Twenty-first century glacial and hydrological change in the Virkisjökull Glacier Observatory, Iceland

by

Jonathan Duncan Mackay

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

School of Geography, Earth and Environmental Sciences
College of Life and Environmental Sciences
University of Birmingham
September 2019

ABSTRACT

This thesis uses climate, glacio-hydrological models (GHMs) and groundwater models to advance understanding of: 1) twenty-first century climate change impacts on glacier-fed river flow regime and proglacial groundwater dynamics at the Virkisjökull Glacier Observatory in Iceland; and 2) uncertainties associated with model projections which underpin this understanding. The research is split into three studies. Study 1 tests a novel, signature-based Limits of Acceptability framework for constraining structural uncertainties in GHMs. The framework successfully identifies deficiencies in different melt and runoff-routing model structures, but cannot identify a population of acceptable model structures. Study 2 uses an ensemble of regional climate projections and GHMs to project changes in 25 characteristics (signatures) of river flow regime up to 2100. The results show that the magnitude, timing and variability of river flow are sensitive to climate change and that projection uncertainties stem from incomplete knowledge of future climate and glacio-hydrological processes. The dominant uncertainty source, however, is signature-specific. Study 3 includes a proglacial groundwater model into the climate-GHM model chain and shows that climate change will perturb intra-annual groundwater level timing and variability leading to changes in groundwater-surface water interactions. Uncertainties in groundwater projections primarily stem from future climate uncertainty.

ACKNOWLEDGEMENTS

This research was supported by a NERC studentship awarded via the Central England NERC Training Alliance. It would not have been possible without the guidance from my supervisors including Nick Barrand, David Hannah and Stefan Krause (University of Birmingham), Jez Everest and Chris Jackson from the British Geological Survey (BGS) and Tolly Aðalgeirsdóttir from the University of Iceland (UoI).

This thesis builds on the legacy of data collection at the Virkisjökull Glacier Observatory led by the BGS since 2009. Countless staff members and collaborators have collected and analysed field data to develop new catchment process understanding which underpins much of the modelling work presented in this thesis. Their willingness to share these data and their expertise has made this research possible. For that, I am extremely grateful. In particular, Andrew Black (University of Dundee) provided guidance on river gauging and also contributed to the development of the river gauging QA strategy which undoubtedly improved the reliability of the modelling results. Outcomes from lively discussions and debates with Alan MacDonald and Brighid Ó Dochartaigh (BGS) about the role of groundwater in the catchment hydrology fed directly into the formulation of the groundwater model used in chapter 6. My fieldwork campaigns would not have been possible without the assistance of Heiko Buxel (BGS) and our local hosts at Svínafell, Pálína Þorsteinsdóttir and Óli Sigurðsson.

Other researchers who kindly provided advice and shared data include Þóranna Pálsdóttir and Tómas Jóhannesson (Icelandic Meteorological office) who provided weather station data from the national network. Louise Steffensen Schmidt (UoI) provided ICRA precipitation data. Joaquin Munoz-Cobo Belart (UoI) provided the historic ice DEMs. Jim Freer and Gemma Coxon (University of Bristol) and Ida Westerberg (IVL Swedish Environmental Research Institute) provided guidance on implementing the limits of acceptability approach used in chapter 4. Finally, I would like to thank my family, friends and colleagues for their continued support over the last four years.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

ANOVA analysis of variance.

ASG automatic stream gauge.

AWS automatic weather station.

CDF cumulative distribution function.

CMIP Coupled Model Intercomparison Project.

CORDEX COordinated Regional Climate Downscaling EXperiment.

DEM digital elevation model.

DS downscaling.

DTGWL depth to groundwater level.

ECDF empirical cumulative distribution function.

EQM equidistant quantile mapping.

FDC flow duration curve.

GCM general circulation model.

GHM glacio-hydrological model.

GPR ground penetrating radar.

GW-SW groundwater-surface water.

ICRA atmospheric reanalysis project for Iceland.

IMO Icelandic Meteorological Office.

IPCC Intergovernmental Panel on Climate Change.

LIA Little Ice Age.

LOA limits of acceptability.

MA moving average.

MODIS Moderate Resolution Imaging Spectroradiometer.

OAT one at a time.

PDF probability density function.

RCM regional climate model.

RCP representative concentration pathway.

RES radio echo sounding.

RMSE root mean squared error.

ROR runoff-routing model.

TIM temperature-index model.

VGO Virkisjökull glacier observatory.

CHAPTER 1:

INTRODUCTION

1.1 Background and rationale

Mountain watersheds are referred to as the worlds water towers (Viviroli and Weingartner, 2004; Viviroli et al., 2007), partly because they receive large quantities of precipitation relative to adjacent lowlands, but also because they regulate runoff through the accumulation and melt of snow and ice, providing a stable source of fresh water to downstream end users. It is estimated that more than one-sixth of the Earth's population uses meltwater from mountain glaciers and snow for their water supply (Barnett et al., 2005). Of these, 370 million people live in river basins where glacier meltwater alone contributes at least 10% of seasonal river discharge (Schaner et al., 2012).

Meltwater from the Himalayas contributes 60% of surface runoff to the Indus River basin, providing energy (hydroelectric power) and food (irrigation) security to the basin's 215 million inhabitants (Immerzeel et al., 2010). Glacial meltwater from the tropical Andes provides 15% of the total domestic and industrial water supply to the urban centres of La Paz and El Alto in Bolivia and Huaraz in Peru (Buytaert et al., 2017; Soruco et al., 2015). It has also facilitated the development of the large-scale Chavimochic and Chinecas agricultural projects which receive more than 50% of their water supply from glacier melt during 10-year return period rainfall droughts (Buytaert et al., 2017; Carey et al., 2014). In the European Alps and Northern Europe, hydroelectric dams are used to exploit the reliable supply of mountain runoff provided by the seasonal melt of snow and ice. In Norway and Iceland, meltwater contributes 20% and 91% of the total hydroelectric output respectively (Milner et al., 2017). Glaciers also provide an important water source in regions where water-related tensions between upstream and downstream users are high, such as in the central Asian states of the Aral Sea basin where meltwater from the Tian Shan and Pamir mountains underpin regional food and energy security (Orlove, 2009).

Meltwater rivers emanating from glaciated mountain systems also house some of the world's most pristine freshwater ecosystems which include diverse populations of microbial and algal communities, macroinvertebrates and fish (Jacobsen et al., 2012; Milner et al., 2017). These rivers typically have a high beta-diversity (along stream) and gamma-diversity (regional) due to the combined water inputs from snow and ice melt and other water sources such as rainfall and groundwater which result in highly heterogeneous physiochemical water properties (Brown and Hannah, 2008; Brown et al., 2006; Ward et al., 1999). In the Gulf of Alaska, the World's most productive wild salmon fisheries are dependent on the buffering effect of seasonal meltwater runoff inputs to rivers which help to maintain river flow regulate water temperature during the summer months (Beamer et al., 2017; Schoen et al., 2017).

Given the pivotal role of meltwater in providing socio-economic and ecological well-being to downstream communities and ecosystems, there is rising concern over the potential impacts of projected glacier and snow line retreat over the twenty-first century (Carey et al., 2016). The most recent analysis of global glaciological and geodetic datasets showed that between 2006 and 2016, the mass balance of glaciers has consistently been negative with a global glacier mass loss of 335 $\pm$ 144 Gt y^{-1} (equivalent to 0.92 $\pm$ 0.39 mm y-1 sea level rise) (Zemp et al., 2019). Snow coverage is also receding globally, with the largest reductions in the Northern Hemisphere where June snow coverage receded by almost 15% per decade between 1979 and 2012 (Vaughan et al., 2013). This global snow and ice retreat is directly linked to rising air temperature and the most recent Intergovernmental Panel on Climate Change (IPCC) assessment indicates that mean global surface air temperature will be 0.6-4.4 °C warmer by the end of the century compared the recent past (1986 and 2005) (Collins et al., 2013). Air temperature rises are expected to be up to 4.5 times the global average in high-latitude regions due to reductions in surface albedo near the poles as snow and ice continue to retreat (Gosseling, 2017). Recent global glacier projection studies indicate that total global glacier volume will reduce by 25-64% by 2100 (Huss and Hock, 2015; Radić et al., 2014; Shannon et al., 2019).

A priority for the scientific community, therefore, is to facilitate the development of

appropriate adaptation strategies that reduce vulnerability and increase resilience of populations and ecosystems that will be affected by changes in water cycling due to glacier and snow line retreat over the coming decades (Vuille, 2015). Key to developing these strategies is an ability to forecast probable changes in strategically important hydrological stores and fluxes e.g. snow, ice, river flow and groundwater storage over equivalent timescales which must be underpinned by a sound scientific understanding of the processes driving changes in glaciated river basin water cycles (Buytaert and Domzalski, 2015). Mathematical computer models of global climate (IPCC, 2018) and models that simulate glacial and hydrological (glacio-hydrological) processes (Schaefli et al., 2014; Schulla, 2015) provide a means to generate these forecasts with mathematical equations that are based on our best approximation of system behaviour. Climate and glacio-hydrological model projections are at the forefront of industrial (IHA, 2019; Thorsteinsson and Björnsson, 2012) and intergovernmental water resource planning for glaciated mountain regions (ICIMOD, 2010; UNESCO, 2017). Typically, climate model projections are used as boundary conditions to drive glacio-hydrological models (GHMs) that simulate the accumulation and ablation of snow and ice as well as catchment water cycling (e.g. Garee et al., 2017; Huss and Hock, 2018; Shea and Immerzeel, 2016). These model projections underpin current understanding of glaciological and hydrological change in glaciated watersheds, but this understanding remains incomplete. For example, almost all projection studies in glaciated regions have focussed on projecting changes in surface runoff and downstream river discharge, but ignore how the propagation of meltwater through the wider hydrosphere might impact other potentially important hydrological stores such as groundwater even though these are known to form an important component of water cycling in glaciated catchments (Buytaert et al., 2017; Vincent et al., 2019). Additionally, though, these projections of glacial and hydrological change are subject to considerable uncertainty due to inadequacies in the formulation, parameterisation and boundary conditions of the models, yet these are rarely quantified adequately which undermines their robustness and prevents the dissemination of uncertainties to decision makers (Buytaert and Domzalski, 2015; Vuille, 2015).

This thesis aims to address these challenges by utilising state-of-the-art climate projections with numerical GHMs and groundwater models to project twenty-first century glacial and hydrological change at the Virkisjökull glacier observatory (VGO), a glaciated catchment in southern Iceland. Specifically, the thesis will assess changes in glacier-fed river flow regime and downstream groundwater dynamics. The VGO was selected for the focus of this research for a number of reasons. Firstly, a range of observation data have been collected at the site that can be used to parameterise and evaluate computer models with. These data cover key drivers and fluxes of glacial and hydrological change in the catchment including climate, glacier mass balance, meltwater runoff and proglacial groundwater level dynamics. Secondly, the VGO includes a proglacial alluvial aquifer which forms a significant groundwater store and is known to interact with meltwater channels (MacDonald et al., 2016; Ó Dochartaigh et al., 2019) making it ideal for investigating the propagation of meltwater through the wider hydrological system. There are also key societal benefits to studying glacial and hydrological change in Iceland. Tourism is one of three key economic sectors in Iceland of which glaciers are a major attraction (Jóhannesson and Huijbens, 2010; Welling et al., 2015). Meltwater runoff from Iceland's glaciers also forms an important national resource, particularly for the hydropower industry which meets more than 70% of Iceland's electricity demand (Jóhannesson et al., 2007). Finally, flood waters from glacier-fed rivers can cause major disruption to Iceland's road network (Björnsson and Pálsson, 2008; Jóhannesson et al., 2006).

1.2 Research gaps

Throughout this thesis, four primary research gaps are identified and addressed to advance understanding of glacial and hydrological change in glaciated watersheds under climate change and advance understanding of uncertainty associated with model projections which inform this understanding.

Research gap 1: A framework for GHM intercomparison and selection

A major challenge in developing water supply adaptation policies in glaciated watersheds is accounting for the uncertainties in future projections of hydrological variables which are known to be large for glaciated river basins (Huss et al., 2014). These uncertainties stem, in part, from an incomplete understanding of the glacio-hydrological processes that govern the evolution of these watersheds (e.g. glacier mass balance Jobst et al., 2018; Ragettli et al., 2013) and an inability to represent them at the level of detail that these processes are known to operate at (Beven, 2016). This has led to the development of an array of different simplified model structures that purport to simulate the same glacio-hydrological processes, but which adopt different process assumptions and levels of complexity that are typically more reflective of observation data availability and computational resource constraints. Model intercomparison studies undertaken have shown no clear approach to distinguish which (if any) model is appropriate for the purpose of making projections (Gabbi et al., 2014; Irvine-Fynn et al., 2014; Reveillet et al., 2017). The choice of appropriate model structures and parameterisations for key glacio-hydrological processes, therefore, remains a key source of hydrological projection uncertainty in glaciated river basins. *There is a need for a model selection framework that better discriminates between competing model structure hypotheses so that these uncertainties can be better constrained and quantified.*

Research gap 2: Twenty-first century changes in glacier-fed river flow regime

Water cycling in glaciated mountain regions is complex and distinctive from other types of watersheds. Water is accumulated and stored as snow and ice and released according to melt cycles which operate over diurnal, seasonal and decadal timescales (Jansson et al., 2003). Liquid water may also be stored in snow and firn or in glacier crevasses, fractures, tunnels and cavities (Cuffey and Paterson, 2010). The drainage efficiency of these reservoirs is highly dynamic and typically increases through the melt season as meltwater develops preferential flow pathways through snow and ice (Machguth et al., 2018; Rada and Schoof, 2018; Willis, 2005). The loss of this unique water storage behaviour under

twenty-first century climate change could, therefore, cause changes to the magnitude, timing and variability of downstream river flows over a range of timescales which will impact the energy sector (Carvajal et al., 2017; Gaudard et al., 2014; Laghari., 2013) and agriculture (Baraer et al., 2015; Carey et al., 2014; McDowell and Hess, 2012; Nolin et al., 2010). These changes could also threaten the sustainability of sensitive alpine river ecosystems (Bunn and Arthington, 2002; Clausen and Biggs, 2000; Naiman et al., 2008; Puckridge et al., 1998). However, an analysis of the range of potential characteristic shifts in river flow regime has yet to be undertaken. *There is a need to project the range of characteristic changes in glacier-fed river flow regime that might be expected under twenty-first century climate change.*

Research gap 3: Sources of uncertainty in glacier-fed river flow regime projections

Such projections of glacier-fed river flow regime will be uncertain, partly due to inadequacies in GHMs, but also due to gaps and errors in the observation data used to drive and evaluate models, an issue which is especially important in relatively inaccessible mountain catchments, but rarely considered in practice (Beniston et al., 2017). The issue of uncertainty will be exacerbated when considering climate change forcing given the imperfections in the structure and boundary conditions of climate models (Giorgi et al., 2009). Quantifying the propagation of uncertainty from these different sources in the model chain would provide a basis for assigning more robust levels of confidence to river flow regime projections. Such analyses have been undertaken in the past (Addor et al., 2014; Giuntoli et al., 2015; Vetter et al., 2017, e.g.), but none have investigated uncertainty across a full range of characteristic changes in river flow regime. *There is a need to quantify the contribution to uncertainty in projections of different characteristics of glacier-fed river flow regime from different sources in the model chain.*

Research gap 4: Twenty-first century changes in proglacial groundwater dynamics

Water drainage dynamics in glaciated mountain watersheds are also controlled by variable geological conditions which may include low permeability mountain bedrock and higher permeability deposits such as moraines and alluvium (Wilson and Guan, 2004). These high permeability materials can form significant aquifers in the foreland areas of glaciated catchments (La Frenierre and Mark, 2014; Vincent et al., 2019). Some recent studies have shown that proglacial aquifers can receive significant recharge from meltwater runoff emanating from neighbouring mountains as it flows along river channels which actively exchange water with the aquifer (Liljedahl et al., 2017; Ó Dochartaigh et al., 2019). Other studies have shown that, further downstream, groundwater baseflow to glacier-fed rivers can exceed contributions from meltwater (Andermann et al., 2012; Kobierska et al., 2015). In this respect, aquifers serve to slow down the transfer of meltwater to rivers and can therefore help to sustain river flow outside of the melt season (Andermann et al., 2012; Jódar et al., 2017). This process of storage and release could serve to buffer seasonal changes to the magnitude and timing of glacier-fed river flows due to glacier retreat (Markovich et al., 2016) and could also make groundwater stores more strategically important in the absence of significant ice and snow stores in the future (Taylor, 2013). Even so, future changes in water storage in proglacial aquifers, the contribution of groundwater to glacier-fed river flow and how these changes might be linked to changes in runoff from neighbouring glaciated watersheds has yet to be investigated. *There is a need to project how groundwater storage dynamics and groundwater-surface water (GW-SW) exchanges will evolve under twenty-first century climate change in proglacial aquifers.*

1.3 Research objectives

To address each of the identified research gaps, the thesis has the following research objectives:

1. Implement a novel GHM comparison and selection framework to undertake a rigor-

ous evaluation of multiple GHM structures.

