increases in the inundation extent for the median- and high-emission scenarios for each ensemble member. The x-axis shows the prediction interval of the climate models CM (lower: CM with the lowest predicted flood magnitude; median: CM with the 6th highest predicted flood magnitude; upper: CM with the highest predicted flood magnitude). The y-axis is the percentile of the flood frequency analysis ensemble of the respective climate model.

7.3.4 The effect of the morphological evolution

The morphological evolution of the fan topography has low effects on the predicted inundation extents of the flood hazards (Table 7.7). However, the topography defines the pathways of the flood water and hence the morphological evolution potentially alters the locations of the inundation and the inundation depths. These aspects are analysed in this section by comparing the spatial flood hazard patterns of the ensemble predicted for the initial and projected topographies.

The large-scale patterns of the inundation extents remain similar for both topographies for all scenarios (Figures 7.15 and 7.16). For most of the inundated areas throughout the fan, the inundation is predicted for both topographies. Differences occur mostly on the local scale of a few thousand square metres. The differences in the inundation extents are most pronounced for CM_{lower}. which predicts inundations in the Island for the initial topography but not for the projected topographies. Furthermore, CM_{lower} predicts inundations in the north of Tikapur city only for the initial topography, and the inundations in the South-Western parts of the fan extend further in the initial topography for the 1-in-50-years and 1-in-100-years events.

The median and upper prediction interval CM_{median} and CM_{upper} predict a decrease in the inundation extent along the Eastern and Western edges of the flooded area (Figures 7.15 and 7.16). The extent decreases locally in the north and east of Tikapur city, particularly for the 1-in-10-years event. Furthermore, parts of the Eastern floodplain are inundated by the most 1-in-100-years event of the projected scenarios for CM_{upper} but only for the initial topography. Hence, the inundated areas spread further to the edges for the initial topography.

The connectivity of the Island increases for the projected topographies of CM_{median} and CM_{upper}. The increase of the inundation extent on the Island is most prominent for the 1-in-10-years events of the medium- and high-emission scenarios (Figure 7.15). These increases are predicted in the intersection of the main branches and the Island which indicates a better connectivity. Furthermore, the inundation extent increases in the Southern parts of the Island also at greater distance (3-4 km) to the branches. However, for the largest parts of the fan, inundation is predicted for both topographies and small-scale variations are predicted throughout the fan.

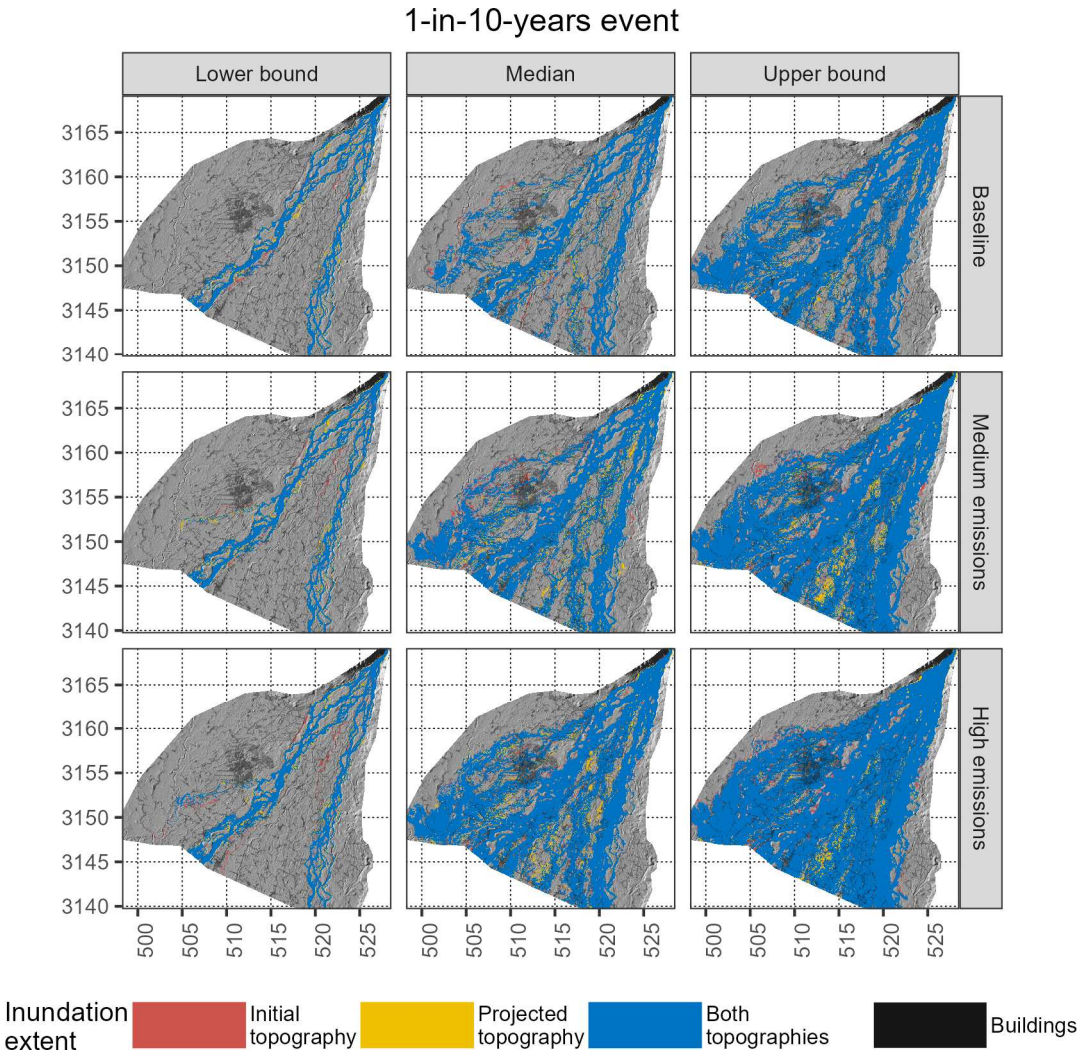


Figure 7.8: The inundation extents (inundation depth ≥ 0.1 m) predicted for the initial and projected topographies for the 1-in-10-years event. Further details are provided in the description of Fig. 14.

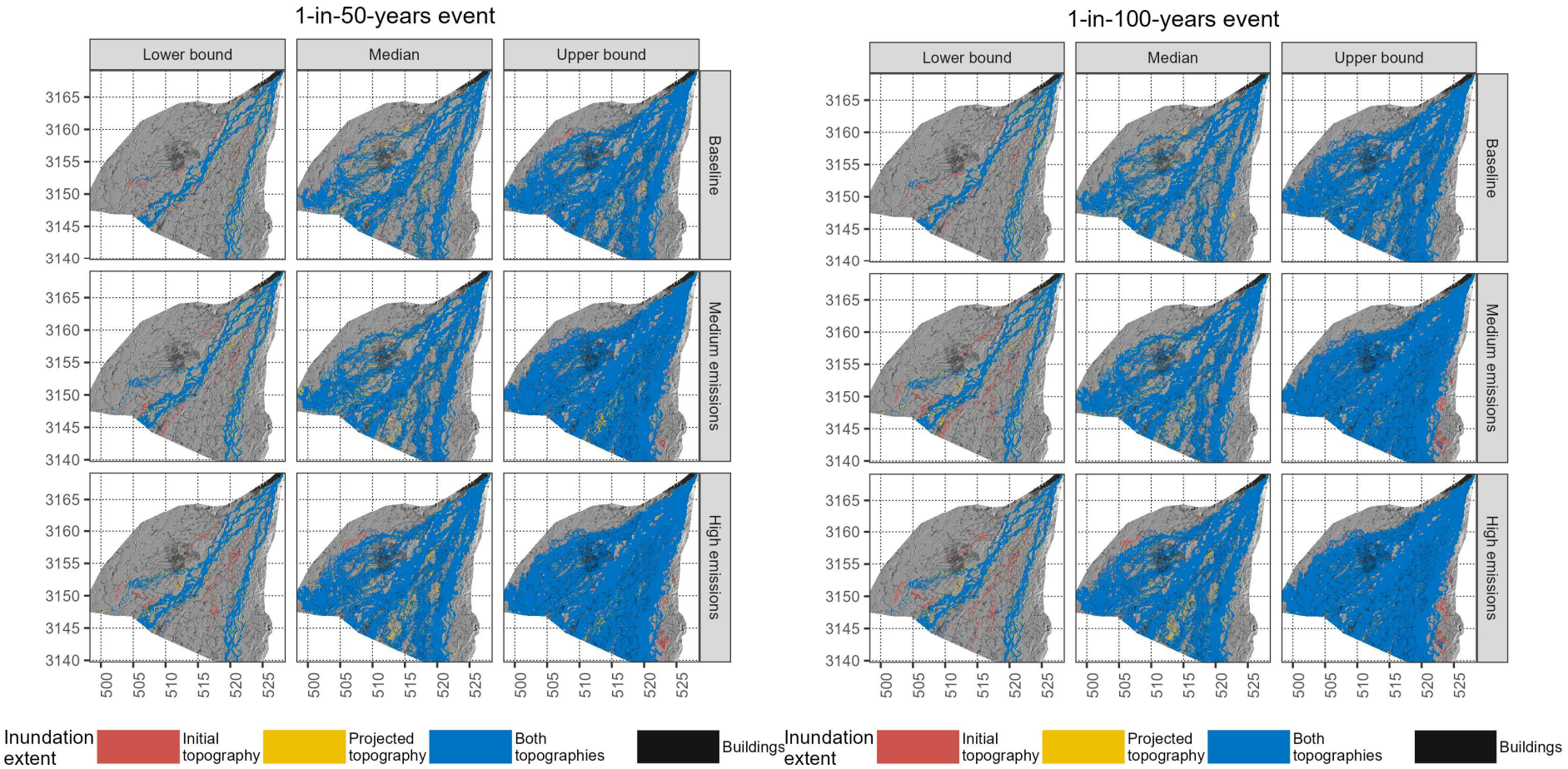


Figure 7.9: The inundation extents (inundation depth ≥ 0.1 m) predicted for the initial and projected topographies for the 1-in-50-years event (left) and for the 1-in-100-years event (right). The columns show the predictions for CM_{lower}, CM_{median}, and CM_{upper}. The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km.

The morphological evolution alters the inundation depth throughout the fan. These inundation depth differences for the initial and projected topographies are shown in Figures 7.17 and 7.18. The highest differences (≥ 1.5 m) are predicted in the channels whereas areas of increased and decreased depths are in proximity to each other. Generally, the areas with decreased inundation depth for the projected topographies dominate throughout the fan. However, the influence of the morphological evolution on the inundation depth is complex and varies between the scenarios, the prediction interval, and the return periods.

The influence of the topography on the inundation depth varies between the scenario which is illustrated by the inundation depths of the median predictions (CM_{median}) of the 1-in-100-years event (Figure 7.18). In the baseline scenario, the inundation depths are higher for the projected topography in the Island and most of the Western floodplain. Decreasing depths in the projected topographies are predicted along a floodplain channel in the north and east of Tikapur city, and a section in the depression in the south-west of the fan. These patterns change in the median-emission scenario SSP245 for which the inundation depths also decrease along floodplain channels in the South of Tikapur city. The change from increased to decreased inundation depths for the projected scenarios is also predicted along channels in the Island. For the high-emission scenario SSP585, the inundation depth differences change further and lower depths for the projected topographies are predicted for most areas throughout the fan. Furthermore, the magnitude of the differences increases in the high-emission scenario SSP585. The depth differences in the floodplain are mostly between ±0.1 m in the baseline and medium-emission scenarios, but in SSP585 decreasing depths between -0.1 m and -0.5 m are predicted in extended areas throughout the fan (median predictions).

The prediction interval also affects the impact of the morphological evolution on the inundation depths. This is illustrated by the simulations of the 1-in-100-years event of the medium-emission scenario SSP245 (Figure 7.18). The inundation depths increase for the projected topographies in the depression for the CM_{median} predictions while these decrease for the CM_{upper} predictions. The comparison of the lower prediction interval CM_{lower} is hampered by the low inundation extent.

The sensitivity of the topography also varies for the return periods. The inundation depth decreases for the projected topography in the depression for the 1-in-10-years event for the CM_{median} of the median-emission scenario SSP245 (Figure 7.17). This ensemble member predicts an increase in the inundation depth for the 1-in-50-years and 1-in-100-years events at this particular location (Figure 7.18).

The morphological evolution alters the predicted inundation depths throughout the fan. The inundation depths increase or decrease locally. Generally, it appears that lower inundation depths are predicted for the projected topographies for larger areas with increasing flood flow (i.e. with higher return periods, emissions, and the prediction interval). However, a consistent increase is predicted by all simulations of CM_{median} and CM_{upper} in a section of the Island 3-4 km east of the Western branch which stretches parallel to the channel in a north-south direction until the downstream boundary (Figures 7.17 and 7.18).

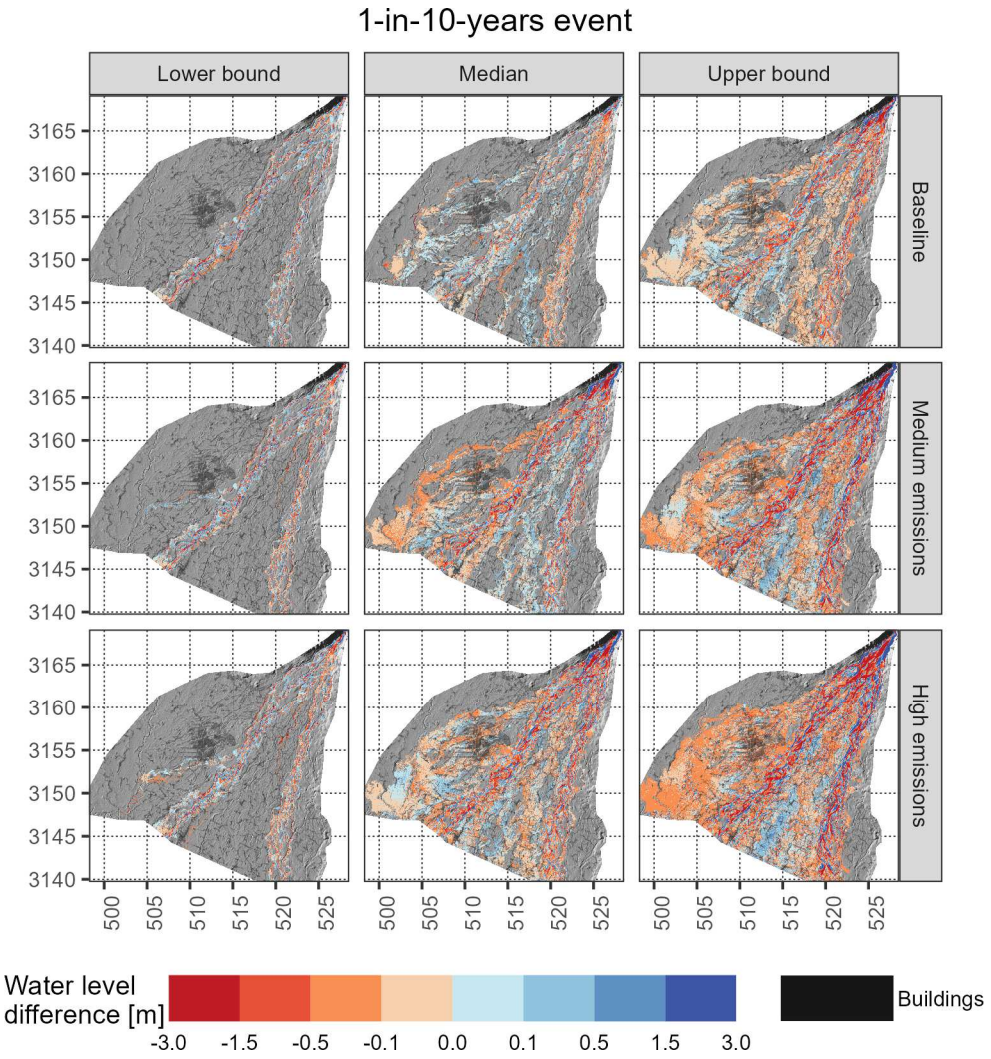


Figure 7.10: The differences in inundation depths between the predictions for the initial and projected topographies for the 1-in-10-years event. Positive values indicate that the depth is higher for the projected topographies. The depths are capped at ± 3 m and may exceed these values locally. Note that the scale is not linear. The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km.

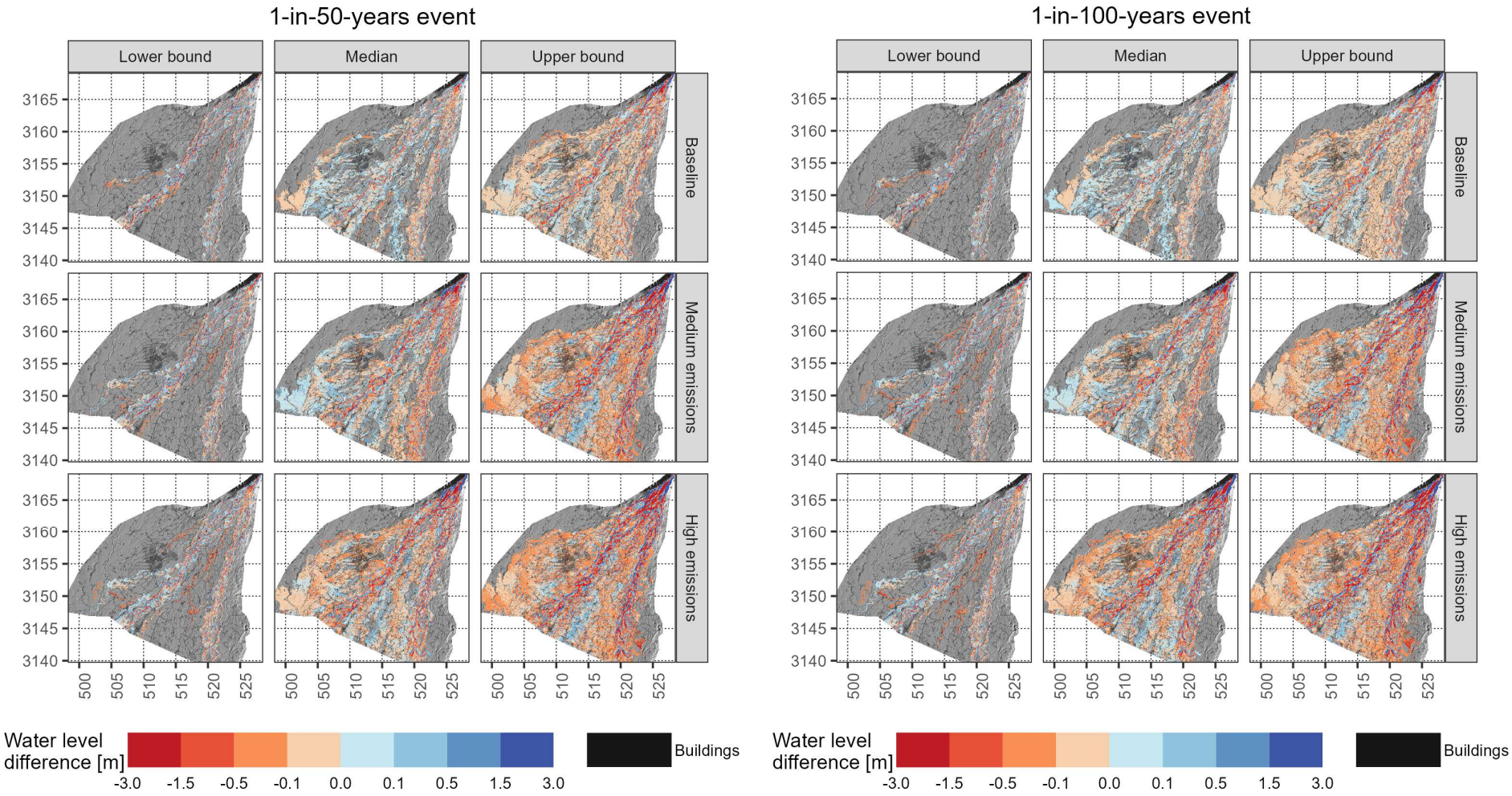


Figure 7.11: The differences in inundation depths between the predictions for the initial and projected topographies for the 1-in-50-years event (left) and the 1-in-100-years event (right). Positive values indicate that the depth is higher for the projected topographies. The depths are capped at ± 3 m and may exceed these values locally. Note that the scale is not linear. The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km.

The morphological evolution influences the inundation extent and the inundation depth. This influence varies with the flood discharge into the fan (Figure 7.19). For the lowest simulated flows (< 6,000 m^3/s), the predicted inundation extent is higher and the depth is lower for the projected topographies than for the initial topographies. For flows between 6,000 – 9,500 m^3/s, the influence of the morphological evolution is not systematic because for some ensemble members greater inundation extents are predicted for the initial topography while greater inundation extents are predicted for the projected topography for other members. Above this flow, 5 – 16 km^2 greater inundation extents and lower depths are predicted for the projected topographies. The difference in the predicted inundation extent decreases with the flow rate and similar extents are predicted for flows ≥ 15,000 m^3/s. For the highest simulated flows (≥ 30,000 m^3/s), larger inundation extents are predicted for the initial topography. However, the median inundation depth is 0.01 – 0.08 m lower for the projected topographies for all simulations with flows ≥ 10,000 m^3/s independent of whether lower or greater inundation extents are predicted for the projected topographies (Figure 7.19).

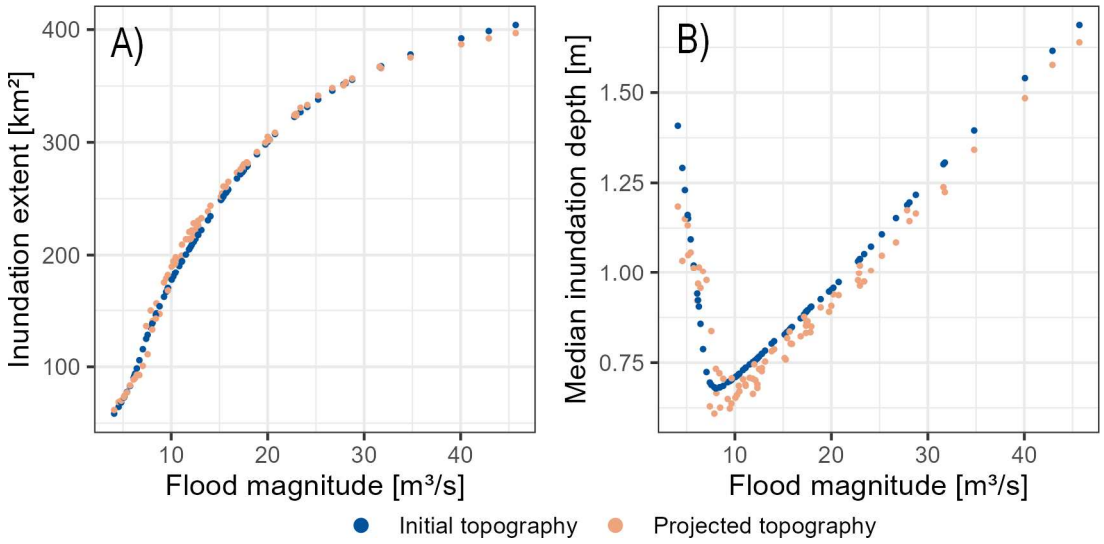


Figure 7.12: The influence of morphological evolution on the inundation characteristics. Figure A shows the relationship between the flood magnitude and the inundation extent and Figure B shows the relationship between flood magnitude and the median inundation depth. These relationships include the predictions of all ensemble members and return periods presented in Figures 7.11 and 7.12.

7.4 Discussion of the predicted flood hazards in the Karnali fan

The previous section presented the predictions of the hydrodynamic simulations of the 1-in-10-years, 1-in-50-years, and 1-in-100-years flood events for the observed record and the climate scenarios. This section discusses these predictions with a focus on their reliability by comparing them with other studies (Section 7.4.1), the changes for the projected climates (Section 7.4.2), the sources of uncertainty (Section 7.4.3), and the sensitivity of the flood hazards to changes in the topography (Section 7.4.4).

7.4.1 Comparison with other studies

The performance assessment of the simulation of the flood hazard characteristics is hampered by the lack of observations of the inundation extent and depths of past flood events. Instead, the results are compared with the ones of other studies to investigate whether the predicted characteristics are realistic. For this, predictions are chosen that match the return period or flood discharge of the simulations of Meteor Consortium (2019) and Dingle *et al.* (2020a) (Figure 7.20).

The simulations in this study predict the inundation of the Island which is consistent with the other studies (Figure 7.20). The flood water enters this Island at the northern edge and traverses it until the downstream border of the modelling domain. This is consistent with the simulations of the other studies and the spatial patterns of the inundations agree between the studies despite local deviations. Most of the flood water drains in the south-eastern direction in the lower half of the Island which is consistent with the predictions of Meteor Consortium (2019) (Figure 7.20 A and B). Dingle *et al.* (2020a) predict the inundation of larger areas in the Island than this study which is likely to be attributed to the spatial representation in the model (structured mesh with increasing size with distance to the branches) (Figure 7.20 C and D). Flood water enters the Island at several locations from the Eastern and Western branches which is also predicted by Meteor Consortium (2019). The inundation maps of Dingle *et al.* (2020a) predict connections of channel and floodplain throughout the length of the Island for both branches. The high flood risk predicted in the simulations is complemented by Duwal *et al.* (2023) who predict moderate to very high flood susceptibility throughout the Island.

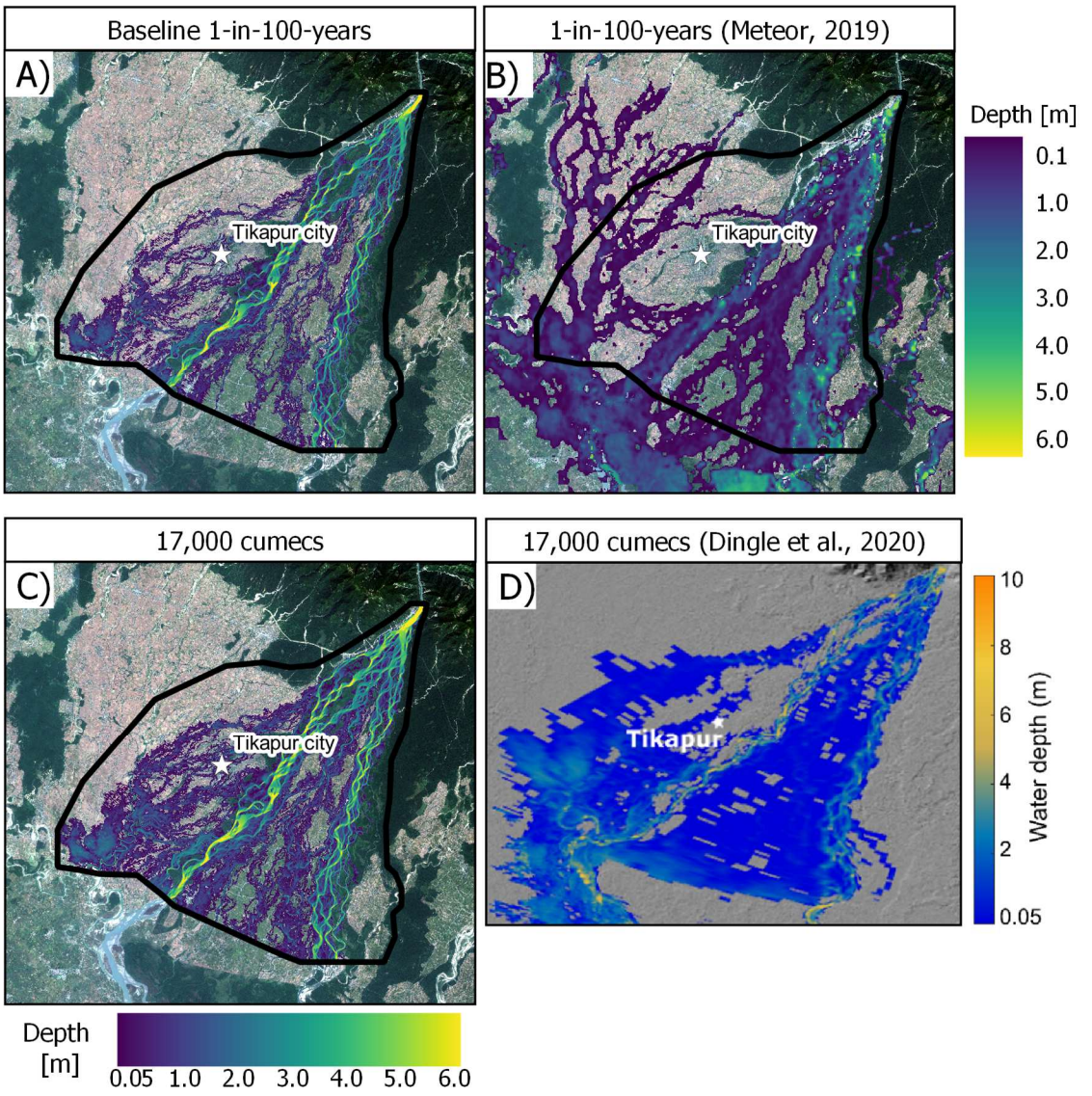


Figure 7.20: Comparison of the predicted flood hazard maps with other studies. Figure A shows the predicted flood inundation of the 1-in-100-years event of the baseline (initial topography). Figure B shows the flood inundation of the 1-in-100-years event of Meteor Consortium (2019). Figures A and B share the same legend. Figures C and D show the predicted flood inundation at 17,000 m^3/s of this study (Figure C) and Dingle et al. (2020a, Figure S3 C).

The Western floodplain is flood-prone but the location where the flood water enters the floodplain varies in the studies. The Western floodplain is inundated during the 1-in-100-years event in all studies. In this study, the overbank flow drains the Western floodplain in the south-western direction and accumulates in a topographic depression at the south-western border of the modelling domain. These spatial patterns are similar to the ones predicted by Meteor Consortium (2019) and Dingle *et al.* (2020a). However, the locations where the Western branch starts flooding differ between the studies:

- On the fan head (~11 km downstream of the inlet) from where it drains in the south-western direction until the depression. This location is predicted in this study and the one of Dingle *et al.* (2020a).
- North of Tikapur forest (~15 km downstream of the inlet). This is the entry point of flood water into the Western floodplain for the predictions of Meteor Consortium (2019) but is not predicted by the other studies.
- In a section in the south-west of Tikapur city (~20 km downstream of the inlet). This location is predicted in this study and by Dingle *et al.* (2020a). However, the flood water drains in the south-western direction in this study while it drains parallel to the Western branch in the predictions of Dingle *et al.* (2020a).
- In the lower fan (~ 25 km downstream of the inlet). This location is predicted in each study. However, more locations of overbank flow into the Western floodplain in the lower fan are predicted in this study and by Dingle *et al.* (2020a) than by Meteor Consortium (2019).

The predictions of this study are uncertain in forests and downstream of these forested areas which concerns the Western floodplain in the south and south-west of Tikapur city. The drainage in the south-western direction is only predicted at this location in this study. Dingle *et al.* (2020a) predict some inundation which may be the reoccupation of a previously abandoned channel rather than floodplain flows. The different predictions of all studies arise from the utilisation of different topographic datasets and procedures to remove the forest cover. The predicted pattern in this study may result from an underestimation of the ground elevation during the statistical forest removal procedure. The forested areas along the Eastern border of the modelling domain are not affected by this because these are located high above the channel and are neither flooded in this study nor in any other of the evaluated studies.