2. Implement a climate-GHM model chain to project twenty-first century changes in different characteristics of river flow regime.

3. Determine the relative contribution of individual model chain components to uncertainty in river flow regime projections.

4. Incorporate a distributed groundwater model into the climate-GHM model chain to project twenty-first century changes in proglacial groundwater level dynamics and GW-SW exchanges.

These research objectives are closely aligned with three of the most important scientific problems in hydrology identified in a recent community-led initiative involving 230 scientists (Blöschl et al., 2019) (Fig. 1.1).

1.4 Thesis structure

Following on from this introduction, chapter 2 introduces the VGO study site. Chapter 3 presents the overarching research design and provides information on the models and datasets used to meet the research objectives. Chapters 4, 5 and 6 contain the majority of the research undertaken in this thesis and each serves to address the identified research gaps. Finally, chapter 7 concludes the research by synthesising the main findings from the thesis before identifying areas for future work.

1.5 Summary

This chapter has introduced the main research gaps and objectives identified for this thesis. A more detailed background review and justification for the identified gaps and objectives is provided in each of the corresponding main research chapters. The next chapter presents an overview of VGO study site.

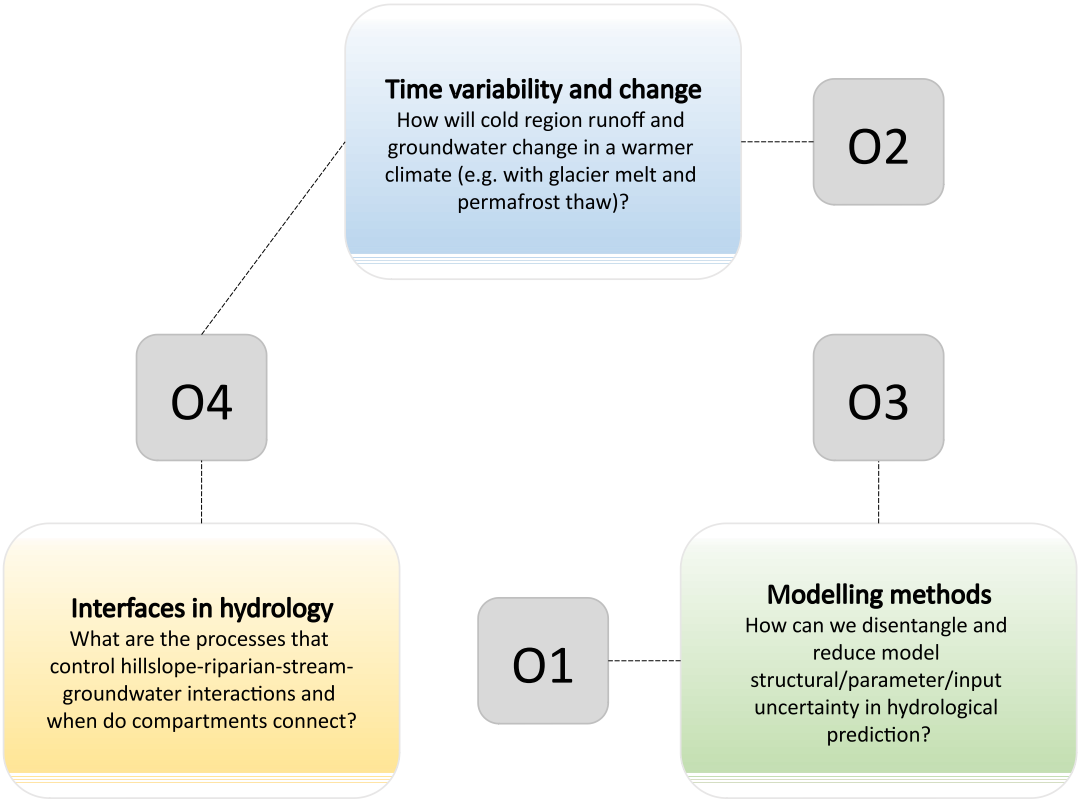

Figure 1.1: Links between thesis objectives (grey boxes) and important scientific problems in hydrology (coloured boxes) identified by Blöschl et al. (2019).

CHAPTER 2:

VIRKISJÖKULL GLACIER OBSERVATORY

2.1 Introduction

The VGO is situated in south-east Iceland on the southern edge of Europe's largest ice cap, Vatnajökull (Fig. 2.1a) and has been operated and funded by the British Geological Survey since 2009, with additional funding provided by the UK Natural Environment Research Council under a Capital Grant. The observatory was initially set up to investigate local climate drivers of short-term landscape evolution of the proglacial foreland, but then evolved to study past, present and future glacier evolution and its impact on proglacial hydrology and hydrogeology. It includes the glaciated Virkisá River basin which is situated on the western side of the high relief, ice-capped Öræfajökull stratovolcano. It also includes a proglacial groundwater catchment situated on a lowland sandur floodplain which forms a significant groundwater store (Ó Dochartaigh et al., 2019) and is hydrologically connected to the upstream mountainous river basin by the Virkisá River (MacDonald et al., 2016).

A range of instrumentation have been installed at the site including three off-ice automatic weather stations, six on-ice GPS stations, an ablation stake network, four seismic stations, an automatic stream gauge and eight peizometers in the proglacial groundwater catchment. These data in conjunction with additional data collected from subannual field campaigns have provided extensive conceptual understanding of the glacial and hydrological behaviour of the VGO including: 1) glacier fluctuations from the early-mid Holocene to present day (Bradwell et al., 2013; Everest et al., 2017); 2) mechanisms for the recent (post-2000) rapid retreat of the glacier (Phillips et al., 2014); 3) patterns of glacier ice ablation (Flett, 2016); 4) meltwater transport through the glacier (Flett et al., 2017; MacDonald et al., 2016); and 5) interactions between meltwater runoff and sandur groundwater (MacDonald et al., 2016; Ó Dochartaigh et al., 2019). This chapter describes

key characteristics of the VGO and draws on research previously undertaken at the study site.

2.2 Geology and topography

The VGO sits approximately 50 km south-east of Iceland's active rift zone and drains an area of 22.2 km^2, 60% of which is glacier-covered (Fig. 2.1c). The bedrock geology is dominated by very young (0.78 Ma), normally magnetised subglacial basic and intermediate volcanic rocks from the Brunhes chron (Roberts and Gudmundsson, 2015). The topography is controlled by the steeply sided conical form of Öræfajökull which rises from near sea level to the east, south and south-west, where it is bounded by steep cliffs, up to the Hvannadalshnjúkur summit, Iceland's highest peak, at approximately 2100 m asl. At the summit, a 14 km^2 ice-capped caldera exists, the edge of which forms the uppermost boundary of the Virkisá basin. This caldera forms the main ice accumulation zone from which ice flows radially downslope via nine outlet glaciers. Two of these glaciers, Virkisjökull and Falljökull, which together comprise a twin-lobed outlet glacier (herein referred to as Virkisjökull) of the Öræfajökull ice cap, provide one of the primary drainage channels for accumulated ice at the summit (Björnsson and Pálsson, 2008) (see Fig. 2.2a). Virkisjökull has been in a phase of retreat since 1991 (IGS, 2017) and the immediate foreland in the vicinity of the terminus is characterised by extensive areas of remnant glacier ice buried under outwash sand and gravels which are 1–4 m in thickness. This zone of buried ice now forms the bed of an expanding proglacial lake (Fig. 2.2b). The forelands in the vicinity of the lake support nested moraines composed of sand, gravel and boulders, formed by small glacier fluctuations over seasonal and shorter timescales since 1990 (Everest et al., 2017). Broader zones of sharp-crested moraines up to 80 m high extend from the lower reaches of Öræfajökull onto the sandur floodplain to the south. These moraines are composed of well-rounded cobbles, sand, gravel and silt and are thought to have formed in the last 200-300 years which would coincide with the Little Ice Age (LIA) maximum c. 1870-1890 (Everest et al., 2017; Thompson, 1988). Further south still are

a number of isolated moraine mounds which probably relate to an early-mid Holocene advance (5000–6000 BP) (Everest et al., 2017). These have well developed soils and are vegetated, unlike most of the Virkisá River basin and surrounding area where the steep-sided valley walls and the relatively recent glacial maximum at the end of the LIA mean that there is limited soil development. Where thin soils have developed, vegetation is dominated by mosses, sparse grass and shrubs such as dwarf willow and birch.

The proglacial sandur (Fig. 2.2a) is an active glacial outwash floodplain which forms part of the world's largest sandur: Skeiðarársandur (Fig. 2.1b). The sandur is made up of deposits from actively shifting meltwater streams as well as frequent (c. 5 per century) jökulhlaups - glacial outburst floods - which have distributed huge quantities of loosely consolidated, moderately to poorly sorted, dominantly medium- to coarse-grained glaciofluvial sand, gravel and cobbles across the floodplain (Robinson et al., 2008). Two of the largest jökulhlaups were caused by explosive eruptions of Öræfajökull in 1362 and 1727, the only recorded eruptions in recent history. Here, flows as high as 1×10^5 m^3 s^{-1} inundated the foreland with sediment-rich meltwater (Roberts and Gudmundsson, 2015). Geophysical evidence from Tromino® passive seismic surveys indicate that 2 km downstream of the lake outlet, the aquifer is between 60 and 100 m thick which then increases to between 100 and 150 m thick 3.5 km downstream of the lake outlet. The Virkisá groundwater catchment is estimated to store between 1-2% of the water stored in the Virkisjökull glacier (Ó Dochartaigh et al., 2019).

2.3 Virkisjökull

Virkisjökull is a retreating glacier (Bradwell et al., 2013; IGS, 2017; Phillips et al., 2014) with a high mass balance gradient (Björnsson et al., 1998; Björnsson and Pálsson, 2008; Flett, 2016) and is one of the highest mass turnover glaciers in Europe (Dyurgerov et al., 2002). Mass balance measurements on the glacier are limited. Annual snow cores taken at the summit of Öræfajökull between 1993 and 1998 showed net accumulation of up to 7.8 m w.e. y^{-1} (Guðmundsson, 2000) while ablation stake measurements collected

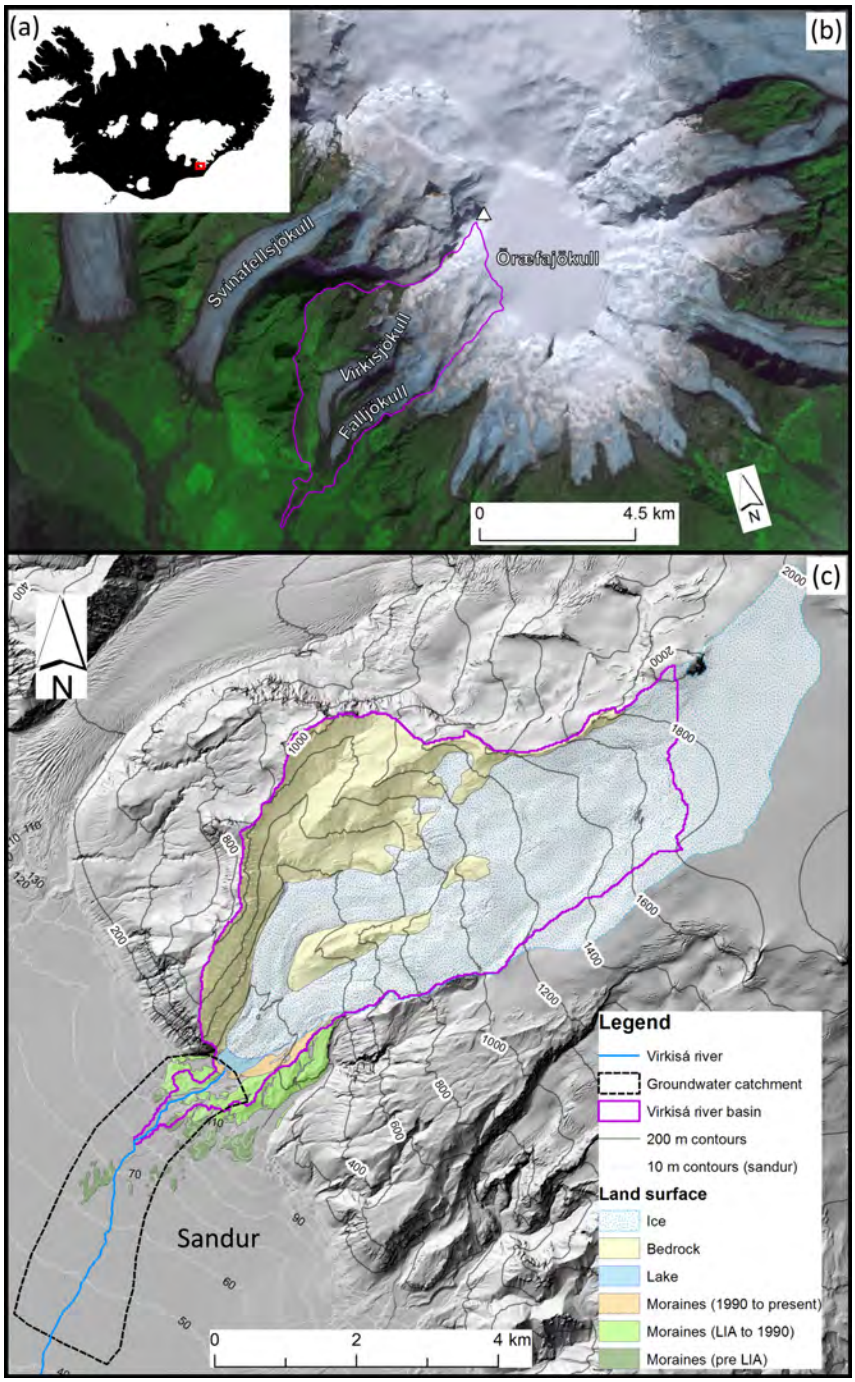

Figure 2.1: Location of the VGO in Iceland (a); on Öræfajökull (b); and detailed topographical map of study area including the Virkisá River and its corresponding basin area, major land surface types and the approximate Virkisá groundwater catchment adapted from Ó Dochartaigh et al. (2019) (c). Note the basin area here was derived using the hydraulic potential gradient method as applied previously to outlet glaciers of Vatnajökull (Bjornsson, 1982).

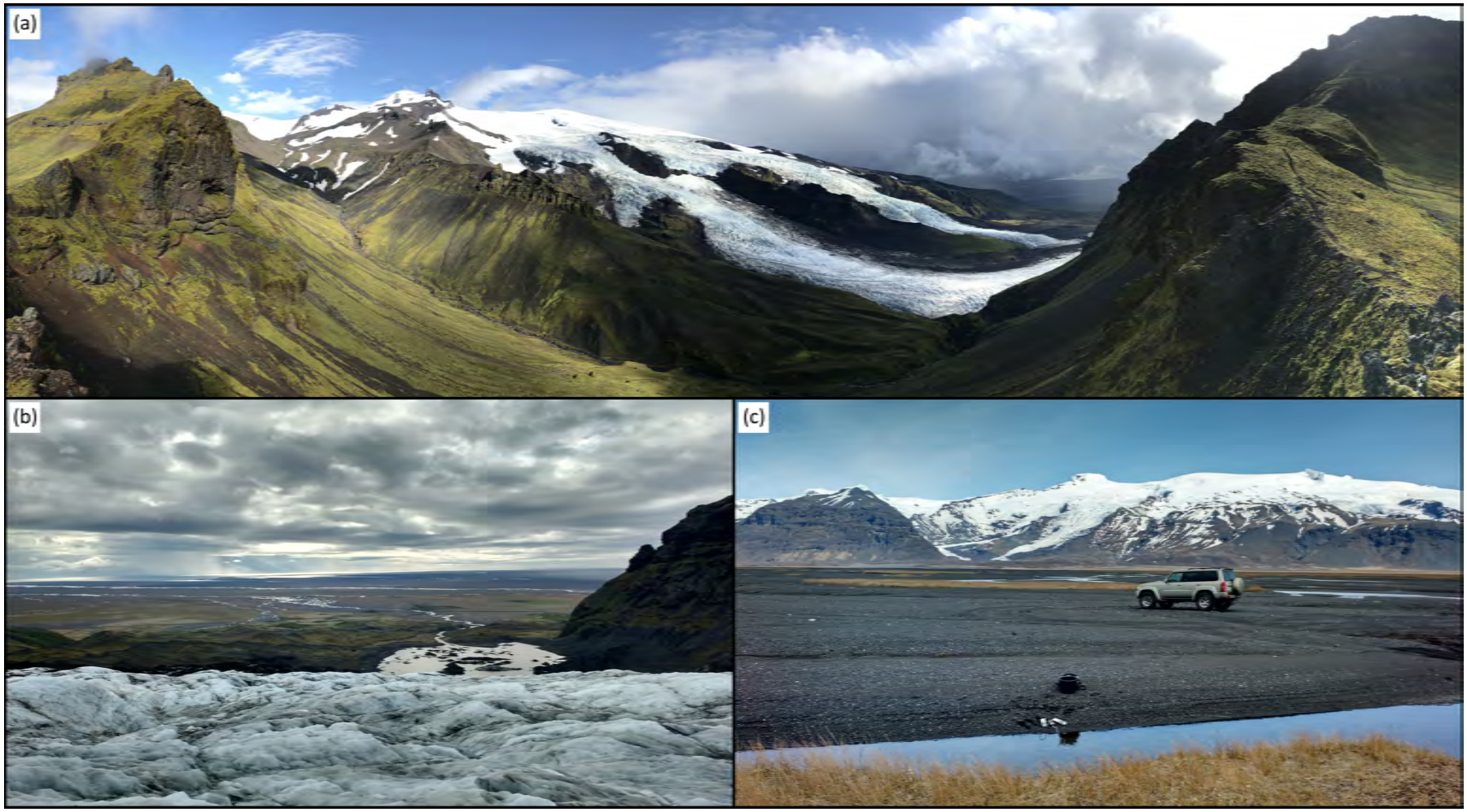

Figure 2.2: Photos of VGO including: panorama of Virkisjökull glacier looking south-east from western boundary of the Virkisá River basin, August 2018 (a); proglacial lake feeding the Virkisá River with surrounding moraine fields and Skeiðarársandur in the distance looking south-west from the glacier tongue, August 2016 (b); and lower boundary of the VGO looking north-east towards Virkisjökull from the sandur floodplain.

between 2012 and 2014 showed annual ice melt of 4.8 m w.e. y^{-1} below the ice fall up to 8.9 m w.e. y^{-1} near the ice margin (Flett, 2016). This high mass balance gradient is typical of many glaciers in Iceland and makes them some of the most sensitive to climate warming in the world where mass balance sensitivities range from -0.6 to -3 m w.e. $°C^{-1}$ (Björnsson et al., 2013). Records of specific mass balance for Vatnajökull show a shift from approximately 1.0 m w.e. to -2 m w.e. between 1992 and 2010, a period in which mean annual air temperature rose by approximately 1.0 °C (Björnsson et al., 2013). Inter-annual variability in mass balance of Iceland's glaciers are also known to be influenced by other factors including snowfall which can dramatically reduce melt rates by increasing surface albedo as well as volcanic eruptions where ash deposition has the opposite effect (Björnsson et al., 2013).

Bradwell et al. (2013) used high resolution photography along with geomorphological interpretation of the moraine field and lichenometric dating to derive a detailed history of Virkisjökull length variations extending back to 1912. They found evidence for two historical phases of retreat (1935-1945 and 1990-present) which correspond to periods of rapid climate warming. The most recent phase of retreat and its relation to air temperature can be seen in Fig. 2.3a which shows the annual frontal movement of the Falljökull arm of the glacier since 1958. From 2005, the rate of retreat has increased substantially, exceeding 30 m y^{-1} for all but one year and showing a record annul retreat of 110 m y^{-1} in 2015. Bradwell et al. (2013) noted that no annual push moraines have been formed since 2005 indicating that this recent phase of rapid retreat could be linked to some change in ice flow behaviour. Phillips et al. (2014) installed a series of GPS sensors along both arms of Virkisjökull to investigate the current ice flow regime. They found that while the accumulated ice at the summit of Öræfajökull continues to flow down slope at a rate of 72 m y^{-1} on average, the ice at the front of the glacier has become detached and is now in a phase of passive down wasting. Furthermore, fractures in the ice indicate that the continued forward motion of accumulated ice is now being thrust over the immobile ice in the terminal zone. The reason for this detachment is not known, although Phillips et al. (2014) note evidence of a previously pressurised subglacial hydrological network that is

no longer active.

Figure 2.3b shows the most recent complete ice thickness distribution of Virkisjökull derived by Magnússon et al. (2012) and Fig. 2.3c show the two corresponding cross sections of ice and bedrock topography through along the Virkisjökull and Falljökull arms of the glacier. The glacier shows several regions with distinct topography. At the summit where ice elevations exceed 1800 m asl, the ice is relatively flat and greater than 400 m thick in the deepest sections of the ice-filled caldera. This planar summit forms the main accumulation zone for the glacier ice which then drains in a south-westerly direction down the steep sides of Öræfajökull with a slope between 0.15 and 0.5 in the steepest sections. This fast flowing, relatively thin band of ice experiences high longitudinal stresses resulting in a highly crevassed section approximately 4 km long (see Fig. 2.2a). Along this section the glacier splits at the surface into two distinct arms as it flows around a bedrock ridge which is the principal source of distinctive supraglacial debris bands that hug the ridge and accumulate at the front of the glacier (see Fig. 2.1b). From the base of the crevassed section, the glacier flows through an overdeepened valley which, like many of the outlet glaciers that drain the Öræfajökull ice cap, have been progressively carved into the bedrock over time. Here ice is greater than 100 m thick.

2.4 Climate

Iceland is situated just south of the Arctic Circle between 63 and 67 degrees north and experiences a maritime climate characterised by cool summers and mild winters with year-round precipitation. The Icelandic climate is dictated, to a large degree, by the meeting of cold and warm air masses at the polar front (Einarsson, 1984). Here, the North Atlantic Current flows along the southern and western coast bringing with it mild Atlantic air which comes into contact with cold Arctic air from the north. At this front, a semi-permanent low pressure system exists known as the Icelandic low, which is associated with frequent cyclone activity, particularly in the winter when contrasts between tropical and polar air masses are largest. These cyclones typically approach from the south-west,

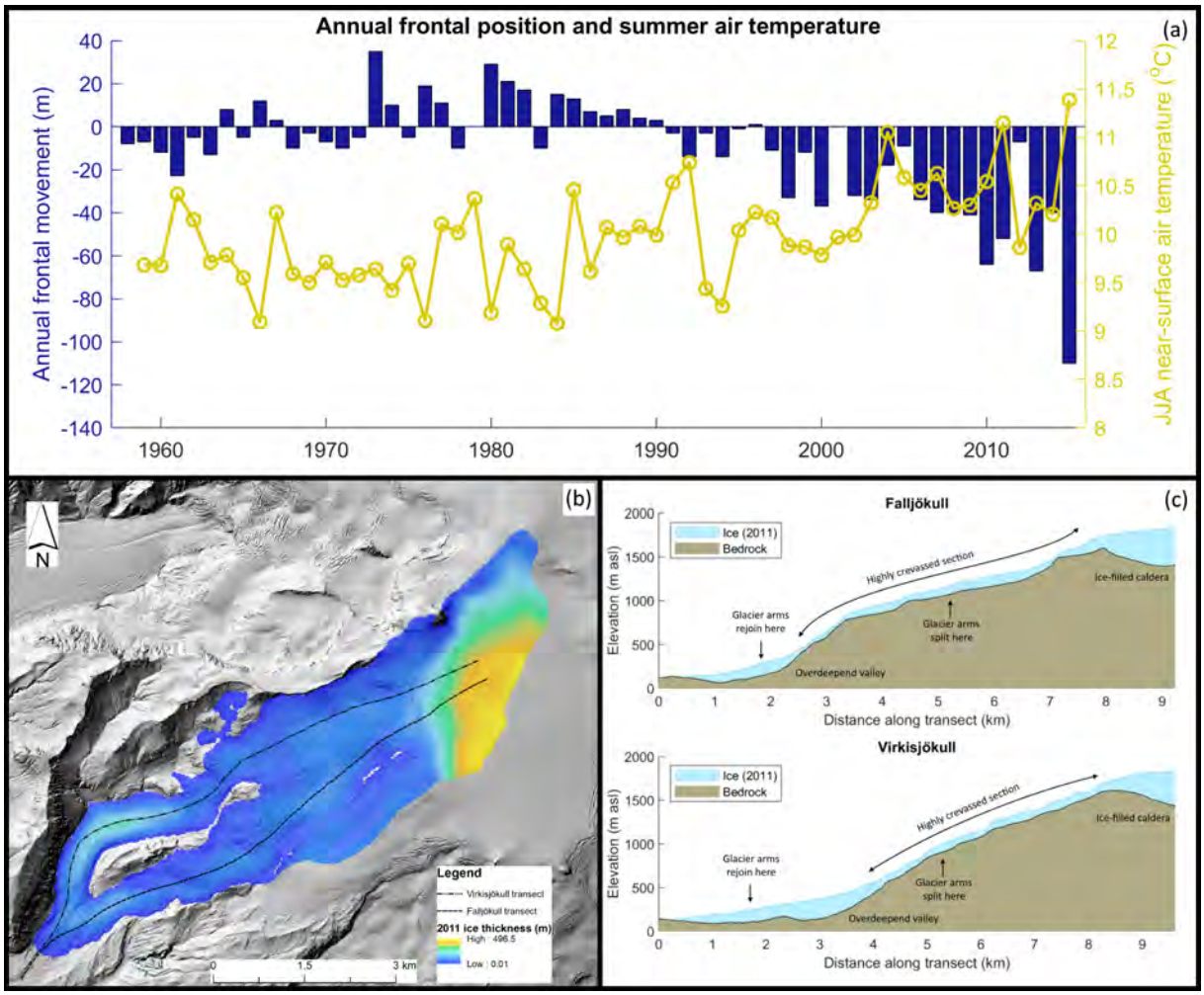

Figure 2.3: Historic observations of Virkisjökull geometry including: annual evolution of the Falljökull front position between 1958 and 2015 (IGS, 2017) (a); ice thickness distribution in 2011 (b); and cross-section of ice and bedrock topography along transects through the Virkisjökull and Falljökull arms of the glacier.

bringing with them strong winds and intense rainfall. As such, the VGO is situated in one of the wettest and windiest parts of Iceland.

Figure 2.4a shows the mean annual total precipitation distribution over the VGO taken from the atmospheric reanalysis project for Iceland (ICRA) data (further details provided in section 3.4.1). The prevailing north-easterly wind, which blows moist air over Öræfajökull, induces a strong lateral precipitation gradient due to orographic lift with >8000 mm y^{-1} of precipitation falling at the summit of Öræfajökull and <1500 mm y^{-1} falling in the sandur region. Seasonally, the precipitation is substantially lower between April and August than the rest of the year (blue bars in Fig. 2.4b).

Mean monthly air temperature on the sandur (yellow line, Fig. 2.4b) shows that at lower elevations, temperature peaks in July (10.7 °C) and is lowest in January (0.7 °C). Regional variations in near-surface air temperature in Iceland are controlled by topography to a large extent (Crochet and Jóhannesson, 2011). Analysis of weather station data conducted by Flett (2016) indicates that the average temperature lapse rate at the VGO is -0.44 °C 100 m^{-1}.

2.5 Hydrology and hydrogeology

The Virkisá River flow is controlled to a large extent by the seasonal input of water from snow and ice melt. Flett (2016) analysed measurements of river flow, ice melt and snow line location to characterise seasonal variations in water inputs to the river. They conclude that the main melt season occurs between May and September, but that snow melt likely contributes to runoff year-round given that the snow line elevation rarely exceeds 1200 m asl and the occurrence of above-freezing days throughout the year at lower elevations. Indeed, analysis of ^{2}H and ^{18}O stable isotope compositions of river water revealed glacier melt is the dominant source of water, accounting for up to 80% of wintertime river flow (MacDonald et al., 2016).

The highest flows in the Virkisá River occur between June and August (8.7 m^3 s^{-1} on average) when near-surface air temperature is highest and the lowest flows occur in

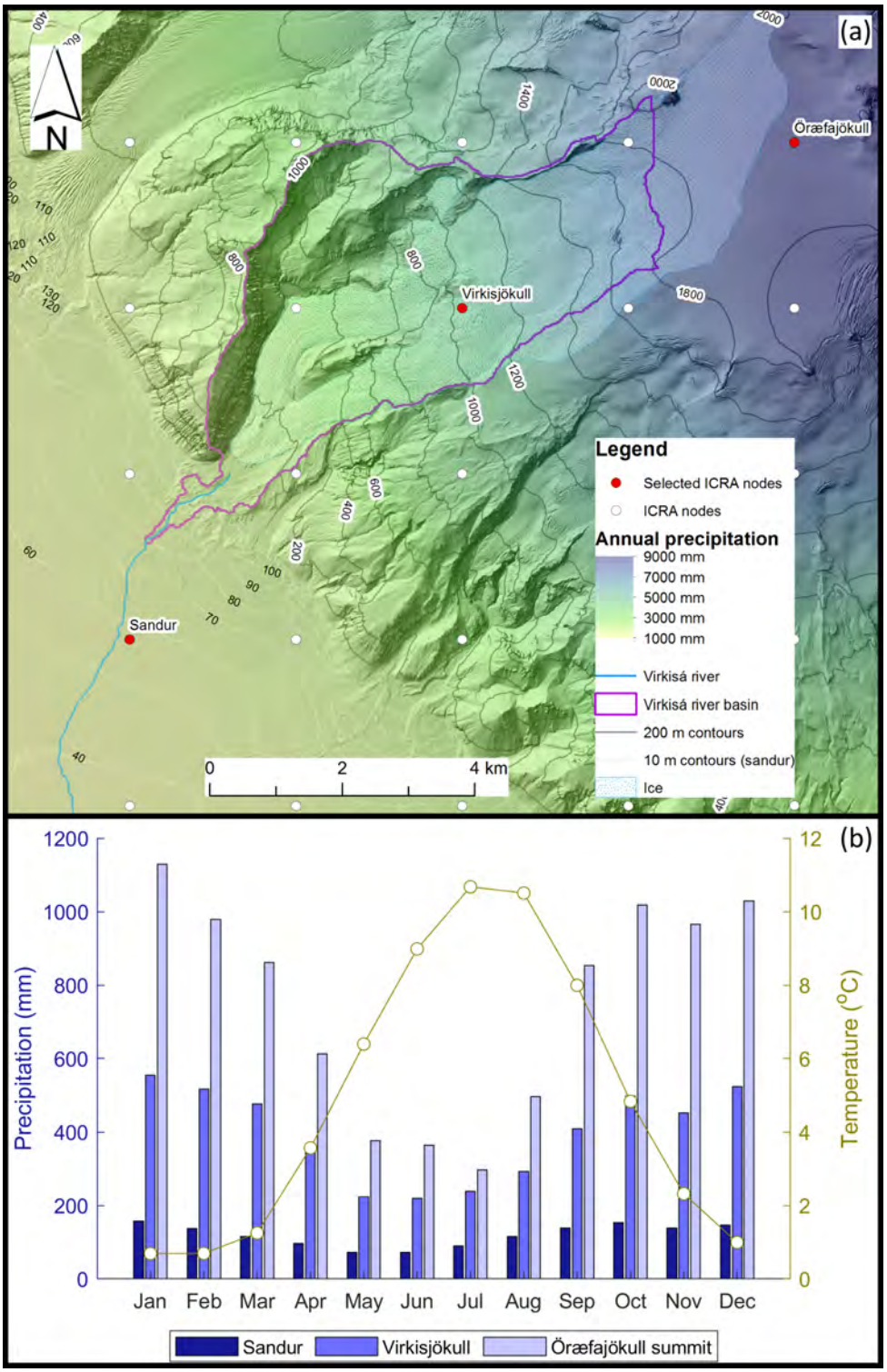

Figure 2.4: Mean annual total precipitation distribution (1980-2016) over the VGO (a) and monthly precipitation distribution (blue bars) at three selected locations (red dots in a) and monthly mean temperature (yellow line) from the Fagurhólsmýri weather station situated on the sandur 12 km south of the study site at 10 m asl (b).

early-spring between March and April (1.7 m^3 s^{-1} on average) before the initiation of the main melt season. River flow responds rapidly to runoff events with an average lag time between peak rainfall and runoff of 10.5 h (Flett, 2016). This rapid response time can be partly explained by the catchment's small size. Additionally, ground penetrating radar (GPR) surveys conducted on the glacier ice and remnant buried ice immediately in front of the glacier indicate the presence of an extensive conduit network (Phillips et al., 2013; Phillips et al., 2014) and subsequent tracer tests from moulins in the main ablation zone beneath the ice fall indicate that this conduit network provides an efficient drainage system for runoff with mean flow velocities of 0.58 m s^{-1} at the end of the melt season (Flett et al., 2017). Hydrograph recession analysis has also shown that this drainage system remains open throughout the year, providing efficient drainage channels that can accommodate large runoff inputs (MacDonald et al., 2016). Accordingly, the Virkisá River is highly responsive to daily and even sub-daily variations in rainfall and melt throughout the year.