The Western floodplain and areas in the south-eastern corner of the modelling domain are prone to floodings of smaller rivers originating in the southern flanks of the first Siwalik mountain range (Figure 7.20 B; Duwal *et al.*, 2023). These rivers have not been considered in this study and hence the inundations may be larger and deeper if these contributing rivers and the Karnali River flood simultaneously. Particularly in the depression, the actual inundations would be deeper than the predicted ones because three Siwalik rivers confluence in this area (Meteor Consortium, 2019).

7.4.2 The climate change impact on the flood hazard

The hydrodynamic simulations provide robust evidence that flood hazards intensify with emissions for all return periods. The inundation extent of the 1-in-100-years event is projected to increase by 29% for the medium-emission scenario SSP245 and by 51% for the high-emission scenario SSP585 at the end of the century (2060 – 2099) (Table 7.3). The inundation extent of the more frequent events are projected to increase stronger, and this increase is 55% (SSP245) and 94% (SSP585) for the 1-in-10-years event (median predictions).

The impact of increased climate forcing on flood hazards is illustrated by the projected increase in flood frequencies. The inundation extent of the 1-in-10-years event projected for SSP245 (217 km^2) is similar to the one predicted for the baseline 1-in-100-years event (217 km^2). The inundation extent of the SSP585 1-in-10-years event (273 km^2) equals the one predicted for the SSP245 1-in-100-years event (276 km^2) (Table 7.3 and Figure 7.5). This increase in the projected frequencies emphasises the intensification of the flood hazards with the emissions. The scale of the projected increases of the inundation extent agrees with the projected increases in the flood frequencies where the baseline 1-in-100-years event is projected to occur every 11 years for SSP245 and every 3 years for SSP585 (Table 5.11). The scale of the projected increases in flood frequencies agrees with regional projections which project return periods of 5-25 years (SSP245) and 2-25 years (SSP585) for the Ganga Plain (Hirabayashi *et al.*, 2021) (median predictions).

The confidence is high that the flood hazards will intensify with the emissions in the Karnali fan because increases in the inundation extent are projected for the full prediction range of the ensemble (Table 7.3). These projected increases are lower for the lower and upper boundaries (CM_{lower} and CM_{upper}) of the prediction interval than for the median predictions (CM_{median}). The causes for the lower projected increases are different CM_{lower} and CM_{upper}. For CM_{lower}, the cause is related to the lower projected increases in the flood magnitudes (Figure 5.37 and Table 5.10). This cannot be the case for CM_{upper} because the projected increases in the flood magnitude are greater for CM_{upper} than for CM_{median} (Tables 5.9 and 5.10). However, this greater projected increase in the flood magnitude does not translate into greater inundation extents. This lower projected increase in the inundation extent is caused by the topographic characteristics of the fan.

The topography divides the Karnali fan into areas that are very prone to floods and areas that are not inundated even during the highest simulated flows (≥ 40,000 m^3/s). The

longitudinal gradient is larger than the lateral gradient in the Island and the Western floodplain, and hence overbank flow drains on the floodplain to the downstream study area boundary rather than draining back into the Karnali River (Figure 7.6). As a consequence, large areas are inundated at comparatively low flows (< 15,000 m^3/s). The areas along the eastern model boundary are superelevated above the channel, and the areas along the north-western model boundary are protected by the south-western elevation gradient and these areas remain dry even for the highest simulated events. This polarisation of high-risk and no-risk areas causes a deceleration in the inundation extent with increasing flood magnitude ≥ 15,000 m^3/s (Figure 7.19) which explains why the projected relative increase in the inundation extents decreases with the return periods and prediction interval. However, the inundations get deeper as the increase in the flood magnitude is higher than the increase in the inundation extent (Figures 7.8 and 7.9). This is illustrated by the increase in the median inundation depths of the upper prediction interval which increases from 1.09 m in the baseline scenario to 1.60 m (SSP245) and 1.66 m (SSP585) for the 1-in-100-years event (Table 7.4). Hence, the inundation extent is not a sufficient metric to capture the full extent of the intensification of the projected flood hazards.

The projected intensification of the flood hazard characteristics (inundation extent and depth) affects the most densely populated regions of the fan. The 1-in-100-years event of the baseline scenario affects 8% of the populated area which increases to 20% (SSP245) and 32% (SSP585) at the end of the century (Figure 7.10). The densely populated municipalities Geruwa (Northern Island), Rajapur (Southern Island), and Tikapur (Western floodplain) are the most affected regions. In these municipalities, the affected populated area increases from 22 – 27% in the baseline scenario to 37 – 49% (SSP245) and 52 – 69% (SSP585). Furthermore, the topographic depression along the south-western border of the study area is sensitive to increased flooding. Aside the projected increase in the affected population, this population is also exposed to more dangerous flows with higher damage potential (Table 7.5) (median predictions).

The combination of the increased inundation extent and depth decreases safe non-flooded areas and access to these areas. Such safe zones are particularly important because many areas may be inundated from multiple directions. The Geruwa and Rajapur municipalities on the Island are surrounded by both Karnali branches in all directions. The population of the Western floodplain is exposed to floodings from multiple directions. Several Siwalik rivers traverse the floodplain West or within the study area and drain into the Mohana River which flows along the India-Nepal border in the South-Eastern direction and drains into the Karnali 2-3 km South of the study area. Hence the

municipalities of the Western floodplain are surrounded by water from the Western branch to the East, by the Siwalik rivers to the West, and by the Mohana River to the South (Meteor Consortium, 2019; Duwal *et al.*, 2023). Areas that are located ≥ 4-5 km from the Siwalik foothills are furthermore surrounded by flood flows of the Karnali River to the North (Figures 7.6, 7.7, and 7.20). If the centre of the flood-triggering rainfall is located over the Siwalik range, as for the 2014 flood event, the Karnali River and Siwalik rivers experience simultaneous floodings (MacClune *et al.*, 2015; Shrestha *et al.*, 2015a). In such a situation, large areas of the fan are surrounded to all sides by water which hampers the evacuation of the population from the risk areas, and the access of rescue teams to the risk areas. The projected increase in the inundation extent and the water levels intensifies this situation (Figure 7.8).

It is further worth noting that the population in the Terai is projected to increase due to population growth and migration which has not been considered in this study (UNFPA Nepal, 2017). Hence, the affected population will increase not only due to the intensification of the flood hazards but also due to the population dynamics. The projected increase in inundation extent and depth potentially increases agricultural damages which is the main source of income for 80% of the population (Dixit *et al.*, 2007; ADPC, 2010; Perera *et al.*, 2015). Furthermore, the projected changes potentially alter the water quality and reduce access to clean water (Hannah *et al.*, 2020; Kosow *et al.*, 2022).

7.4.3 The uncertainty in the predictions

The areas at risk of flooding vary largely for the ensemble predictions. For example, the inundation extent of the1-in-100-years event of the baseline varies between 13 – 59% (Figure 7.5). The inundations of CM_{lower} are constrained to the channels and a small area in the south of Tikapur city, while most of the Island and the Western floodplain (except for the Northern areas) are inundated for CM_{upper} (Figure 7.6). This large prediction range is caused by the uncertainty propagation through the modelling cascade. The sources of uncertainty arise from the hydrological model, the Flood Frequency Analysis (FFA), the variation in the climate ensemble, and the morphological evolution. Furthermore, the morphodynamic model and the hydrodynamic model add uncertainty to the predictions (e.g. parameter uncertainty). The uncertainty of these models has not been quantified to not increase the ensemble size further and is not assessed in this study. The sources of the assessed uncertainty are the hydrological uncertainty (from the hydrological model and FFA) and the climatic uncertainty from the CMIP6 ensemble.

The hydrological uncertainty (hydrological model and FFA) has the highest prediction interval with a mean range of the inundation extent of 136 km^2 for the 1-in-10-years event (Table. 7.7). This uncertainty arises from the hydrological modelling and is then amplified by the FFA. The significant contribution of the hydrological modelling uncertainty is illustrated by the prediction range of the baseline AMAX P_{90} which is 5,800 m^3/s and hence only slightly lower than the prediction range of the 1-in-10-years flood magnitude of 6,800 m^3/s (Table 5.15). However, this difference increases with the return period and hence the contribution of the FFA to the hydrological uncertainty increases with the return period to 163 km^2 (Table 7.7).

The climate uncertainty describes the fraction of the uncertainty arising from the variation in the climate input (temperature, precipitation) of the 12 CMIP6-ensemble members. This uncertainty is the second largest source of uncertainty with a mean variation of 120 – 124 km^2 (Table 7.7). While it is lower than the flood magnitude uncertainty it remains an important source of uncertainty. This uncertainty increases with the emission scenarios. This is illustrated by the increasing range in the inundation extent of the 1-in-100-years event from 109 km^2 in the baseline scenario to 128 km^2 for SSP245 and 153 km^2 for SSP585 (median flood magnitudes and initial topography) (Figure 7.13). This increasing uncertainty range is caused by an increasing discrepancy in the projections of the climate members (Tables 5.12 – 5.14).

The majority of climate models project an increase in the 1-in-100-years flood magnitude of 37 – 69% for SSP245 (8/12 members) and 68 – 143% for SSP585 (9/12 members) while the lowest members predict a lower 12% increase (SSP245) and a decrease of 12% (SSP585) (Tables 5.13 and 5.14). The climate models predicting lower changes are the ones which capture the flood seasonality of the monsoon system poorly and are less probable realisations (Chapter 5). Note that the lower climate boundary still predicts an increase in the inundation extent for SSP585 because the lower climate boundary is represented by different climate models for the baseline and projected scenarios (Table 7.3).

The inundations predicted for the simulated records underpredict the ones for the observed records but the difference decreases with the prediction interval. This is indicated by the comparison of the simulation with the flood magnitudes of the baseline and observed flood magnitudes (Figure 7.11 and Table 7.6). This underestimation arises from deficits in the process description of the hydrological model which cannot reproduce the full extent of the fast rainfall-runoff conversion, the parameter transfer from the observed climate data to the CMIP6 data, and the bias of the CMIP6 models.

The differences between the baseline and observed simulation are largest for the lower prediction interval (CM_{lower}). The poor performance of this interval results from the combination of the climate models and the hydrological predictions. The hydrological ensemble underpredicts the flood peaks and this underprediction is largest for CM_{lower}. Furthermore, the simulated peaks are lower for the CMIP6 climate forcing than for the observed forcing and this combination results in inundation extents that are restricted to the channel even for the 1-in-100-years event. The underestimations decrease with the prediction interval and CM_{upper} of the baseline ensemble reproduces the patterns of the observed simulations well, both in terms of inundation extent and depth (Figure 7.11 and Table 7.6). Hence, the lower boundary is unbehavioural and vastly underestimates the flood hazards but the CM_{median} and CM_{upper} are behavioural representations and reproduce the spatial flood hazard characteristics as also indicated by the comparison with other studies (Figure 7.20).

The variation in the inundation extents of the climate scenarios (85 km^2 for the 1-in-100-years event) is lower than the variation introduced by the flood magnitudes and climate ensemble. Nonetheless, the confidence is high that the flood hazards intensify with emissions. All ensemble combinations project an increase for the projected climate and this increase is higher for the high-emission scenario except for one member for the 1-in-10-years flood (CM_{lower} 97.5th FM) (Figure 7.14). The high agreement of all models provides a strong signal that floods intensify with emissions. The highest increases (≥ 90%) are projected for CM_{median} $FM_{2.5}$ while the increases are lower for the higher FM and CM intervals. The reason for this is the topography where the inundation extent increases more slowly with increasing magnitude (see the previous section).

The main contributions to the uncertainty are the hydrological model and the climate projections. This is consistent with the uncertainty analysis for flood projections of Giuntoli *et al.* (2018) who predict that both sources are important contributors to the overall uncertainty and that the uncertainty of the hydrological model can exceed the one of the climate models in catchments for which melt processes are important components of the hydrological cycle. It also agrees with the uncertainty analysis of global high flow (Q_{95}) projections of Giuntoli *et al.* (2015) who predict that the hydrological model contributes most to the uncertainty in the Central Himalayas. However, it is worth noting that both studies investigate the structural uncertainty by applying an ensemble of hydrological models while this research applies an ensemble of parameter sets of a single hydrological model. The increasing hydrological uncertainty with the return period is reasonable as the most extreme flows are most difficult to predict by the hydrological model, these events are most affected by the internal climate variability, the FFA is

extrapolated beyond the 40-year record, and are more sensitive to the slope of the flood frequency curves (Kjeldsen *et al.*, 2014; Huang *et al.*, 2017; Giuntoli *et al.*, 2018, 2021; Li *et al.*, 2018).

This research followed a “one model one vote“ approach which regards each ensemble member as an equally likely predictor of the potential flood hazard. However, the poor performance of the ensemble members (both climate and hydrological ensemble) indicates that the predictions of some members are inferior to others. The ranking of the ensemble members based on their agreement with the historical observations may improve the projections (Padrón *et al.*, 2019; Giuntoli *et al.*, 2021).

The climate and flood magnitude uncertainty is caused by the variation of the flood discharge into the Karnali fan (the morphological impact will be assessed later). However, the topographic representation also introduces uncertainty. The DTM was generated without ground-truth data and contains noise which affects the predicted flood hazard characteristics (Chapter 6). Certain landscape features are particularly affected by noise and these features are:

- The channel representation – main channels: The bed elevations of the Karnali branches are unknown and are interpolated from the elevation along the edges of the inundated channel sections. This uncertainty of the bed elevation affects the channel capacity and hence the bankfull flow rate might be under- or overestimated.
- The channel representation – floodplain: The floodplains are intersected by small channels (< 10 m width) and these are difficult to represent at the 25 m modelling resolution (see Figure 7.1 for the location of some of these rivers). In some river sections, the banks are covered by trees and hence there is no information about the bank and bed elevation. In such areas, the channel elevation might be overestimated and the channels superelevated to the adjacent floodplain. This also causes a disturbance in the longitudinal gradient. As a consequence, these small channels cannot drain the floodplains as efficiently and might cause false local floodings.
- Forested areas: The forest elevation is unknown and might be over- or underestimated. This is particularly relevant for the forested riverbank of the Western branch (see Tikapur forest in Figure 7.2). The models predict overbank flow in this location which is not predicted in other studies (Meteor Consortium, 2019; Dingle *et al.*, 2020a). Hence, this elevation might be underestimated leading to false inundation prediction in this area and the areas along the downstream flow pathway.

- Buildings: The noise caused by buildings is particularly relevant for Tikapur city. In this area, not all buildings may have been removed and the ground elevation might be overestimated in some sections. The inundation patterns in the city should not be interpreted on the local scale of hundreds of metres because artefacts of the buildings remain in the topographic data.

The similar inundation patterns predicted in this research and other studies indicate that the DEM depicts the large-scale topography well (Meteor Consortium, 2019; Dingle *et al.*, 2020a). However, the predictions should not be used on the local scale of hundreds of metres. Such high-resolution flood hazard mapping requires a more detailed topographic representation based on ground surveys and a higher modelling resolution.

7.4.4 The sensitivity towards morphological changes

The morphological evolution of the fan topography alters the flood hazard characteristics but this response is complex and varies spatially and between the ensemble members. The ensemble variation in the inundation extent of the fan is 5 – 7 km^2 and is very low compared to the variation introduced by the climate flood magnitude ensembles (Table 7.7). However, the topographic changes alter the flow pathways and hence change the spatial patterns of inundation extent and depth. This predicted response to the morphodynamic evolution varies spatially and between the ensemble members.

Both sets of simulations predict inundations at similar locations for most of the fan and differences are predicted on the local scale of a few thousand square metres. The morphological evolution increases the connectivity of the Island and the lower Western floodplain (20 – 30 km distance to the inlet) to the Karnali branches which results in locally increased inundation extents and depths (Figures 7.17 and 7.18). This higher connectivity potentially results from the erosion of the irrigation gates which control the upstream influx to the Island, and the erosion of embankments along the main channels particularly in the lower fan which are represented in the initial topography. Contrarily, the connectivity to the channels decreases in sections along the North-Western and South-Eastern study area boundaries which causes the decrease of the inundation extent and depth. Decreasing inundation depths are predicted for larger areas with increasing flood discharge into the fan (i.e. return period, prediction interval, emission scenario).

The comparison of the simulations with the initial and projected topographies does not indicate a clear trend that the morphological evolution intensifies or weakens the flood hazard. The total inundation extent increases for the projected topographies for flood flows between 9,500 – 15,000 m^3/s but above this magnitude, the extents are similar for both topographies. For the highest simulated flood flows (≥ 30,000 m^3/s), the inundation extent decreases for the projected topographies (Figure 7.19). The simulations do not suggest that the morphological evolution causes an intensification of the flood hazard through channel aggradation as observed for the Koshi River system (Sinha, 2008; Dixit, 2009; Sinha *et al.*, 2014). However, the simulations do not indicate a weakening of the flood hazards through channel degradation which would increase the channel capacity.

The response of the flood hazards to morphological changes is complex and varies for the topographies predicted by the ensemble members. This complexity is illustrated by the variation of the predicted inundation depths in the depression (see Figure 7.2 for the location of the depression). The inundation depths decrease for the projected topographies (compared to the simulations with the initial topography) of the baseline and the high-emission scenario SSP585 but increase for the topography of the medium-emission scenario SSP245 for the 1-in-50-years simulation of CM_{median}. Furthermore, the predictions also vary between the simulated flood events in this area. The CM_{median} predicts a decrease in the depths for the 1-in-10-years event for SSP245. However, for the 1-in-50-years and the 1-in-100-years events, a relative increase compared to the initial topography is predicted (Figures 7.17 and 7.18). The CM_{median} is represented by the ACCESS-CM2 member for the 1-in-10-years event, and by the INM-CM5-0 member for the 1-in-50-years and 1-in-100-years events. The most likely explanation for this shift from decreased to increased inundation depths is that the morphological evolution predicted for SSP245 for the ACCESS-CM2 member causes a decrease of the overland flow in the North-East of Tikapur city and hence less water is delivered to the depression, while this decrease is not predicted for the morphological evolution of the INM-CM5-0 member.

The projected topographies have an increased channel capacity and more efficient drainage of the small channels in the floodplain. The increased channel capacity is indicated by the decrease in the inundation extent and an increase in the median inundation depth for flood flows between 6,000 – 9,500 m^3/s (Figure 7.19). The bankfull flow rate increases with the channel capacity and hence the inundation extent decreases. Consequently, the median inundation depth increases because the water is kept within the deeper channels. The more efficient drainage is indicated by the predictions of higher flood magnitudes ≥ 18,000 m^3/s. In this range, similar inundation

extents are predicted for the initial and projected topography, but the water levels are lower for the projected topographies. This is most pronounced for the most extreme flood flows ≥ 30,000 m^3/s for which the inundation extent and the depths are lower for the projected topographies. However, the depths should increase if the same amount of water inundates a smaller area. The lower depths for similar or lower inundation extent are caused by the more efficient drainage of flood water and hence the water is routed faster to the downstream boundary which reduces both, the inundation extent and depth.

The improved drainage raises the question of how much of the variation is caused by the morphological evolution and how much by the noise in the initial DEM. The small floodplain channels contain noise in the initial channel (see previous section). Such noise introduces obstacles in the longitudinal direction and increases the inundation depth in the adjacent floodplain because less water can drain through the channel. The morphological simulations erode such obstacles which increases the drainage efficiency for the projected topographies. The inundation depths of the projected topographies decrease compared to the initial topography along a floodplain channel in the Island flowing from the upstream beginning of the Island to the downstream boundary in the South-East, and along smaller channels North, Central, and South of Tikapur city (Figures 7.1 and 7.18) which indicates that water can drain more efficiently in these channels in the projected topographies. However, this decrease is more evident for the upper prediction interval and the projected climate scenarios. The reason for this is that these small floodplain channels are inundated more frequently for ensemble members with high flood flows (i.e. higher climate interval, hydrological interval and emission scenario). Thus, more noise is removed by the morphodynamic modelling of these members which explains why particularly the upper prediction interval drains more efficiently.

The noise in the main channels of the Karnali has been removed in the DEM generation process and these channels drain efficiently even in the initial topography. However, the riverbeds have been interpolated and are flat in the lateral direction. Changes in the channel capacity can be, at least to some extent, caused by erosional and depositional processes to generate a more realistic bed topography.

7.5 Conclusions

A hydrodynamic model was applied with the flood magnitudes (Chapter 5) and the topographies (Chapter 6) predicted for an ensemble of projected climates to predict the spatial characteristics of the flood hazards at the end of the century (O5). The model was also applied with the initial topography to evaluate the sensitivity to the morphological change. The predictions provide realistic flood pathways in the Island and the Western floodplain which agrees with other studies (Meteor Consortium, 2019; Dingle *et al.*, 2020a). This agreement indicates that the hydrodynamic model is behavioural and can be used to predict potential flood hazards.

The flood hazards are projected to intensify with emissions with high confidence. The Karnali fan will experience more frequent flooding in a larger area. The inundation extent of the baseline (1975 – 2014) 1-in-100-years event (39%) is exceeded by the 1-in-10-years event of the medium-emission scenario SSP245 (40%) at the end of the century (2060 – 2099). The extent projected for the 1-in-10-years event of the high-emission scenario SSP585 (50%) is similar to the one of the 1-in-100-years event of the medium-emission scenario SSP245 which indicates the high impact of increased climate forcing on the flood hazards (median predictions). An increase in the inundation extent and depth with the emissions is predicted for all ensemble members and hence the confidence is high that the flood hazards in the Karnali River intensify with climate change.

The areas most prone to the projected intensification of flood hazards are densely populated. The most flood-prone areas are located in the Island (Geruwa and Rajapur municipalities), and the Western floodplain (Tikapur municipality). For the current climatic conditions, 19% of the populated area is inundated by the 1-in-100-years event. This is projected to increase to 33% for the medium-emission scenario SSP245 and to 46% for the high-emission scenario SSP585 (median predictions).

The projected increase in the inundation extent and the inundation depths reduce the availability and access to save not inundated areas. The evacuation and emergency response is complicated by the branching of the Karnali River and the drainage of the Siwalik by smaller rivers traversing the Terai. The flood-prone areas in the Island and Western floodplain are surrounded by the Karnali River and Siwalik rivers from all sides. The Karnali River and the Siwalik rivers potentially flood simultaneously if the centre of the rainfield is located over the Siwaliks such as during the 2014 flood event (MacClune

et al., 2015). Therefore, flood adaptation and mitigation plans need to assess the risk of the Karnali and the Siwalik rivers jointly.

The main sources of uncertainty are the hydrological uncertainty arising from the hydrological ensemble (Chapter 4) and the Flood Frequency Analysis (Chapter 5), and the climatic uncertainty arising from the variation in the climate projections of the CMIP6 ensemble. The increasing climate uncertainty with emissions is problematic because it means that the flood hazards will intensify but are more difficult to predict and prepare for.

The predictions of the lower uncertainty boundary underestimate the flood hazards and predict only minor floodings. This underestimation arises from the underestimation of the flood discharge into the fan (hydrological uncertainty) and the underestimation of the flood-triggering rainfall events (climatic uncertainty). The hydrological ensemble underpredicts the flood flows (Chapter 4) and the climate models with the lowest predicted flood magnitudes are the ones that are inferior in capturing the flood seasonality of the Karnali River system (Chapter 5). The reduction of the hydrological ensemble and/or the climate ensemble (e.g. by redefining the selection criteria of the ensembles) could improve the performance of the predictions for the lower interval.

The morphological evolution has a low predicted impact on the overall hazard but affects the local flood hazard characteristics. The morphological evolution increases the connectivity of the Island by the erosion of man-made features such as irrigation gates and embankments. This leads to an increase in the inundation extent on the local scale of a few thousand square metres in the Island. Contrary, the inundated areas reduce in areas at the edges of the predicted inundations in the Western and Eastern floodplains. However, the response of the flood hazard to the morphological evolution is complex and varies between the ensemble members but no simulation indicates that this evolution is a driver of increased or weakening of flood hazards.

The quantification of the sensitivity of flood hazards to morphological evolution is hampered by noises in the DEM and uncertainties in the morphodynamic model. Particularly, the poor representation of the geometry of the small channels draining the floodplain hampers this impact assessment. The predicted decreases in the inundation depths throughout large parts of the fan are merely a result of the improved drainage by the removal of noise by the morphodynamic simulations. The morphodynamic model underpredicts the lateral erosion and, thus, the lateral channel migration which potentially alters the flow pathways and the flood hazards. Furthermore, the predicted

fan head aggradation increases the topographic gradient between channel and floodplain and this increases the risk of channel avulsions which would largely alter the flood hazards but is not depicted by the morphodynamic model (Chapter 6).

The low predicted sensitivity should not be interpreted in the way that the morphological evolution is not important for the evolution of flood hazards. Rather, the impact cannot be quantified with the available datasets and needs to be further investigated. Such an investigation needs topographic data with a good topographic representation of the riverbed and the floodplain channels. Furthermore, repetitive surveys of the topography would enable a better parameterisation of the morphodynamic model to further improve the model predictions.

The hydrodynamic simulations have contributed to a better understanding of how sensitive flood hazards are to climatic and morphological changes. These have updated our knowledge about the climate change impact to the latest climate projections and have greatly improved the spatial resolution of projected flood hazards in the Karnali fan which were previously only available from coarse global modelling. This chapter presents the first flood inundation mapping for projected topographies for fans in mountainous foreland and underlines the necessity of accounting for morphological changes when predicting potential future flood hazards in alluvial fans. Furthermore, the variation in the hydrodynamic predictions improves our understanding of the uncertainty sources and propagation through the modelling cascade.

8 Research summary and conclusions

Section 8.1 summarises the main findings of the modelling framework and the individual research stages. Section 8.2 presents the main sources of uncertainty in this research, followed by recommendations for future studies to decrease these uncertainties and improve the modelling framework (Section 8.3), contributions of this research to knowledge (Section 8.4), and final remarks in the last Section 8.5.

8.1 Summary of this research

The aim of this research was the prediction of the evolution of flood hazards until the end of the 21st century for the Central Himalayan Karnali River in Nepal. To address this aim, a geospatial modelling framework was developed that combines environmental modelling techniques on the catchment scale and floodplain scale to predict the projected climate change impacts on the flood discharge, flood magnitudes, and flood frequencies (catchment scale), to predict the morphological evolution of the Karnali fan for the projected flood discharge (floodplain scale), and to map the spatial flood hazard characteristics for the projected flood magnitudes and topographies (floodplain scale).

Stage 1 of this research (Chapter 4) addressed the first objective (O1): **To establish a hydrological model that replicates the hydrological system of the mountainous catchment and reproduces the observed discharge at the mountain outlet.** For this, the hydrological model SPHY (Terink *et al.*, 2015a) was calibrated and validated for the Karnali catchment. A multi-criteria approach was used to identify 64 parameter sets that provide behavioural simulations of the catchment's hydrology. This hydrological ensemble has a median R^2 of 0.84, median Nash-Sutcliffe efficiency (NSE) of 0.82, and median PBIAS of 14% for the validation period indicating that it reproduces the hydrological system well. The confidence in the simulations is supported by the seasonality of the runoff composition which matches our knowledge of the hydrology of Central Himalayan River systems. The hydrological ensemble was the foundation for the climate change impact assessment on catchment hydrology in Chapter 5 (O2 and O3).

Stage 2 of this research applied the hydrological model parameter sets ensemble with probabilistic climate projections of 12 CMIP6 climate models for three scenarios, the baseline which represents the current climatic conditions (1975 – 2014), the medium-emission scenario SSP245 and the high-emission scenario SSP585 (both 2020 – 2099)

and met the second objective (O2): **to predict the flood discharge at the mountain outlet for the projected climates from an ensemble of climate models until the end of the century using the hydrological model established in O1**. A Flood Frequency Analysis (FFA) was then applied and met the third objective (O3): **to quantify the flood frequencies and magnitudes at the mountain outlet for the flood discharge projected in O2**.

These simulations provide strong evidence that the flood hazards in the Karnali River intensify with the emissions. The hydrological modelling projects an increase in the 30-year mean flood flows (O2) by 38% for the medium-emission scenario SSP245 and 80% for the high-emission scenario SSP585 at the end of the century (2070 – 2099) compared to the baseline (1985 – 2014) (median projections). This projected increase translates into higher flood magnitudes for the projected climates (O3). The FFA projects an increase of the flood magnitude of the 1-in-100-years event by 40% for the medium-emission scenario SSP245 and by 79% for the high-emission scenario SSP585 at the end of the century (2060 – 2099) compared to the baseline (1975 – 2014) (median projections). Consequently, the flood frequencies increase and the FFA projects that the current 1-in-100-year flood magnitude will occur every 11 years in the medium-emission scenario SSP245, and every 3 years in the high-emission scenario SSP585 (median projections).

The impact of the increased climate forcing of the high-emission scenario manifests in the second half of the century. The projections of both scenarios are similar in the near future (2020 – 2059) and decouple in the far future (2060 – 2099). The ensemble median flood magnitude of the 1-in-100-years event increases by 23% (SSP245) and by 26% (SSP585) in the near future. In the far future, the rate of the projected increase decelerates to 40% for the medium-emission scenario and the flood hazards stabilise, even though on a higher level compared to the baseline. In the case of high emissions, the rate of change accelerates to 79%, and particularly large flood events are predicted after the year 2080. Hence, the projections indicate that future flood hazards will be of higher magnitude and occur more frequently than currently, but that the reduction of greenhouse gas emissions limits this hazard intensification.

The consensus of the flood hazard intensification is high among the climate model ensemble but the uncertainty increases alongside the time and emissions. This research used an ensemble of 12 climate models rather than a subset of the extremes (i.e. cold-wet, cold-dry, warm-wet, warm-dry) to analyse the probabilities of the projections. The majority of models predict higher increases while few models predict low increases or decreases. The 1-in-100-years flood magnitude (2060 – 2099) is projected to increase

between 37 – 69% for the medium-emission scenario SSP245 (8/12 models) and between 68 – 143% for the high-emission scenario (9/12 models) (median predictions). This agreement of the climate models provides high confidence that the flood hazards intensify. However, the CMIP6 variation of the projected changes increases (SSP245: +10% to +69%, SSP585: -12% to +144%) indicating that the uncertainty in the projected flood hazard increases with the emissions. This means that flood hazards intensify with increasing emissions but become more difficult to predict and, thus, more difficult to anticipate and prepare for.

The projected intensification of the flood hazards is driven by the intensification of heavy rainfall events during the monsoon season. Rainfall-runoff is the most important source of flood water contributing > 75% to the mean flood event and this importance increases with time, emissions and flow magnitude. The largest simulated events are composed of over 90% rainfall-runoff. Baseflow, snow and glacier melt can contribute significantly to individual events of lower magnitudes but the projected intensification of the flood hazards is driven to more than 90% by rainfall-runoff increases. The simulations do not predict changes in the seasonality or duration of flood events.

Stage 3 (Chapter 6) applied the morphodynamic CAESAR-LISFLOOD model (Bates et al., 2010; Coulthard et al., 2013) to meet the fourth objective O4: **to predict the morphological evolution of the Karnali fan until the end of the century for the projected flood discharge (O2) from morphodynamic modelling**. This stage was challenged by the lack of data to determine the sediment delivery to the fan and calibrate and validate the model. As a consequence, this stage was designed as an experiment to better understand the fan behaviour and the controls of the morphological evolution.