The proglacial lake at the glacier terminus forms the headwater of the Virkisá River, the only drainage pathway for melt and rainfall from the Virkisá River basin. The river flows south-westerly, firstly through 800 m of bedrock controlled section which is flanked on either side by moraines. Beyond this, the river continues its route towards the coast over the vast and gently sloping sandur floodplain: a shallow, unconfined aquifer with an average surface gradient of 0.017. Groundwater in the aquifer is recharged by rainfall, snow melt and river losses. The water table typically resides between 0 and 4.4 m of the ground surface promoting interactions between groundwater and surface features. Indeed, an extensive water sample field campaign using stable isotopes ^{2}H and ^{18}O to distinguish between meltwater and precipitation indicates that glacial meltwater influences groundwater within 500 m of the river (see groundwater catchment in Fig. 2.1c) with more than 25% of groundwater in this part of the aquifer sourced from meltwater (Ó Dochartaigh et al., 2019). Additionally, groundwater provides a reliable source of runoff to the river ∼2 km downstream of the lake outlet both through direct exchange at the river bed and via a large network of springs and seeps that discharge back to the river. In winter, when

meltwater runoff is relatively low, 15–20% of river water in this lower section of sandur originates from groundwater (MacDonald et al., 2016).

Pumping tests conducted at eight boreholes drilled into the top 15 m of sandur estimate transmissivity between 100-2500 m^2 d^{-1} with a median value of 600 m^2 d^{-1}. Additionally, surface permeability has been estimated from Guelph permeameter experiments and particle size analysis which showed a median surface hydraulic conductivity of 35 m d^{-1} with an interquartile range of 25 - 40 m d^{-1} (Ó Dochartaigh et al., 2019). The permeability of the deeper sandur aquifer has not been measured. The bedrock is poorly fractured and pumping tests indicate its transmissivity is very small ($<$0.25 m^2 d^{-1}).

2.6 Summary

This chapter has given an overview of the key characteristics of the VGO study site including the geological setting, the historic evolution of Virkisjökull, the dominant climate and current understanding of the hydrological and hydrogeological behaviour of the catchment based on site investigations undertaken since 2009. The next chapter introduces the overall research design of this thesis and provides more detail on available observation data for the VGO and the model codes used to meet the research objectives identified in chapter 1.

CHAPTER 3:

RESEARCH DESIGN, DATA AND MODELS

3.1 Introduction

The research undertaken in the subsequent chapters uses a variety of environmental data including observation data obtained by direct measurement in the field, remotely sensed field data and climate model outputs. These data are used to drive and evaluate simulations from two principal model codes which are used throughout this thesis. They include one glacio-hydrological model code, GHM++ and one groundwater model code, MODFLOW-NWT. This chapter presents the research design which outlines how the models and data have been used to meet the research objectives identified in Chapter 1. It then provides more detail of the model codes and environmental datasets which have been used to undertake the research in this thesis.

3.2 Research design

Chapters 4, 5 and 6 contain the bulk of the experimental research undertaken in this thesis, each of which has been designed to meet the identified research objectives and to feed newly gained knowledge and newly generated datasets directly into the subsequent chapters (Fig. 3.1).

Research objective 1 (chapter 4): Implement a novel GHM comparison and selection framework to undertake a rigorous evaluation of multiple GHM structures

Chapter 4 presents a model intercomparison study which compares the efficiency of different structural representations of two key glacio-hydrological processes: i) snow and ice melt; and ii) runoff-routing which are implemented in the GHM++ glacio-hydrological

model code (detailed below). This chapter implements a limits of acceptability (LOA) framework in which the glacio-hydrological characteristics of the study site are characterised using 33 metrics (signatures) derived from historic observations of ice melt, snow coverage and river discharge. The uncertainty of each set of observations is harnessed to define LOA, a set of criteria used to objectively evaluate the efficiency of different GHM structures and parameterisations and identify model deficiencies. Specifically, this chapter evaluates different structural representations of snow/ice melt and runoff-routing.

Research objective 2 (chapter 5): Implement a climate-GHM model chain to project twenty-first century changes in different characteristics of river flow regime

Chapter 5 draws on the findings from chapter 4 and uses the signature-based LOA framework to define an ensemble of the most behavioural GHM structural configurations and parameterisations which are then driven with an ensemble of state-of-the-art climate projections to project twenty-first century changes in the Virkisá River flow regime. This chapter employs signatures to quantify changes in different characteristics of river flow regime change (magnitude, timing and variability).

Research objective 3 (chapter 5): Determine the relative contribution of individual model chain components to uncertainty in river flow regime projections

Within chapter 5, a decomposition of the projection uncertainties is also undertaken to quantify the relative contributions of different components of the climate-GHM model chain to projection uncertainty across the signatures. In total, five contributors to projection uncertainty are considered including the future greenhouse gas concentrations, climate model and climate projection downscaling. Contributions from snow/ice melt and runoff-routing structure uncertainty in the GHM make up the final two components.

Research objective 4 (chapter 6): Incorporate a distributed groundwater model into the climate-GHM model chain to project twenty-first century changes in proglacial groundwater level dynamics and GW-SW exchanges

Finally, Chapter 6 uses a subsample of the climate and river flow ensemble projections from Chapter 5 as boundary conditions for a distributed MODFLOW-NWT groundwater model of the proglacial sandur. The groundwater model projections are used to evaluate impacts of twenty-first century climate and river flow regime change on proglacial groundwater level dynamics and GW-SW exchanges.

Chapter 7 synthesises the main conclusions from each of the research chapters and identifies research areas for future work.

3.3 Model codes

3.3.1 GHM++ glacio-hydrological model

Key to achieving the research objectives was using a GHM code that is: i) computationally efficient, allowing for the simulation of large ensembles; and ii) easily customisable, allowing for the incorporation of different model structures that can be used interchangeably. It was decided that to guarantee these requirements, a new GHM code should be written. This decision was made for a number of reasons. Firstly, the advantage of having total familiarisation with a model code allows one to customise that code with relative ease while reducing the potential for introducing errors to simulations. Similarly, given the variety of model evaluation data used (see section 3.5 below), a deep understanding of the model code helps to maximise commensurability of model simulations with observation data. Finally, developing a new code allows for the prioritisation of model complexity (and computational demand) for those catchment properties and processes deemed most important, thereby removing potentially redundant and computationally expensive model characteristics.

C++ was deemed the ideal programming language for developing the GHM, firstly

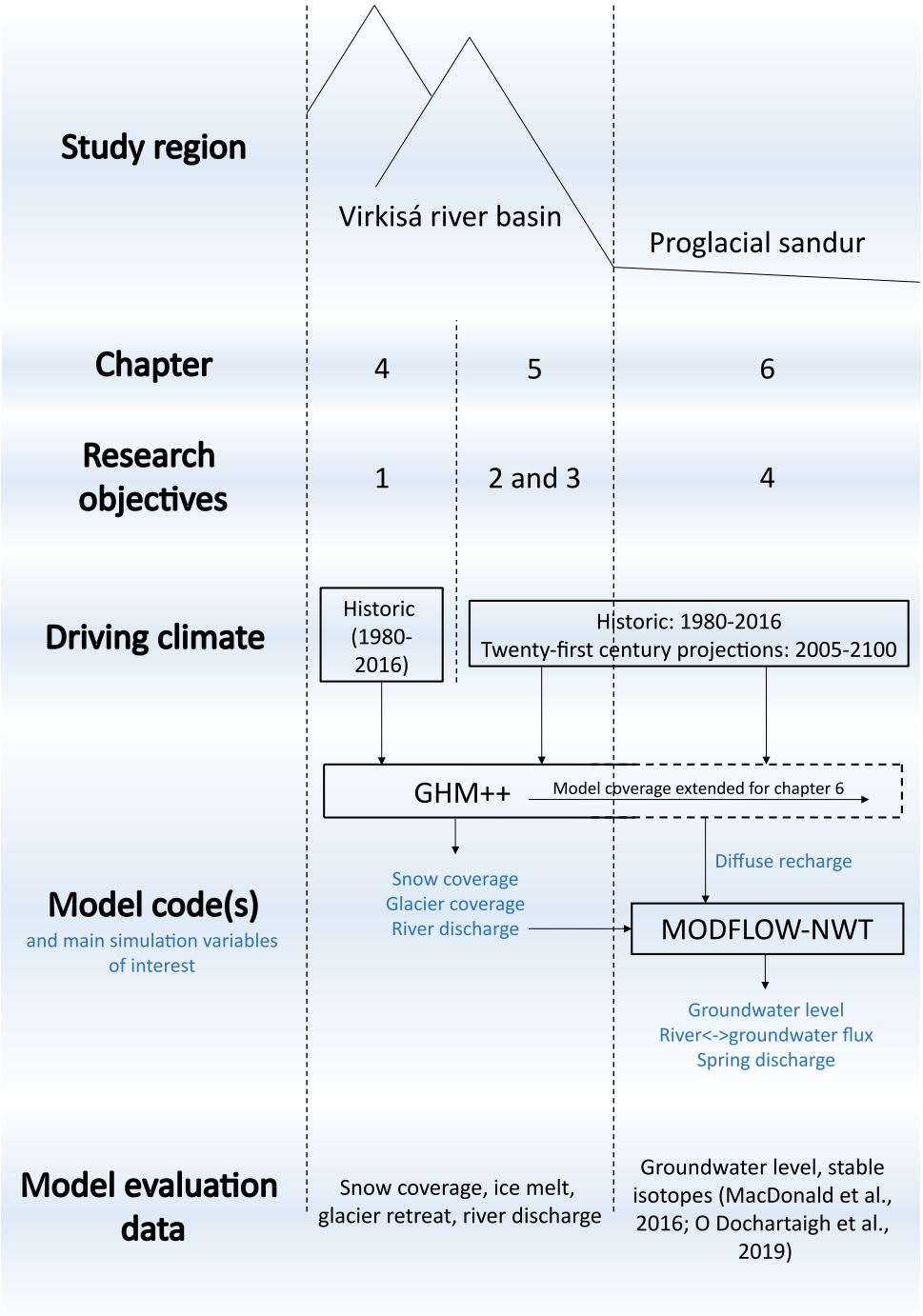

Figure 3.1: Overview of thesis research design which shows how the study region, driving climate and evaluation data and model codes are used across the chapters to meet the research objectives.

because it is relatively low level with small overheads, allowing for one to write highly efficient code with minimal processor and memory demand. Additionally, it uses an object-oriented programming paradigm which allows one to design model codes with a modular structure. By doing so, processes within the model code can be arranged as a series of interacting objects allowing one to easily add or modify individual process representations without the need to extensively alter the overarching code and potentially introduce model errors.

The model code is hereafter refereed to as GHM++. The code is driven by gridded climate time series data and simulates a number of key catchment processes including the accumulation and melt of snow and ice, snow drift, glacier advance and retreat, soil infiltration and evapotranspiration and runoff-routing through the snow, ice and overland. A detailed description of the model structure and equations are provided in Chapter 4 and associated appendices.

3.3.2 MODFLOW-NWT groundwater model

The U.S. Geological Survey MODFLOW-NWT groundwater model code (Niswonger et al., 2011), a variant of the core MODFLOW-2005 code, was used to simulate groundwater level dynamics and GW-SW exchanges in the proglacial sandur. MODLFOW-NWT simulates nonsteady groundwater flow in three dimensions over a distributed grid of cells using a finite difference approach. It can also simulate external stresses including recharge, rivers and surface springs. A key advantage of the MODFLOW family of model codes is their modular structure and use of 'packages' to represent different processes. The ability to switch between packages allows the user to tailor the structure of the model to their specific requirements. Accordingly, the code has been used for a wide range of applications including contaminant transport modelling, parameter estimation and simulation of coupled groundwater/surface water systems (Hariharan and Uma Shankar, 2017; McDonald and Harbaugh, 2003). MODFLOW models have also been applied to cold-region and mountain-fed shallow alluvial aquifers (Allen et al., 2004; Huntington and Niswonger, 2012; Okkonen and Kløve, 2011; Scibek et al., 2007).

MODFLOW-NWT is a Newton-Raphson formulation of the MODFLOW-2005 code which has been designed specifically for simulating unconfined groundwater flow problems. The Newton solver mitigates model convergence problems due to drying and wetting of cells when representing unconfined aquifers where the water table moves between model layers. This improved model stability is particularly beneficial for future projection studies where simulated groundwater levels may deviate outside of the simulation range during historical periods.

3.4 Climate data

Continuous climate data are required as boundary conditions for GHM++ to drive the glacial and hydrological process simulations. A range of data sources were available including historic weather station observations and gridded precipitation data covering the years 1980-2016 inclusive as well as future climate projections up to 2100.

3.4.1 Historic climate data (1980-2016)

Weather station observations

Measurements of climate variables at the VGO are available from three off-ice, telemetered automatic weather stations (AWSs). These are situated at 156 m asl (AWS1), 444 m asl (AWS3) and 805 m asl (AWS4) (Fig. 3.2) and have been operational since September 2009, 2010 and 2011 respectively. They measure meteorological variables including air temperature, rainfall, wind speed and wind direction every 15 minutes (see Appendix A.1 for raw data). AWS1 is also equipped with a cosine-corrected pyranometer which measures incident solar radiation. To measure rainfall, AWS1 uses a tipping-bucket rain gauge, while AWS3 and AWS4 are equip with Vaisala RAINCAP® technology: acoustic sensors that measure the impact of individual raindrops which are subsequently converted to rainfall volume. All data are discontinuous for several reasons. Firstly, hardware has on occasion malfunctioned, especially during extreme weather conditions e.g. strong winds have broken the wind-vanes and anemometers. Secondly, the long, dark and cold

winters can result in loss of power from the solar-powered battery packs. Finally, freezing temperatures result in erroneous precipitation measurements as none of the AWSs are designed to measure snowfall. Additionally, the tipping bucket gauge can accumulate large amounts of snowfall which must melt before it is able to measure rainfall accurately. Similarly, the acoustic sensors are prone to icing over which can affect their operation. Accordingly, all days where the previous three days were not above 0°C were considered unreliable and removed from the database. Figure 3.3 shows the number of days of usable data at each weather station for each month. For air temperature (Fig. 3.3a) and wind speed (Fig. 3.3b) AWS1 has the most usable data due to the fact it was installed one and two years prior to AWS3 and AWS4 respectively. It is also the only station with incident solar radiation data (Fig. 3.3c). For rainfall, AWS4 has the most data during the summer months between June and August, but the prevalence of sub-zero conditions at higher elevations means that it has less than 16 days of usable data for all months between November and April. AWS3 has more days of data during the cooler months, but AWS1 consistently has the most available rainfall data outside of the summer months.

In addition to the AWSs, the Fagurhólsmýri weather station operated by the Icelandic Meteorological Office (IMO) approximately 12 km south of the VGO has daily measurements of temperature dating back to 1949 (Appendix A.2) and therefore provides long-term variations in regional near-surface air temperature.

Gridded precipitation

Given the lack of continuous precipitation data (particularly snowfall) within the study region, and the known orographic influence of Öræfajökull on precipitation patterns, it was deemed necessary to obtain an additional, reliable gridded precipitation product which covers the VGO.

Recently, the IMO initiated the atmospheric reanalysis project for Iceland (ICRA) to produce a historic record of gridded meteorological variables for the whole of Iceland for use in glacier mass balance simulations and hydrological studies (Nawri et al., 2017).