The morphodynamic simulations indicate that the fan evolution is, at the current state and for the simulated time (40 – 80 years), driven by the redistribution of fan sediments and not by the sediment delivery from the upstream catchment. The delivered sediments are deposited on the fan head within 5 – 10 km of the mountain outlet and the downstream evolution is caused by the redistribution of the fan material. The geomorphic activity increases with the projected flows throughout the fan. The fan head aggradation alters the bifurcation and, therefore, the downstream water availability, and increases the risk of channel avulsions because the topographic between the channel and the Western floodplain increases. The simulations predict local channel degradation at the confluence of braid channels and channel sections with decreasing channel width, and the channel aggradation downstream of these sections. Furthermore, the simulations predict lateral bank erosion and channel aggradation which leads to the development of wider and shallower channels. The geomorphic changes scale with the flood discharge

and hence greater changes (both degradation and aggradation) are predicted for the members with higher predicted flood discharge.

These experimental simulations emphasise the necessity to assess the impact of river engineering on the morphological evolution and that this assessment should ideally consider the projected flow conditions. For example, the construction of embankments constrains the tendency to increase the channel width which might lead to the local channel degradation and downstream aggradation which in turn decreases the channel capacity, alters the flood hazard and increases the risk of channel avulsions. The intensification of the geomorphic activity with increasing flow indicates that the construction of structures with long (i.e. decades) design life needs to consider projected flow rather than the past flow.

Stage four (Chapter 7) concluded the modelling framework by applying the hydrodynamic LISFLOOD-FP model (Bates and De Roo, 2000; Bates *et al.*, 2010) and met the fifth objective (O5): **to map the spatial flood hazard characteristics (inundation extent and depth) for the projected flood magnitudes (O3) and topographies (O4).** The spatial flood hazard characteristics predicted for the historical (observed) flood magnitude and topography showed good agreement with previous studies which indicates that the model is behavioural and can be used to map the changes of the flood hazards for the projected flood magnitudes and topographies.

The hydrodynamic simulations provide strong evidence that the flood hazards intensify with the emissions. The 1-in-100-years flood event inundates 39% of the Karnali fan in the baseline (1975 – 2014). This extent increases to 51% for the medium-emission scenario SSP245 and to 60% for the high-emission scenario SSP585 (2060 – 2099) (median predictions). The impact of the emissions is also indicated by the fact that the inundation extent of the SSP245 1-in-10-years event is similar to the one of the baseline 1-in-100-years event, and that the SSP585 1-in-10-years event has the same inundation extent as the SSP245 1-in-100-years event. This increase in the flood frequencies means that current extreme events will occur frequently in the future, even in the medium-emission scenario but that high emissions will intensify this trend even further. The inundations are projected to increase for all predictions which emphasises the high confidence in the sensitivity the flood hazards to climate change.

The fan topography constrains the pathways of the flood flow and divides the fan into areas at high risk and areas at no risk. The elevation gradient of the Western floodplain and the Island between the two Karnali branches is steeper in the longitudinal direction than in the lateral direction. In these areas, the overbank flow that enters the floodplain

drains the study area on the floodplain and does not flow back into the main channels. On the contrary, the areas along the Eastern border of the fan and at the foot of the Siwalik hills are superelevated over the channels and are not inundated by the highest simulated flood events of ≥ 40,000 m^3/s. This topographic setting causes a steep increase in the inundation extent with flood discharge until flows < 15,000 m^3/s. The additional water of higher discharges increases the inundation depths stronger than the inundation extent, although the extent still increases.

The areas most exposed to floods and most sensitive to the projected intensification are the densely populated areas in the Island and the Western floodplain. The populated area inundated during the 1-in-100-years event increases from 19% in the baseline to 33% for the medium-emission scenario SSP245, and to 46% for the high-emission scenario SSP585 (2060 – 2099) (median predictions). The inundation depths increase alongside the inundation extent from 0.41 m in the baseline to 0.51 m for SSP245 and 0.57 m for SSP585 (median water level of the median predictions) and hence the population is exposed to more damaging and dangerous floods. Furthermore, the projected intensification of the inundation extent and inundation depths increases the damage potential to agriculture and livestock which is the main source of income of the population.

Large areas of the fan are exposed to flooding from multiple directions which threatens the access to and escape from the areas at risk. The municipalities Geruwa, Rajapur (Island) and Tikapur (Western floodplain) are surrounded by floodwater of the Karnali. Furthermore, the fan population is also exposed to flooding from smaller Siwalik rivers which have not been studied in this research but add additional flood water to the fan, whereas this interaction depends on the spatial and temporal characteristics of the rainfall event. However, flood management needs to consider all contributing rivers because the simultaneous flooding of the Karnali River and the Siwalik rivers would lead to larger and deeper inundations as predicted by this study. The population in large parts of the fan would be locked in by flood water from all directions limiting the access to safe areas, and escape routes, and hampering the emergency response.

This research provides no clear evidence that the morphodynamic evolution intensifies or decreases the flood hazards. The projected topographies alter the flow pathways leading to changes in the inundation extent on the local scale of a few thousand of square metres. Generally, the connectivity of the Island to the Karnali branches is projected to increase which is caused by the erosion of water and flood management structures and leads to a slight increase in the inundation extent in the Island. Contrary, the Eastern and Western inundation boundaries decrease in some areas. However, the

predicted effect of the morphological evolution on the flood hazard is low and superimposed by the uncertainty in the topographic data and needs further investigation.

This research presented the first application of a modelling framework that assesses the joint impact of climatic and geomorphic changes on flood hazards in a sediment-rich, mountainous catchment. These simulations indicate that the potential future flood hazards change significantly from the current flood hazards and that long-term flood risk management needs to consider these potential changes for the development of effective adaptation and mitigation strategies. Although the effect of morphological changes on the flood hazard patterns was low compared to the climate-change-induced increase in flood magnitudes, it alters the flow pathways to a degree that varies between the climate models and scenarios. This indicates that the prediction of potential future flood hazards requires the integrative simulation of climate, catchment hydrology, morphological evolution and flood hazards. The modelling framework provides a tool for such a comprehensive analysis.

The developed framework is transferable to other mountainous and sediment-rich river systems both within and beyond the Himalayas. However, it may be adjusted or extended to the catchment-specific requirements. For example, the impact of land-use changes on flood flows may be estimated by adding land-use scenarios to the hydrological simulations in catchments which are characterised by rapid socio-economic changes (e.g. urbanisation). Furthermore, the framework could be extended by simulating the sediment delivery from the hillslopes to the channel for different climate scenarios to better constrain the sediment inflow boundary condition. As such, the developed modelling framework can be used as a blueprint for the prediction of potential future flood hazards and may be adapted to the specific characteristics of the study area.

8.2 The sources of uncertainty in the modelling framework

The modelling framework links a set of environmental and statistical models and each of these models adds uncertainty that propagates to the next stages. The uncertainty arises from the incomplete knowledge of the modelled system (the theoretical knowledge and its translation into the model) and the incomplete knowledge of the modelled area (the catchment and fan). The developed framework was designed to incorporate this uncertainty, but not all uncertainty sources could be quantified due to restrictions in the computational resources, data, and science's understanding of the hydrological processes, and pathways within the Karnali fan system. This section discusses these quantified and not-quantified uncertainties.

8.2.1 Uncertainty in the catchment modelling of Stages 1 and 2

The uncertainty in the hydrological simulations of Stages 1 and 2 arises from the scarcity of data particularly in the high-mountainous parts of the catchment, the process representation of the hydrological model, the model parameterisation and the parameter transfer from the observed climate data to the simulated climate of the CMIP6 ensemble.

The hydrological model underestimates the flood discharge because the model cannot reproduce the fast nature of the rainfall-runoff conversion of the Karnali River to the full extent and hence a fraction of the flood discharge drains on consecutive days. This process representation deficit leads to the underestimation of the flood magnitudes, the flood discharge for the prediction of the morphological evolution, and the underestimation of the spatial flood hazard characteristics by the hydrodynamic model.

The parameter uncertainty was quantified from ensemble modelling based on the exploration of 64 parameter sets with different combinations of 15 parameters using the GLUE framework. The parameters that control the allocation of rainfall to rainfall-runoff and baseflow have the largest effect on the predictions, while the ensemble variation in the snowmelt and glacier melt contribution is comparatively low. Furthermore, the runoff coefficient that controls the speed of transport within the channel network and thus the timing of the hydrograph is a sensitive parameter. The uncertainty analysis in the last Stage 4 indicates that the parameter uncertainty of the hydrological modelling contributes a significant proportion to the total uncertainty. The Regional Sensitivity Analysis (RSA) and the comparison of the parameterisation and the runoff contribution indicate that the parameter uncertainty is mainly caused by a few parameters which

control the runoff routing (within the channel network) and the allocation into rainfall-runoff and baseflow.

Generally, the simulation of Central Himalayan catchment hydrology is challenging because of the pronounced seasonality of the climate and hydrological regime. This seasonality cannot be captured to the full extent by global (temporal dimension) parameters. For example, the precipitation correction factor and the *alphaInf* parameter (fraction of rainfall during the hour of most intensity) cannot depict the characteristics of both, monsoon and non-monsoon season precipitation. The model calibration favoured the performance during the monsoon season which resulted in systematic errors during the non-monsoon season (i.e. the overprediction of peak flows in the winter and pre-monsoon seasons). Furthermore, the use of global (spatial dimension) cannot capture the heterogeneity of the hydrological conditions in the catchment. Instead, the parameters are averaged over the catchment which adds another source of uncertainty.

It is further worth noting that the hydrological ensemble was calibrated with discharge observations with inherent uncertainties. The discharge observations are based on sparse stage-discharge observations during low to medium flow conditions and hence the observations of high discharges are particularly uncertain. This uncertainty was estimated from Bayesian stage-discharge rating curve fitting and considered in the calibration process by using a deviation of the eGLUE approach. However, the lack of in-situ data of the cross-section at the gauging station and the poor temporal resolution of the stage-discharge observations pose a challenge for the uncertainty estimation of the discharge observations.

The transfer of the parameter from the observed climate to the simulated climate of the CMIP6 ensemble adds further uncertainty. The seasonality and composition of the discharge have a high agreement between both sets of simulations which indicates that the model behaviour remained similar. However, the removal of the precipitation correction factor and its interaction with other parameters reduced the discharge in the CMIP6 simulations leading to the underestimation of the flood discharge in the FFA, the morphodynamic simulations and the hydrodynamic simulations.

The lack of climate observations in the high mountains and along the slopes and ridges throughout the catchment hampers the downscaling of the remote sensing observations and climate models. This affects both temperature and precipitation:

- The temperature is downscaled with a lapse rate generated from observations up to 3,500 masl. This lapse rate is extrapolated to above 7,000 masl which led to the overestimation of the temperatures at higher elevations, and the overestimation of

snowmelt. The transfer of this lapse rate to the CMIP6 data, and the parameter transfer increased the glacier melt predictions.

- The precipitation bias was estimated from the calibration of a correction factor. This factor was removed for the CMIP6 simulations because this data was bias-corrected by Mishra et al. (2020). However, the lack of observations in the catchment hampers the bias-correction, and the lower predicted discharge for the CMIP6 simulations indicates that this dataset underestimates the precipitation.

8.2.2 Uncertainty in the morphodynamic simulations of Stage 3

The simulation of the morphodynamic fan evolution is the most uncertain part of the modelling framework and this uncertainty is most difficult to quantify because:

- The morphodynamic system involves different processes that interact on different spatial and temporal scales and we lack the theoretical understanding to construct models that can accurately quantify these processes and their interactions (Mosselman, 2012; Hardy, 2013; Ancey, 2020a).
- The lack of data on the channel geometries (main channel and floodplain channels), the sediment delivery from the Himalayas, and the past morphological changes challenge the determination of the boundary conditions and identification of the ideal model parameterisation.
- The complexity of the computations and the resulting long computational runtimes prevent the simulations of large ensembles to estimate the parameter uncertainty which is especially problematic because the model could not be calibrated to the conditions of the Karnali fan. The simulated ensemble considers the climatic uncertainty (12 CMIP6 models) and the hydrological uncertainty ($P_{2.5}$, P_{50}, $P_{97.5}$ of the hydrological ensemble) but all members use the same parameter set which was determined from the literature and initial testing.

These limitations only allow for experimental simulations to better understand the geomorphic system which have limited predictive conclusiveness for the projected flood hazards. Paradoxically, while this is the most uncertain stage of this research, it adds little uncertainty to the prediction intervals because this uncertainty cannot be quantified.

8.2.3 Uncertainty in the hydrodynamic simulations of Stage 4

The uncertainties of all models are combined in the hydrodynamic simulations of Stage 4. The high agreement of the inundation pathways with other studies indicates that the hydrodynamic simulations are behavioural and represent the large-scale flood hazard characteristics. There might be small-scale uncertainty introduced by the hydrodynamic simulations which has not been quantified. However, it is assumed that this uncertainty is low compared to the uncertainty propagating from the previous models considering the physics-based flow model, the low number of parameters, the agreement of the used parameters with the literature, and the agreement of the inundation characteristics with other studies.

The hydrological uncertainty arising from the hydrological modelling and the FFA and the climate uncertainty arising from the variation in the climate projections from the CMIP6 ensemble are the main uncertainty sources. The comparison of the spatial hazard patterns between the baseline predictions and the ones of the observed flood magnitudes indicates that the lower uncertainty boundary underestimated the flood hazard but that the median and upper prediction intervals are reasonable representations of the flood hazards.

A further source of uncertainty is inaccuracies in the Digital Terrain Model (DTM) that represents the fan topography. The estimation of the ground elevation of forests along the riverbanks (particularly west of Tikapur city) and the channel bed of the Karnali branches affect the channel capacity and might cause the under- or overestimation of the bankfull flow rate. The poor representation of small channels draining the floodplains, particularly in sections with vegetated riverbanks, which introduce a flow barrier in the longitudinal direction, leads to a local overestimation of the inundation extent and depth because it reduces the drainage of the floodwater. Such flow barriers have been removed by the morphodynamic simulations which increased the drainage efficiency. This introduces an artificial discrepancy between the initial and projected DTMs which impedes an accurate assessment of the impact of the morphological evolution on the flood hazards.

The linking of environmental and statistical models introduces uncertainty that propagates through the modelling cascade. The main sources of uncertainty arise from the hydrological modelling and the climate projections but not all uncertainty could be quantified due to restrictions in the available data and the computational resources. Furthermore, not all future developments are considered by the modelling framework (e.g. land use change, urbanisation, river engineering). However, the good agreement

of the simulated and observed behaviour indicates that the developed modelling framework provides reasonable projections, that flood hazards will intensify, and that flood risk management needs to consider the evolution of flood hazards.

8.3 Recommendations

The developed modelling framework may be refined by altering or extending it or increasing the database of the river system. Additional datasets for a better determination of the boundary conditions potentially improve the modelling performance in each stage of the modelling framework. Each of these areas for improvement is covered in the following sections.

The following alterations of the environmental models may improve the predictions of the modelling framework:

- Hydrological modelling: The implementation of seasonal parameters (e.g. monsoon and non-monsoon seasons) could improve the prediction of the runoff seasonality of Central Himalayan river systems. This would increase the required sample size in the calibration process and it is, therefore, recommended to focus on those parameters that affect the rainfall-runoff conversion.
- Hydrological modelling: It is recommended to investigate the potential of coupling continuous and event-based hydrological models to predict flood discharge. The continuous model can be applied to predict long-term processes (e.g. glacier dynamics, snow dynamics, groundwater re- and discharge) and provide the boundary conditions (e.g. soil moisture, snowmelt runoff, glacier melt runoff, baseflow) for the event-based model that is used to predict the rainfall-runoff conversion for the flood-triggering rainfall events.
- Hydrological modelling: The extrapolation of the lapse rate from the observations in the hills to the high mountains is not recommended. Instead, the lapse rate at unobserved elevations could be calibrated against the snow cover and glacier mass balance observations.
- Hydrological modelling: The model was calibrated for the observed historical conditions. It is worth investigating to calibrate the model for the simulated baseline climate data. Such a calibration could aim to maximize the fit between the simulated and observed distributions of different hydrological variables (e.g. Majone *et al.*, 2022; Wu *et al.*, 2022). This direct calibration for the CMIP6 climate data would remove the uncertainty added by the parameter transfer between different climate datasets.
- Morphodynamic modelling: The ensemble of the morphodynamic modelling could be reduced which provides resources to simulate the parameter uncertainty. The evolution was predicted for each of the 12 CMIP6 members for three discharge percentiles to account for the non-linear nature of the morphodynamic system.

However, the model predicted a somewhat linear response. Therefore, it could be valuable to reduce the climate ensemble in Stage 3 and instead include multiple parameter sets to estimate the parameter uncertainty. The reduced ensemble size could also be used to apply a more complex model with an improved physical process representation.

- Morphodynamic modelling: The morphodynamic simulations in this thesis were experimental, exploratory predictions and further studies are required to address the following questions: How much physical complexity is necessary and feasible for ensemble predictions of the large fans in the Terai Plain? This question could be a first step to establishing guidelines for predicting the long-term impact of the evolution of flood hazards in these environments. How can we best depict the heterogeneous environment of the fan and mountain gauge? The use of global parameters to represent the gauge and the fan causes inaccuracies at the fan inlet which affect the prediction until the downstream border. An investigation of the representation of this heterogeneity may improve future predictions. How many flow events need to be simulated to capture the evolution? This research only simulated the evolution caused by flood flows which may not be adequate to capture the evolution to the full extent, particularly further downstream where the grain sizes decrease.
- Modelling framework: The predictions of the lower prediction interval underpredicted the observed behaviour (i.e. flood discharge, inundation extent). The performance could be improved by redefining the selection criteria of the climate data (e.g. by considering the flood seasonality) and the hydrological ensemble (e.g. by focusing more strongly on the flood discharge predictions). Furthermore, better-performing members could be assigned a higher leverage on the ensemble predictions because the members which better reproduce the historical behaviour are likely more reliable in predicting the future (Padrón *et al.*, 2019; Giuntoli *et al.*, 2021).
- Modelling framework: The modelling framework could be extended thematically to consider more factors. For example, the effect of landuse change could be considered by adding different landuse scenarios to the hydrological modelling in Stage 2. These scenarios could be integrated into the climate scenarios which also include socio-economic pathways that could be used to develop the landuse scenarios. This integration has the benefit that the ensemble size would not increase but it would not be possible to separate the impact of the landuse change and climate change on the flood hazard.
- Modelling framework: The modelling framework could be extended spatially to include the Siwalik rivers that add floodwater to the floodplain. This extension

challenges the calibration of the hydrological model because these small rivers have different hydrological characteristics and need a separate calibration. However, the flood hazard mapping for flood risk management and flood forecasting simulations must consider these rivers because otherwise the inundation extent and depth may be underestimated when the rainfield is located over the Siwaliks.

The following datasets could improve our understanding of flood hazards and catchment hydrology and improve the predictions by providing additional information to apply and validate environmental models:

- Temperature data (high potential): The extension of the temperature station network to higher elevations (> 3,500 masl) would improve our understanding of the mountain climate and more accurately represent temperatures in environmental modelling (hydrological and glaciological models). These high elevation temperature datasets would be useful for a better determination of lapse rates and to improve the bias-correction of remote sensing and reanalysis products which would decrease the uncertainty in the temperature representation in environmental modelling. The extension of the station network has a high potential to improve our water management, flood risk management and climate adaptation.
- Precipitation data (high potential): The small-scale variability of the precipitation and the high-mountainous precipitation patterns are not well understood due to the concentration of the precipitation station network along the plains and valleys. So far, the information about high-mountainous precipitation is ambiguous; the precipitation in gridded precipitation products decreases with elevation (> 4,000 masl) but this lower precipitation cannot sustain the observed mass balance of glaciers (Winiger *et al.,* 2005; Immerzeel *et al.,* 2015; Lutz and Immerzeel, 2016). However, an accurate representation of the precipitation in the high mountains is crucial for hydrological, glaciological and climate modelling. The extension of the station network to the high mountains would improve the modelling predictions which would benefit water management, flood risk management, and climate adaptation. Furthermore, dense measurements along cross-sections covering the ridges, slopes and valleys would potentially improve our understanding of the small-scale variability of precipitation patterns and would expand our knowledge about the climatology and hydrology in the Himalayas.Soil moisture data (medium potential): The simulation of flood peaks is sensitive towards the parameters that control the distribution of rainfall between rainfall-runoff and baseflow. Soil moisture data could be used to better parameterise hydrological models and have the potential to decrease the uncertainty in peak flow predictions.

- Discharge data (low potential): The discharge data includes considerable uncertainty due to the large range of flows and the sampling during low – medium flow conditions. The discharge uncertainty is, to this date, only a side note in most hydrological studies in this region. The publication of the discharge uncertainty, stage-discharge samples, stage-discharge rating curves and cross-section profiles is a cheap and easy-to-implement way to communicate this uncertainty. It would enable the calibration of hydrological models with the limits of the acceptability approach.
- High-resolution Digital Terrain Models (high potential): A time series of DTMs (stereo) satellite or drone imagery of the fans before and after events with high geomorphic activity would be very useful to better understand the morphological processes in the fan, the sediment delivery during single events and could be used to constrain the sediment influx boundary condition and/or to calibrate and validate morphodynamic models. The choice of the carrier system depends on the financial resources and the area of interest. Drone-based DTMs have a higher spatial resolution than satellite-based DTMs and represent the ground elevation (e.g. in forested areas) more accurately. These DTMs can accurately represent the channel topography if the drone is equipped with green LIDAR. For large study areas such as the Karnali fan, satellite-based DTMs can be a cost-effective alternative, particularly for the application of morphodynamic models because these models are constrained by the computational resources and hence DTMs need to be downscaled for the application in large areas. However, ground surveys (i.e. ground elevation, channel geometry) need to be integrated when generating a DTM from satellite imagery.

8.4 Contribution to knowledge

The developed modelling framework extends existing frameworks to predict future flood hazards and links climate change and sediment dynamics which is a novel approach. Commonly, potential flood hazard changes are predicted either for different climatic conditions (e.g. Hirabayashi *et al.,* 2013; 2021; Dankers *et al.,* 2014; Wijngaard *et al.,* 2017; Huang *et al.,* 2017) or topographies (e.g. Pender *et al.,* 2016; Dingle *et al.,* 2020a). This thesis presents to the best knowledge of the author the first framework that relates climate, catchment hydrology and geomorphological processes to predict future flood hazards and contributed to the process understanding and methodological advancements.

Contribution to new knowledge and process understanding

The hydrological simulations of the Karnali River (Chapter 4) complement our knowledge of the hydrology of Central Himalayan River systems. It provides insight into the contribution of hydrological processes to the discharge (e.g. snow melt and glacier melt) and their seasonal and annual variation. These simulations also improve our understanding of the flood peak generation and how the spatial and temporal characteristics of the precipitation events shape the flood hydrographs.

The hydrological simulations with the climate projections (Chapter 5) update our knowledge about the climate change impact on the hydrology of Central Himalayan river systems by applying an ensemble of the latest generation of CMIP6 climate projections on the catchment scale. Furthermore, this study improves our knowledge about the implications of climate uncertainty on catchment hydrology by applying probabilistic ensemble modelling (12 climate models) instead of climate model subsets (i.e. cold-dry, cold-wet, warm-dry, warm-wet). These simulations contribute to a deeper understanding of how climatic changes alter the hydrological processes for flood peak generation and how these trends evolve until the end of the century and establish that the projected increases in flood runoff generation are driven by increases in rainfall intensity.

The morphodynamic simulations (Chapter 6) extend our knowledge of the climate change impact on Himalayan River systems and sediment-rich mountainous river systems more broadly. This research presents, to the best knowledge of the author, the first application of a morphodynamic model using flow projections of probabilistic climate modelling to predict potential future topographies for different climate scenarios. These simulations indicate that the morphological evolution of alluvial fans is affected by

greenhouse gas emissions and, thus, deepen our understanding of the climate change impact on mountainous river systems.

The morphodynamic simulations (Chapter 6) improve our understanding of the behaviour of the Karnali fan, its potential evolution until the end of the 21st century, and the factors which control this evolution. Furthermore, these simulations provide evidence that the construction of embankments constrains the river and potentially increases the downstream flood risk.

The flood hazard mapping (Chapter 7) complements our knowledge of the spatial flood inundation patterns of current flood hazards and provides novel information about areas at risk of potential future flooding. Previously, information about flood projections was only available from global flood inundation mapping with coarse resolution and this research updates this information to the latest generation of climate models and a fine 25 m resolution.

The flood hazard mapping (Chapter 7) has improved our knowledge of how morphological changes potentially alter flood hazards and has identified regions that may be at higher risk due to the morphological evolution. The wide range of simulated inflows has further improved our understanding of how flood waves propagate through the Karnali fan. These simulations emphasise the necessity to consider the potential future flood hazards regarding both, climatic and morphological changes for effective flood risk management.

Methodological advancements

The hydrological modelling of the historical flood flows (Chapter 4) has advanced our understanding of how to predict flood flows in these Central Himalayan river systems. The peak flow predictions are sensitive to the parameters that control the allocation of rainfall between rainfall-runoff and baseflow. Furthermore, the ensemble variation decreases with increasing flood flow because more rainfall is converted into rainfall-runoff. This information may help for model selection and parameterisation to improve the simulation of flood flows from hydrological modelling.

The hydrological simulations (Chapters 4 and 5) complemented our understanding of the challenges of simulating the hydrology of Central Himalayan catchments with current models (e.g. global parameters) and assessing the climate change impact on the catchment hydrology in mountainous river systems (e.g. parameter transfer between datasets).

The simulations of the morphodynamic evolution of the Karnali fan for probabilistic flow projections (Chapter 6) have identified challenges in predicting the potential evolution of large alluvial fans such as the representation of the heterogenous fan characteristics and the conflict between model complexity and climate uncertainty. This information may help develop further research projects to better understand and simulate fan environments.

A hydrodynamic model was applied with projected topographies and flood flows for probabilistic climate scenarios (Chapter 7). These simulations advanced our understanding of the uncertainty sources and propagation through modelling cascades which may improve the linkage of environmental simulation techniques in future research projects.

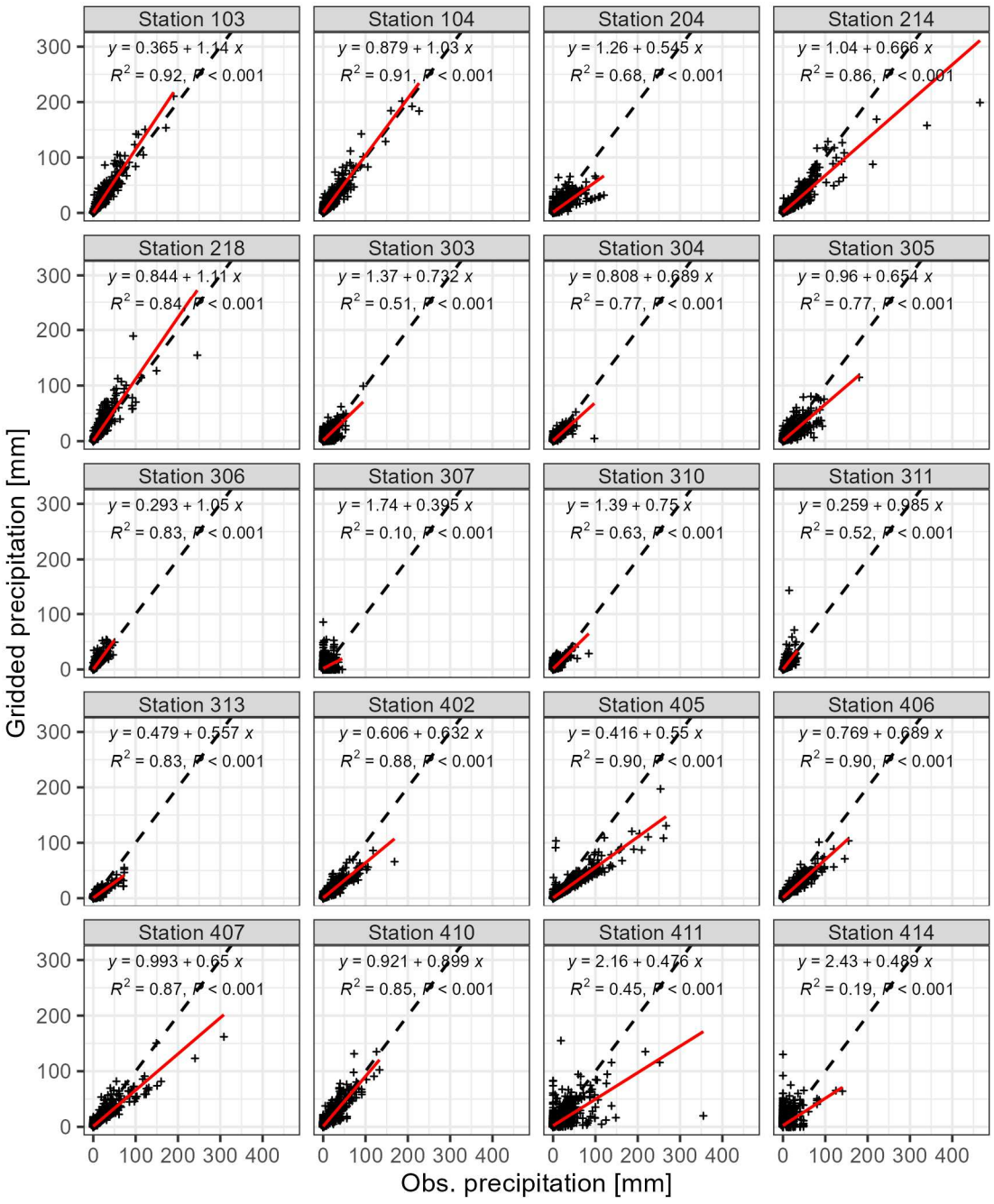
A novel approach was developed to generate Digital Terrain Models for point clouds based on stereo satellite data with lower point density than airborne LiDAR (Chapter 6). This approach needs refinement and the integration of ground-truth datasets but is potentially applicable for a wide range of disciplines that study earth surface deformation processes.

8.5 Final remarks

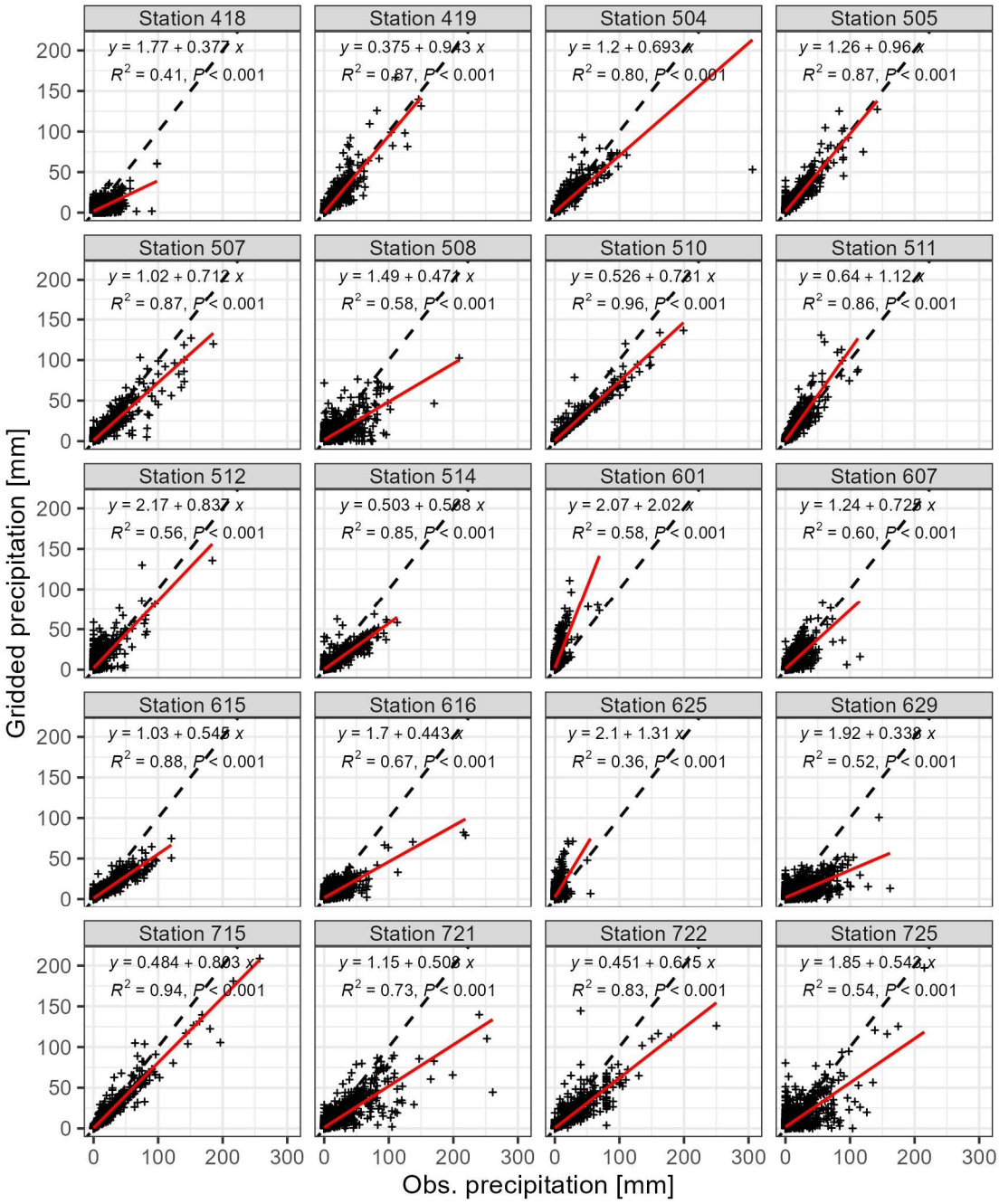
I would like to conclude this thesis with the following two statements that are my personal opinion.

It takes a global effort to limit climate change and it is our responsibility, particularly for the industrial countries, to reduce the emissions as quickly and as much as possible. The flood hazard projections in this and other studies provide very strong evidence that climate change intensifies these hazards. From the analysis presented in this thesis, the population of the Karnali fan are projected to experience the current 1-in-100-years flood event every 2 – 7 years in the worst-case scenario (high-emission scenario, 2060 – 2099). It is questionable whether and how the people could sustain their livelihoods under such extreme conditions particularly because Nepal is one of the poorest countries in the world and the climate change impact is neither restricted to the Terai Plain nor to flood hazards. We need to move beyond our national, cultural, historical and ideological differences to tackle the climate crisis. We owe this to the Nepalese and global communities at risk who have never been the driver of this crisis but are strongly threatened by the consequences.

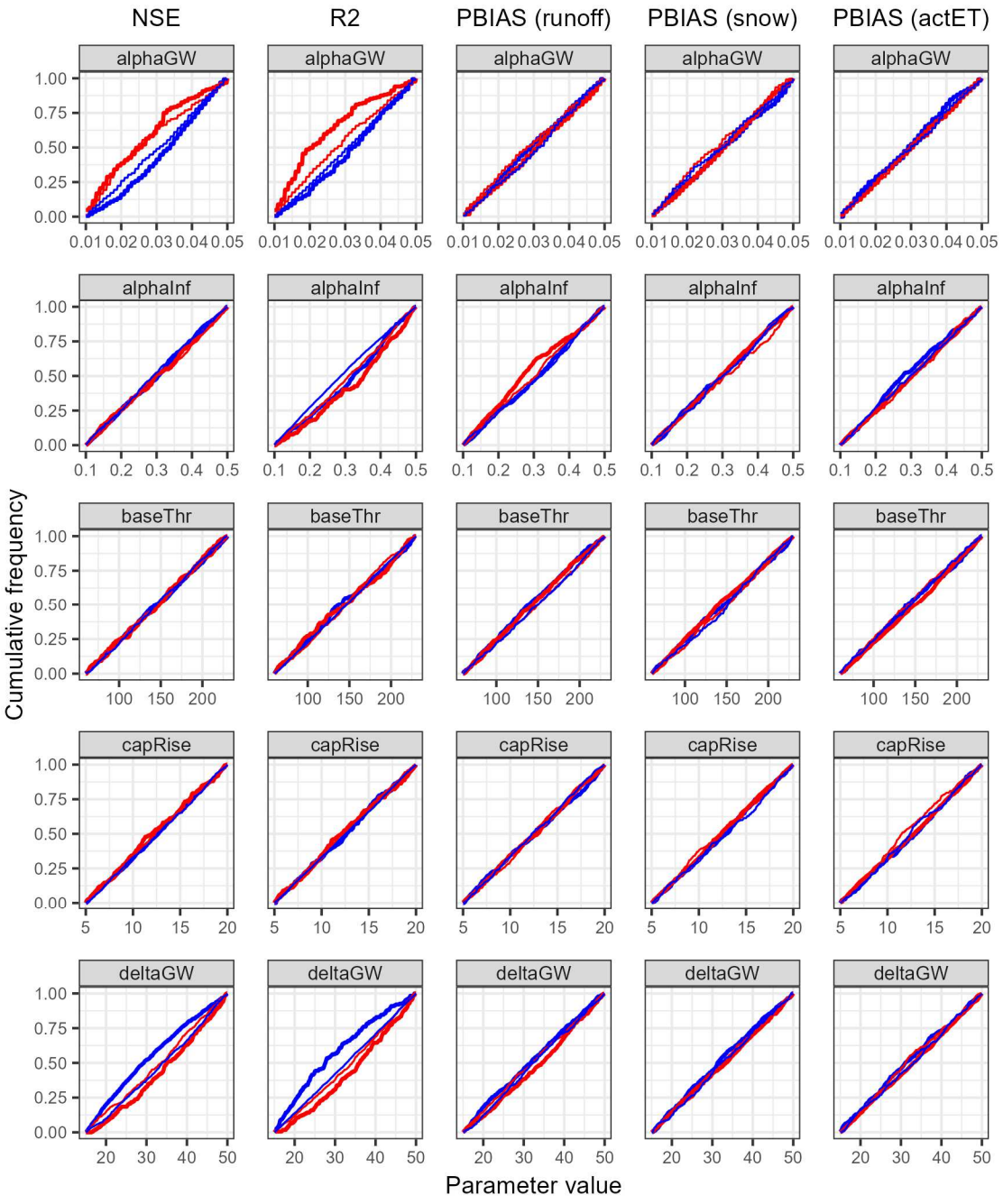
The rivers sourced in the Himalayas connect countries with heterogeneous political, religious and cultural backgrounds and may be used to foster confrontation or promote collaboration. Examples of conflict potentials are the construction of dams and river rerouting that reduce the water availability in the downstream country, or the construction of embankments along national borders that intensify flood hazards in the upstream country. Climate change increases the hydrological extremes and with that the conflict potential. However, the water resources can only be effectively managed on a supranational scale that integrates the interests of all members in this large, densely-populated and diverse region, particularly in light of climate change. For example, flood forecasting systems would benefit from meteorological observations in the Nepalese Himalayas. Thus, these rivers have the potential to connect people and societies. It remains to see whether climate change will unite or divide this region that hosts more than one-tenth of the population of the world.



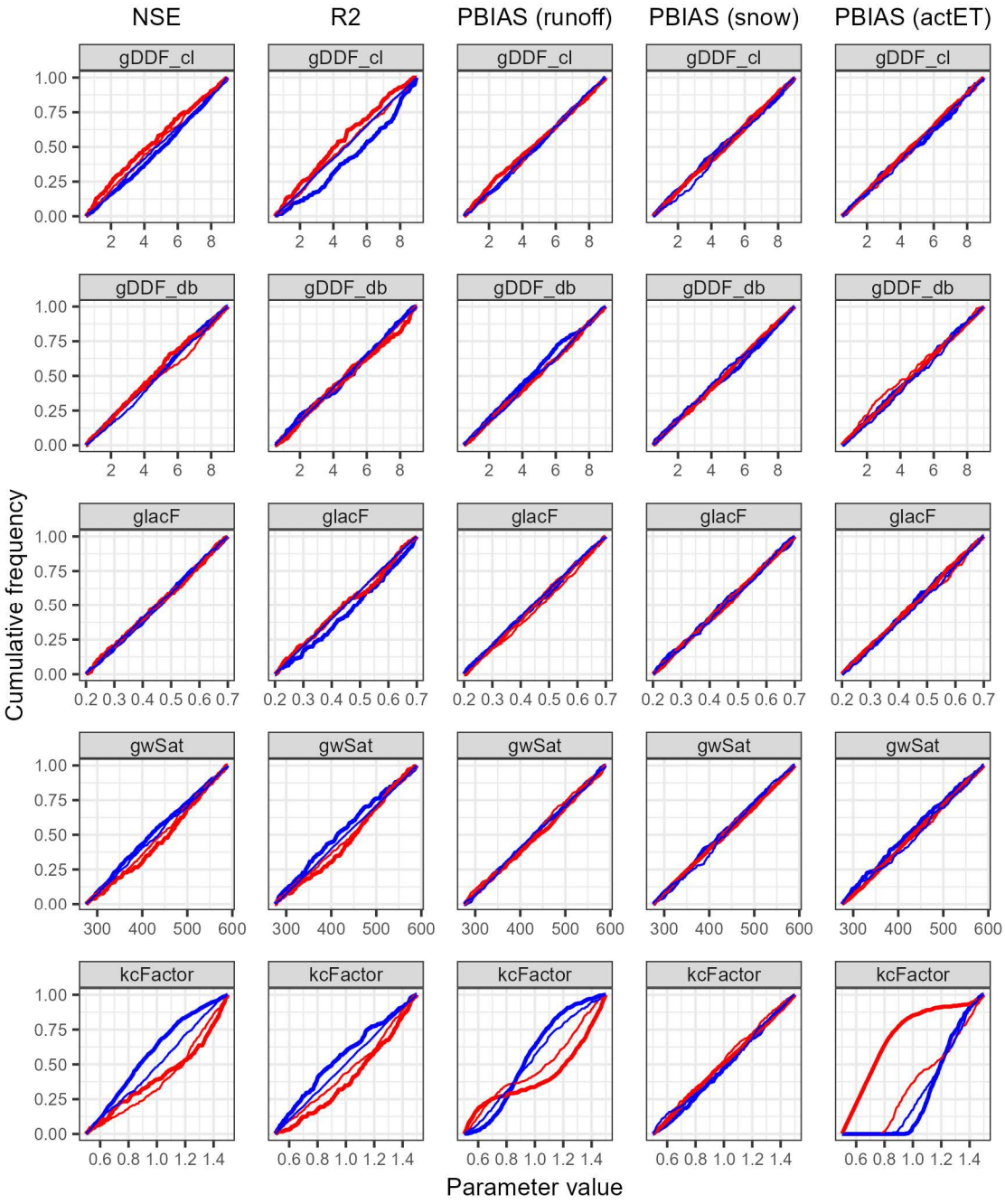
Appendix 4.1: Comparison of gridded daily precipitation and DHM observations for 20 precipitation gauges in the region – Part 1. The daily gridded precipitation is based on monthly GPM data which was diaggregated to the daily resolution following Arias-Hidalgo et al. (2013). The investigated period is 2000 - 2009. The red line indicates the linear regression and the black dashed line indicates the line of the perfect fit. The location of the gauges are presented in Figure 2.1.



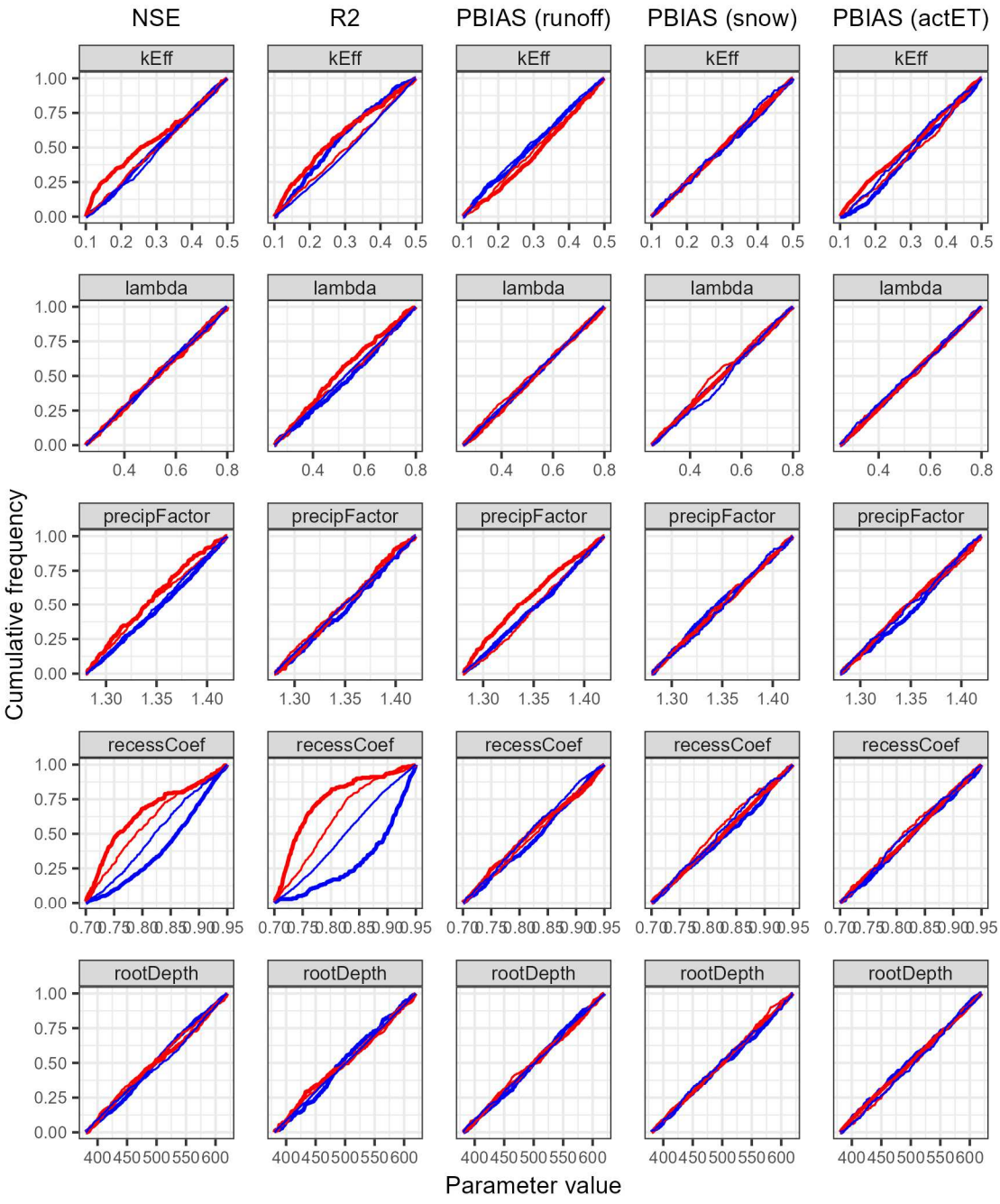
Appendix 4.2: Comparison of gridded daily precipitation and DHM observations for 20 precipitation gauges in the region – Part 2. The daily gridded precipitation is based on monthly GPM data which was diaggregated to the daily resolution following Arias-Hidalgo et al. (2013). The investigated period is 2000 - 2009. The red line indicates the linear regression and the black dashed line indicates the line of the perfect fit. The location of the gauges are presented in Figure 2.1.



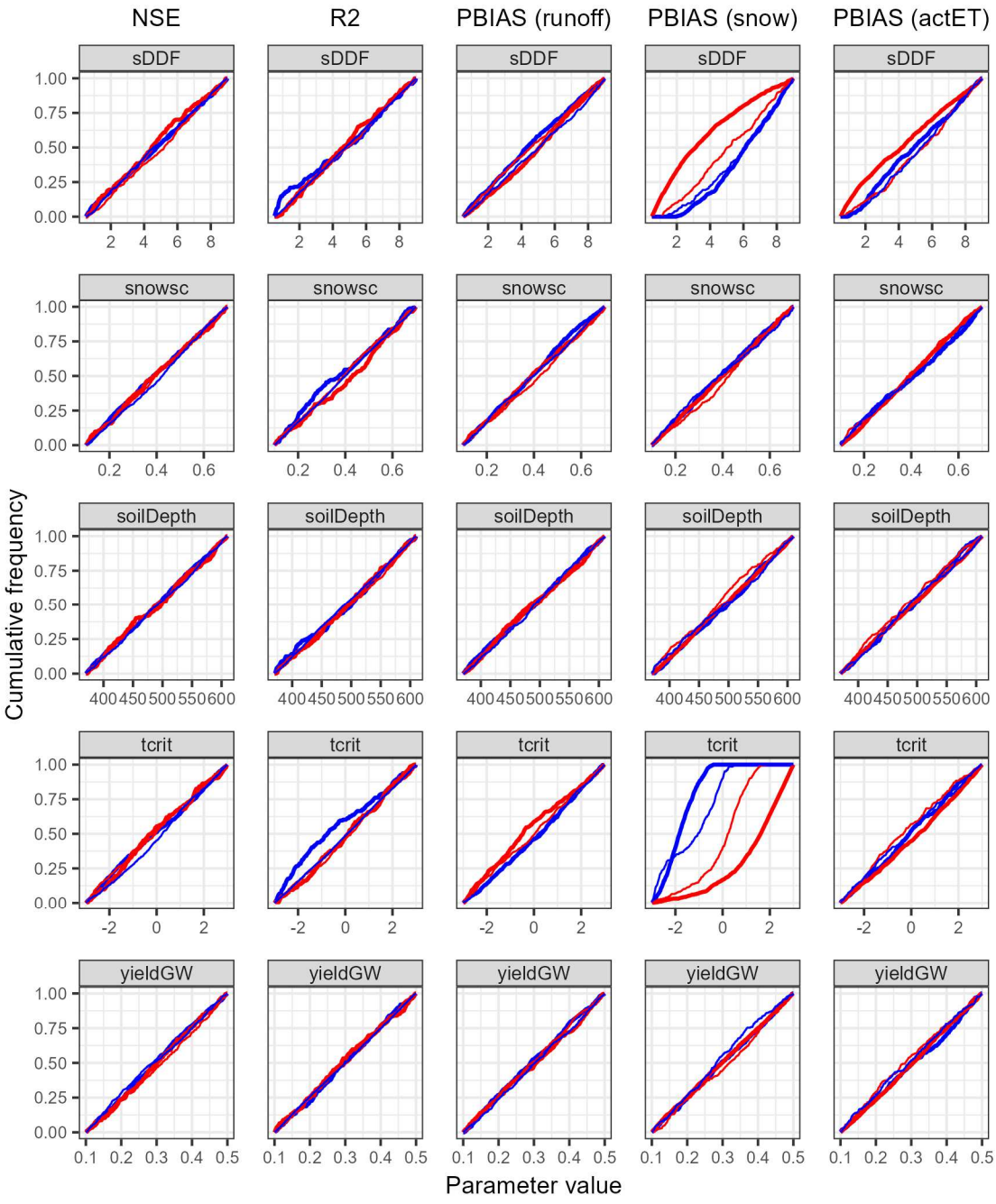
Appendix 4.3: Regional Sensitivity Analysis (RSA) of the all parameters and performance measures – Part 1. The red line indicates the cumulative frequency of the non-behavioural realisations and the blue line is the cumulative frequency of the behavioural simulations.



Appendix 4.4: Regional Sensitivity Analysis (RSA) of the all parameters and performance measures – Part 2. The red line indicates the cumulative frequency of the non-behavioural realisations and the blue line is the cumulative frequency of the behavioural simulations.



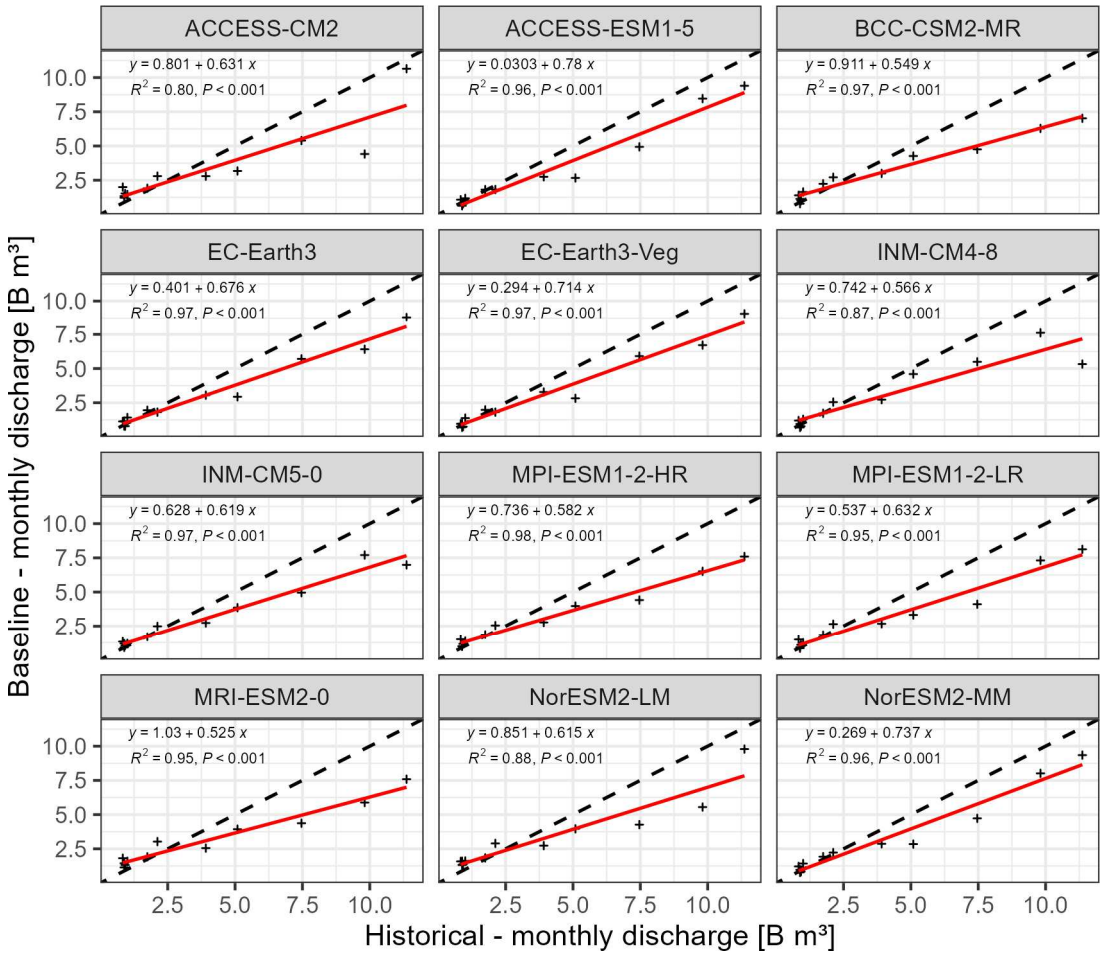
Appendix 4.5: Regional Sensitivity Analysis (RSA) of the all parameters and performance measures – Part 3. The red line indicates the cumulative frequency of the non-behavioural realisations and the blue line is the cumulative frequency of the behavioural simulations.



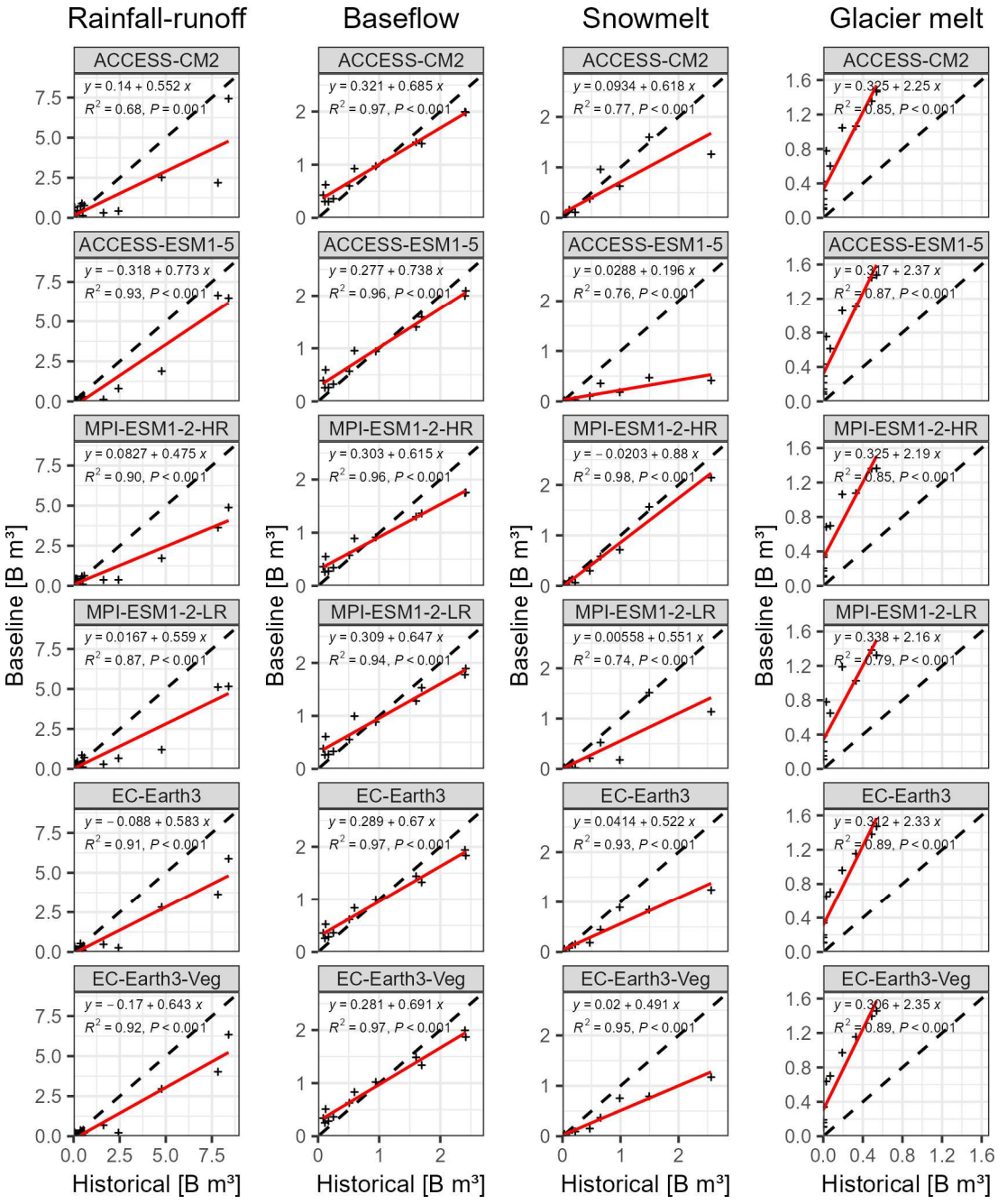
Appendix 4.6: Regional Sensitivity Analysis (RSA) of the all parameters and performance measures – Part 4. The red line indicates the cumulative frequency of the non-behavioural realisations and the blue line is the cumulative frequency of the behavioural simulations.

Appendix 4.7: The percentage bias of the actual evapotranspiration each land cover class.

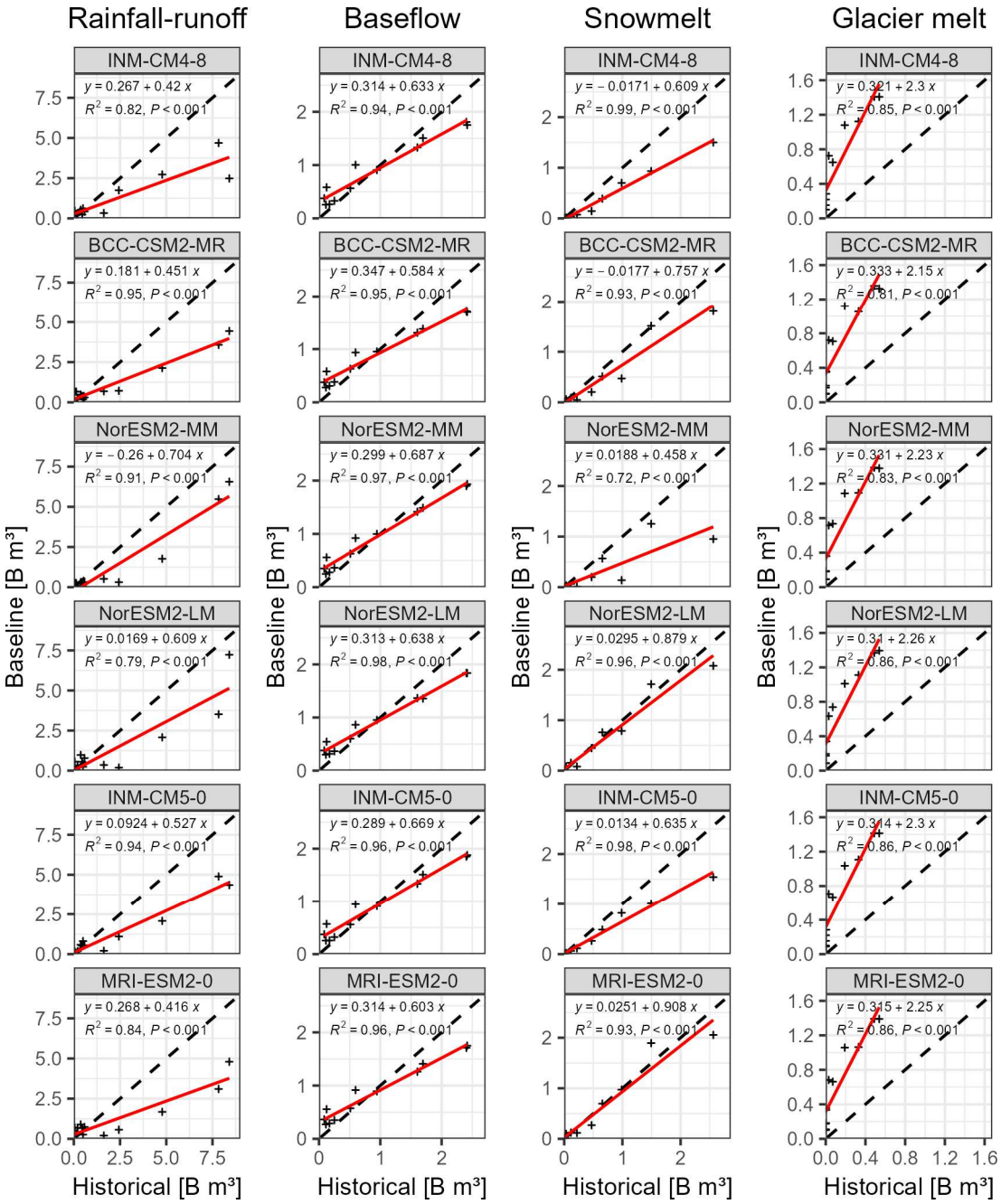
ID	Land cover	Period	PBIAS				
			Min	P_{10}	Median	P_{90}	Max
11	Irrigated croplands	Calibration	-18	-17	-12	-2	20
		Validation	-27	-26	-22	-13	7
14	Rainfed croplands	Calibration	-19	-16	-12	-1	21
		Validation	-27	-24	-20	-11	9
20	Cropland / vegetation	Calibration	-18	-16	-11	-3	20
		Validation	-26	-24	-20	-13	7
30	Vegtation / cropland	Calibration	-17	-16	-11	-3	19
		Validation	-25	-23	-20	-12	7
40	Semi-decidous forest	Calibration	-17	-15	-11	-2	20
		Validation	-22	-21	-17	-8	11
50	Closed decidous forest	Calibration	-19	-19	-13	-4	18
		Validation	-26	-26	-21	-13	7
70	Closed evergreen forest	Calibration	-18	-15	-10	0	25
		Validation	-26	-22	-18	-8	14
100	Mixed forest	Calibration	-18	-16	-11	-1	21
		Validation	-25	-23	-19	-10	9
110	Forest / grassland	Calibration	-16	-12	-8	2	27
		Validation	-25	-21	-17	-8	13
120	Grassland / forest	Calibration	-18	-15	-11	0	25
		Validation	-26	-23	-19	-10	12
130	Shrubland	Calibration	-16	-16	-12	-1	20
		Validation	-23	-22	-18	-9	10
140	Herbaceous vegetation	Calibration	-22	-19	-13	-8	0
		Validation	-18	-15	-8	-3	1
200	Bareland	Calibration	-20	-17	-8	-3	0
		Validation	-17	-13	-5	1	4
210	Water	Calibration	-3	-1	3	6	9
		Validation	0	6	9	12	15
220	Snow and Ice	Calibration	-25	-21	-8	-1	2
		Validation	-22	-17	-4	3	8



Appendix 5.1: The comparison of the mean monthly discharge predicted for the historical simulations (Chapter 4) and the baseline simulations for each CMIP6 ensemble member (Chapter 5). The hydrological ensemble is aggregated by the median. The dashed line is the line of perfect fit and the red line is the linear regression model.