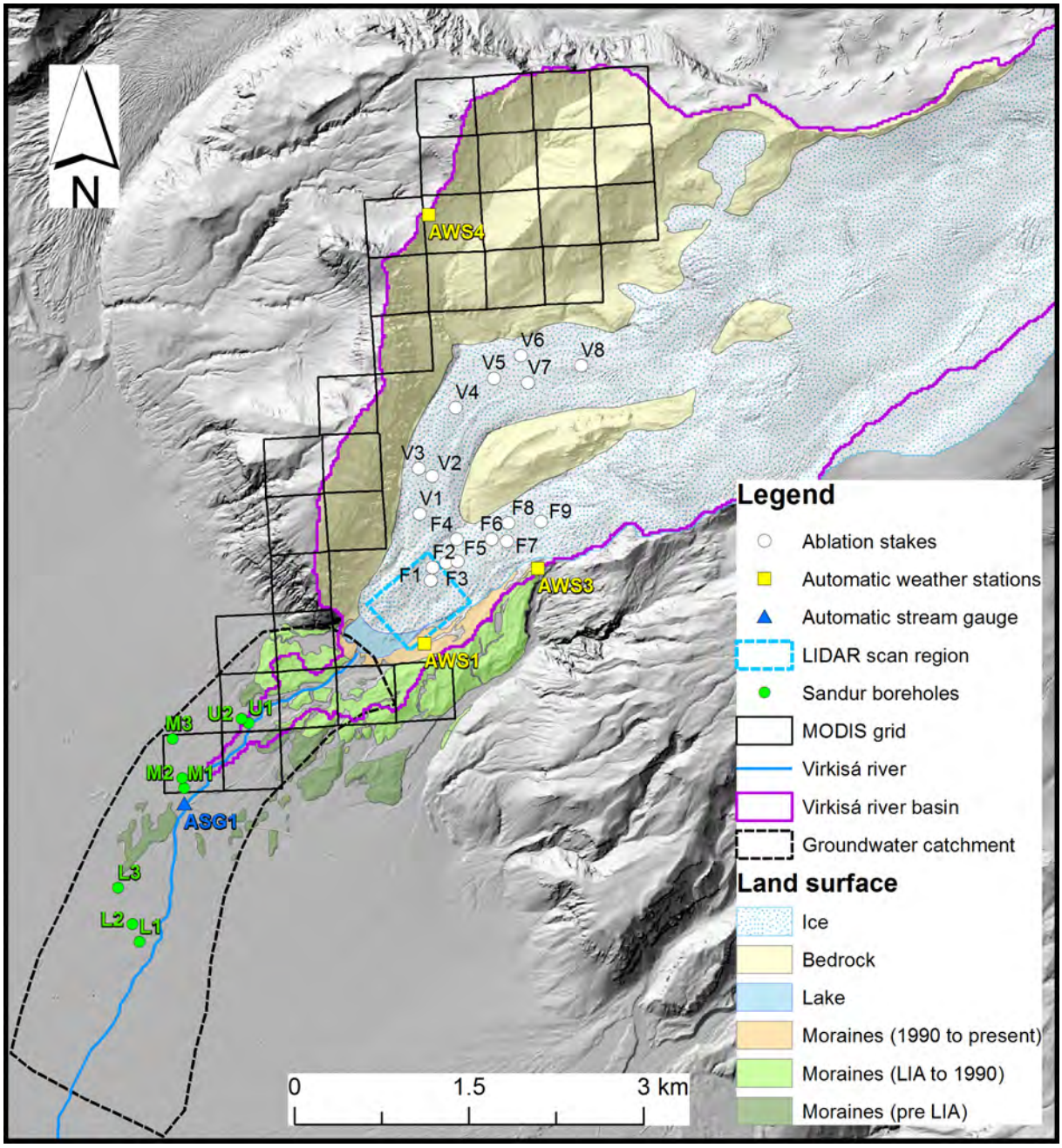

Figure 3.2: Location of primary instrumentation and datasets used throughout the research.

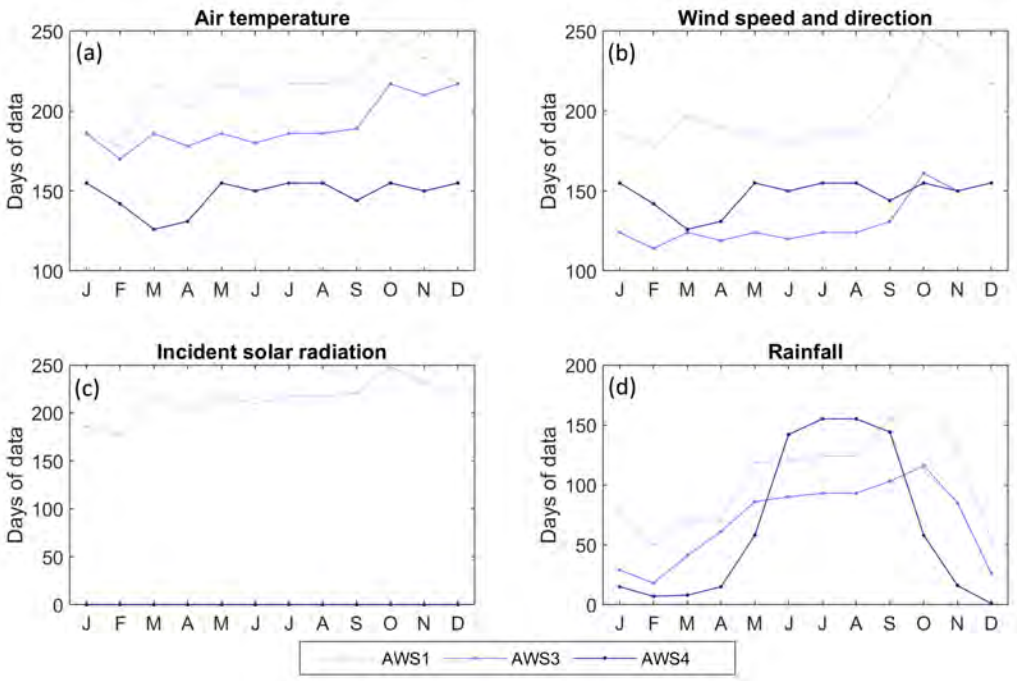

Figure 3.3: Number of days of air temperature (a), wind speed and direction (b), incident solar radiation (c) and rainfall (d) data available for each month from the three AWS installed in the VGO.

These data are produced using the state-of-the-art HARMONIE-AROME mesoscale numerical weather prediction model (Bengtsson et al., 2017) forced by the latest ECMWF ERA-Interim reanalysis product from 1980-2016 inclusive. As part of this they produce hourly precipitation at a spatial resolution of 2.5 km which, when compared to data from 64 weather stations in contrasting topographic settings across Iceland, have shown to match measurements closely and represent a significant improvement over older gridded precipitation products for Iceland (Crochet et al., 2007). These data are therefore considered the most accurate gridded precipitation product currently available (Nawri et al., 2017) and were thus adopted for this study.

It is important to note that Nawri et al. (2017) stress biases may still be large at individual gauges, especially in the winter. Therefore, these data should be used with careful consideration of these biases.

3.4.2 Twenty-first century projections (2005-2100)

Projections of future climate from general circulation models (GCMs) underpin the physical science basis of climate change (Stocker et al., 2013). Since 1995, virtually all climate modelling centres around the world have coordinated their modelling efforts through the Coupled Model Intercomparison Project (CMIP). CMIP provides a community-based infrastructure to support model intercomparison and enable the scientific community to analyse outputs from GCMs in a systematic fashion. As of 2019, the sixth phase of the programme was being undertaken, but the experiments remain incomplete. As such, the experiments conducted under the fifth phase (CMIP5) remain the most up-to-date climate projection data.

A key drawback of using GCM outputs is that the simulations are typically done at a resolution of a few 100 of kms which prevents them from capturing the effects of complex topography and land-surface characteristics. Their performance also deteriorates when looking at higher-order climate statistics such as variability and extremes (Giorgi et al., 2009). To circumvent this problem, regional climate models (RCMs) are increasingly being implemented by the climate modelling community to dynamically downscale GCM

simulations and provide higher-resolution climate information. In 2009, the World Climate Research Programme (WCRP) initiated the COordinated Regional Climate Downscaling EXperiment (CORDEX) (Jacob et al., 2014) to provide a framework to evaluate and benchmark RCM performance and uncertainty . The CORDEX climate projections are based on an ensemble of RCMs forced with an ensemble of GCM simulations from CMIP5 (Taylor et al., 2012) and they currently represent the state-of-the-art for understanding climate change in Iceland (Gosseling, 2017).

Iceland is covered by two of the CORDEX regional model domains: EURO-CORDEX and ARCTIC-CORDEX. A recent review by Gosseling (2017) concludes that the EURO-CORDEX domain is preferable for Iceland as it includes projections at a higher 0.11°C spatial resolution and a larger ensemble of GCM-RCM combinations allowing better representation of climate model uncertainty. Accordingly, the 0.11° EURO-CORDEX data were adopted for this research (Fig. 3.4).

The 0.11° EURO-CORDEX simulations span the years 1950-2100 with simulations up to 2005 constituting the 'recent past' where influences such as atmospheric composition, solar forcing and emissions are imposed based on observations. From 2006, three future representative concentration pathways (RCPs) were imposed on the models which represent different land use and greenhouse gas emission/concentration pathways. These include RCP2.6, RCP4.5 and RCP8.5 which represent an additional radiative forcing by 2100 relative to pre-industrial values of +2.6, +4.5 and +8.5 W m^{-2} respectively (Vuuren et al., 2011). All simulations are available at 3-hourly to 3-monthly resolution; however the 3-hourly simulations were only produced using four GCM-RCMs while daily to seasonal simulations were produced using 15. Given the intent of this study to analyse projection uncertainty, it was decided that the daily data were most suitable. Table 3.1 shows the six GCMs and seven RCMs which make up the 15 unique GCM-RCMs.

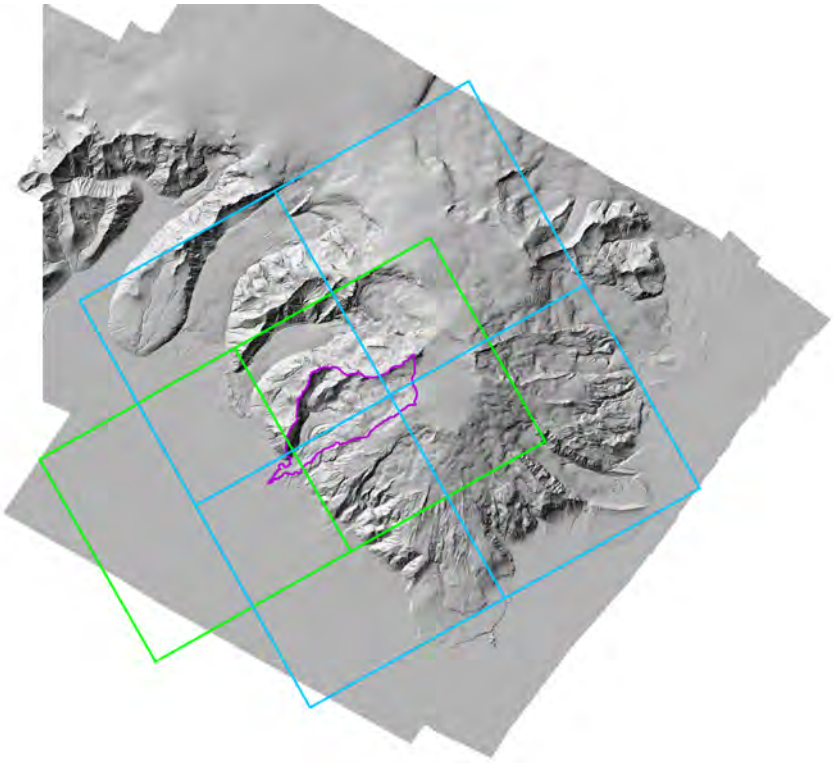

Figure 3.4: EURO-CORDEX 0.11° RCM grid lines. RCM nodes are situated at grid line intersects. All RCMs utilise the green grid except for REMO2009 which uses the blue grid.

Table 3.1: List of GCMs and RCMs used in this study.

Model name	Institution	Type	Driving GCMs
CNRM-CM5	National Centre for Meteorological Research	GCM	-
EC-EARTH	Europe-wide consortium	GCM	-
IPSL-CM5A-MR	Institut Pierre-Simon Laplace	GCM	-
HadGEM2-ES	Met Office Hadley Centre	GCM	-
MPI-ESM-LR	Max Planck Institute for Meteorology	GCM	-
NorESM1-M	Norwegian Climate Center	GCM	-
CCLM4-8-17	Climate Limited-area Modelling Community	RCM	CNRM-CM5, EC-EARTH, HadGEM2-ES, MPI-ESM-LR
ALADIN53	National Centre for Meteorological Research	RCM	CNRM-CM5
RCA4	Swedish Meteorological and Hydrological Institute, Rossby Centre	RCM	CNRM-CM5, EC-EARTH, HadGEM2-ES, MPI-ESM-LR
HIRHAM5	Danish Meteorological Institute	RCM	EC-EARTH, NorESM1-M
RACMO22E	Royal Netherlands Meteorological Institute	RCM	EC-EARTH, HadGEM2-ES
WRF331F	Institut Pierre Simon Laplace and Institut National de l Environnement industriel et des RISques	RCM	IPSL-CM5A-MR
REMO2009	Helmholtz-Zentrum Geesthacht, Climate Service Center, Max Planck Institute for Meteorology	RCM	MPI-ESM-LR

3.5 Observation data for model evaluation

3.5.1 Gridded snow coverage

Gridded snow coverage data from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) (Riggs and Hall, 2015) have been widely used in the evaluation of GHMs (Frey and Holzmann, 2015; Hanzer et al., 2016; Pellicciotti et al., 2012; Valentin et al., 2018). These data have been archived since 2000 as part of the MOD10A1 product and consist of daily 500 m gridded maps of snow cover extent with values ranging between 0 and 1 which relate to the proportion of the ground that is snow covered. The snow cover extent is calculated using the normalized difference snow index (NDSI): a spectral band ratio that utilises differences in infrared and visible spectral bands to identify snow covered features on Earth's surface. This method of determining snow coverage has shown to be robust when compared against observations of snow coverage (Salomonson and Appel, 2004).

As part of the MOD10A1 product, a basic estimate of the data quality is calculated as a means to avoid measurements affected by cloud cover and poor light conditions. For this study, only those data that achieved a QA score of 'good' or 'best' were used. This precluded the use of data collected between September to February presumably because of reduced daylight hours and increased cloud cover during these months. Of the remaining data $<5\%$ achieved a QA score of 'good' or 'best'. Furthermore, the MOD10A1 product cannot distinguish between snow and ice-covered regions, so only data that covered ice-free parts of the catchment have been used (Fig. 3.2) which limited the analysis up to a maximum elevation of just under 1200 m asl.

3.5.2 Ablation

Direct measurements of ablation rates on the glacier were undertaken by Flett (2016) by installing ablation stakes at 17 locations at elevations ranging between 174 - 459 m asl (Fig. 3.2). They periodically measured ablation rates between September 2012 and May

Table 3.2: Mean annual ablation rates measured at 17 ablation stakes between September 2012 and September 2013.

Stake ID	Elevation (m asl)	Ablation rate (m w.e. y^{-1})
F1	174	7.3
F2	185	7.9
F3	202	8.5
F4	213	7.3
F5	218	8.9
F6	283	8.7
F7	304	7.5
F8	315	6.6
F9	359	6.7
V1	217	8.0
V2	251	6.7
V3	252	7.9
V4	321	7.2
V5	358	6.6
V6	378	6.1
V7	387	5.8
V8	459	4.8

2014, capturing the entire 2013 melt season over the lower section of Virkisjökull and recorded annual ice melt of 4.8 m w.e. y^{-1} below the ice fall and 8.9 m w.e. y^{-1} near the ice margin (Table 3.2).

In addition, researchers at the British Geological Survey have undertaken annual high resolution (sub-metre) terrestrial lidar scans of the proglacial region between 2009 and 2016 including ice at the front of the glacier (cyan dashed box in Fig. 3.2). The scans were carried out using two scanners: The Riegl LPM-i800HA which is a medium to long-range scanner which can scan up to 800m with an accuracy of ±15mm and the Riegl VZ1000, a higher precision scanner which is able to scan up to 1200m with an accuracy of ±5mm. Point-cloud data collected from both scanners were oriented using dGPS measurements of the scanner and back-sights positioned on the scan area and then converted into digital elevation models (DEMs) using the InnovMetrics Polyworks® software package. Given that ice flow is negligible at the front of the glacier, these data could provide an additional indicator of ablation rates. Figure 3.5 shows the difference in surface elevation recorded

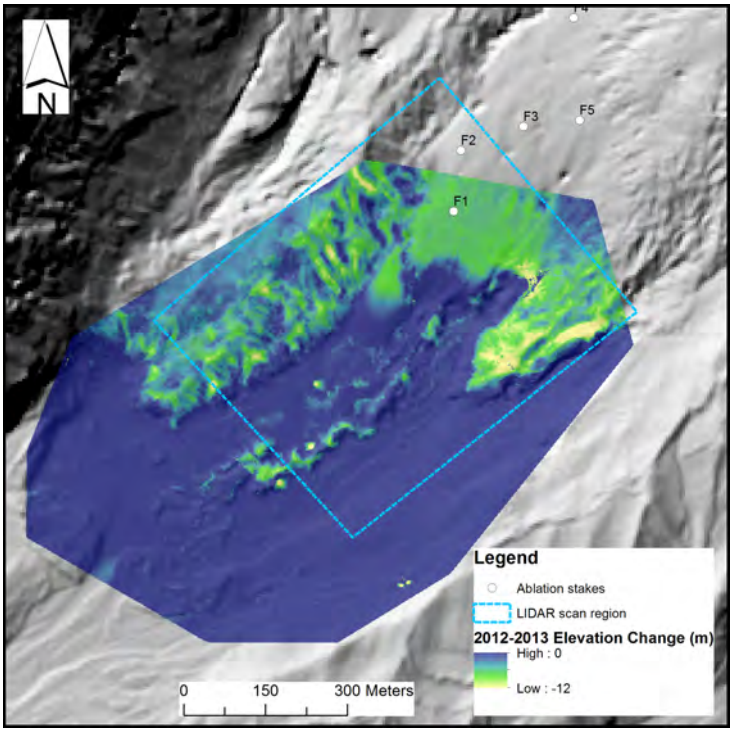

Figure 3.5: Elevation change in proglacial region between two lidar scans taken in September 2012 and September 2013.

between two scans taken in September 2012 and September 2013. When converted to units of m w.e. (assuming an ice density of 900 kg m^{-3}), there is a 0.9 m w.e. y^{-1} (13%) difference between the estimated ablation from the lidar scan and the F1 ablation stake which is likely due to scale differences between the two measurement approaches.

3.5.3 Glacier geometry

Accurate DEMs of Iceland's main glaciers were produced using airborne lidar technology as part of an initiative by the IMO and Institute of Earth Sciences, University of Iceland to establish a baseline for monitoring future changes in glacier geometry. The surveys were carried out between 2008 and 2011 using an Optech ALTM 3100 laser scanner and are available as a 5 m gridded dataset. The scans of Öræfajökull and its outlet glaciers were undertaken in 2011 (Jóhannesson et al., 2013) and they are the most accurate measurement of the ice geometry and surrounding landscape currently available.

Earlier maps of ice surface topography have been derived for some glaciated regions

in Iceland using stereo photogrammetric processing of aerial photography and satellite imagery (Gudmundsson et al., 2011). Hannesdóttir et al. (2015a) constructed six historic (1890-2002) DEMs of Öræfajökull's outlet glaciers using a range of glacial extent indicators including field geomorphology, aerial imagery and differential GPS measurements on the ice surface. They did not, however, include Virkisjökull in their analysis due to limited data availability. More recently, Magnússon et al. (2016) used photogrammetric methods to derive six historic DEMs of the Drangajökull ice cap in Iceland between 1946 and 2005 using aerial images from the National Land Survey of Iceland archives and the Loftmyndir ehf aerial photography company. Upon request, the authors produced an equivalent historic DEM of Virkisjökull from aerial images taken in 1988 at 20 m resolution. The DEM was created by implementing a four step approach in the ERDAS IMAGINE® Photogrammetry software. The steps include: i) orientation of the images; ii) stereo matching; iii) manual edits of the DEMs; and iv) orthorectification of aerial photographs. The orthorectified aerial photographs were then used to define the glacier margin.

In addition to the 1988 and 2011 DEMs, Magnússon et al. (2012) derived a map of Öræfajökull's bedrock topography based on radio echo sounding (RES) profiles and point measurements carried out between 1991 and 2012. Upon request, the authors supplied the map as a digital 20 m raster file. They note that the bedrock elevations of the caldera at the summit of Öræfajökull are well constrained by RES profile data. Point measurements are mainly focussed in the lower parts of Öræfajökull including the relatively flat ablation zone of Virkisjökull near the glacier terminus. For the steep and crevassed regions (including the crevassed sections outlined in Fig. 2.3b), there are no observation data due to access difficulties. Accordingly, Magnússon et al. (2012) used a linear relationship between the slope and ice thickness to define ice thickness in these sections. They note, however, that the linear relationship breaks down where surface slope $<8°$ making ice thickness most uncertain in these areas of the highly crevassed zone.

Figures 3.6a and b show the estimated ice thickness in 1988 and 2011 respectively taken as the difference between the DEM and the bedrock topography. There has been more than 100 m lowering of ice at the front of the glacier between 1988 and 2011 and

almost no change in ice thickness over a large section of glacier covering the summit down to the beginning top of the bedrock ridge that separates the two glacier arms (Fig. 3.6c). The total volume change between 1988 and 2011 was ∼0.3 km^3.

3.5.4 River discharge

River stage measurements

Continuous hourly river stage data are available from an automatic stream gauge (ASG) installed by the British Geological Survey in September 2011. ASG1 is situated at the Icelandic national Route 1 road bridge 2 km downstream of the lake outlet on the Virkisá River (Fig. 3.2). The bridge is built on two central concrete piers with wooden trestle supports at either end (Fig. 3.7a). A central island made up of sandur material has formed around one of the concrete piers (Figs. 3.7a and d) which separates two water-filled channels. Flow is generally constrained to the larger channel (up to 22 m wide) to the east while flow in the smaller channel (up to 15 m wide) to the west is insignificant. It should be noted, however, that as water levels rise, overspill from the main eastern channel feeds the smaller channel to the west. Even so, the pooling of water here is thought to derive mainly from groundwater which maintains surface water temperatures above freezing even when the main channel freezes over (see Fig. 3.7b). During exceptionally high flows, water levels can exceed the central bank material and the entire channel transmits river water.

Stilling wells with submerged pressure transducers and water temperature sensors (log every 15 minutes) are installed on the eastern wooden trestle and on the downstream face of one of the concrete piers (hereafter referred to as the eastern and central stilling well respectively). Additionally, a camera is mounted next to the river which takes photos of the channel three times a day (9:00, 12:00 and 15:00 UTC). Given that the river is prone to freezing over during the winter months, the photographic archive and temperature data were used to remove these periods from the river stage time series.

Significant efforts were made to ensure the reliability of the river stage data which depends both of the time of year (e.g. due to freezing) and flow conditions in the river.

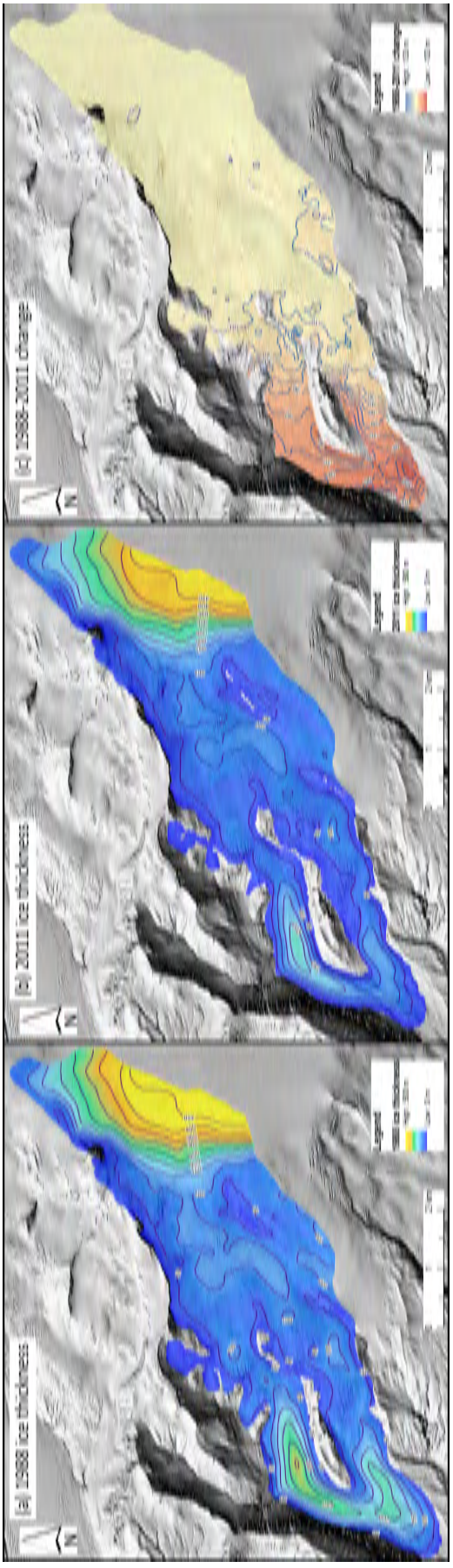

Figure 3.6: Estimated ice thickness in 1988 (a) and 2011 (b) derived by differencing available DEMs and bedrock topography map. Also shown is the change in ice thickness between 1988 and 2011 (c).

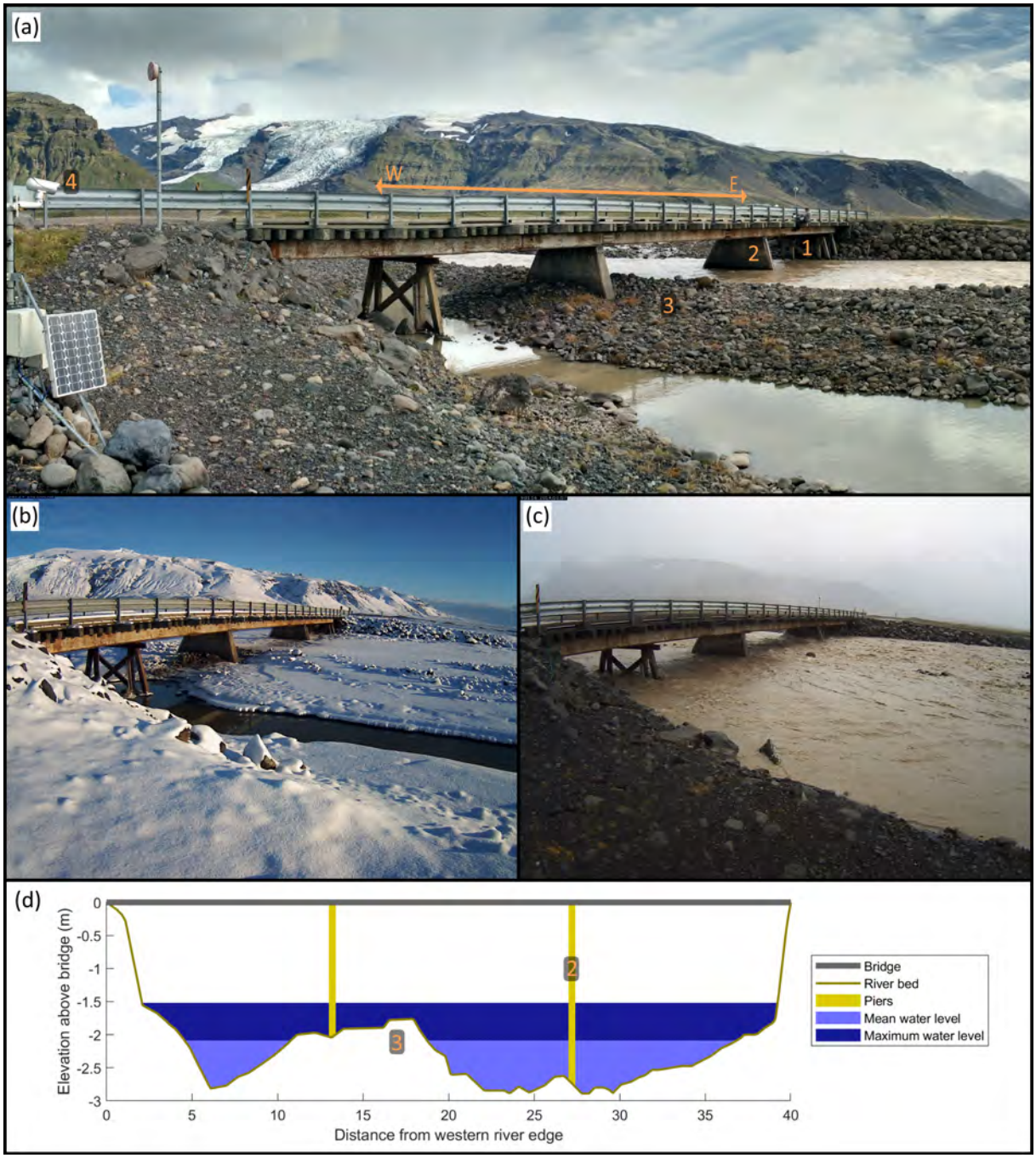

Figure 3.7: Icelandic national Route 1 road bridge where ASG1 is located (a). Annotated are the locations of the eastern (1) and western (2) stilling wells, the central island (3) and the mounted camera (4). Also shown are pictures from the mounted camera during the exceptionally high flows, October 2014 (b), during freezing conditions, February 2015 (c) and a cross section profile of the river bed and water levels at the bridge based on a survey undertaken in March 2015 (d).

The eastern stilling well is reliable during moderate and high flows, but can dry when water levels are low and/or freeze during the winter months. The central stilling well is reliable during low flow and cold periods as it is positioned in the deeper section of river channel, but hydraulic drag effects are observed to cause local lowering of the water level downstream of the pier as flow velocities increase making measurements unreliable during high flows. Accordingly, the flow records were combined based on their known uncertainties to form a single river stage time series (Fig. 3.8a). Prior to harmonising the stage data, a comparison of the logged water levels at both wells revealed discrepancies of up to 40 cm. Accordingly, these discrepancies were analysed in detail and corrected for (see Appendix B). A key conclusion from the analysis was that data prior to 2013 should not be used as the discrepancies between the logged levels at the two stilling wells are difficult to characterise and thus correct for. Additionally, significant bridge maintenance works were undertaken by the highway authority in October 2012 resulting in a major diversion of the main river channel making measurements at this time more unreliable.

River flow gauging

River flow gaugings have been undertaken periodically at the bridge since September 2011 by collaborators at the University of Dundee using the velocity-area method (Herschy, 1999). For the majority of gaugings, an Ott C-31 current meter and electronic counter unit were used. Flett (2016) used gaugings taken in 2011, 2013 and 2014 to construct a stage-discharge rating curve based on the widely-applied power law equation:

$$Q = a(c + h)^b \tag{3.1}$$

where Q is river discharge (m^3 s^{-1}), h is the stage (m) and a, b and c are calibration parameters.

Given the significant alteration to the river channel in 2012, gaugings from 2011 were not used in this study, but additional data collected between January 2015 and March 2017 were available, giving a total of 43 gaugings. In addition, it was also deemed appropriate to experiment with using a two and three-section power law approach (McMillan

and Westerberg, 2015), given the irregular geometry of the river bed and the periodic activation of the western section of the river channel. It was found that a two-section equation produced a better fit than that of Eq. 3.1, but that the three-section approach provided no additional benefit. Accordingly, the two-section power-law equation was deemed the most parsimonious approach and has been used throughout this research. It has the form:

$$Q = \begin{cases} a_1(c_1 + h)^{b_1} & h < br_1 \\ a_2(c_2 + h)^{b_2} & h \geq br_1 \end{cases} \tag{3.2}$$

where br_i is the breakpoint on the rating curve (m). The following constraint is placed to endure continuity at breakpoints:

$$a_2 = \frac{a_1(br_1 + c_1)^{b_1}}{(br_1 + c_2^{b_2})} \tag{3.3}$$

A shift in the rating was noted between the end of 2014 and beginning of 2015 which is likely due to the movement of river bed material which is known to become mobile during high flow events such as that which occurred in October 2014. Accordingly, two separate rating curves were derived (Fig 3.8b). Figure 3.8c shows the river discharge time series derived using these rating curves.

3.5.5 Groundwater level

Eight boreholes were drilled into the sandur over a six day period between 17$^{\text{th}}$ and 22$^{\text{nd}}$ August 2012. They form three transects leading away from the Virkisá River channel including an upper (U1,U2), middle (M1,M2,M3) and lower (L1,L2,L3) transect (Fig. 3.2). The boreholes were drilled to between 9 and 15 m deep and installed as piezometers, with 88 mm diameter uPVC plain casing and a 3–6 m length of 0.5 mm slotted well screen. The completed piezometer depths range from 8.2 to 14.9m below ground level (Table 3.3).