Appendix 5.2: The comparison of the mean monthly component runoff for the historical simulations (Chapter 4) and the baseline simulations for each CMIP6 ensemble member (Chapter 5) – Part 1. The hydrological ensemble is aggregated by the median. The dashed line is the line of perfect fit and the red line is the linear regression model



Appendix 5.3: The comparison of the mean monthly component runoff for the historical simulations (Chapter 4) and the baseline simulations for each CMIP6 ensemble member (Chapter 5) – Part 2. The hydrological ensemble is aggregated by the median. The dashed line is the line of perfect fit and the red line is the linear regression model.

Appendix 5.4: The predicted 30-year mean annual maximum flows (AMAX) for each CMIP6 member.

Model	Scenario	Year	30-year AMAX [1,000 m^3/s]			Change to baseline [%]		
			Median	$P_{2.5}$	$P_{97.5}$	Median	$P_{2.5}$	$P_{97.5}$
ACCESS-CM2	Baseline	2014	6.2	4.4	7.9	-	-	-
ACCESS-CM2	SSP245	2059	8.4	6.2	10.7	36	39	35
ACCESS-CM2	SSP585	2059	7.9	5.7	10	28	28	27
ACCESS-CM2	SSP245	2099	8.9	6.2	11.5	44	40	45
ACCESS-CM2	SSP585	2099	10.5	7.3	13.7	70	65	72
ACCESS-ESM1-5	Baseline	2014	7.1	4.9	9.5	-	-	-
ACCESS-ESM1-5	SSP245	2059	9.1	6.3	12.5	29	29	31
ACCESS-ESM1-5	SSP585	2059	10.6	7.2	14.7	50	46	54
ACCESS-ESM1-5	SSP245	2099	11.1	7.4	15.5	57	51	62
ACCESS-ESM1-5	SSP585	2099	16.3	11.1	22.1	131	125	132
BCC-CSM2-MR	Baseline	2014	4.9	3.3	6.8	-	-	-
BCC-CSM2-MR	SSP245	2059	5.4	3.6	7.6	10	10	12
BCC-CSM2-MR	SSP585	2059	5.6	3.7	7.9	14	12	15
BCC-CSM2-MR	SSP245	2099	5.9	3.9	8.4	21	18	23
BCC-CSM2-MR	SSP585	2099	7.5	4.7	10.6	52	41	55
EC-Earth3	Baseline	2014	5.9	4	8.1	-	-	-
EC-Earth3	SSP245	2059	8.2	5.6	11.3	39	39	39
EC-Earth3	SSP585	2059	8.7	6	11.7	47	49	44
EC-Earth3	SSP245	2099	9.2	6.3	12.6	55	56	54
EC-Earth3	SSP585	2099	14.3	9.8	19.4	142	143	139
EC-Earth3-Veg	Baseline	2014	5.9	4	8.1	-	-	-
EC-Earth3-Veg	SSP245	2059	7.4	5	10.1	25	24	25
EC-Earth3-Veg	SSP585	2059	8	5.4	11.1	36	34	36
EC-Earth3-Veg	SSP245	2099	10.2	6.7	14.2	72	65	75
EC-Earth3-Veg	SSP585	2099	14.3	9.3	19.6	142	131	141
INM-CM4-8	Baseline	2014	5.4	3.7	7.6	-	-	-
INM-CM4-8	SSP245	2059	6.9	4.8	9.6	27	29	27
INM-CM4-8	SSP585	2059	7	4.7	9.8	29	27	29
INM-CM4-8	SSP245	2099	6.8	4.8	9.3	25	29	23
INM-CM4-8	SSP585	2099	8.5	5.9	11.6	58	60	53
INM-CM5-0	Baseline	2014	5.9	4	8.1	-	-	-
INM-CM5-0	SSP245	2059	7	4.9	9.8	20	21	22
INM-CM5-0	SSP585	2059	7.4	5.1	10.1	27	27	25
INM-CM5-0	SSP245	2099	7.5	5.1	10.1	28	27	25
INM-CM5-0	SSP585	2099	8.9	5.8	12.5	52	44	55
MPI-ESM1-2-HR	Baseline	2014	4.9	3.3	6.5	-	-	-
MPI-ESM1-2-HR	SSP245	2059	4.8	3.3	6.4	-2	-1	-2
MPI-ESM1-2-HR	SSP585	2059	4.7	3.2	6.4	-4	-3	-3
MPI-ESM1-2-HR	SSP245	2099	5.2	3.4	7.1	6	1	9
MPI-ESM1-2-HR	SSP585	2099	6.6	4.1	9.1	35	24	39
MPI-ESM1-2-LR	Baseline	2014	5.4	3.6	7.3	-	-	-
MPI-ESM1-2-LR	SSP245	2059	5.3	3.5	7.2	-2	-3	-2
MPI-ESM1-2-LR	SSP585	2059	5.6	3.7	7.7	5	5	6
MPI-ESM1-2-LR	SSP245	2099	5.5	3.6	7.6	3	1	3
MPI-ESM1-2-LR	SSP585	2099	5.6	3.8	7.6	5	7	3

Continuation of appendix 5.4

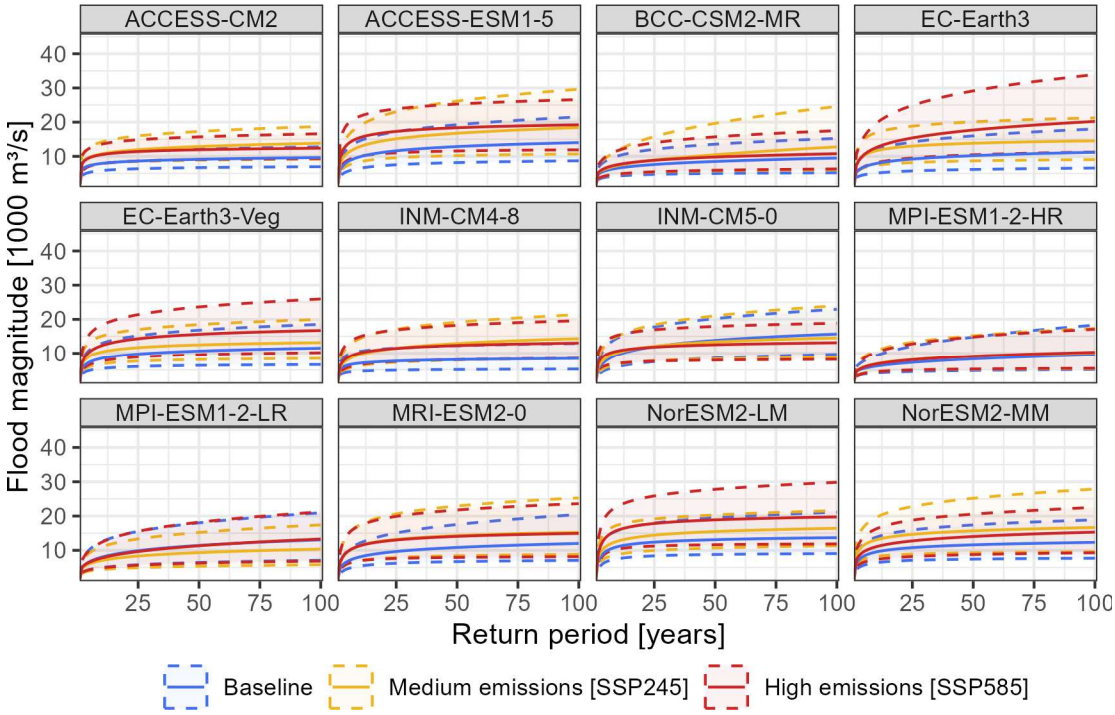
Model	**Scenario**	**Year**	**30-year AMAX [1,000 m^3/s]**			**Change to baseline [%]**		
			Median	$P_{2.5}$	$P_{97.5}$	Median	$P_{2.5}$	$P_{97.5}$
ACCESS-CM2	Baseline	2014	6.2	4.4	7.9	-	-	-
ACCESS-CM2	SSP245	2059	8.4	6.2	10.7	36	39	35
ACCESS-CM2	SSP585	2059	7.9	5.7	10	28	28	27
ACCESS-CM2	SSP245	2099	8.9	6.2	11.5	44	40	45
ACCESS-CM2	SSP585	2099	10.5	7.3	13.7	70	65	72
ACCESS-ESM1-5	Baseline	2014	7.1	4.9	9.5	-	-	-
ACCESS-ESM1-5	SSP245	2059	9.1	6.3	12.5	29	29	31
ACCESS-ESM1-5	SSP585	2059	10.6	7.2	14.7	50	46	54
ACCESS-ESM1-5	SSP245	2099	11.1	7.4	15.5	57	51	62
ACCESS-ESM1-5	SSP585	2099	16.3	11.1	22.1	131	125	132
BCC-CSM2-MR	Baseline	2014	4.9	3.3	6.8	-	-	-
BCC-CSM2-MR	SSP245	2059	5.4	3.6	7.6	10	10	12
BCC-CSM2-MR	SSP585	2059	5.6	3.7	7.9	14	12	15
BCC-CSM2-MR	SSP245	2099	5.9	3.9	8.4	21	18	23
BCC-CSM2-MR	SSP585	2099	7.5	4.7	10.6	52	41	55

Appendix 5.5: The AMAX percentiles for each CMIP6 ensemble member. The AMAX P_{50} is statistically exceeded once every two years and the AMAX P_{100} is the highest simulated flood event in the 40-year record. The hydrological ensemble is aggregated by the median. The periods are 1975 – 2014 (Hist) for the baseline and 2020 – 2059 (NF) and 2060 – 2099 (FF) for the projected scenarios.

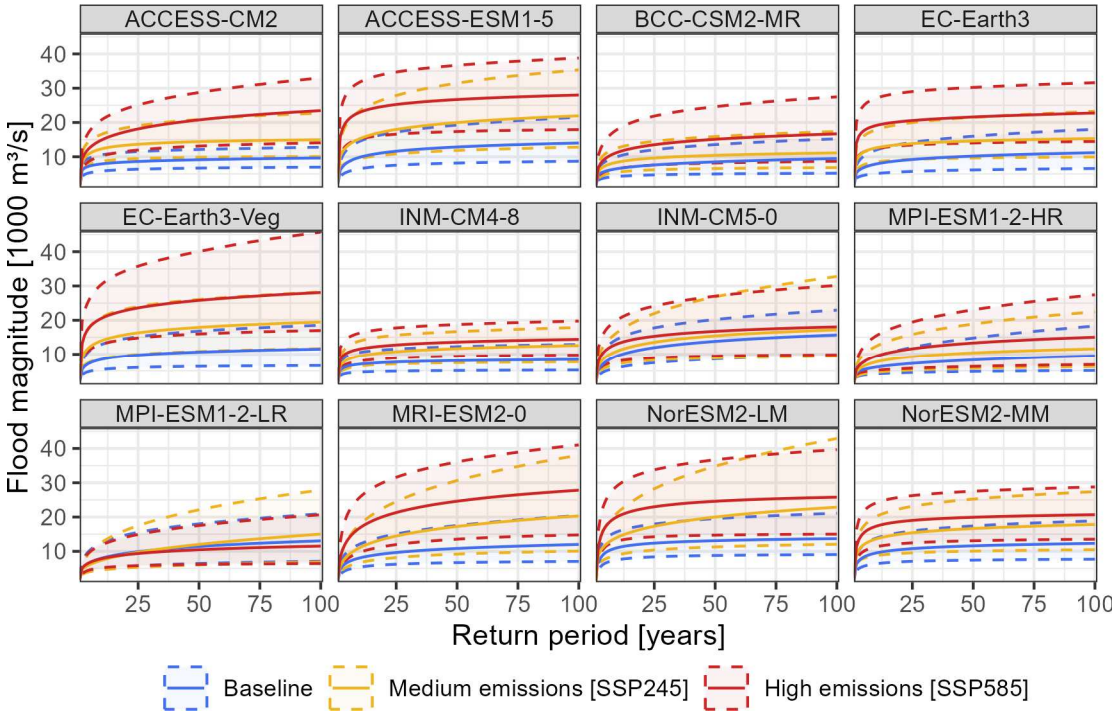
Model	**Scenario**	**Per-iod**	**AMAX Percentile [1,000 m^3/s]**					**Change to baseline [%]**		
			P_{50}	P_{80}	P_{90}	P_{95}	P_{100}	P_{80}	P_{95}	P_{100}
ACCESS-CM2	Baseline	Hist	6	7.3	7.8	8.2	9.5	-	-	-
ACCESS-CM2	SSP245	NF	8.1	9.5	10.7	11.7	13.6	30	43	43
ACCESS-CM2	SSP585	NF	8	9.3	10.7	11.1	11.9	27	35	25
ACCESS-CM2	SSP245	FF	7.9	11.2	11.8	12.6	15.4	53	54	62
ACCESS-CM2	SSP585	FF	9.5	12.1	15.1	17.3	23.4	66	111	146
ACCESS-ESM1-5	Baseline	Hist	5.8	8.7	9.5	10.8	13	-	-	-
ACCESS-ESM1-5	SSP245	NF	8	11	12.2	15	18.3	26	39	41
ACCESS-ESM1-5	SSP585	NF	9.6	13	15.5	15.6	18.7	49	44	44
ACCESS-ESM1-5	SSP245	FF	10.1	12.4	15.3	16.6	22.6	43	54	74
ACCESS-ESM1-5	SSP585	FF	15.6	20.1	21.7	24.3	28.1	131	125	116
BCC-CSM2-MR	Baseline	Hist	4.8	5.6	6.4	7.4	9.8	-	-	-
BCC-CSM2-MR	SSP245	NF	5.2	6.1	6.6	8.2	14.9	9	11	52
BCC-CSM2-MR	SSP585	NF	5.2	6.3	7.5	8.8	10.1	13	19	3
BCC-CSM2-MR	SSP245	FF	5.5	7.5	8.5	8.9	10.9	34	20	11
BCC-CSM2-MR	SSP585	FF	6.8	9.4	10.4	12.4	16.6	68	68	69
EC-Earth3	Baseline	Hist	5.3	6.9	7.6	8.1	11.2	-	-	-
EC-Earth3	SSP245	NF	7.6	10.2	12.3	12.8	13.4	48	58	20
EC-Earth3	SSP585	NF	7.4	9.5	11.6	15.2	20.5	38	88	83
EC-Earth3	SSP245	FF	8.9	11.2	12.2	13.7	14.7	62	69	31
EC-Earth3	SSP585	FF	14.3	17.4	17.9	20.1	23.8	152	148	113
EC-Earth3-Veg	Baseline	Hist	5.4	7.1	8.6	9	11.3	-	-	-
EC-Earth3-Veg	SSP245	NF	6.6	9.6	10.1	10.5	12.9	35	17	14
EC-Earth3-Veg	SSP585	NF	6.9	10.4	11.9	13.3	16.4	47	48	45
EC-Earth3-Veg	SSP245	FF	8.9	11.3	14.1	15.9	17.7	59	77	57
EC-Earth3-Veg	SSP585	FF	13.2	16.6	19.4	23.9	31.5	134	166	179
INM-CM4-8	Baseline	Hist	5.2	6.3	6.9	7.1	8.7	-	-	-
INM-CM4-8	SSP245	NF	5.7	7.5	9.9	11.1	12.4	19	56	43
INM-CM4-8	SSP585	NF	6.2	8.5	9.7	10.9	11.7	35	54	35
INM-CM4-8	SSP245	FF	6.4	8.1	9.3	10.8	11.7	29	52	35
INM-CM4-8	SSP585	FF	8	10.1	11.4	11.8	15.1	60	66	74
INM-CM5-0	Baseline	Hist	4.9	7.4	9.9	11.8	12.1	-	-	-
INM-CM5-0	SSP245	NF	5.4	8.8	9.3	10.1	13.6	19	-14	12
INM-CM5-0	SSP585	NF	6.7	9.1	10.4	10.9	12.6	23	-8	4
INM-CM5-0	SSP245	FF	6	9.3	10.9	11.1	21	26	-6	74
INM-CM5-0	SSP585	FF	8.2	11.1	12.3	13.1	20.4	50	11	69
MPI-ESM1-2-HR	Baseline	Hist	4.5	5.3	6	6.8	11.2	-	-	-
MPI-ESM1-2-HR	SSP245	NF	4.5	5.6	6.8	8.3	9.7	6	22	-13
MPI-ESM1-2-HR	SSP585	NF	4.4	5.5	6	8.5	9.7	4	25	-13
MPI-ESM1-2-HR	SSP245	FF	4.6	6.5	7.4	8.3	12.5	23	22	12
MPI-ESM1-2-HR	SSP585	FF	5.3	8.1	9.9	11.7	13.3	53	72	19

Contiunation of appendix 5.5.

Model	Scenario	Per-iod	AMAX Percentile [1,000 m^3/s]					Change to baseline [%]		
			P_{50}	P_{80}	P_{90}	P_{95}	P_{100}	P_{80}	P_{95}	P_{100}
MPI-ESM1-2-LR	Baseline	Hist	5	6.7	8.1	10.1	10.8	-	-	-
MPI-ESM1-2-LR	SSP245	NF	4.8	6.1	7.1	8.3	9.4	-9	-18	-13
MPI-ESM1-2-LR	SSP585	NF	5	6.4	7.3	9.2	12.9	-5	-9	19
MPI-ESM1-2-LR	SSP245	FF	4.8	5.7	7.4	10.3	13.6	-15	2	26
MPI-ESM1-2-LR	SSP585	FF	4.9	6.7	7.8	8.6	11.4	0	-15	6
MRI-ESM2-0	Baseline	Hist	4.7	7	7.5	9	12.3	-	-	-
MRI-ESM2-0	SSP245	NF	6.1	9.2	11.2	12	14.8	31	33	20
MRI-ESM2-0	SSP585	NF	6.5	9.7	11.4	11.7	13.5	39	30	10
MRI-ESM2-0	SSP245	FF	5.5	8.4	11.1	13.5	21.7	20	50	76
MRI-ESM2-0	SSP585	FF	7.8	14.6	17	18.8	25.2	109	109	105
NorESM2-LM	Baseline	Hist	6.7	9.5	11.2	11.4	14	-	-	-
NorESM2-LM	SSP245	NF	7.8	10.5	13.1	14.1	15.2	11	24	9
NorESM2-LM	SSP585	NF	9.3	13.2	16.3	16.9	19.7	39	48	41
NorESM2-LM	SSP245	FF	9	11.3	13.9	15	28.5	19	32	104
NorESM2-LM	SSP585	FF	11.2	18.2	20.1	22.6	24.8	92	98	77
NorESM2-MM	Baseline	Hist	6.2	8.1	9.6	10.3	11.7	-	-	-
NorESM2-MM	SSP245	NF	8.8	10.9	12.7	14.7	18.3	35	43	56
NorESM2-MM	SSP585	NF	7	8.9	10.7	12.8	14.5	10	24	24
NorESM2-MM	SSP245	FF	9.1	11.6	13.9	14.9	18.5	43	45	58
NorESM2-MM	SSP585	FF	11.1	15	16.4	18.9	20.3	85	84	74



Appendix 5.6: The flood frequency curves of the CMIP6 ensemble members for the near future (2020 – 2059). The solid line indicates the median and the dashed lines indicate the 2.5th and 97.5th percentiles of the flood frequency analysis. The flood frequency curves of the baseline (1975 – 2014) are provided for reference.



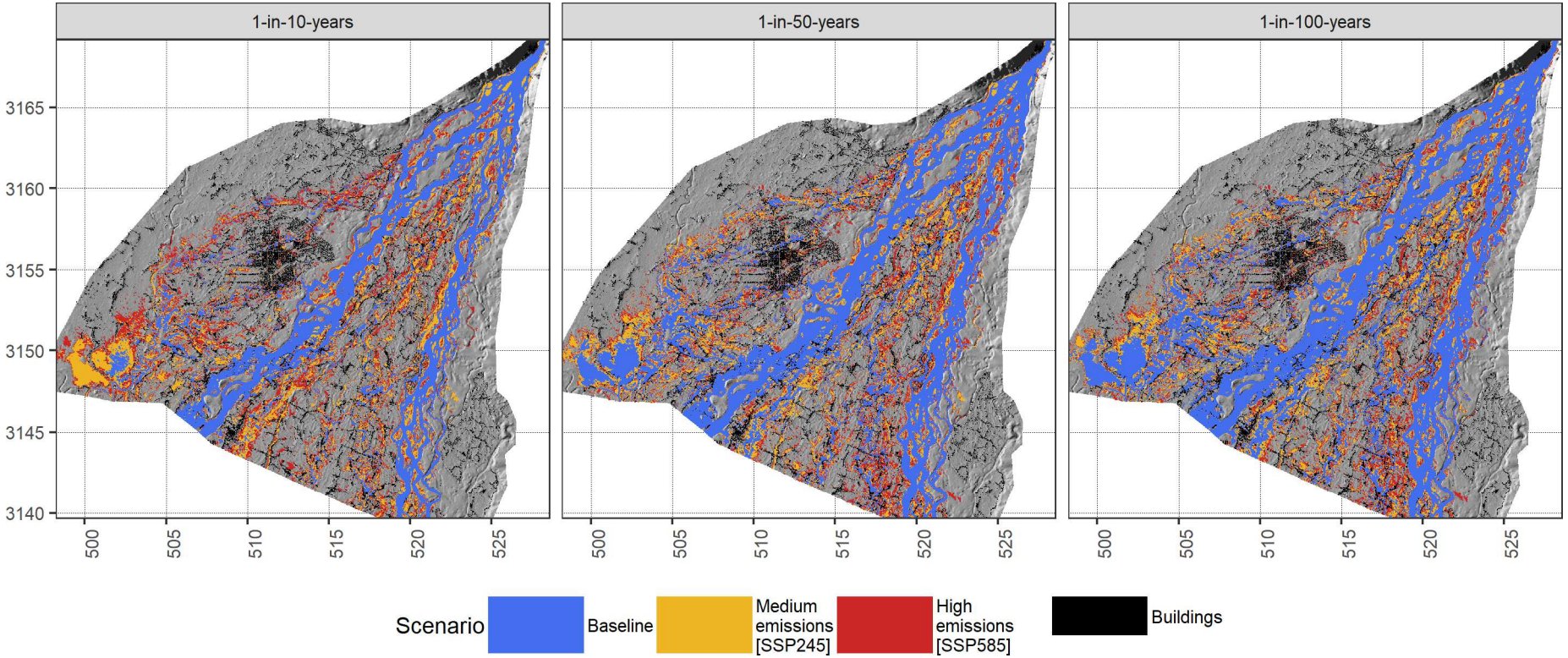
Appendix 5.7: The flood frequency curves of the CMIP6 ensemble members for the far future (2060 – 2099). The solid line indicates the median and the dashed lines indicate the 2.5th and 97.5th percentiles of the flood frequency analysis. The flood frequency curves of the baseline (1975 – 2014) are provided for reference.

Appendix 5.8: The 1-in-100-years flood magnitudes predicted for each CMIP6 ensemble member.

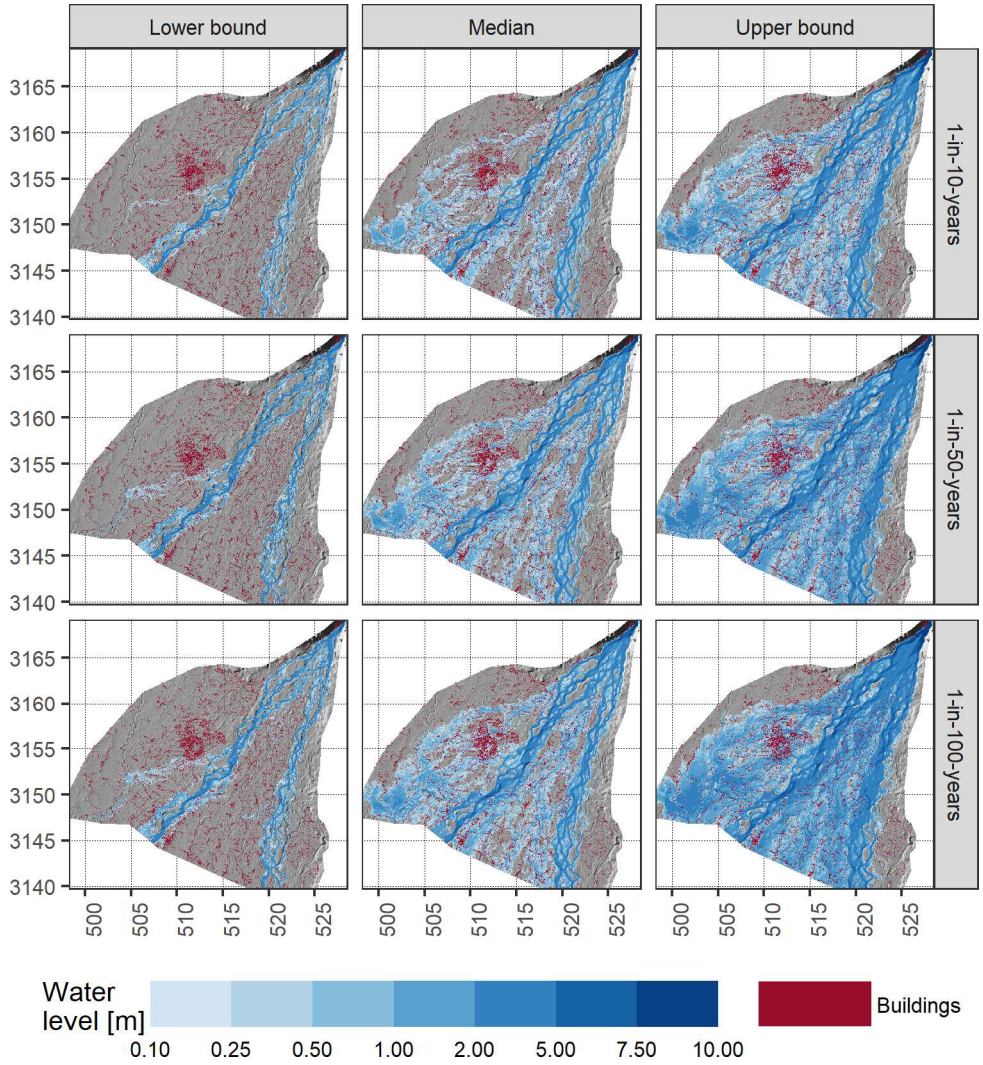
Model	Scenario	Year	Flood magnitude [1,000 m^3/s]			Change to baseline [%]		
			Median	$P_{2.5}$	$P_{97.5}$	Median	$P_{2.5}$	$P_{97.5}$
ACCESS-CM2	Baseline	2014	9.7	7	12.8	-	-	-
ACCESS-CM2	SSP245	2059	15	10.1	22.7	55	44	78
ACCESS-CM2	SSP585	2059	13.8	9.2	18.7	43	32	46
ACCESS-CM2	SSP245	2099	23.5	14.1	33	143	102	158
ACCESS-CM2	SSP585	2099	12.4	9.4	16.6	28	34	30
ACCESS-ESM1-5	Baseline	2014	14	8.7	21.5	-	-	-
ACCESS-ESM1-5	SSP245	2059	21.9	12.8	35.3	57	47	64
ACCESS-ESM1-5	SSP585	2059	18.4	10.7	29.6	31	23	38
ACCESS-ESM1-5	SSP245	2099	28	17.9	38.8	100	106	80
ACCESS-ESM1-5	SSP585	2099	19.2	11.9	26.6	37	37	24
BCC-CSM2-MR	Baseline	2014	9.5	5.2	15.2	-	-	-
BCC-CSM2-MR	SSP245	2059	11.1	6.8	17.4	17	32	14
BCC-CSM2-MR	SSP585	2059	12.8	6.1	24.6	34	17	61
BCC-CSM2-MR	SSP245	2099	16.7	8.7	27.5	76	67	80
BCC-CSM2-MR	SSP585	2099	10.8	6.3	17.5	13	22	15
EC-Earth3	Baseline	2014	11.2	6.6	18	-	-	-
EC-Earth3	SSP245	2059	15.3	9.9	23.3	37	51	29
EC-Earth3	SSP585	2059	14.5	9.1	21.2	30	38	18
EC-Earth3	SSP245	2099	22.8	14.5	31.6	103	120	76
EC-Earth3	SSP585	2099	20.2	11.3	33.9	81	71	88
EC-Earth3-Veg	Baseline	2014	11.5	6.6	18.5	-	-	-
EC-Earth3-Veg	SSP245	2059	19.5	11.7	28.2	69	78	52
EC-Earth3-Veg	SSP585	2059	13.2	8.4	20	14	28	8
EC-Earth3-Veg	SSP245	2099	28.1	17	45.7	144	159	147
EC-Earth3-Veg	SSP585	2099	16.7	10.2	26	45	55	40
INM-CM4-8	Baseline	2014	8.4	5.2	12.9	-	-	-
INM-CM4-8	SSP245	2059	12.6	7.7	17.9	50	46	39
INM-CM4-8	SSP585	2059	14.3	8.4	21.4	70	61	66
INM-CM4-8	SSP245	2099	14.4	9.7	19.8	71	86	53
INM-CM4-8	SSP585	2099	13	8.5	19.6	55	62	52
INM-CM5-0	Baseline	2014	15.7	9.7	23	-	-	-
INM-CM5-0	SSP245	2059	17.2	9.3	32.8	10	-3	43
INM-CM5-0	SSP585	2059	14.5	8.6	24	-7	-11	5
INM-CM5-0	SSP245	2099	18.1	9.8	30.2	15	2	31
INM-CM5-0	SSP585	2099	13.1	8.1	18.9	-16	-16	-18
MPI-ESM1-2-HR	Baseline	2014	9.5	5.1	18.3	-	-	-
MPI-ESM1-2-HR	SSP245	2059	11.6	6.3	22.5	22	23	23
MPI-ESM1-2-HR	SSP585	2059	10.1	5.3	17.3	6	4	-5
MPI-ESM1-2-HR	SSP245	2099	15	6.9	27.4	58	34	50
MPI-ESM1-2-HR	SSP585	2099	10.3	5.5	17.1	8	7	-7
MPI-ESM1-2-LR	Baseline	2014	13.1	7.1	20.9	-	-	-
MPI-ESM1-2-LR	SSP245	2059	15	6.9	27.9	15	-3	33
MPI-ESM1-2-LR	SSP585	2059	10.3	5.8	17.4	-21	-18	-17
MPI-ESM1-2-LR	SSP245	2099	11.5	6.4	20.7	-12	-10	-1
MPI-ESM1-2-LR	SSP585	2099	13.3	6.9	21.2	2	-3	1