Groundwater level in each piezometer was measured at 15 minutes intervals using In-Situ Inc. Rugged Troll 100 non-vented pressure transducers at 7–8.4 m depth which were corrected for air pressure based on measurements from two In-Situ Rugged Barometer

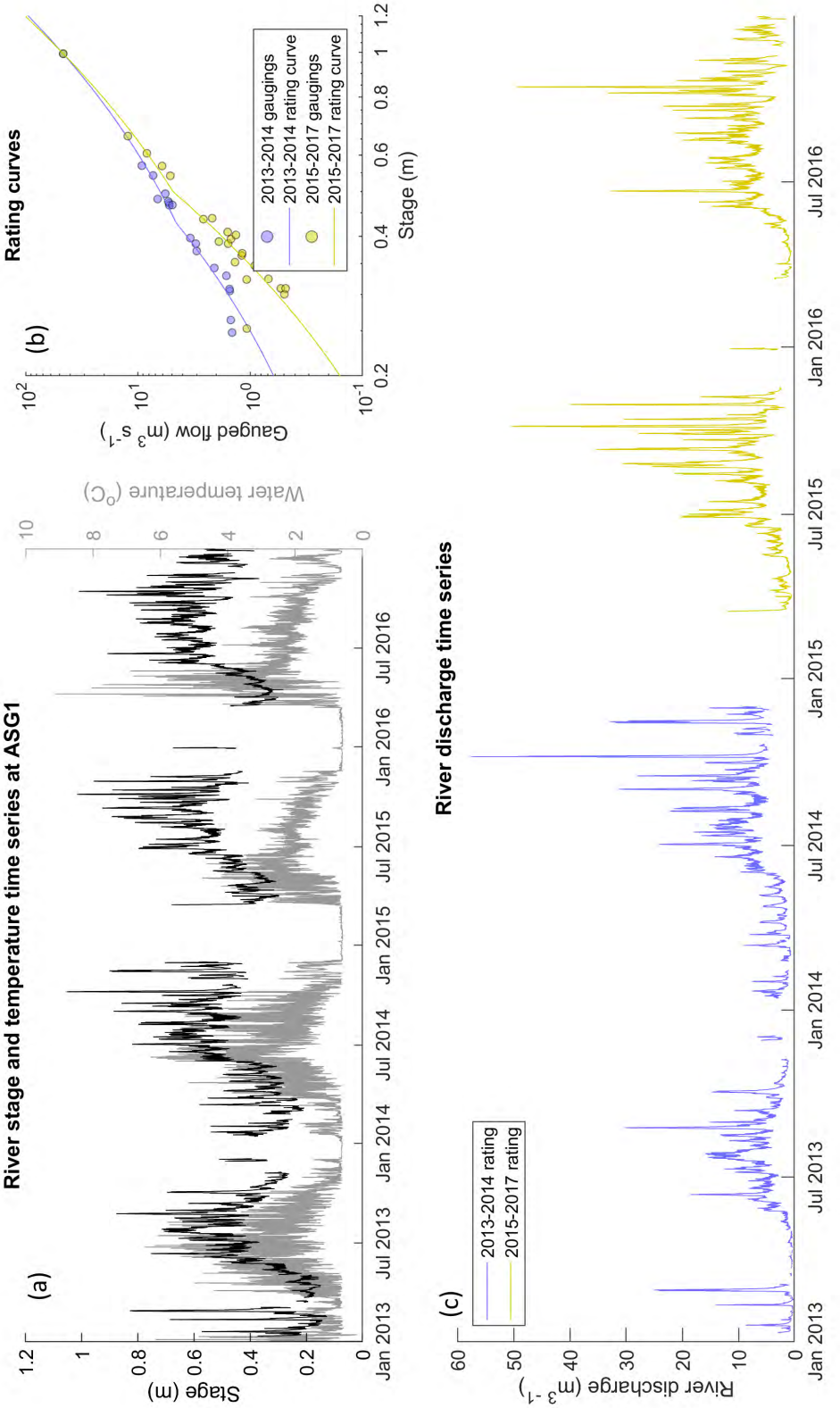

Figure 3.8: Harmonised hourly average river stage and temperature time series at ASG1 (a), rating curves from gaugings plotted on log-log scale (b) and subsequent conversion of continuous stage measurements to river discharge time series (c). Note, periods where the river was frozen over have been removed from the time series.

Table 3.3: Summary of sandur boreholes adapted from Ó Dochartaigh et al. (2012).

Borehole ID	Date completed	Completed depth (m bgl)	Depth of screen(s) (m bgl)	Mean level (m asl)	Level range (m)
U1	17/08/2012	14.4	10.4-13.4	88.6	1.5
U2	18/08/2012	14.6	8.6-9.6; 11.6-12.6	88.7	1.9
M1	19/08/2012	14.9	10.95-13.95	73.8	1.9
M2	20/08/2012	14.75	11.75-13.75	74.1	2.4
M3	21/08/2012	14.68	7.7-13.7	76.1	3.6
L1	21/08/2012	11.98	5.05-11.05	55.1	1.0
L2	22/08/2012	8.23	4.25-7.25	56.6	1.0
L3	22/08/2012	8.42	4.45-7.45	58.8	1.6

Trolls. Figure 3.9 shows the groundwater level time series obtained from each borehole. Note the groundwater level fluctuations at borehole M3 are at least 50% larger than at the other boreholes. Discussions with A. MacDonald confirmed that these are likely due to an ephemeral surface water channel which has been seen to flow in the immediate vicinity of the borehole causing discrete focussed groundwater recharge (MacDonald, 2019, personal communication, 28 February). Accordingly, it was decided not to use these data in this study given that they are not representative of regional groundwater level fluctuations.

3.6 Summary

This chapter has detailed the models and datasets that will be used throughout this thesis to meet the research objectives identified in Chapter 1. The GHM++ glacio-hydrological model will be used in chapters 4 and 5 to simulate accumulation and melt of snow and ice, snow drift, ice flow, soil infiltration and evapotranspiration and runoff-routing through the snow, ice and overland in the Virkisá River basin over historic (1980-2016) and future (up to 2100) time periods. It will also be used in chapter 6 to provide diffuse recharge and river flow boundary conditions for a MODFLOW-NWT groundwater model of the proglacial sandur aquifer. Both models will be used in conjunction with available catchment observation data to aid model parameterisation and evaluation. The next chapter is the first of the main research chapters. Here a signature-based LOA framework is implemented to compare the efficiency of different structural representations of snow/ice melt and runoff-routing in the GHM++ glacio-hydrological model code.

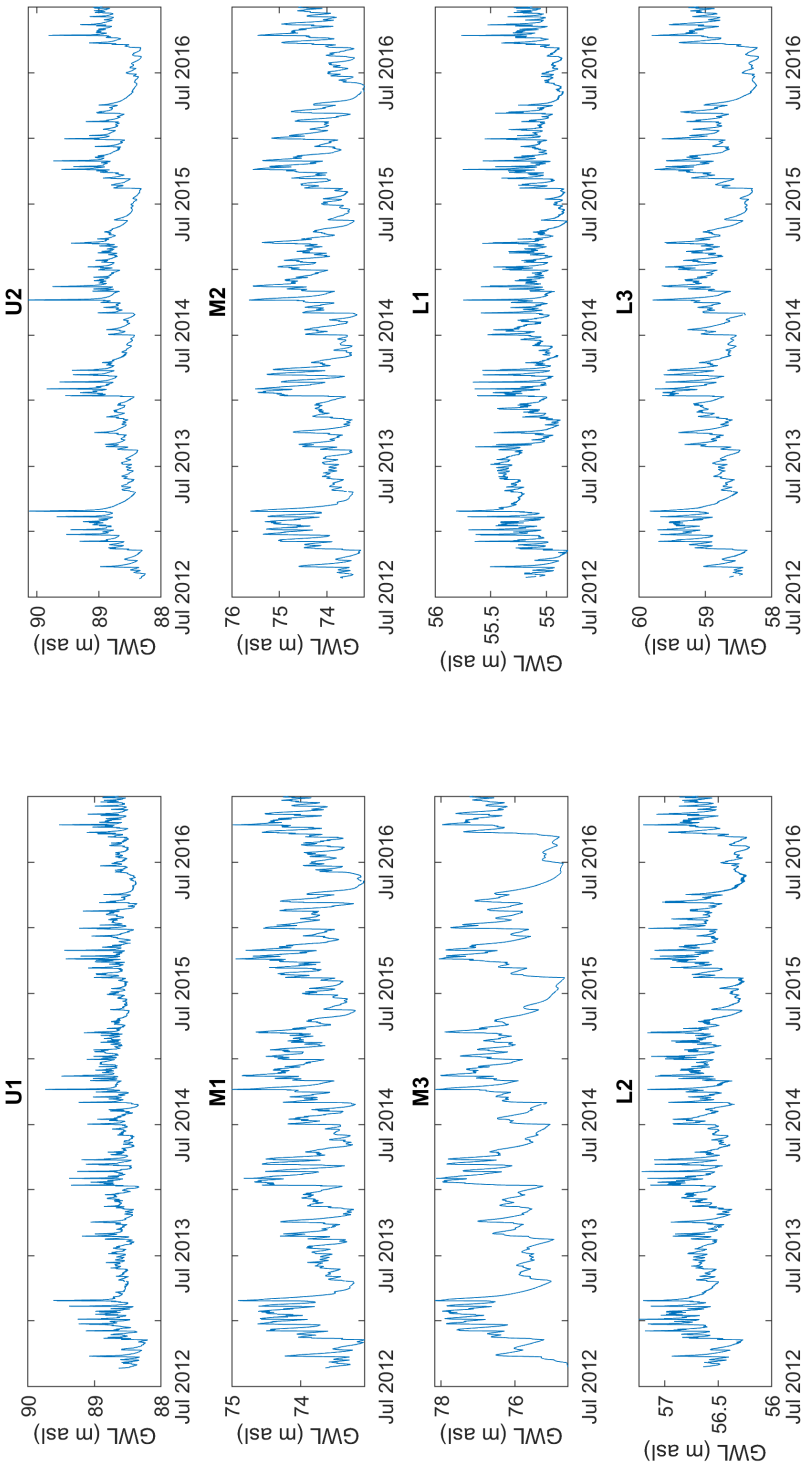

Figure 3.9: Groundwater level time series data for eight observation boreholes.

CHAPTER 4:

APPLICATION OF A LIMITS OF ACCEPTABILITY FRAMEWORK FOR GLACIO-HYDROLOGICAL MODEL COMPARISON AND SELECTION

4.1 Introduction

This chapter aims to meet the first research objective of this thesis by implementing a signature-based LOA model intercomparison framework to compare the efficiency of different snow/ice melt model structures and runoff-routing model structures within the GHM++ glacio-hydrological model code. A range of signatures are derived from observations of ice melt, snow coverage and river discharge to quantify different aspects of the Virkisá River basin's glacio-hydrological behaviour. The uncertainty of each set of observations is then harnessed to define the LOA, a set of criteria used to objectively evaluate the acceptability of different GHM structures and parameterisations. The findings and methodology developed from this study will be used to inform the model calibration and selection procedure for the future projection study presented in the next chapter.

The material presented in this chapter has been published in The Cryosphere (Mackay et al., 2018). For this, JM designed and undertook all aspects of the methodology including the fieldwork, GHM theory and code, construction of bias-corrected climate time series and the definition and calculation of glacio-hydrological signatures and their LOA. He also led the analysis, interpretation and writing of the manuscript. All co-authors (NB, DH, SK, CJ, JE and GA) contributed to the fieldwork design, provided guidance on the choice of signatures and GHM structures and contributed to the final edits of the manuscript.

4.2 Background

Computational GHMs allow us to develop an understanding of how future climate change will affect river flow regimes in glaciated watersheds (Lutz et al., 2014; Radić and Hock, 2014; Ragettli et al., 2016; Singh et al., 2016b). A variety of GHM codes exist (e.g. Boscarello et al., 2014; Ciarapica and Todini, 2002; Huss et al., 2008b; Schaefli et al., 2014; Schulla, 2015), each of which include a number of model components that represent two broad groups of processes: i) glaciological mass balance: the accumulation and ablation of snow and ice; and ii) hydrological water balance: the storage and release of melt and rainfall through snow, ice, overland and the subsurface. The exact form that these model components should take, both in terms of their governing equations (structure) and numerical constants (parameterisation) is not known. Physically based models which solve equations derived from first principles, typically over a distributed grid, are our closest approximation of the 'true' structure. However, limited parameterisation data and computer resources often preclude the use of such complex models, particularly in remote mountainous regions where data are scarce and where the inclusion of extra complexity does not guarantee better predictions (e.g. Gabbi et al., 2014).

Simplified process models offer an alternative. They are faster to run and employ fewer parameters that are typically calibrated to available observation data. They are based on, but do not necessarily adhere to, physical laws and as such their mathematical structure is somewhat unconstrained and may be biased towards a particular scientist's own perceptions and understanding of environmental processes. This has led to the development of a variety of competing model structures which purport to simulate the same process, but which have been derived from different process hypotheses. For example, a number of simplified 'index' model structures of snow and ice melt exist. The classical temperature-index model (TIM) simulates melt as a linear piecewise function of temperature only and ignores other drivers of melt which can be found in the physically based surface energy balance equation (Braithwaite, 1995). The TIM hypothesis can be justified because of the influence temperature has on the total energy balance of ice and snow,

particularly in temperate climates (Aðalgeirsdóttir et al., 2011; Guðmundsson et al., 2009; Ohmura, 2001). So-called 'enhanced' TIM structures have also been proposed which include added levels of complexity with the purpose of providing more accurate estimates of melt. These have accounted for perturbations in melt caused by topographic shading (Hock, 1999), surface albedo characteristics (Oerlemans, 2001; Pellicciotti et al., 2005) and more recently, debris cover (Carenzo et al., 2016).

Similarly a number of simplified representations of processes governing the hydrological water balance have been used in GHMs. Arguably, the equations that govern the routing (transport) of runoff are most important in relation to river flow predictions in glaciated river basins, as storage characteristics of ice and snow strongly influence river flow regimes over a range of timescales (Jansson et al., 2003). The concept of linear reservoirs is the most widely adopted approach for runoff-routing in glaciated basins (Gao et al., 2017; Hanzer et al., 2016; Zhang et al., 2015). A linear reservoir lumps all of the interacting, non-linear and non-stationary components of water transmission within a predefined area (e.g. a watershed) into a single leaky bucket. Despite its simplicity, the linear reservoir has shown to be remarkably versatile at capturing the storage-discharge characteristics of glaciated river basins around the world (Farinotti et al., 2012; Hock and Jansson, 2005; Woul et al., 2006). This is partly because the concept lends itself to structural modifications that can represent different glacio-hydrological systems. For example, Hanzer et al. (2016) hypothesised that the snow pack, firn layer, glacier ice and the region free from ice all exhibit unique runoff-discharge responses and advocate the use of four linear reservoirs in parallel to distinguish between these units. In contrast, Hannah and Gurnell (2001) were able to accurately reproduce river discharge time series collected from the outlet of the Taillon glacier in the French Pyrénées using only two linear reservoirs to route meltwater through the snowpack and ice separately, while Boscarello et al. (2014) were able to accurately simulate the observed river discharge time series in the Rhone watershed using a single linear reservoir to route all rainfall and melt runoff simultaneously.

The availability of multiple, presumably plausible, simplified model structures presents

somewhat of a dilemma to glaciologists and hydrologists as they are left with some uncertainty about how processes should be represented in their models. For the purpose of river discharge predictions, this problem is particularly pertinent as there are competing structures for two fundamental controls on these predictions: snow and ice melt and runoff-routing. One approach to mitigate this is to determine the 'optimum' structure that best captures the observation data. Structural optimisation of simplified runoff-routing routines has largely been ignored in glacio-hydrological contexts, although Hannah and Gurnell (2001) provide one notable exception where they used hydrograph recession limb analysis to identify the most appropriate linear reservoir configuration for a glaciated river basin. More studies have sought to optimise and compare simplified models of melt. Gabbi et al. (2014) applied four different TIMs to Rhonegletscher, Switzerland. They found that all achieved a similar goodness-of-fit to six years of ablation stake data, but that the inclusion of a solar radiation term provided the most accurate predictions of multi-decadal measurements of ice volume change. Irvine-Fynn et al. (2014) applied six different TIMs to the High-Arctic Midtre Lovénbreen glacier and compared them against a two-year dataset of seasonal ablation stake data. They found some minor improvements when various levels of complexity were introduced to the classical (temperature-only) TIM, however simulations were generally comparable between competing model structures. More recently, a comparison of four TIMs applied to four glaciers in the French Alps by Reveillet et al. (2017) found no clear evidence that using an enhanced TIM over the classical temperature-only approach provided better simulations when compared to a 17-year dataset of ablation stake measurements. Mosier et al. (2016) used a multi-criterion evaluation approach to compare the performance of different conceptual melt model structures. They compared seven competing melt model structures in two glaciated catchments in Alaska to ablation stake, river discharge and remotely-sensed snow coverage data. They found that no single model was best across all of the observation datasets, but the inclusion of a snow cold-content representation consistently produced the best goodness-of-fit scores over the evaluation data.