Contiunation of appendix 5.8

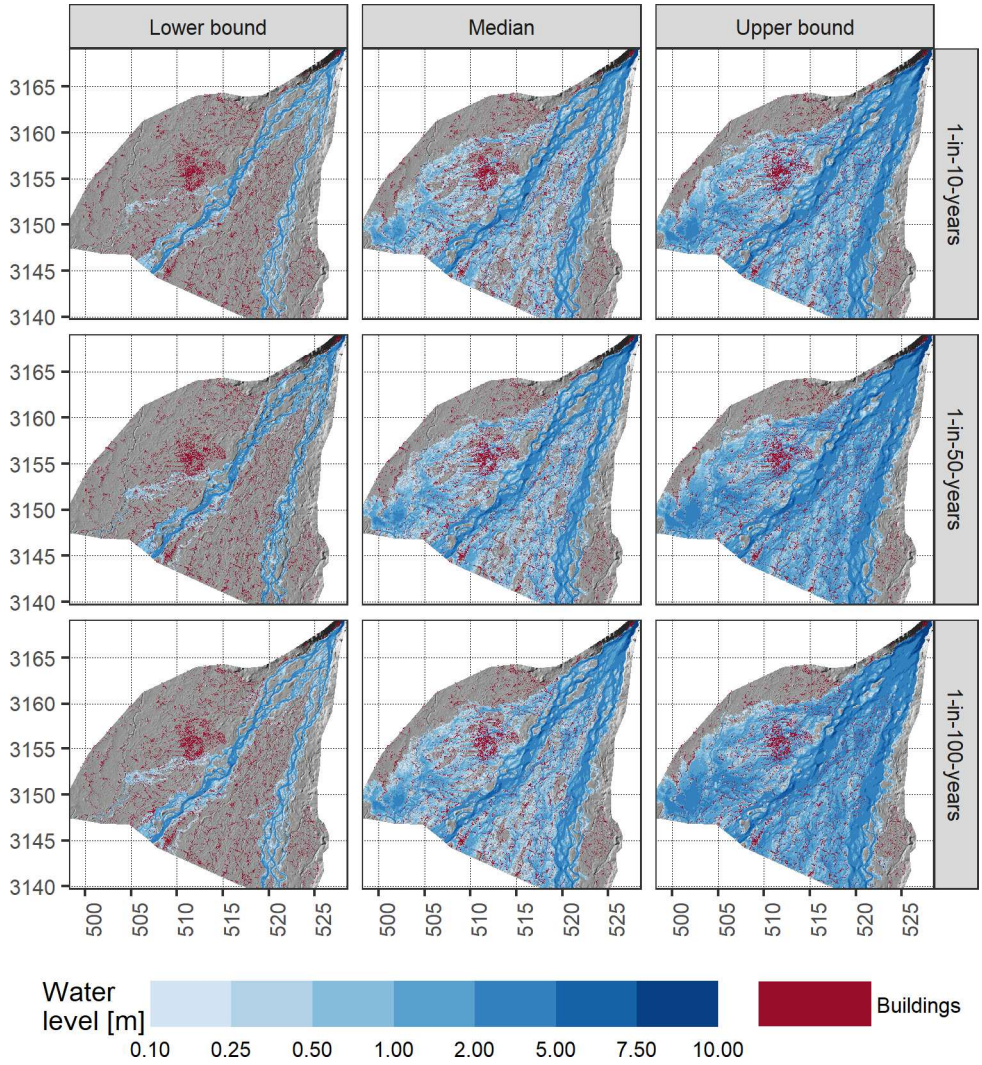
Model	**Scenario**	**Year**	**Flood magnitude [1,000 m^3/s]**			**Change to baseline [%]**		
			Median	$P_{2.5}$	$P_{97.5}$	Median	$P_{2.5}$	$P_{97.5}$
ACCESS-CM2	Baseline	2014	12	7.1	20.4	-	-	-
ACCESS-CM2	SSP245	2059	20.3	10.1	38	69	43	86
ACCESS-CM2	SSP585	2059	15.3	8.8	25.3	27	25	24
ACCESS-CM2	SSP245	2099	27.8	14.8	41	132	109	101
ACCESS-CM2	SSP585	2099	14.9	8.1	23.6	24	15	16
ACCESS-ESM1-5	Baseline	2014	13.7	9.1	21.1	-	-	-
ACCESS-ESM1-5	SSP245	2059	22.9	12.1	42.9	67	33	103
ACCESS-ESM1-5	SSP585	2059	16.4	11.3	21.6	20	25	2
ACCESS-ESM1-5	SSP245	2099	25.8	15	39.7	88	66	88
ACCESS-ESM1-5	SSP585	2099	19.7	11.9	29.9	44	31	41
BCC-CSM2-MR	Baseline	2014	12.3	7.7	18.9	-	-	-
BCC-CSM2-MR	SSP245	2059	17.8	10.5	27.4	45	36	45
BCC-CSM2-MR	SSP585	2059	16.7	9.6	27.9	35	24	48
BCC-CSM2-MR	SSP245	2099	20.7	13.5	28.8	68	76	52
BCC-CSM2-MR	SSP585	2099	15.3	9.3	22.5	24	20	19



Appendix 7.1: The inundation extent (inundation depth ≥ 0.5 m) for the median predictions (CM_{median}) of the three CMIP6 scenarios for the 1-in-10-years, 1-in-50-years and 1-in-100-years flood events. The buildings are obtained from OSM (2023). The coordinate system is UTM Zone 44 N in km.



Appendix 8.2: The maximum water level (≥0.1 m) predicted for the medium-emission scenario SSP245. The left column shows the predictions of the lower boundary of the prediction interval (CM_{lower}), the centre column the ones of the median predictions (CM_{median}), and the right column the ones of the upper boundary of the prediction interval (CM_{upper}). The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km. Note that the scale is non-linear.



Appendix 7.3: The maximum water level (≥0.1 m) predicted for the high-emission scenario SSP585. The left column shows the predictions of the lower boundary of the prediction interval (CM_{lower}), the centre column the ones of the median predictions (CM_{median}), and the right column the ones of the upper boundary of the prediction interval (CM_{upper}). The buildings are obtained from OSM (2023) and the coordinate system is UTM Zone 44 N in km. Note that the scale is non-linear.

References

Adhikari, T.R., Baniya, B., Tang, Q., Talchabhadel, R., Gouli, M.R., Budhathoki, B.R. and Awasthi, R.P. (2023) 'Evaluation of post extreme floods in high mountain region: A case study of the Melamchi flood 2021 at the Koshi River Basin in Nepal', *Natural Hazards Research* [Preprint]. Available at: https://doi.org/10.1016/j.nhres.2023.07.001.

ADPC (2010) 'Nepal Hazard Risk Assessment'. Asian Disaster Preparedness Center. Available at: https://www.adpc.net/igo/category/ID276/doc/2013-b27Iym-ADPC-NHRA_Report.pdf (Accessed: 9 April 2023).

Ahad, N.A. and Yahaya, S.S.S. (2014) 'Sensitivity analysis of Welch's t-test', *AIP Conference Proceedings*, 1605(1), pp. 888–893. Available at: https://doi.org/10.1063/1.4887707.

Ahmad, F., Uddin, M.M., Goparaju, L., Dhyani, S.K., Oli, B.N. and Rizvi, J. (2021) 'Tree suitability modeling and mapping in Nepal: a geospatial approach to scaling agroforestry', *Modeling Earth Systems and Environment*, 7(1), pp. 169–179. Available at: https://doi.org/10.1007/s40808-020-00922-7.

Ahmad, M.I., Sinclair, C.D. and Spurr, B.D. (1988) 'Assessment of flood frequency models using empirical distribution function statistics', *Water Resources Research*, 24(8), pp. 1323–1328. Available at: https://doi.org/10.1029/WR024i008p01323.

Ali, S., Eum, H.-I., Cho, J., Dan, L., Khan, F., Dairaku, K., Shrestha, M.L., Hwang, S., Nasim, W. and Khan, I.A. (2019) 'Assessment of climate extremes in future projections downscaled by multiple statistical downscaling methods over Pakistan', *Atmospheric Research*, 222, pp. 114–133.

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. (1998) 'Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56', *Fao, Rome*, 300(9), p. D05109.

Al-Qurashi, A., McIntyre, N., Wheater, H. and Unkrich, C. (2008) 'Application of the Kineros2 rainfall–runoff model to an arid catchment in Oman', *Journal of Hydrology*, 355(1–4), pp. 91–105.

Ancey, C. (2020a) 'Bedload transport: a walk between randomness and determinism. Part 1. The state of the art', *Journal of Hydraulic Research*, 58(1), pp. 1–17. Available at: https://doi.org/10.1080/00221686.2019.1702594.

Ancey, C. (2020b) 'Bedload transport: a walk between randomness and determinism. Part 2. Challenges and prospects', *Journal of Hydraulic Research*, 58(1), pp. 18–33. Available at: https://doi.org/10.1080/00221686.2019.1702595.

Ancey, C. and Pascal, I. (2020) 'Estimating Mean Bedload Transport Rates and Their Uncertainty', *Journal of Geophysical Research: Earth Surface*, 125(7), p. e2020JF005534. Available at: https://doi.org/10.1029/2020JF005534.

Andermann, C., Crave, A., Gloaguen, R., Davy, P. and Bonnet, S. (2012) 'Connecting source and transport: Suspended sediments in the Nepal Himalayas', *Earth and Planetary Science Letters*, 351, pp. 158–170.

Andermann, C., Longuevergne, L., Bonnet, S., Crave, A., Davy, P. and Gloaguen, R. (2012) 'Impact of transient groundwater storage on the discharge of Himalayan rivers', *Nature Geoscience*, 5(2), pp. 127–132. Available at: https://doi.org/10.1038/ngeo1356.

Anderson, T.W. and Darling, D.A. (1952) 'Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes', *The Annals of Mathematical Statistics*, 23(2), pp. 193–212.

Apel, H., Merz, B. and Thieken, A.H. (2008) 'Quantification of uncertainties in flood risk assessments', *International Journal of River Basin Management*, 6(2), pp. 149–162. Available at: https://doi.org/10.1080/15715124.2008.9635344.

Arias-Hidalgo, M., Bhattacharya, B., Mynett, A. and Van Griensven, A. (2013) 'Experiences in using the TMPA-3B42R satellite data to complement rain gauge measurements in the Ecuadorian coastal foothills', *Hydrology and Earth System Sciences*, 17(7), pp. 2905–2915.

Arino, O., Ramos Perez, J.J., Kalogirou, V., Bontemps, S., Defourny, P. and Van Bogaert, E. (2012) 'Global Land Cover Map for 2009 (GlobCover 2009)', *© European Space Agency (ESA) & Université catholique de Louvain (UCL)*. PANGAEA. Available at: https://doi.org/10.1594/PANGAEA.787668.

Arnáez, J., Lana-Renault, N., Lasanta, T., Ruiz-Flaño, P. and Castroviejo, J. (2015) 'Effects of farming terraces on hydrological and geomorphological processes. A review', *CATENA*, 128, pp. 122–134. Available at: https://doi.org/10.1016/j.catena.2015.01.021.

Arnold, J.G., Engel, B.A. and Srinivasan, R. (1993) 'A continuous time, grid cell watershed model', *Application of advanced information technologies for management of natural resources*, pp. 17–19.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R. and Prat, O.P. (2015) 'PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies', *Bulletin of the American Meteorological Society*, 96(1), pp. 69–83.

Bajracharya, A.R., Bajracharya, S.R., Shrestha, A.B. and Maharjan, S.B. (2018) 'Climate change impact assessment on the hydrological regime of the Kaligandaki Basin, Nepal', *Science of The Total Environment*, 625, pp. 837–848. Available at: https://doi.org/10.1016/j.scitotenv.2017.12.332.

Ballard, C., Bulygina, N., Cluckie, I., Dangerfield, S., Ewen, J., Frogbrook, Z., Geris, J., Henshaw, A., Jackson, B., Marshall, M., McIntyre, N., Pagella, T., Park, J.-S., O'Connell, E., O'Donnell, G., Reynolds, B., Sinclair, F., Solloway, I., Thorne, C. and Wheater, H. (2013) *Land use management effects on flood flows and sediments - guidance on prediction*. London: CIRIA.

Bates, P.D. and De Roo, A.P.J. (2000) 'A simple raster-based model for flood inundation simulation', *Journal of Hydrology*, 236(1), pp. 54–77. Available at: https://doi.org/10.1016/S0022-1694(00)00278-X.

Bates, P.D., Horritt, M.S. and Fewtrell, T.J. (2010) 'A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling', *Journal of Hydrology*, 387(1), pp. 33–45. Available at: https://doi.org/10.1016/j.jhydrol.2010.03.027.

Bergström, S. (1976) *Development and Application of a Conceptual Runoff Model for Scandinavian Catchments*. Department of Water Resources Engineering, Lund Institute of Technology, University of Lund (Bulletin : Series A - Dept. of Water Resources Engineering, Lund Institute of Technology). Available at: https://books.google.co.uk/books?id=vRyeQAAACAAJ.

Berthet, L., Andréassian, V., Perrin, C. and Javelle, P. (2009) 'How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments', *Hydrology and Earth System Sciences*, 13(6), pp. 819–831. Available at: https://doi.org/10.5194/hess-13-819-2009.

Bérubé, S., Brissette, F. and Arsenault, R. (2022) 'Optimal hydrological model calibration strategy for climate change impact studies', *Journal of Hydrologic Engineering*, 27(3), p. 04021053.

Beven, K. and Binley, A. (1992) 'The future of distributed models: Model calibration and uncertainty prediction', *Hydrological Processes*, 6(3), pp. 279–298. Available at: https://doi.org/10.1002/hyp.3360060305.

Beven, K. and Binley, A. (2014) 'GLUE: 20 years on', *Hydrological Processes*, 28(24), pp. 5897–5918. Available at: https://doi.org/10.1002/hyp.10082.

Beven, K. and Freer, J. (2001) 'Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology', *Journal of Hydrology*, 249(1), pp. 11–29. Available at: https://doi.org/10.1016/S0022-1694(01)00421-8.

Beven, K.J. (2012) *Rainfall - Runoff Modelling: The Primer*. Second Edition. Wiley-Blackwell.

Bhandari, D., Khadka, P., Pokharel, L.N., Kumal, B., Ghimire, G. and Uprety, M. (2018) 'Nepal flood 2017: Wake up call for effective preparedness and response'.

Bharati, L., Jayakody, P., Gurung, P. and Bhattarai, U. (2014) 'The Projected Impact of Climate Change on Water Availability and Development in the Koshi Basin, Nepal', *Mountain Research and Development*, 34(2), pp. 118–130. Available at: https://doi.org/10.1659/MRD-JOURNAL-D-13-00096.1.

Bhardwaj, A., Ziegler, A.D., Wasson, R.J. and Chow, W.T. (2017) 'Accuracy of rainfall estimates at high altitude in the Garhwal Himalaya (India): A comparison of secondary precipitation products and station rainfall measurements', *Atmospheric Research*, 188, pp. 30–38.

Bhattacharya, T., Khare, D. and Arora, M. (2020) 'Evaluation of reanalysis and global meteorological products in Beas river basin of North-Western Himalaya', *Environmental Systems Research*, 9(1), pp. 1–29.

Bhattarai, S., Zhou, Y., Shakya, N.M. and Zhao, C. (2018) 'Hydrological modelling and climate change impact assessment using HBV light model: a case study of Narayani River Basin, Nepal', *Nature Environment and Pollution Technology*, 17(3), pp. 691–702.

Bolch, T., Shea, J.M., Liu, S., Azam, F.M., Gao, Y., Gruber, S., Immerzeel, W.W., Kulkarni, A., Li, H., Tahir, A.A., Zhang, G. and Zhang, Y. (2019) 'Status and Change of the Cryosphere in the Extended Hindu Kush Himalaya Region', in P. Wester, A.

Mishra, A. Mukherji, and A.B. Shrestha (eds) *The Hindu Kush Himalaya Assessment: Mountains, Climate Change, Sustainability and People*. Cham: Springer International Publishing, pp. 209–255. Available at: https://doi.org/10.1007/978-3-319-92288-1_7.

Bontemps, S., Defourny, P., Van Bogaert, E., Team, E., Arino, O., Kalogirou, V. and Perez, J.R. (2011) 'GLOBCOVER 2009 Products Description and Validation Report'.

Bookhagen, B. (2012) 'Himalayan groundwater', *Nature Geoscience*, 5(2), pp. 97–98. Available at: https://doi.org/10.1038/ngeo1366.

Bookhagen, B. and Burbank, D.W. (2010) 'Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge', *Journal of Geophysical Research: Earth Surface*, 115(F3). Available at: https://doi.org/10.1029/2009JF001426.

Bracken, L.J., Turnbull, L., Wainwright, J. and Bogaart, P. (2015) 'Sediment connectivity: a framework for understanding sediment transfer at multiple scales', *Earth Surface Processes and Landforms*, 40(2), pp. 177–188. Available at: https://doi.org/10.1002/esp.3635.

Burn, D.H. (2003) 'The use of resampling for estimating confidence intervals for single site and pooled frequency analysis / Utilisation d'un rééchantillonnage pour l'estimation des intervalles de confiance lors d'analyses fréquentielles mono et multi-site', *Hydrological Sciences Journal*, 48(1), pp. 25–38. Available at: https://doi.org/10.1623/hysj.48.1.25.43485.

Burt, T.P. and Allison, R.J. (2010) 'Sediment Cascades in the Environment: An Integrated Approach', in *Sediment Cascades*, pp. 1–15. Available at: https://doi.org/10.1002/9780470682876.ch1.

Caretta, A.M.M.A., Arfanuzzaman, R.B.M., Morgan, S.M.R. and Kumar, M. (2022) 'Water. In: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change'.

CARITAS India (2013) 'Rapid Assessment Report of Landslide and Flash Flood in Uttarakhand'. Available at: https://reliefweb.int/report/india/rapid-assessment-report-landslide-and-flash-flood-uttarakhand (Accessed: 4 September 2023).

Carvalho, D., Rafael, S., Monteiro, A., Rodrigues, V., Lopes, M. and Rocha, A. (2022) 'How well have CMIP3, CMIP5 and CMIP6 future climate projections portrayed the recently observed warming', *Scientific Reports*, 12(1), p. 11983.

CBS (2017) 'Statistical Year book of Nepal - 2017'. Central Bureau of Statistics, Government of Nepal.

CBS (2022) *Preliminary Data of National Population and Housing Census 2021*. Kathmandu: Central Bureau of Statistics. Available at: https://censusnepal.cbs.gov.np/Home/Details?tpid=5&dcid=a82d0762-6f49-4752-a02d-19fdaddc55ed&tfsid=17 (Accessed: 24 October 2023).

Chandel, V.S. and Ghosh, S. (2021) 'Components of Himalayan River Flows in a Changing Climate', *Water Resources Research*, 57(2), p. e2020WR027589. Available at: https://doi.org/10.1029/2020WR027589.

Chen, C.-M., Hollingsworth, J., Clark, M., Zekkos, D., Chamlagain, D., Bista, S., Siwakoti, A. and West, J. (2023) 'The 2021 Melamchi Flood: A massive erosional cascade in the Himalayan Mountains of central Nepal', *PREPRINT (Version 1)* [Preprint]. Available at: https://doi.org/10.21203/rs.3.rs-2766739/v1.

Chhetri, R., Pandey, V.P., Talchabhadel, R. and Thapa, B.R. (2021) 'How do CMIP6 models project changes in precipitation extremes over seasons and locations across the mid hills of Nepal?', *Theoretical and Applied Climatology*, 145(3), pp. 1127–1144. Available at: https://doi.org/10.1007/s00704-021-03698-7.

Chow, V. (1959) *Open-channel Hydraulics*. Blackburn Press (McGraw-Hill civil engineering series).

C-L Manual (2023) 'The CAESAR-LISFLOOD Wiki'. Available at: https://sourceforge.net/p/caesar-lisflood/wiki/Home/ (Accessed: 26 September 2023).

Clark, M.P., Kavetski, D. and Fenicia, F. (2011) 'Pursuing the method of multiple working hypotheses for hydrological modeling', *Water Resources Research*, 47(9). Available at: https://doi.org/10.1029/2010WR009827.

Clark, M.P., Kavetski, D. and Fenicia, F. (2012) 'Reply to comment by K. Beven et al. on "Pursuing the method of multiple working hypotheses for hydrological modeling"', *Water Resources Research*, 48(11). Available at: https://doi.org/10.1029/2012WR012547.

Clarke, R.T. (1999) 'Uncertainty in the estimation of mean annual flood due to rating-curve indefinition', *Journal of Hydrology*, 222(1), pp. 185–190. Available at: https://doi.org/10.1016/S0022-1694(99)00097-9.

Collier, E. and Immerzeel, W.W. (2015) 'High-resolution modeling of atmospheric dynamics in the Nepalese Himalaya', *Journal of Geophysical Research: Atmospheres*, 120(19), pp. 9882–9896. Available at: https://doi.org/10.1002/2015JD023266.

Cook, K.L., Andermann, C., Gimbert, F., Adhikari, B.R. and Hovius, N. (2018) 'Glacial lake outburst floods as drivers of fluvial erosion in the Himalaya', *Science*, 362(6410), pp. 53–57. Available at: https://doi.org/10.1126/science.aat4981.

Coulthard, T.J., Macklin, M.G. and Kirkby, M.J. (2002) 'A cellular model of Holocene upland river basin and alluvial fan evolution', *Earth Surface Processes and Landforms*, 27(3), pp. 269–288. Available at: https://doi.org/10.1002/esp.318.

Coulthard, T.J., Neal, J.C., Bates, P.D., Ramirez, J., de Almeida, G.A.M. and Hancock, G.R. (2013) 'Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: implications for modelling landscape evolution', *Earth Surface Processes and Landforms*, 38(15), pp. 1897–1906. Available at: https://doi.org/10.1002/esp.3478.

Coxon, G., Freer, J., Westerberg, I.K., Wagener, T., Woods, R. and Smith, P.J. (2015) 'A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations', *Water Resources Research*, 51(7), pp. 5531–5546. Available at: https://doi.org/10.1002/2014WR016532.

Crawford, H.H. and Linsley, R.K. (1966) *Digital Simulation in Hydrology: Stanford Watershed Model IV*. Technical Report 39. Stanford: Department of Civil and Environmental Engineering. Stanford University.

Crozier, M.J. (2010) 'Deciphering the effect of climate change on landslide activity: A review', *Geomorphology*, 124(3–4), pp. 260–267.

Dahal, P., Shrestha, M.L., Panthi, J. and Pradhananga, D. (2020) 'Modeling the future impacts of climate change on water availability in the Karnali River Basin of Nepal Himalaya', *Environmental Research*, 185, p. 109430. Available at: https://doi.org/10.1016/j.envres.2020.109430.

Dalrymple, T. (1960) *Flood-frequency analyses, Manual of Hydrology: Part 3*. Report 1543A. Available at: https://doi.org/10.3133/wsp1543A.

Dankers, R., Arnell, N.W., Clark, D.B., Falloon, P.D., Fekete, B.M., Gosling, S.N., Heinke, J., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D. (2014) 'First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble', *Proceedings of the National Academy of Sciences*, 111(9), pp. 3257–3261. Available at: https://doi.org/10.1073/pnas.1302078110.

De Boer, F. (2016) *HiHydroSoil: A High Resolution Soil Map of Hydraulic Properties. Version 1.2*. Report FutureWater:134. Wageningen: FutureWater.

De Roo, A.P.J., Wesseling, C.G. and Van Deursen, W.P.A. (2000) 'Physically based river basin modelling within a GIS: the LISFLOOD model', *Hydrological Processes*, 14(11–12), pp. 1981–1992. Available at: https://doi.org/10.1002/1099-1085(20000815/30)14:11/12<1981::AID-HYP49>3.0.CO;2-F.

Deltares (2023) *Delft3D-FLOW. Simulation of multi-dimensional hydrodynamic flows and transport phenomena, including sediments. User Manual. Version 4.5*. Delft: Deltares.

Devi, G.K., Ganasri, B.P. and Dwarakish, G.S. (2015) 'A Review on Hydrological Models', *INTERNATIONAL CONFERENCE ON WATER RESOURCES, COASTAL AND OCEAN ENGINEERING (ICWRCOE'15)*, 4, pp. 1001–1007. Available at: https://doi.org/10.1016/j.aqpro.2015.02.126.

DFRS (2014) *Terai Forests of Nepal*. Kathmandu: Department of Forest Research and Survey, Government of Nepal (Forest Resource Assessment Nepal Project, 1).

Dhakal, S. (2013) 'Flood Hazard in Nepal and New Approach of Risk Reduction', *International Journal of Landslide and Environment*, 1(1), pp. 13–14.

Dhami, B., Himanshu, S.K., Pandey, A. and Gautam, A.K. (2018) 'Evaluation of the SWAT model for water balance study of a mountainous snowfed river basin of Nepal', *Environmental Earth Sciences*, 77(1), p. 21. Available at: https://doi.org/10.1007/s12665-017-7210-8.

DHM (2018) 'Standard Operating Procedure for Flood Early Warning System in Nepal'. Department of Hydrology and Meteorology, Gouvernment of Nepal.

Di Baldassarre, G., Schumann, G., Bates, P.D., Freer, J.E. and Beven, K.J. (2010) 'Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches', *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, 55(3), pp. 364–376.

Dijkshoorn, K. and Huting, J. (2009) *Soil and Terrain database for Nepal*. 2009/01. Wageningen: ISRIC - World Soil Information, p. 29. Available at: http://www.isric.org.

Dilley, M., Chen, R.S., Deichmann, U., Lerner-Lam, A.L. and Arnold, M. (2005) *Natural Disaster Hotspots: a global risk analysis*. The World Bank (Disaster Risk Management). Available at: https://doi.org/10.1596/0-8213-5930-4.

Dingle, E.H. (2018) *River dynamics in the Himalayan foreland basin*. University of Edinburgh. Available at: http://hdl.handle.net/1842/31285 (Accessed: 11 September 2023).

Dingle, E.H., Attal, M. and Sinclair, H.D. (2017) 'Abrasion-set limits on Himalayan gravel flux', *Nature*, 544(7651), pp. 471–474. Available at: https://doi.org/10.1038/nature22039.

Dingle, Elizabeth H, Creed, M., Sinclair, H., Gautam, D., Gourmelen, N., Borthwick, A. and Attal, M. (2020a) 'Dynamic flood topographies in the Terai region of Nepal', *Earth Surface Processes and Landforms*, 45(13), pp. 3092–3102.

Dingle, E.H., Sinclair, H.D., Attal, M., Milodowski, D.T. and Singh, V. (2016) 'Subsidence control on river morphology and grain size in the Ganga Plain', *American Journal of Science*, 316(8), pp. 778–812.

Dingle, Elizabeth H, Sinclair, H.D., Venditti, J.G., Attal, M., Kinnaird, T.C., Creed, M., Quick, L., Nittrouer, J.A. and Gautam, D. (2020b) 'Sediment dynamics across gravel-sand transitions: Implications for river stability and floodplain recycling', *Geology*, 48(5), pp. 468–472.

Dixit, A. (2009) 'Kosi Embankment Breach in Nepal: Need for a Paradigm Shift in Responding to Floods', *Economic and Political Weekly*, 44(6), pp. 70–78.

Dixit, A., Upadhya, M., Pokhrel, A., Dixit, K.M., Rai, D.R. and Devkota, M. (2007) 'Flood Disaster Impacts and Responses in Nepal Tarai's Marginalised Basins', in *Working with the Winds of Change. Towards Strategies for Responding to the Risks Associated with Climate Change and other Hazards*. 2nd edn. Kathmandu: ProVention Consortium, Institute for Social and Environmental Transition-International and Institute for Social and Environmental Transition-Nepa, pp. 119–157. Available at: http://www.proventionconsortium.org/themes/default/pdfs/winds_of_change.pdf (Accessed: 4 September 2023).

Douville, H., Raghavan, K., Renwick, J., Allan, R.P., Arias, P.A., Barlow, M., Cerezo-Mota, R., Cherchi, A., Gan, T.Y., Gergis, J., Jiang, D., Khan, A., Pokam Mba, W., Rosenfeld, D., Tierney, J. and Zolina, O. (2021) 'Water Cycle Changes', *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Available at: https://doi.org/10.1017/9781009157896.010.

Duwal, S., Liu, D. and Pradhan, P.M. (2023) 'Flood susceptibility modeling of the Karnali river basin of Nepal using different machine learning approaches', *Geomatics, Natural Hazards and Risk*, 14(1), p. 2217321.

DWIDP (2014) 'Annual Disaster Review 2013'. Department of Water Induced Disaster Prevention, Government of Nepal.

DWIDP (2015) ‘Annual Disaster Review 2014’. Department of Water Induced Disaster Prevention, Government of Nepal.

Egozi, R. and Ashmore, P. (2008) ‘Defining and measuring braiding intensity’, *Earth surface processes and landforms*, 33(14), pp. 2121–2138.

Einstein, H.A. (1950) *The bed-load function for sediment transportation in open channel flows*. US Department of Agriculture.

Elalem, S. and Pal, I. (2015) ‘Mapping the vulnerability hotspots over Hindu-Kush Himalaya region to flooding disasters’, *Weather and Climate Extremes*, 8, pp. 46–58.

EM-DAT/CRED (2021) ‘EM-DAT: The Emergency Events Database’. Available at: www.emdat.be (Accessed: 31 October 2021).

FAO (2011) ‘Major hydrological basins of the world’. Available at: https://data.apps.fao.org/catalog/iso/7707086d-af3c-41cc-8aa5-323d8609b2d1 (Accessed: 20 October 2023).

FAO (2022) ‘Major rivers of the world’. Available at: https://data.apps.fao.org/catalog/iso/76e4aacc-b89e-4091-831f-63986fe029f9 (Accessed: 20 October 2023).

Farinotti, D., Huss, M., Fürst, J.J., Landmann, J., Machguth, H., Maussion, F. and Pandit, A. (2019) ‘A consensus estimate for the ice thickness distribution of all glaciers on Earth’, *Nature Geoscience*, 12(3), pp. 168–173.

Feeney, C.J., Chiverrell, R.C., Smith, H.G., Hooke, J.M. and Cooper, J.R. (2020) ‘Modelling the decadal dynamics of reach-scale river channel evolution and floodplain turnover in CAESAR-Lisflood’, *Earth Surface Processes and Landforms*, 45(5), pp. 1273–1291.

Ferguson, R. and Church, M. (2009) ‘A critical perspective on 1-D modeling of river processes: Gravel load and aggradation in lower Fraser River’, *Water Resources Research*, 45(11). Available at: https://doi.org/10.1029/2009WR007740.

Ferguson, R.I. (1987) ‘Accuracy and precision of methods for estimating river loads’, *Earth Surface Processes and Landforms*, 12(1), pp. 95–104. Available at: https://doi.org/10.1002/esp.3290120111.

Feyisa, G.L., Meilby, H., Fensholt, R. and Proud, S.R. (2014) ‘Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery’, *Remote sensing of environment*, 140, pp. 23–35.

Gabet, E.J., Burbank, D.W., Pratt-Sitaula, B., Putkonen, J. and Bookhagen, B. (2008) ‘Modern erosion rates in the High Himalayas of Nepal’, *Earth and Planetary Science Letters*, 267(3), pp. 482–494. Available at: https://doi.org/10.1016/j.epsl.2007.11.059.

Gain, A.K., Immerzeel, W., Sperna Weiland, F. and Bierkens, M. (2011) ‘Impact of climate change on the stream flow of the lower Brahmaputra: trends in high and low flows based on discharge-weighted ensemble modelling’, *Hydrology and Earth System Sciences*, 15(5), pp. 1537–1545.

Gan, Y., Duan, Q., Gong, W., Tong, C., Sun, Y., Chu, W., Ye, A., Miao, C. and Di, Z. (2014) ‘A comprehensive evaluation of various sensitivity analysis methods: A case

study with a hydrological model', *Environmental Modelling & Software*, 51, pp. 269–285. Available at: https://doi.org/10.1016/j.envsoft.2013.09.031.

Gardner, R.A.M. and Gerrard, A.J. (2003) 'Runoff and soil erosion on cultivated rainfed terraces in the Middle Hills of Nepal', *Applied Geography*, 23(1), pp. 23–45. Available at: https://doi.org/10.1016/S0143-6228(02)00069-3.

Gariano, S.L. and Guzzetti, F. (2016) 'Landslides in a changing climate', *Earth-Science Reviews*, 162, pp. 227–252.

Gautam, M.R. and Acharya, K. (2012) 'Streamflow trends in Nepal', *Hydrological Sciences Journal*, 57(2), pp. 344–357. Available at: https://doi.org/10.1080/02626667.2011.637042.

Ghimire, S.K., Higaki, D. and Bhattarai, T.P. (2013) 'Estimation of Soil Erosion Rates and Eroded Sediment in a Degraded Catchment of the Siwalik Hills, Nepal', *Land*, 2(3), pp. 370–391. Available at: https://doi.org/10.3390/land2030370.

Giorgi, F. (2010) 'Uncertainties in climate change projections, from the global to the regional scale', in. *EPJ Web of conferences*, EDP Sciences, pp. 115–129.

Giuntoli, I., Prosdocimi, I. and Hannah, D.M. (2021) 'Going beyond the ensemble mean: Assessment of future floods from global multi-models', *Water Resources Research*, 57(3), p. e2020WR027897.

Giuntoli, I., Vidal, J.-P., Prudhomme, C. and Hannah, D.M. (2015) 'Future hydrological extremes: the uncertainty from multiple global climate and global hydrological models', *Earth System Dynamics*, 6(1), pp. 267–285.

Giuntoli, I., Villarini, G., Prudhomme, C. and Hannah, D.M. (2018) 'Uncertainties in projected runoff over the conterminous United States', *Climatic change*, 150, pp. 149–162.

Gupta, H.V., Beven, K.J. and Wagener, T. (2005) 'Model Calibration and Uncertainty Estimation', in *Encyclopedia of Hydrological Sciences*. Available at: https://doi.org/10.1002/0470848944.hsa138.

Gupta, H.V., Kling, H., Yilmaz, K.K. and Martinez, G.F. (2009) 'Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling', *Journal of hydrology*, 377(1–2), pp. 80–91.

Hall, D.K. and Riggs, G.A. (2007) 'Accuracy assessment of the MODIS snow products', *Hydrological Processes: An International Journal*, 21(12), pp. 1534–1547.

Hall, J.W. and Penning-Rowsell, E.C. (2010) 'Setting the Scene for Flood Risk Management', in *Flood Risk Science and Management*, pp. 1–16. Available at: https://doi.org/10.1002/9781444324846.ch1.

Hannah, D.M., Kansakar, S.R., Gerrard, A. and Rees, G. (2005) 'Flow regimes of Himalayan rivers of Nepal: nature and spatial patterns', *Journal of Hydrology*, 308(1–4), pp. 18–32.

Hannah, D.M., Lynch, I., Mao, F., Miller, J.D., Young, S.L. and Krause, S. (2020) 'Water and sanitation for all in a pandemic', *Nature Sustainability*, 3(10), pp. 773–775.

Hardy, R.J. (2013) '2.11 Process-Based Sediment Transport Modeling', in J.F. Shroder (ed.) *Treatise on Geomorphology*. San Diego: Academic Press, pp. 147–159. Available at: https://doi.org/10.1016/B978-0-12-374739-6.00036-1.

Hargreaves, G.H. and Samani, Z.A. (1985) 'Reference crop evapotranspiration from temperature', *Applied engineering in agriculture*, 1(2), pp. 96–99.

Harvey, A. (2018) 'Alluvial Fans', in *Reference Module in Earth Systems and Environmental Sciences*. Elsevier. Available at: https://doi.org/10.1016/B978-0-12-409548-9.11066-8.

Hawker, L., Uhe, P., Paulo, L., Sosa, J., Savage, J., Sampson, C. and Neal, J. (2022) 'A 30 m global map of elevation with forests and buildings removed', *Environmental Research Letters*, 17(2), p. 024016.

Hengl, T., De Jesus, J.M., MacMillan, R.A., Batjes, N.H., Heuvelink, G.B., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G. and Walsh, M.G. (2014) 'SoilGrids1km—global soil information based on automated mapping', *PloS one*, 9(8), p. e105992.

Heo, J.-H., Kho, Y.W., Shin, H., Kim, S. and Kim, T. (2008) 'Regression equations of probability plot correlation coefficient test statistics from several probability distributions', *Journal of Hydrology*, 355(1), pp. 1–15. Available at: https://doi.org/10.1016/j.jhydrol.2008.01.027.

Heo, J.-H., Shin, H., Nam, W., Om, J. and Jeong, C. (2013) 'Approximation of modified Anderson–Darling test statistics for extreme value distributions with unknown shape parameter', *Journal of Hydrology*, 499, pp. 41–49. Available at: https://doi.org/10.1016/j.jhydrol.2013.06.008.

Herschy, R. (1993) 'The stage-discharge relation', *Special Issue - Open Channel Measurements*, 4(1), pp. 11–15. Available at: https://doi.org/10.1016/0955-5986(93)90005-4.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S. (2013) 'Global flood risk under climate change', *Nature Climate Change*, 3(9), pp. 816–821. Available at: https://doi.org/10.1038/nclimate1911.

Hirabayashi, Y., Tanoue, M., Sasaki, O., Zhou, X. and Yamazaki, D. (2021) 'Global exposure to flooding from the new CMIP6 climate model projections', *Scientific Reports*, 11(1), p. 3740. Available at: https://doi.org/10.1038/s41598-021-83279-w.

Hock, R. (2003) 'Temperature index melt modelling in mountain areas', *Journal of hydrology*, 282(1–4), pp. 104–115.

Holden, J. (2008) *An Introduction to Physical Geography and the Environment*. Second Edition. Pearson Prentice Hall. Available at: https://books.google.co.uk/books?id=gSp-w49ks8sC.

Hornberger, G.M. and Spear, R.C. (1981) 'Approach to the preliminary analysis of environmental systems', *J. Environ. Manage.; (United States)*, 12:1. Available at: https://www.osti.gov/biblio/6396608.

Horritt, M. (2006) 'A methodology for the validation of uncertain flood inundation models', *Journal of Hydrology*, 326(1–4), pp. 153–165.

Horritt, M.S. and Bates, P.D. (2002) 'Evaluation of 1D and 2D numerical models for predicting river flood inundation', *Journal of Hydrology*, 268(1), pp. 87–99. Available at: https://doi.org/10.1016/S0022-1694(02)00121-X.

Hosking, J.R.M. (1990) 'L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics', *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(1), pp. 105–124.

Hosking, J.R.M. and Wallis, J.R. (1997) *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9780511529443.

Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985) 'Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments', *Technometrics*, 27(3), pp. 251–261. Available at: https://doi.org/10.2307/1269706.

Huang, S., Kumar, R., Flörke, M., Yang, T., Hundecha, Y., Kraft, P., Gao, C., Gelfan, A., Liersch, S., Lobanova, A., Strauch, M., van Ogtrop, F., Reinhardt, J., Haberlandt, U. and Krysanova, V. (2017) 'Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide', *Climatic Change*, 141(3), pp. 381–397. Available at: https://doi.org/10.1007/s10584-016-1841-8.

Huffman, G., J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E., Sorooshian, S., Tan, J. and Xan, P. (2019) *NASA Global Precipitation Measurement (GPM) - Integrated Multi-satellitE Retrievals for GPM (IMERG)*. Greenbelt: National Aeronautics and Space Administration, p. 38.

Huffman, G., J., Bolvin, D.T., Nelkin, E. and Adler, R.F. (2016) 'TRMM (TMPA) Precipitation L3 1 day 0.25 degree x 0.25 degree V7'. Goddard Earth Sciences Data and Information Services Center (GES DISC). Available at: https://doi.org/10.5067/TRMM/TMPA/DAY/7.

Huffman, G.J., Adler, R.F., Bolvin, D.T. and Nelkin, E.J. (2010) 'The TRMM multi-satellite precipitation analysis (TMPA)', *Satellite rainfall applications for surface hydrology*, pp. 3–22.

Hurtrez, J.-E., Lucazeau, F., Lavé, J. and Avouac, J.-P. (1999) 'Investigation of the relationships between basin morphology, tectonic uplift, and denudation from the study of an active fold belt in the Siwalik Hills, central Nepal', *Journal of Geophysical Research: Solid Earth*, 104(B6), pp. 12779–12796. Available at: https://doi.org/10.1029/1998JB900098.

ICIMOD (2008) 'Outline boundary of Hindu Kush Himalayan Region'. Available at: https://rds.icimod.org/ (Accessed: 18 October 2023).

ICIMOD (2011) 'The physiographic regions of Nepal.' Available at: https://rds.icimod.org/.

Immerzeel, W.W., Pellicciotti, F. and Bierkens, M.F.P. (2013) 'Rising river flows throughout the twenty-first century in two Himalayan glacierized watersheds', *Nature Geoscience*, 6(9), pp. 742–745. Available at: https://doi.org/10.1038/ngeo1896.

Immerzeel, W.W., Petersen, L., Ragettli, S. and Pellicciotti, F. (2014) 'The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas', *Water Resources Research*, 50(3), pp. 2212–2226.

Immerzeel, W.W., Van Beek, L., Konz, M., Shrestha, A. and Bierkens, M. (2012) 'Hydrological response to climate change in a glacierized catchment in the Himalayas', *Climatic change*, 110, pp. 721–736.

Immerzeel, W.W., Van Beek, L.P. and Bierkens, M.F. (2010) 'Climate change will affect the Asian water towers', *science*, 328(5984), pp. 1382–1385.

Immerzeel, W.W., Wanders, N., Lutz, A.F., Shea, J.M. and Bierkens, M.F.P. (2015) 'Reconciling high-altitude precipitation in the upper Indus basin with glacier mass balances and runoff', *Hydrology and Earth System Sciences*, 19(11), pp. 4673–4687. Available at: https://doi.org/10.5194/hess-19-4673-2015.

Ingham, D.B. and Ma, L. (2005) 'Fundamental Equations for CFD in River Flow Simulations', in *Computational Fluid Dynamics*, pp. 17–49. Available at: https://doi.org/10.1002/0470015195.ch2.

IPCC (2023) *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/9781009157896.

ISO (2007) *Hydrometry - Measurement of liquid flow in open channels using current-meters or floats (ISO 748:2007)*. BS EN ISO 748:2007. Geneva,Switzerland: International Organisation for Standardisation, p. 58.

Ives, J.D. and Messerli, B. (1989) *The Himalayan dilemma: reconciling development and conservation*. Routledge.

Jalbert, J., Mathevet, T. and Favre, A.-C. (2011) 'Temporal uncertainty estimation of discharges from rating curves using a variographic analysis', *Journal of Hydrology*, 397(1), pp. 83–92. Available at: https://doi.org/10.1016/j.jhydrol.2010.11.031.

Jarvis, A., Reuter, H., Nelson, A. and Guevara, E. (2008) 'Hole-filled SRTM for the globe Version 4'. Available at: http://srtm.csi.cgiar.org (Accessed: 5 September 2023).

Jehanzaib, M., Shah, S.A., Yoo, J. and Kim, T.-W. (2020) 'Investigating the impacts of climate change and human activities on hydrological drought using non-stationary approaches', *Journal of Hydrology*, 588, p. 125052. Available at: https://doi.org/10.1016/j.jhydrol.2020.125052.

Jirka, G.H. and Uijttewaal, W.S.J. (2004) *Shallow Flows: Research Presented at the International Symposium on Shallow Flows, Delft, Netherlands, 2003*. CRC Press. Available at: https://doi.org/10.1201/9780203027325.

Karssenberg, D. (2002) 'The value of environmental modelling languages for building distributed hydrological models', *Hydrological Processes*, 16(14), pp. 2751–2766.

Karssenberg, D., Burrough, P.A., Sluiter, R. and de Jong, K. (2001) 'The PCRaster software and course materials for teaching numerical modelling in the environmental sciences', *Transactions in GIS*, 5(2), pp. 99–110.

Kennedy, E.J. (1983) *Computation of continuous records of streamflow*. Report 03-A13. Available at: https://doi.org/10.3133/twri03A13.

Khanal, N., Shrestha, M. and Ghimire, M. (2007) 'Preparing for Flood Disaster. Mapping and Assessing Hazard in the Ratu Watershed, Nepal.' International Centre for Integrated Mountain Development.

Khanal, S., Lutz, A.F., Kraaijenbrink, P.D.A., van den Hurk, B., Yao, T. and Immerzeel, W.W. (2021) 'Variable 21st Century Climate Change Response for Rivers in High Mountain Asia at Seasonal to Decadal Time Scales', *Water Resources Research*, 57(5), p. e2020WR029266. Available at: https://doi.org/10.1029/2020WR029266.

Kiang, J.E., Gazoorian, C., McMillan, H., Coxon, G., Le Coz, J., Westerberg, I.K., Belleville, A., Sevrez, D., Sikorska, A.E., Petersen-Øverleir, A., Reitan, T., Freer, J., Renard, B., Mansanarez, V. and Mason, R. (2018) 'A Comparison of Methods for Streamflow Uncertainty Estimation', *Water Resources Research*, 54(10), pp. 7149–7176. Available at: https://doi.org/10.1029/2018WR022708.

Kjeldsen, T.R., Lamb, R. and Blazkova, S. (2014) 'Uncertainty in flood frequency analysis', in K. Beven and J. Hall (eds) *Applied Uncertainty Analysis for Flood Risk Management*. Imperial College Press, pp. 153–197.

Kleinhans, M.G. (2010) 'Sorting out river channel patterns', *Progress in Physical Geography: Earth and Environment*, 34(3), pp. 287–326. Available at: https://doi.org/10.1177/0309133310365300.

Kleinhans, M.G., Ferguson, R.I., Lane, S.N. and Hardy, R.J. (2013) 'Splitting rivers at their seams: bifurcations and avulsion', *Earth Surface Processes and Landforms*, 38(1), pp. 47–61. Available at: https://doi.org/10.1002/esp.3268.

Kleinhans, M.G., Jagers, H.R.A., Mosselman, E. and Sloff, C.J. (2008) 'Bifurcation dynamics and avulsion duration in meandering rivers by one-dimensional and three-dimensional models', *Water Resources Research*, 44(8). Available at: https://doi.org/10.1029/2007WR005912.

Kokkonen, T., Koivusalo, H., Jakeman, T. and Norton, J. (2006) 'Construction of a degree-day snow model in the light of the ten iterative steps in model development'.

Komi, K., Neal, J., Trigg, M.A. and Diekkrüger, B. (2017) 'Modelling of flood hazard extent in data sparse areas: a case study of the Oti River basin, West Africa', *Journal of Hydrology: Regional Studies*, 10, pp. 122–132.

Kosow, H., Kirschke, S., Borchardt, D., Cullmann, J., Guillaume, J.H., Hannah, D.M., Schaub, S. and Tosun, J. (2022) 'Scenarios of water extremes: Framing ways forward for wicked problems', *Hydrological Processes*, 36(2), p. e14492.

Koutsoyiannis, D. and Montanari, A. (2007) 'Statistical analysis of hydroclimatic time series: Uncertainty and insights', *Water Resources Research*, 43(5). Available at: https://doi.org/10.1029/2006WR005592.

Krause, P. (2001) *Das hydrologische Modellsystem J2000 - Beschreibung und Anwendung in großen Flußgebieten*. Jülich: Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag (Schriften des Forschungszentrums Jülich. Reihe Umwelt / Environment, PreJuSER-37462). Available at: https://juser.fz-juelich.de/record/37462.

Krishnan, R., Shrestha, A.B., Ren, G., Rajbhandari, R., Saeed, S., Sanjay, J., Syed, Md.A., Vellore, R., Xu, Y., You, Q. and Ren, Y. (2019) 'Unravelling Climate Change in the Hindu Kush Himalaya: Rapid Warming in the Mountains and Increasing Extremes', in P. Wester, A. Mishra, A. Mukherji, and Arun Bhakta Shrestha (eds) *The Hindu Kush*

Himalaya Assessment: Mountains, Climate Change, Sustainability and People. Cham: Springer International Publishing, pp. 57–97. Available at: https://doi.org/10.1007/978-3-319-92288-1_3.

Krysanova, V., Vetter, T., Eisner, S., Huang, S., Pechlivanidis, I., Strauch, M., Gelfan, A., Kumar, R., Aich, V. and Arheimer, B. (2017) 'Intercomparison of regional-scale hydrological models and climate change impacts projected for 12 large river basins worldwide—a synthesis', *Environmental Research Letters*, 12(10), p. 105002.

Lane, S., Tayefi, V., Reid, S., Yu, D. and Hardy, R. (2007) 'Interactions between sediment delivery, channel change, climate change and flood risk in a temperate upland environment', *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, 32(3), pp. 429–446.

Lane, S.N. (1998) 'Hydraulic modelling in hydrology and geomorphology: a review of high resolution approaches', *Hydrological Processes*, 12(8), pp. 1131–1150. Available at: https://doi.org/10.1002/(SICI)1099-1085(19980630)12:8<1131::AID-HYP611>3.0.CO;2-K.

Lane, S.N. (2017) 'Natural flood management', *WIREs Water*, 4(3), p. e1211. Available at: https://doi.org/10.1002/wat2.1211.

Lane, S.N., Landström, C. and Whatmore, S.J. (2011) 'Imagining flood futures: risk assessment and management in practice', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1942), pp. 1784–1806. Available at: https://doi.org/10.1098/rsta.2010.0346.

Lang, N., Jetz, W., Schindler, K. and Wegner, J.D. (2022) 'A high-resolution canopy height model of the Earth', *arXiv preprint arXiv:2204.08322* [Preprint].

Langhorst, T. and Pavelsky, T. (2023) 'Global Observations of Riverbank Erosion and Accretion From Landsat Imagery', *Journal of Geophysical Research: Earth Surface*, 128(2), p. e2022JF006774. Available at: https://doi.org/10.1029/2022JF006774.

Le Coz, J. (2014) 'BaRatinAGE User Manual - Controls and estimation of priors.'

Le Coz, J., Renard, B., Bonnifait, L., Branger, F. and Le Boursicaud, R. (2014) 'Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A Bayesian approach', *Journal of Hydrology*, 509, pp. 573–587. Available at: https://doi.org/10.1016/j.jhydrol.2013.11.016.

Leenman, A.S. and Eaton, B.C. (2022) 'Episodic sediment supply to alluvial fans: implications for fan incision and morphometry', *Earth Surface Dynamics 10(6)*, pp 1097 – 1114. Available at: https://doi.org/10.5194/esurf-10-1097-2022

Legates, D.R. and McCabe, G.J. (1999) 'Evaluating the use of "goodness-of-fit" Measures in hydrologic and hydroclimatic model validation', *Water Resources Research*, 35(1), pp. 233–241. Available at: https://doi.org/10.1029/1998WR900018.

Lehner, B., Verdin, K. and Jarvis, A. (2008) 'New global hydrography derived from spaceborne elevation data', *Eos, Transactions American Geophysical Union*, 89(10), pp. 93–94.

Leopold, L.B. and Wolman, M.G. (1957) *River channel patterns: Braided, meandering, and straight*. Report 282B. Washington, D.C., p. 50. Available at: https://doi.org/10.3133/pp282B.

Li, H., Li, X. and Xiao, P. (2016) 'Impact of sensor zenith angle on MOD10A1 data reliability and modification of snow cover data for the Tarim River Basin', *Remote Sensing*, 8(9), p. 750.

Li, H., Sun, J., Zhang, H., Zhang, J., Jung, K., Kim, J., Xuan, Y., Wang, X. and Li, F. (2018) 'What Large Sample Size Is Sufficient for Hydrologic Frequency Analysis?—A Rational Argument for a 30-Year Hydrologic Sample Size in Water Resources Management', *Water*, 10(4). Available at: https://doi.org/10.3390/w10040430.

Lindsay, J.B. (2018) *A new method for the removal of off-terrain objects from LiDAR-derived raster surface models*. Technical Report. Guelph, p. 17. Available at: DOI: 10.13140/RG.2.2.21226.62401.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. and Bergström, S. (1997) 'Development and test of the distributed HBV-96 hydrological model', *Journal of Hydrology*, 201(1), pp. 272–288. Available at: https://doi.org/10.1016/S0022-1694(97)00041-3.

Liu, Y., Freer, J., Beven, K. and Matgen, P. (2009) 'Towards a limits of acceptability approach to the calibration of hydrological models: Extending observation error', *Journal of Hydrology*, 367(1), pp. 93–103. Available at: https://doi.org/10.1016/j.jhydrol.2009.01.016.

Lupker, M., Blard, P.-H., Lavé, J., France-Lanord, C., Leanni, L., Puchol, N., Charreau, J. and Bourlès, D. (2012) '10Be-derived Himalayan denudation rates and sediment budgets in the Ganga basin', *Earth and Planetary Science Letters*, 333, pp. 146–156.

Lupker, M., France-Lanord, C., Lavé, J., Bouchez, J., Galy, V., Métivier, F., Gaillardet, J., Lartiges, B. and Mugnier, J.-L. (2011) 'A Rouse-based method to integrate the chemical composition of river sediments: Application to the Ganga basin', *Journal of Geophysical Research: Earth Surface*, 116(F4). Available at: https://doi.org/10.1029/2010JF001947.

Lutz, A.F. and Immerzeel, W.W. (2016) *Reference Climate Dataset for the Indus, Ganges, and Brahmaputra River Basins*. 2. Kathmandu: International Centre for Integrated Mountain Development, p. 34.

Lutz, A.F., Immerzeel, W.W., Shrestha, A.B. and Bierkens, M.F.P. (2014) 'Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation', *Nature Climate Change*, 4(7), pp. 587–592. Available at: https://doi.org/10.1038/nclimate2237.

MacClune, K., Venkateswaran, K., Dixit, K., Maharjan, R. and Dugar, S. (2015) *Urgent case for recovery: what we can learn from the August 2014 Karnali River floods in Nepal*. ISET-International; Practical Action. Available at: http://isetnepal.org.np/wp-content/uploads/2018/12/risk-nexus-karnali-river-floods-nepal-july-2015.pdf (Accessed: 4 September 2023).

Maharjan, S.B., Steiner, J.F., Shrestha, A.B., Maharjan, A., Nepal, S., Shrestha, Mandira, Bajracharya, B., Rasul, G., Shrestha, Maxim, Jackson, M. and Gupta, N. (2021) *The Melamchi flood disaster: Cascading hazard and the need for multihazard risk management*. Kathmandu: International Centre for Integrated Mountain Development, p. 21.

Majone, B., Avesani, D., Zulian, P., Fiori, A. and Bellin, A. (2022) 'Analysis of high streamflow extremes in climate change studies: how do we calibrate hydrological models?', *Hydrology and Earth System Sciences*, 26(14), pp. 3863–3883.

Mangini, W., Viglione, A., Hall, J., Hundecha, Y., Ceola, S., Montanari, A., Rogger, M., Salinas, J.L., Borzì, I. and Parajka, J. (2018) 'Detection of trends in magnitude and frequency of flood peaks across Europe', *Hydrological Sciences Journal*, 63(4), pp. 493–512. Available at: https://doi.org/10.1080/02626667.2018.1444766.

Martel, J.-L., Mailhot, A., Brissette, F. and Caya, D. (2018) 'Role of natural climate variability in the detection of anthropogenic climate change signal for mean and extreme precipitation at local and regional scales', *Journal of Climate*, 31(11), pp. 4241–4263.

Massei, N., Kingston, D.G., Hannah, D.M., Vidal, J.-P., Dieppois, B., Fossa, M., Hartmann, A., Lavers, D.A. and Laignel, B. (2020) 'Understanding and predicting large-scale hydrological variability in a changing environment', *Proceedings of the International Association of Hydrological Sciences*, 383, pp. 141–149.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L. and Gomis, M. (2021) 'Climate change 2021: the physical science basis', *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*, 2.

McKay, M.D., Beckman, R.J. and Conover, W.J. (1979) 'A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code', *Technometrics*, 21(2), pp. 239–245. Available at: https://doi.org/10.2307/1268522.

McMahon, T.A. and Peel, M.C. (2019) 'Uncertainty in stage–discharge rating curves: application to Australian Hydrologic Reference Stations data', *Hydrological Sciences Journal*, 64(3), pp. 255–275. Available at: https://doi.org/10.1080/02626667.2019.1577555.

McMillan, H., Krueger, T. and Freer, J. (2012) 'Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality', *Hydrological Processes*, 26(26), pp. 4078–4111. Available at: https://doi.org/10.1002/hyp.9384.

Merritt, W.S., Letcher, R.A. and Jakeman, A.J. (2003) 'A review of erosion and sediment transport models', *The Modelling of Hydrologic Systems*, 18(8), pp. 761–799. Available at: https://doi.org/10.1016/S1364-8152(03)00078-1.

Meteor Consortium (2019) 'Nepal Flood maps: Fluvial Defended, Undefended and Pluvial for various return periods.' Available at: https://maps.meteor-project.org/map/flood-npl/#7/27.915/84.046 (Accessed: 7 October 2023).

Meyer-Peter, E. and Müller, R. (1948) 'Formulas for bed-load transport', in. *IAHSR 2nd meeting, Stockholm, appendix 2*, IAHR.

Miller, J.E. (1984) *Basic concepts of kinematic-wave models*. Report 1302. Available at: https://doi.org/10.3133/pp1302.

Milliman, J.D. and Syvitski, J.P.M. (1992) 'Geomorphic/Tectonic Control of Sediment Discharge to the Ocean: The Importance of Small Mountainous Rivers', *The Journal of Geology*, 100(5), pp. 525–544.

Mishra, V., Bhatia, U. and Tiwari, A.D. (2020) 'Bias-corrected climate projections for South Asia from Coupled Model Intercomparison Project-6', *Scientific Data*, 7(1), p. 338. Available at: https://doi.org/10.1038/s41597-020-00681-1.

Miura, K. (2011) ‘An Introduction to Maximum Likelihood Estimation and Information Geometry’, *Interdisciplinary Information Sciences*, 17(3), pp. 155–174. Available at: https://doi.org/10.4036/iis.2011.155.

MoF (2021) ‘Exonomic Survey 2020/21’. Ministry of Finance, Government of Nepal.

MoHA (2009) ‘Nepal Disaster Report: The Hazardscape and Vulnerability’. Ministry of Home Affairs, Government of Nepal.

MoHA (2018) ‘Disaster Risk Reduction National Strategic Plan of Action 2018 - 2030’. Ministry of Home Affairs, Government of Nepal.

Monteith, J.L. (1965) ‘Evaporation and environment’, in. *Symposia of the society for experimental biology*, Cambridge University Press (CUP) Cambridge, pp. 205–234.

Moriasi, D., Arnold, J., W. Van Liew, M., L. Bingner, R., D. Harmel, R. and L. Veith, T. (2007) ‘Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations’, *Transactions of the ASABE*, 50(3), pp. 885–900. Available at: https://doi.org/10.13031/2013.23153.

Moriasi, D., Gitau, M., Pai, N. and Daggupati, P. (2015) ‘Hydrologic and Water Quality Models: Performance Measures and Evaluation Criteria’, *Transactions of the ASABE*, 58(6), pp. 1763–1785. Available at: https://doi.org/10.13031/trans.58.10715.

Moriasi, D.N., Gitau, M.W., Pai, N. and Daggupati, P. (2015) ‘Hydrologic and water quality models: Performance measures and evaluation criteria’, *Transactions of the ASABE*, 58(6), pp. 1763–1785.

Morin, G.P., Lavé, J., France-Lanord, C., Rigaudier, T., Gajurel, A.P. and Sinha, R. (2018) ‘Annual Sediment Transport Dynamics in the Narayani Basin, Central Nepal: Assessing the Impacts of Erosion Processes in the Annual Sediment Budget’, *Journal of Geophysical Research: Earth Surface*, 123(10), pp. 2341–2376. Available at: https://doi.org/10.1029/2017JF004460.

Mosselman, E. (2012) ‘Modelling Sediment Transport and Morphodynamics of Gravel-Bed Rivers’, in *Gravel-Bed Rivers*, pp. 101–115. Available at: https://doi.org/10.1002/9781119952497.ch9.

Muhammad, S. and Thapa, A. (2019) ‘Improved MODIS TERRA/AQUA composite Snow and glacier (RGI6. 0) data for High Mountain Asia (2002–2018), PANGAEA’.

Muhammad, S. and Thapa, A. (2020) ‘An improved Terra–Aqua MODIS snow cover and Randolph Glacier Inventory 6.0 combined product (MOYDGL06*) for high-mountain Asia between 2002 and 2018’, *Earth System Science Data*, 12(1), pp. 345–356.

Murray, A.B. and Paola, C. (1994) ‘A cellular model of braided rivers’, *Nature*, 371(6492), pp. 54–57. Available at: https://doi.org/10.1038/371054a0.

Murray, A.B. and Paola, C. (2003) ‘Modelling the effect of vegetation on channel pattern in bedload rivers’, *Earth Surface Processes and Landforms*, 28(2), pp. 131–143. Available at: https://doi.org/10.1002/esp.428.

Nash, J.E. and Sutcliffe, J.V. (1970) ‘River flow forecasting through conceptual models part I—A discussion of principles’, *Journal of hydrology*, 10(3), pp. 282–290.

Nathan, R. and Ball, J. (2019) 'Approaches to Flood Estimation', in J. Ball, R. Nathan, Weeks, W, Weinmann, E, M. Retallick, and I. Testoni (eds) *Australian Rainfall and Runoff: A Guide to Flood Estimation*. ARR 2019. Commonwealth of Australia.

National Center for Atmospheric Research (2023) 'The Climate Data Guide: CPC Unified Gauge-Based Analysis of Global Daily Precipitation'. Available at: https://climatedataguide.ucar.edu/climate-data/cpc-unified-gauge-based-analysis-global-daily-precipitation (Accessed: 6 September 2023).

Neal, J., Villanueva, I., Wright, N., Willis, T., Fewtrell, T. and Bates, P. (2012) 'How much physical complexity is needed to model flood inundation?', *Hydrological Processes*, 26(15), pp. 2264–2282. Available at: https://doi.org/10.1002/hyp.8339.

Neitsch, S.L., Arnold, J.G., Kiniry, J.R. and Williams, J.R. (2011) *Soil and water assessment tool theoretical documentation version 2009*. Texas Water Resources Institute.

Nepal, S. (2012) *Evaluating Upstream-Downstream Linkages of Hydrological Dynamics in the Himalayan Region*. University of Jena.

Nepal, S. (2016) 'Impacts of climate change on the hydrological regime of the Koshi river basin in the Himalayan region', *Journal of Hydro-environment Research*, 10, pp. 76–89. Available at: https://doi.org/10.1016/j.jher.2015.12.001.

Nepal, S., Krause, P., Flügel, W.-A., Fink, M. and Fischer, C. (2014) 'Understanding the hydrological system dynamics of a glaciated alpine catchment in the Himalayan region using the J2000 hydrological model', *Hydrological Processes*, 28(3), pp. 1329–1344. Available at: https://doi.org/10.1002/hyp.9627.

Nicholas, A. (2013) '2.12 Morphodynamic Modeling of Rivers and Floodplains', in J.F. Shroder (ed.) *Treatise on Geomorphology*. San Diego: Academic Press, pp. 160–179. Available at: https://doi.org/10.1016/B978-0-12-374739-6.00037-3.

Nicholas, A.P. (2005) 'Cellular modelling in fluvial geomorphology', *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, 30(5), pp. 645–649. Available at: https://doi.org/10.1002/esp.1231.

Nicholas, A.P. (2013) 'Modelling the continuum of river channel patterns', *Earth Surface Processes and Landforms*, 38(10), pp. 1187–1196. Available at: https://doi.org/10.1002/esp.3431.

Nicholas, A.P. and Quine, T.A. (2007) 'Modeling alluvial landform change in the absence of external environmental forcing', *Geology*, 35(6), pp. 527–530. Available at: https://doi.org/10.1130/G23377A.1.

Nicholas, A.P., Thomas, R. and Quine, T.A. (2006) 'Cellular Modelling of Braided River form and Process', in *Braided Rivers*, pp. 137–151. Available at: https://doi.org/10.1002/9781444304374.ch6.

NPC (2017) 'Nepal Flood 2017 Post Flood Recovery Needs Assessment', *National Planning Commission, Government of Nepal: Kathmandu, Nepal* [Preprint].

NSO (2022) *Prelimiary Report of National Population 2021*. Kathmandu: National Statistical Office, Government of Nepal. Available at: https://censusnepal.cbs.gov.np/Home/Details?tpid=5&tfsid=17.

O'Callaghan, J.F. and Mark, D.M. (1984) 'The extraction of drainage networks from digital elevation data', *Computer vision, graphics, and image processing*, 28(3), pp. 323–344.

Odoni, N., Alan (2007) *Exploring Equifinality in a Landscape Evolution Model*. University of Southhampton.

Okura, Y., Neupane, S. and Rana, B. (2020) *Avoiding a perfect storm: COVID-19 and floods in Nepal. Findings from Community Disaster Management Committees.* Flood Resilience Alliance. Available at: https://reliefweb.int/report/nepal/avoiding-perfect-storm-covid-19-and-floods-nepal-findings-community-disaster-management#:~:text=COVID%2D19%20is%20a%20serious,movement%20of%20Nepali%20migrant%20workers. (Accessed: 4 September 2023).

Omani, N., Srinivasan, R., Smith, P.K. and Karthikeyan, R. (2017) 'Glacier mass balance simulation using SWAT distributed snow algorithm', *Hydrological Sciences Journal*, 62(4), pp. 546–560.

O'Neill, B.C., Tebaldi, C., Van Vuuren, D.P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F. and Lowe, J. (2016) 'The scenario model intercomparison project (ScenarioMIP) for CMIP6', *Geoscientific Model Development*, 9(9), pp. 3461–3482.

Padrón, R.S., Gudmundsson, L. and Seneviratne, S.I. (2019) 'Observational constraints reduce likelihood of extreme changes in multidecadal land water availability', *Geophysical Research Letters*, 46(2), pp. 736–744.

Pan, Y., Wang, W. and Shi, W. (2019) 'Assessment of CPC global daily surface air temperature (CPC-T2m) analysis', *Assessment*, 22, p. 24.

Pandey, V.P., Dhaubanjar, S., Bharati, L. and Thapa, B.R. (2020) 'Spatio-temporal distribution of water availability in Karnali-Mohana Basin, Western Nepal: Hydrological model development using multi-site calibration approach (Part-A)', *Journal of Hydrology: Regional Studies*, 29, p. 100690.

Papanicolaou, A. (Thanos) N., Elhakeem, M., Krallis, G., Prakash, S. and Edinger, J. (2008) 'Sediment transport modeling review—current and future developments', *Journal of hydraulic engineering*, 134(1), pp. 1–14.

Papanicolaou, A.N., Diplas, P., Dancey, C.L. and Balakrishnan, M. (2001) 'Surface roughness effects in near-bed turbulence: Implications to sediment entrainment', *Journal of Engineering Mechanics*, 127(3), pp. 211–218.

Parker, G., Dhamotharan, S. and Stefan, H. (1982) 'Model experiments on mobile, paved gravel bed streams', *Water Resources Research*, 18(5), pp. 1395–1408.

Parker, G., Klingeman, P.C. and McLean, D.G. (1982) 'Bedload and size distribution in paved gravel-bed streams', *Journal of the Hydraulics Division*, 108(4), pp. 544–571.

Parkes, B. (2015) *Uncertainty in Flood Risk and its Implications for Management*. King's College London.

Parkes, B. and Demeritt, D. (2016) 'Defining the hundred year flood: A Bayesian approach for using historic data to reduce uncertainty in flood frequency estimates', *Journal of Hydrology*, 540, pp. 1189–1208. Available at: https://doi.org/10.1016/j.jhydrol.2016.07.025.

Pattison, I. and Lane, S.N. (2012) 'The relationship between Lamb weather types and long-term changes in flood frequency, River Eden, UK', *International Journal of Climatology*, 32(13), pp. 1971–1989. Available at: https://doi.org/10.1002/joc.2415.

Pearson, C. (2020) *Catchment-scale spatial targeting of ood management measures to reduce ood hazard: An end-to-end modelling approach applied to the East Rapti catchment, Nepal*. Durham University. Available at: http://etheses.dur.ac.uk/13794/ (Accessed: 9 July 2021).

Pearson, C., Reaney, S., Perks, M., Hortobágyi, B., Rosser, N. and Large, A. (2022) 'Identification of floodwater source areas in Nepal using SCIMAP-Flood', *Journal of Flood Risk Management*, 15(4), p. e12840.

Pechlivanidis, I.G., Arheimer, B., Donnelly, C., Hundecha, Y., Huang, S., Aich, V., Samaniego, L., Eisner, S. and Shi, P. (2017) 'Analysis of hydrological extremes at different hydro-climatic regimes under present and future conditions', *Climatic Change*, 141(3), pp. 467–481. Available at: https://doi.org/10.1007/s10584-016-1723-0.

Pechlivanidis, I.G., Jackson, B.M., McIntyre, N. and Wheater, H. (2011) 'Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications', *Global Nest Journal*, 3(3), pp. 193–214.

Pender, D., Patidar, S., Hassan, K. and Haynes, H. (2016) 'Method for incorporating morphological sensitivity into flood inundation modeling', *Journal of Hydraulic Engineering*, 142(6), p. 04016008.

Perera, E., Hiroe, A., Shrestha, D., Fukami, K., Basnyat, D., Gautam, S., Hasegawa, A., Uenoyama, T. and Tanaka, S. (2015) 'Community-based flood damage assessment approach for lower West Rapti River basin in Nepal under the impact of climate change', *Natural Hazards*, 75, pp. 669–699.

Petersen-Øverleir, A. and Reitan, T. (2005) 'Objective segmentation in compound rating curves', *Journal of Hydrology*, 311(1), pp. 188–201. Available at: https://doi.org/10.1016/j.jhydrol.2005.01.016.

Philip, G. and Watson, D.F. (1982) 'A precise method for determining contoured surfaces', *The APPEA Journal*, 22(1), pp. 205–212.

Pitlick, J., Cui, Y. and Wilcock, P.R. (2009) *Manual for computing bed load transport using BAGS (Bedload Assessment for Gravel-bed Streams) Software*. US Department of Agriculture, Forest Service, Rocky Mountain Research ….

Ponce, V., Lohani, A. and Scheyhing, C. (1996) 'Analytical verification of Muskingum-Cunge routing', *Journal of Hydrology*, 174(3–4), pp. 235–241.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H. and Silva, C.E. (2021) 'Mapping global forest canopy height through integration of GEDI and Landsat data', *Remote Sensing of Environment*, 253, p. 112165.

Poulin, A., Brissette, F., Leconte, R., Arsenault, R. and Malo, J.-S. (2011) 'Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin', *Journal of Hydrology*, 409(3–4), pp. 626–636.

Practical Action (2018) *Building Flood Resilience through Livelihood Diversification in Lower Karnali Basin*. Kathmandu: Practical Action, p. 24.

Prudhomme, C., Giuntoli, I., Robinson, E.L., Clark, D.B., Arnell, N.W., Dankers, R., Fekete, B.M., Franssen, W., Gerten, D. and Gosling, S.N. (2014) 'Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment', *Proceedings of the National Academy of Sciences*, 111(9), pp. 3262–3267.

Quick, L., Sinclair, H.D., Attal, M. and Singh, V. (2019) 'Conglomerate recycling in the Himalayan foreland basin: Implications for grain size and provenance', *GSA Bulletin*, 132(7–8), pp. 1639–1656. Available at: https://doi.org/10.1130/B35334.1.

Ranasinghe, R., Ruane, A. and Vautard, R. (2021) 'Chapter 12: Climate chang e information for regional impact and for risk assessment. Climate Change 2021: Th e Physical Science Basis. Contribution of Working Group I to the Sixth Assessmen t Report of the Intergovernmental Panel on Climate Change'.

Rao, R., A. and Hamed, K., H. (eds) (2000) *Flood Frequency Analysis*. First edition. Boca Raton: CRC Press. Available at: https://doi.org/10.1201/9780429128813.

Reaney, S.M. (2022) 'Spatial targeting of nature-based solutions for flood risk management within river catchments', *Journal of Flood Risk Management*, 15(3), p. e12803. Available at: https://doi.org/10.1111/jfr3.12803.

Red Cross and Red Crescent (2013) 'Nepal: Floods and Landslides - Information bulleting n° 1. International Federation of Red Cross and Red Crescent Societies.

Red Cross and Red Crescent (2021) 'Nepal: Monsoon Floods and Landslides - Operation Update #1, DREF Operation n° MDRNP011'. International Federation of Red Cross and Red Crescent Societies.

Reiss, R.D. and Thomas, M. (2007) *Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields*. Third Edition. Basel: Springer Science & Business Media, 2007.

Ren, Y.-Y., Ren, G.-Y., Sun, X.-B., Shrestha, A.B., You, Q.-L., Zhan, Y.-J., Rajbhandari, R., Zhang, P.-F. and Wen, K.-M. (2017) 'Observed changes in surface air temperature and precipitation in the Hindu Kush Himalayan region over the last 100-plus years', *Advances in Climate Change Research*, 8(3), pp. 148–156.

RGI Consortium, . (2017) 'Randolph Glacier Inventory - A Dataset of Global Glacier Outlines, Version 6'. National Snow and Ice Data Center. Available at: https://doi.org/10.7265/4m1f-gd79.

Riggs, G.A., Hall, D.K. and Román, M.O. (2019) 'MODIS snow products collection 6.1 user guide', *National Snow and Ice Data Center: Boulder, CO, USA* [Preprint].

Ritter, M.E. (2003) *The Physical Environment: An Introduction to Physical Geography*. Available at: https://books.google.co.uk/books?id=9KUlnQAACAAJ.

Running, S.W., Mu, Q., Zhao, M. and Moreno, A. (2019) 'MODIS global terrestrial evapotranspiration (ET) product (MOD16A2/A3 and year-end gap-filled MOD16A2GF/A3GF) NASA Earth Observing System MODIS Land Algorithm (for collection 6)', *National Aeronautics and Space Administration, Washington, DC, USA [data set], https://doi. org/10.5067/MODIS/MOD16A2*, 6.

Ruxton, G.D. (2006) ‘The unequal variance t-test is an underused alternative to Student’s t-test and the Mann–Whitney U test’, *Behavioral Ecology*, 17(4), pp. 688–690.

Savi, S., Comiti, F. and Strecker, M.R. (2021) ‘Pronounced increase in slope instability linked to global warming: A case study from the eastern European Alps’, *Earth Surface Processes and Landforms*, 46(7), pp. 1328–1347.

Scanlon, P., Western, A. and Ozbey, N. (2008) ‘Estimating the uncertainty in flow estimates at flow monitoring sites in Victoria using Australian Standard 3778.2. 3’, *Final Report to Victorian Government, Department of Sustainability and Environment* [Preprint].

Scherler, D., Wulf, H. and Gorelick, N. (2018) ‘Global assessment of supraglacial debris-cover extents’, *Geophysical Research Letters*, 45(21), pp. 11–798.

Schumm, S.A., Mosley, M.P. and Weaver, W. (1987) ‘Experimental fluvial geomorphology’.

Schwindt, S., Negeiros, B., Scolari, F. and Barros, R. (2023) ‘TELEMAC. General Introduction and Tutorial Guide.’ Available at: https://hydro-informatics.com/numerics/telemac/telemac.html (Accessed: 18 October 2023).

Scott, C.A., Zhang, F., Mukherji, A., Immerzeel, W., Mustafa, D. and Bharati, L. (2019) ‘Water in the Hindu Kush Himalaya’, in P. Wester, A. Mishra, A. Mukherji, and A.B. Shrestha (eds) *The Hindu Kush Himalaya Assessment: Mountains, Climate Change, Sustainability and People*. Cham: Springer International Publishing, pp. 257–299. Available at: https://doi.org/10.1007/978-3-319-92288-1_8.

Seibert, J., Vis, M., Kohn, I., Weiler, M. and Stahl, K. (2018) ‘Technical note: Representing glacier geometry changes in a semi-distributed hydrological model, Hydrol. Earth Syst. Sc., 22, 2211–2224’.

Seneviratne, S.I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S.M., Wehner, M. and Zhou, B. (2021) ‘Weather and Climate Extreme Events in a Changing Climate’, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. Available at: https://doi.org/10.1017/9781009157896.013.

Sharma, E., Molden, D., Rahman, A., Khatiwada, Y.R., Zhang, L., Singh, S.P., Yao, T. and Wester, P. (2019) ‘Introduction to the Hindu Kush Himalaya Assessment’, in P. Wester, A. Mishra, A. Mukherji, and A.B. Shrestha (eds) *The Hindu Kush Himalaya Assessment: Mountains, Climate Change, Sustainability and People*. Cham: Springer International Publishing, pp. 1–16. Available at: https://doi.org/10.1007/978-3-319-92288-1_1.

Shaw, E.M. (2005) *Hydrology in Practice, Third Edition*. Third Edition. Milton Park: Taylor & Francis.

Shea, J.M. and Immerzeel, W.W. (2016) ‘An assessment of basin-scale glaciological and hydrological sensitivities in the Hindu Kush–Himalaya’, *Annals of Glaciology*.

2016/03/03 edn, 57(71), pp. 308–318. Available at: https://doi.org/10.3189/2016AoG71A073.

Shen, Z., Chen, L. and Chen, T. (2012) 'Analysis of parameter uncertainty in hydrological and sediment modeling using GLUE method: a case study of SWAT model applied to Three Gorges Reservoir Region, China', *Hydrology and Earth System Sciences*, 16(1), pp. 121–132.

Sherman, L.K. (1932) 'Streamflow from rainfall by unit-graph method', *Engineering News Record*, 108, pp. 501–505.

Shields, A. (1936) 'Application of similarity principles and turbulence research to bed-load movement'.

Shrestha, M., Dangol, P., Shrestha, A.B. and Rajkarnikar, G. (2015) 'Real-Time Monitoring and Flood Outlook for Reduced Flood Risks', in *Nepal Disater Report 2015*. Ministry of Home Affairs, Government of Nepal.

Shrestha, M.S., Grabs, W.E. and Khadgi, V.R. (2015) 'Establishment of a regional flood information system in the Hindu Kush Himalayas: challenges and opportunities', *International Journal of Water Resources Development*, 31(2), pp. 238–252.

Shrestha, R.R. and Simonovic, S.P. (2010) 'Fuzzy set theory based methodology for the analysis of measurement uncertainties in river discharge and stage', *Canadian Journal of Civil Engineering*, 37(3), pp. 429–440. Available at: https://doi.org/10.1139/L09-151.

Sinha, R. (2008) 'Kosi: Rising Waters, Dynamic Channels and Human Disasters', *Economic and Political Weekly*, 43(46), pp. 42–46.

Sinha, R., Gupta, A., Mishra, K., Tripathi, S., Nepal, S., Wahid, S. and Swarnkar, S. (2019) 'Basin-scale hydrology and sediment dynamics of the Kosi river in the Himalayan foreland', *Journal of Hydrology*, 570, pp. 156–166.

Sinha, R., Sripriyanka, K., Jain, V. and Mukul, M. (2014) 'Avulsion threshold and planform dynamics of the Kosi River in north Bihar (India) and Nepal: A GIS framework', *Geomorphology*, 216, pp. 157–170. Available at: https://doi.org/10.1016/j.geomorph.2014.03.035.

Sobrino, J.A., Julien, Y. and García-Monteiro, S. (2020) 'Surface temperature of the planet earth from satellite data', *Remote Sensing*, 12(2), p. 218.

Spasojevic, M. and Holly, F.M. (2008) 'Two-and three-dimensional numerical simulation of mobile-bed hydrodynamics and sedimentation', *Sedimentation engineering: processes, measurements, modeling, and practice*, pp. 683–761.

Stephens, C.M., Johnson, F.M. and Marshall, L.A. (2018) 'Implications of future climate change for event-based hydrologic models', *Advances in Water Resources*, 119, pp. 95–110. Available at: https://doi.org/10.1016/j.advwatres.2018.07.004.

Sullivan, R.C., Cook, D.R., Ghate, V.P., Kotamarthi, V.R. and Feng, Y. (2019) 'Improved spatiotemporal representativeness and bias reduction of satellite-based evapotranspiration retrievals via use of in situ meteorology and constrained canopy surface resistance', *Journal of Geophysical Research: Biogeosciences*, 124(2), pp. 342–352.

Sunilkumar, K., Yatagai, A. and Masuda, M. (2019) 'Preliminary evaluation of GPM-IMERG rainfall estimates over three distinct climate zones with APHRODITE', *Earth and Space Science*, 6(8), pp. 1321–1335.

Syvitski, J.P.M., Slingerland, R.L., Burgess, P., Murray, A.B., Wiberg, P., Tucker, G. and Voinov, A. (2010) 'Morphodynamic models : an overview', in C.A. Vionnet, M.H. Garcia, E.M. Latrubesse, and G.M.E. Perillo (eds) *River, Coastal and Estuarine Morphodynamics*. London: Taylor & Francis, pp. 3–20. Available at: https://ezproxy2.utwente.nl/login?url=https://webapps.itc.utwente.nl/library/2010/chap/voinov_mor.pdf.

Talchabhadel, R. (2021) 'Observations and climate models confirm precipitation pattern is changing over Nepal', *Jalawaayu*, 1(1), pp. 25–46. Available at: https://doi.org/10.3126/jalawaayu.v1i1.36448.

Tayefi, V., Lane, S.N., Hardy, R.J. and Yu, D. (2007) 'A comparison of one- and two-dimensional approaches to modelling flood inundation over complex upland floodplains', *Hydrological Processes*, 21(23), pp. 3190–3202. Available at: https://doi.org/10.1002/hyp.6523.

Tellman, B., Sullivan, J.A., Kuhn, C., Kettner, A.J., Doyle, C.S., Brakenridge, G.R., Erickson, T.A. and Slayback, D.A. (2021) 'Satellite imaging reveals increased proportion of population exposed to floods', *Nature*, 596(7870), pp. 80–86. Available at: https://doi.org/10.1038/s41586-021-03695-w.

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D. and Kim, S. (2017) 'Flood inundation modelling: A review of methods, recent advances and uncertainty analysis', *Environmental Modelling & Software*, 90, pp. 201–216. Available at: https://doi.org/10.1016/j.envsoft.2017.01.006.

Terink, W., Lutz, A.F. and Immerzeel, Walter W. (2015) *SPHY v2.0: Spatial Processes in Hydrology - Model theory, installation, and data preparation*. Report FutureWater:142. Wageningen: FutureWater.

Terink, W., Lutz, A.F., Simons, G.W.H., Immerzeel, W.W. and Droogers, P. (2015) 'SPHY v2.0: Spatial Processes in HYdrology', *Geoscientific Model Development*, 8(7), pp. 2009–2034. Available at: https://doi.org/10.5194/gmd-8-2009-2015.

Thapa, S., Sinclair, H.D., Creed, M.J., Mudd, S.M., Attal, M., Borthwick, A.G.L., Ghimire, B.N. and Watson, C.S. (2024) 'The impact of sediment flux and calibre on flood risk in the Kathmandu Valley, Nepal', *Earth Surface Processes and Landforms*, 49(2), pp. 706–727. Available at: https://doi.org/10.1002/esp.5731

Thiemann, M., Trosset, M., Gupta, H. and Sorooshian, S. (2001) 'Bayesian recursive parameter estimation for hydrologic models', *Water Resources Research*, 37(10), pp. 2521–2535. Available at: https://doi.org/10.1029/2000WR900405.

Thomas, R. and Nicholas, A.P. (2002) 'Simulation of braided river flow using a new cellular routing scheme', *Geomorphology*, 43(3), pp. 179–195. Available at: https://doi.org/10.1016/S0169-555X(01)00128-3.

Thorne, C., Wallerstein, N., Soar, P., Brookes, A., Wishart, D., Biedenharn, D., Gibson, S., Little Jr, C., Mooney, D., Watson, C.C., Green, T. and Coulthard, T. (2010) 'Accounting for Sediment in Flood Risk Management', in *Flood Risk Science and Management*, pp. 87–113. Available at: https://doi.org/10.1002/9781444324846.ch5.

Todd, M.C., Andersson, L., Ambrosino, C., Hughes, D., Kniveton, D.R., Mileham, L., Murray-Hudson, M., Raghavan, S., Taylor, R. and Wolski, P. (2011) 'Climate Change Impacts on Hydrology in Africa: Case Studies of River Basin Water Resources', in C.J.R. Williams and D.R. Kniveton (eds) *African Climate and Climate Change: Physical, Social and Political Perspectives*. Dordrecht: Springer Netherlands, pp. 123–153. Available at: https://doi.org/10.1007/978-90-481-3842-5_6.

UN (2023) *Technical dialogue of the first global stocktake.* Geneva,Switzerland: Framework Convention on Climate Change Secretariat, United Nations. Available at: https://unfccc.int/documents/631600 (Accessed: 20 September 2023).

UNFPA Nepal (2017) *Population Situation Analysis of Nepal*. Kathmandu: United Nations Population Fund Nepal, p. 90.

Upreti, B. (2001) 'The Physiographic and Geology of Nepal and Their Bearing on the Landslide Problem', in *Landslide Hazard Mitigation in the Hindu Kush-Himalaya*. Kathmandu: International Centre for Integrated Mountain Development, pp. 31–49.

Upreti, B.N. (1999) 'An overview of the stratigraphy and tectonics of the Nepal Himalaya', *Journal of Asian Earth Sciences*, 17(5), pp. 577–606. Available at: https://doi.org/10.1016/S1367-9120(99)00047-4.

USACE (2020) *Modeler Application Guidance for Steady vs Unsteady, and 1D vs 2D vs 3D Hydraulic Modeling*. Training document TD-41. Davis, United States of America: US Army Corps of Engineers, Insitute for Water Resources.

USACE (2023) *HEC-RAS 2D User's Manual V6.3*. Davis, United States of America: US Army Corps of Engineers, Insitute for Water Resources.

USDA (2014) *National Engineering Handbook: Part 630 - Hydrology*. 630. Washington, D.C.: United States Department of Agriculture, Natural Resources Conservation Service.

Van De Wiel, M.J., Coulthard, T.J., Macklin, M.G. and Lewin, J. (2007) 'Embedding reach-scale fluvial dynamics within the CAESAR cellular automaton landscape evolution model', *Reduced-Complexity Geomorphological Modelling for River and Catchment Management*, 90(3), pp. 283–301. Available at: https://doi.org/10.1016/j.geomorph.2006.10.024.

Van Der Knijff, J.M., Younis, J. and De Roo, A.P.J. (2010) 'LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation', *International Journal of Geographical Information Science*, 24(2), pp. 189–212. Available at: https://doi.org/10.1080/13658810802549154.

Van Genuchten, M.T. (1980) 'A closed-form equation for predicting the hydraulic conductivity of unsaturated soils', *Soil science society of America journal*, 44(5), pp. 892–898.

Vose, D. (2000) *Risk analysis: a quantitative guide*. Second Edition. Chichester: John Wiley & Sons.

Wainwright, J., Parsons, A.J., Cooper, J.R., Gao, P., Gillies, J.A., Mao, L., Orford, J.D. and Knight, P.G. (2015) 'The concept of transport capacity in geomorphology', *Reviews of Geophysics*, 53(4), pp. 1155–1202.

Wang, Y., Wu, N., Kunze, C., Long, R. and Perlik, M. (2019) 'Drivers of Change to Mountain Sustainability in the Hindu Kush Himalaya', in P. Wester, A. Mishra, A. Mukherji, and A.B. Shrestha (eds) *The Hindu Kush Himalaya Assessment: Mountains, Climate Change, Sustainability and People*. Cham: Springer International Publishing, pp. 17–56. Available at: https://doi.org/10.1007/978-3-319-92288-1_2.

Watson, D.F. (1985) 'A refinement of inverse distance weighted interpolation', *Geo-processing*, 2, pp. 315–327.

Weedon, G., Gomes, S., Viterbo, P., Shuttleworth, W.J., Blyth, E., Österle, H., Adam, J., Bellouin, N., Boucher, O. and Best, M. (2011) 'Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century', *Journal of Hydrometeorology*, 12(5), pp. 823–848.

Weedon, G.P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.J. and Viterbo, P. (2014) 'The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data', *Water Resources Research*, 50(9), pp. 7505–7514.

Welch, B.I. (1947) 'The generalisation of student's problems when several different population variances are involved.', *Biometrika*, 34(1–2), pp. 28–35. Available at: https://doi.org/DOI: 10.1093/biomet/34.1-2.28.

Westerberg, I.K., Guerrero, J.-L., Younger, P.M., Beven, K.J., Seibert, J., Halldin, S., Freer, J.E. and Xu, C.-Y. (2011) 'Calibration of hydrological models using flow-duration curves', *Hydrology and Earth System Sciences*, 15(7), pp. 2205–2227. Available at: https://doi.org/10.5194/hess-15-2205-2011.

Wheater, H.S. (2002) 'Progress in and Prospects for Fluvial Flood Modelling', *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 360(1796), pp. 1409–1431.

Wiche, G.J. and Holmes, R.R. (2016) 'Chapter 13 - Streamflow Data', in T.E. Adams and T.C. Pagano (eds) *Flood Forecasting*. Boston: Academic Press, pp. 371–398. Available at: https://doi.org/10.1016/B978-0-12-801884-2.00013-X.

Wijaya, I.P.K., Towashiraporn, P., Joshi, A., Jayasinghe, S., Dewi, A. and Alam, M.N. (2023) 'Climate change-induced regional landslide hazard and exposure assessment for aiding climate resilient road infrastructure planning: a case study in Bagmati and Madhesh Provinces, Nepal', in *Progress in Landslide Research and Technology, Volume 1 Issue 1, 2022*. Springer, pp. 175–184.

Wijngaard, R.R., Lutz, A.F., Nepal, S., Khanal, S., Pradhananga, S., Shrestha, A.B. and Immerzeel, W.W. (2017) 'Future changes in hydro-climatic extremes in the Upper Indus, Ganges, and Brahmaputra River basins', *PLOS ONE*, 12(12), p. e0190224. Available at: https://doi.org/10.1371/journal.pone.0190224.

Wilcock, P.R. and Crowe, J.C. (2003) 'Surface-based transport model for mixed-size sediment', *Journal of hydraulic engineering*, 129(2), pp. 120–128.

Williams, R.D., Brasington, J. and Hicks, D.M. (2016) 'Numerical Modelling of Braided River Morphodynamics: Review and Future Challenges', *Geography Compass*, 10(3), pp. 102–127. Available at: https://doi.org/10.1111/gec3.12260.

Winiger, M., Gumpert, M. and Yamout, H. (2005) ‘Karakorum–Hindukush–western Himalaya: assessing high-altitude water resources’, *Hydrological Processes*, 19(12), pp. 2329–2338. Available at: https://doi.org/10.1002/hyp.5887.

WMO (1989) *Statistical distributions for flood frequency analysis*. Technical Report WMO-No. 718. Geneva,Switzerland: World Meteorological Organization, p. 118. Available at: https://library.wmo.int/idurl/4/33760 (Accessed: 5 September 2023).

WMO (2008) *Hydrology – From Measurement to Hydrological Information*. Technical Report WMO-No. 168. Geneva,Switzerland: World Meteorological Organization, p. 296. Available at: https://library.wmo.int/idurl/4/35804.

World Bank (2023) ‘Population data’. Available at: https://data.worldbank.org/indicator/SP.POP.TOTL (Accessed: 11 September 2023).

Wu, Y., Miao, C., Fan, X., Gou, J., Zhang, Q. and Zheng, H. (2022) ‘Quantifying the uncertainty sources of future climate projections and narrowing uncertainties with bias correction techniques’, *Earth’s Future*, 10(11), p. e2022EF002963.

Xu, J. (2004) ‘Comparison of hydraulic geometry between sand-and gravel-bed rivers in relation to channel pattern discrimination’, *Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group*, 29(5), pp. 645–657.

Yan, K., Pappenberger, F., Umer, Y.M., Solomatine, D.P. and Di Baldassarre, G. (2014) ‘Regional versus physically-based methods for flood inundation modelling in data scarce areas: an application to the Blue Nile’.

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N. and Kitoh, A. (2012) ‘APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges’, *Bulletin of the American Meteorological Society*, 93(9), pp. 1401–1415.

Ziliani, L. and Surian, N. (2016) ‘Reconstructing temporal changes and prediction of channel evolution in a large Alpine river: the Tagliamento River, Italy’, *Aquatic sciences*, 78, pp. 83–94. Available at: https://doi.org/doi:10.1007/s00027-015-0431-6.

Ziliani, L., Surian, N., Botter, G. and Mao, L. (2020) ‘Assessment of the geomorphic effectiveness of controlled floods in a braided river using a reduced-complexity numerical model’, *Hydrology and Earth System Sciences*, 24(6), pp. 3229–3250. Available at: https://doi.org/10.5194/hess-24-3229-2020.

Ziliani, L., Surian, N., Coulthard, T.J. and Tarantola, S. (2013) ‘Reduced-complexity modeling of braided rivers: Assessing model performance by sensitivity analysis, calibration, and validation’, *Journal of Geophysical Research: Earth Surface*, 118(4), pp. 2243–2262. Available at: https://doi.org/10.1002/jgrf.20154.

Zurick, D.N. (1988) ‘Resource needs and land stress in Rapti zone, Nepal’, *The Professional Geographer*, 40(4), pp. 428–444. Available at: https://doi.org/10.1111/j.0033-0124.1988.00428.x.