Clearly, while some studies have provided useful insight into the comparative behaviour

between competing conceptual process hypotheses (particularly for melt), none provide any definitive reasoning for adopting (or not) a particular model structure. Of course, discriminating between competing model structures in this way is made difficult by the fact that observation data used to drive and evaluate models are uncertain and therefore we cannot be sure whether model deficiencies represent inadequacies in the model or the data (Beven, 2016). Beven (2006) argues that because of this uncertainty and because of the fact that all models are by definition imperfect, no one optimum model structure (or parameterisation) exists. Instead, there is an equifinality of 'behavioural' models that make predictions within some predefined acceptability bounds around the observation data that take account of the various sources of modelling uncertainty. Indeed, parameter equifinality is a well recognised phenomenon in conceptual models of snow and ice melt (Finger et al., 2015; Gabbi et al., 2014; Jost et al., 2012; Pellicciotti et al., 2012; Reveillet et al., 2017). If we accept this, then a priority within the glacio-hydrological modelling community should be to establish frameworks that allow us to robustly evaluate model appropriateness and distinguish between behavioural (acceptable) and non-behavioural (unacceptable) structures and parameterisations. Constraining the range of acceptable models is particularly important for glacio-hydrological modelling as it has been shown that model uncertainty can lead to high uncertainty in twenty-first century predictions of river flows in glaciated basins (Huss et al., 2014).

One potential source for inspiration is the hydrological rainfall-runoff modelling community. Their heavy reliance on an ever-expanding choice of conceptual hydrological process models to make river flow predictions prompted Gupta et al. (2008) to discuss the need for a better framework in which to discriminate between these competing process hypotheses. They focussed on the performance metrics used to evaluate how well a model fits a set of observations and noted that there was an over-reliance on metrics that quantify the average performance of a model (e.g root mean squared error and Nash-Sutcliffe efficiency) which reduce information held in observation data down to a single summary statistic. They argue for a multi-criterion, 'diagnostic' approach where more of the relevant information from observation data is extracted so that inadequacies in model

structures and parameterisations can be better identified. Rye et al. (2012) applied such an approach to optimise a distributed surface mass balance model of two glaciers in Svalbard. They used ablation stake data to define three different features of the observations including mass balance at the stake locations, long-term mass balance trend and mass balance gradient. Using a multi-objective optimisation procedure, they identified structural inadequacies relating to how the mass balance gradient was simulated.

Hydrologists are now moving away from traditional metrics of model performance in favour of more diagnostic 'signatures' of hydrological behaviour. These have typically been derived from river flow time series, an integrator of upstream hydrological processes, and they may be as simple as the mean flow (an indicator of water balance) or they can be used to characterise the distribution (e.g. flow percentiles) and timing (e.g. autocorrelation) of flows. They have shown to have more discrimination power than traditional error metrics (Euser et al., 2013; Hrachowitz et al., 2014; Schaefli, 2016; Shafii and Tolson, 2015) and, importantly, it is also possible to take account of their information content (i.e. their uncertainty) so that decisions about model appropriateness can be made within the uncertainties of observation data used to evaluate the model. Here, observation data uncertainty can be used to define quantitative LOA around each signature. Different model structures and parameterisations can then be systematically evaluated for their ability to capture the signatures within their LOA, allowing the modeller to objectively diagnose model deficiencies and make decisions about model appropriateness.

Blazkova and Beven (2009) adopted a LOA framework to constrain the parameters of a distributed hydrological model for flood prediction in the Skalka catchment in the Czech Republic. They derived 114 different river flow signatures and used the LOA to distinguish between acceptable and unacceptable parameter sets. More recently, Coxon et al. (2014) used the LOA framework to evaluate the appropriateness of different hydrological model structures across 24 river basins in the UK. They demonstrated a clear difference in structure acceptability across the catchments and that this was correlated with the catchment geologies. The LOA framework has also recently been applied to an Alpine catchment in order to diagnose deficiencies in the SEHR-ECHO glacio-hydrological model

(Schaefli, 2016). Using a range of river discharge signatures, they identified a multi-year period in river discharge observations that the model could not capture within the LOA.

A signature-based approach within a LOA framework could also be used to compare and diagnose deficiencies in different simplified melt and runoff-routing model (ROR) structures and parameterisations employed in GHMs. For this purpose, signatures need not be derived just from river discharge data, but should also be taken from other observation sources such as ice melt and snow coverage as these have shown to be useful for evaluating the consistency of GHMs across different aspects of glacio-hydrological systems (Finger et al., 2011; Finger et al., 2015; Hanzer et al., 2016; Mayr et al., 2013). By doing so, this framework could facilitate better predictions of river flow regime changes in glaciated river basins; firstly by helping to diagnose deficiencies in GHM structures that require improvement, and secondly, by objectively selecting the range of acceptable model structures and parameterisations so that prediction uncertainty can be better constrained.

4.3 Aims

This study is the first of its kind to apply a signature-based LOA framework for a multi-GHM-structure evaluation. The framework is used to evaluate three commonly used simplified melt model structures and three different ROR structures of varied complexity. By doing so, this study will meet objective 1 outlined in the introduction of this thesis. Specifically, this study has two aims:

1. Investigate the utility of a signature-based LOA framework for diagnosing deficiencies in the different model structures.

2. Determine if the framework can be used to constrain a prior population of model structures and parameterisations down to a smaller population of acceptable models.

The first aim will demonstrate the framework's usefulness for aiding future improvement of simplified process models, while the second aim will determine the framework's usefulness for constraining simulation uncertainty.

4.4 Methodology

4.4.1 Glacio-hydrological model

The GHM++ code was used for all model experiments which is computationally efficient and ideally suited for incorporating different model structures. It consists of a 2D Cartesian grid of equally spaced model nodes. A spatial resolution of 50 m was selected as the best balance between simulation detail and model performance. Hourly observations of precipitation, temperature and incident solar radiation are used to simulate the accumulation of snowfall and the melt of snow, firn and ice across the model domain. The snow redistribution algorithm developed by Huss et al. (2008a) is used to account for snow drift and avalanches based on the curvature and slope of the surface. A soil infiltration and evapotranspiration model developed by Griffiths et al. (2006) solves the water balance for the non-glaciated regions of the study catchment. Excess soil moisture, rainfall and melt are then routed to the catchment outlet via a semi-distributed network of linear reservoir cascades which represent the water storage and release characteristics of the major hydrological pathways in the watershed. The GHM also simulates the evolution of the glacier geometry under periods of sustained negative mass balance using the Δh parametrisation of glacier retreat which has shown to closely reproduce the evolution of Alpine glaciers with results comparable to more complex 3-D finite-element ice flow models (Huss et al., 2010). Details of this and the soil water balance component are given in Appendix C. The following text details the different melt and runoff-routing structures adopted for this study.

Snow and ice melt model structures

Melt of snow and ice is calculated at each model node separately. Snow melt can occur at any node where a snow pack has developed. Ice melt can only occur at ice-covered nodes where the snow pack has completely melted. The mass balance at a given node is the summation of snowfall minus snow and ice melt. The melt model therefore performs

a pivotal role in the function of GHM++ and influences a number of aspects of the model simulations. Firstly, it simulates the hourly melt runoff in the catchment which influences the discharge regime of the river outlet. It also controls the melt of snowfall and therefore its distribution across the catchment. Finally, it determines the overall mass balance of the glacier which influences the long-term evolution of the glacier geometry. Related to this, GHM++ uses the mass balance calculated at each node to determine the equilibrium line altitude (ELA) of the catchment. Here, the average node elevation where the mass balance is close to zero was taken as the ELA and this was updated after each simulation year. A rolling three-year average ELA was used to determine the dividing line between the firn and the bare ice on the glacier.

For this study, three different conceptual models of snow and ice melt were compared, which have been used extensively to simulate melt processes in glaciated regions around the world (Gao et al., 2017; Matthews and Hodgkins, 2016; Nepal et al., 2017; Ragettli et al., 2016; Reveillet et al., 2017). Each can be viewed as a simplification of the full surface energy balance equation which has the form:

$$Q_M = SW\downarrow(1-\alpha) + LW\downarrow - LW\uparrow + Q_H + Q_E + Q_R + Q_G \tag{4.1}$$

where Q_M is the energy available for melt and sublimation (W m^{-2}), α is the surface albedo, $SW\downarrow$ is the incident solar radiation, $LW\downarrow$ and $LW\uparrow$ are the incoming and outgoing longwave radiation balance terms, Q_H and Q_E are the turbulent sensible and latent heat fluxes, Q_R is the sensible heat flux supplied by rainfall and Q_G is conduction of heat from the ground. Here, Q_M can be converted into a unit depth melt rate, M using the thermodynamic properties of water:

$$M = \frac{Q_M}{\rho_w L_f}\beta \tag{4.2}$$

where ρ_w is the density of water (kg m^{-3}), L_f is the latent heat of fusion of water (J kg^{-1}) and β is a scaling factor to achieve melt in units meters water equivalent per hour (m w.e. h^{-1}).

The first melt model structure (TIM_1) employs a classic TIM approach (Braithwaite, 1995) whereby melt is assumed to increase linearly with temperature above a given critical

temperature threshold:

$$M_i = \begin{cases} a_i(T - T_i^*) & T > T_i^* \\ 0 & T \leq T_i^* \end{cases} \tag{4.3}$$

where a (m w.e. $^{\circ}C^{-1}$ h^{-1}) is the temperature factor calibration parameter that converts temperature into melt, T is the near-surface air temperature and T^* is the critical threshold above which melt occurs. This model effectively incorporates all of the energy balance terms into a single calibration parameter, a. To account for the different properties of snow, firn and ice that may bring about different values of a and T^*, these are defined separately so that $i = (snow, firn, ice)$.

The second melt model structure (TIM_2) was originally proposed by Hock (1999) and includes an additional incident solar radiation term to more closely match the full energy balance equation. This model structure accounts for topographic effects such as slope, aspect and shading which an can bring about spatio-temporal variations in melt (Arnold et al., 2006; Pellicciotti et al., 2008). Their enhanced TIM has the form:

$$M_i = \begin{cases} (T - T_i^*)(a_i + b_i \cdot SW_{\downarrow}) & T > T_i^* \\ 0 & T \leq T_i^* \end{cases} \tag{4.4}$$

where b (m^3 w.e. W^{-1} $^{\circ}C^{-1}$ h^{-1}) is an additional radiation factor calibration parameter that converts the measured incident solar radiation, $SW_{\downarrow}$ (W m^2) into a unit melt. To account for the topographic effects at a given location, the incident angle of solar radiation was calculated to scale the measured incoming radiation:

$$SW_{\downarrow} = SW_{\downarrow}^* \cdot \cos\theta \tag{4.5}$$

where $SW_{\downarrow}^*$ is the uncorrected measured incident solar radiation and θ is the incident angle of solar radiation which was calculated at each node in the GHM as:

$$\cos\theta = \cos\beta\cos Z + \sin\beta\sin Z\cos(\Phi_{sun} - \Phi_{slope}) \tag{4.6}$$

where β is the surface slope angle, Z is the solar zenith angle, Φ_{sun} is the solar azimuth angle and Φ_{slope} is the surface aspect.

The SPA algorithm (Reda and Andreas, 2008) was used to determine hourly solar zenith and azimuth angles used in Eq. (4.6). GHM++ also includes a dynamic topographic shading model that determines which model nodes are completely shaded for each simulation hour based on the model domain DEM and the position of the sun in the sky. For those model nodes that are shaded, $SW_{\downarrow} = 0$.

Note that the form of Eq. (4.4) is not congruent with Eq. (4.1) as temperature is used to multiply the short-wave radiation term. Konya et al. (2004) note that this can lead to overestimation of melt during peak temperatures. Accordingly the melt model structure proposed by Pellicciotti et al. (2005) was also used for this study (TIM_3) which is an enhanced TIM in additive form that also incorporates an albedo parameter, α, and therefore more closely resembles the energy balance equation:

$$M_i = \begin{cases} a_i(T - T_i^*) + b_i \cdot SW_{\downarrow}(1 - \alpha_i) & T > T_i^* \\ 0 & T \leq T_i^* \end{cases} \tag{4.7}$$

where b has the units m^3 w.e. W^{-1} h^{-1}. Following Pellicciotti et al. (2005), this melt model structure also includes the dynamic snow albedo algorithm proposed by Brock et al. (2000) which accounts for the drop in snow albedo as it ages using a logarithmic function with the form:

$$\alpha_{snow} = p_1 - p_2 \cdot log_{10} \cdot T_a \tag{4.8}$$

where p_1 is the albedo of fresh snow (set to 0.9), p_2 is an empirical calibration parameter and T_a is the accumulated daily maximum temperature greater than 0°C since snowfall.

For all melt model structures, melt M is converted into a volumetric melt M_v at each node:

$$M_v = M \cdot A \tag{4.9}$$

where A is the model node area. Following Hopkinson et al. (2010) the area of each node is corrected for surface slope:

$$A = \frac{L^2}{\cos \beta} \tag{4.10}$$

where L is the model node length and β is the node surface slope.

Runoff-routing model structures

Runoff includes any rainfall falling on, and melting of the snow and ice as well as excess soil moisture from those areas free of ice and snow. The concept of linear reservoirs was employed to route runoff to the catchment outlet. A linear reservoir can be visualised as a bucket that receives a volumetric inflow and releases it at a rate proportional to its internal water storage. The continuity equation can be written as:

$$\frac{ds}{dt} = i - q \tag{4.11}$$

where s is the storage (m^3), i is the inflow (m^3 h^{-1}) and q is the outflow (m^3 h^{-1}) which was calculated as a linear function of storage:

$$q = \frac{1}{k}s \tag{4.12}$$

where k is mean residence time of the reservoir (h) which accounts for the diffusive effect of storage and release mechanisms within the catchment. Increasing the value of k increases the diffusion effect on the inflow hydrograph. Additional controls on the diffusion and lag effects can be obtained by arranging a cascade of multiple linear reservoirs in series (Ponce, 1989) so that the outflow from the previous reservoir is the inflow for the subsequent reservoir. With this setup, the continuity equation for the jth reservoir of n reservoirs in series, where $j = (1, 2...n)$ can be re-written as:

$$\frac{ds_j}{dt} = \begin{cases} i - q_j & j = 1 \\ q_{j-1} - q_j & j > 1 \end{cases} \tag{4.13}$$

The outflow hydrograph is then taken from q_n.

Common practice when using linear reservoirs is to subdivide the catchment into one or more hydrological response units (HRU) which are thought to have different water storage and release characteristics. For example, the firn, snow and bare ice have generally shown to respond over relatively long, intermediate and short timescales respectively (Hock and Jansson, 2005) and therefore these may be characterised as separate HRUs, although as noted previously, simpler and more complex definitions of HRUs have been defined

in the past. Subsequently, three ROR structures were proposed with different levels of complexity structured around these subdivisions (Fig. 4.1).

The first and simplest ROR structure (ROR_1) uses a single linear reservoir cascade (e.g. see Boscarello et al., 2014) to route the inflow from all runoff sources simultaneously. This structure makes no distinction between the different runoff sources and flow pathways and assumes that all conform to the same storage-discharge relationship.

The second model structure (ROR_2), employs two linear reservoir cascades in parallel (e.g. Hannah and Gurnell, 2001). The first cascade represents the relatively slow percolation of water through the snow and firn HRUs, while the second cascade represents faster flow of water through the bare ice and overland. This approach therefore makes some distinction between the different flow pathways and, by conditioning the parameters so that the snow and firn have a more diffuse response function, it introduces a degree of non-linearity in the discharge response to runoff.

The third ROR structure (ROR_3) has not been used previously. It employs separate linear reservoir cascades to route runoff from the firn, snow, ice and soil HRUs. Here the parameters are conditioned so that the firn is the most diffuse, slowly responding reservoir, followed by the snow and then the ice and soil zones which are considered to be relatively flashy, quickly responding HRUs. This approach also includes some representation of linkages between these various units. Here it is hypothesised that water that flows through the firn, must then flow through the downstream bare ice HRU before it reaches the river. Similarly, water that percolates through the snow pack must also flow via the HRU that it overlies before it reaches the river. There are therefore six different flow pathways that runoff may take before reaching the river outlet (see Fig. 4.1c) and this represents the most complex, non-linear ROR structure.

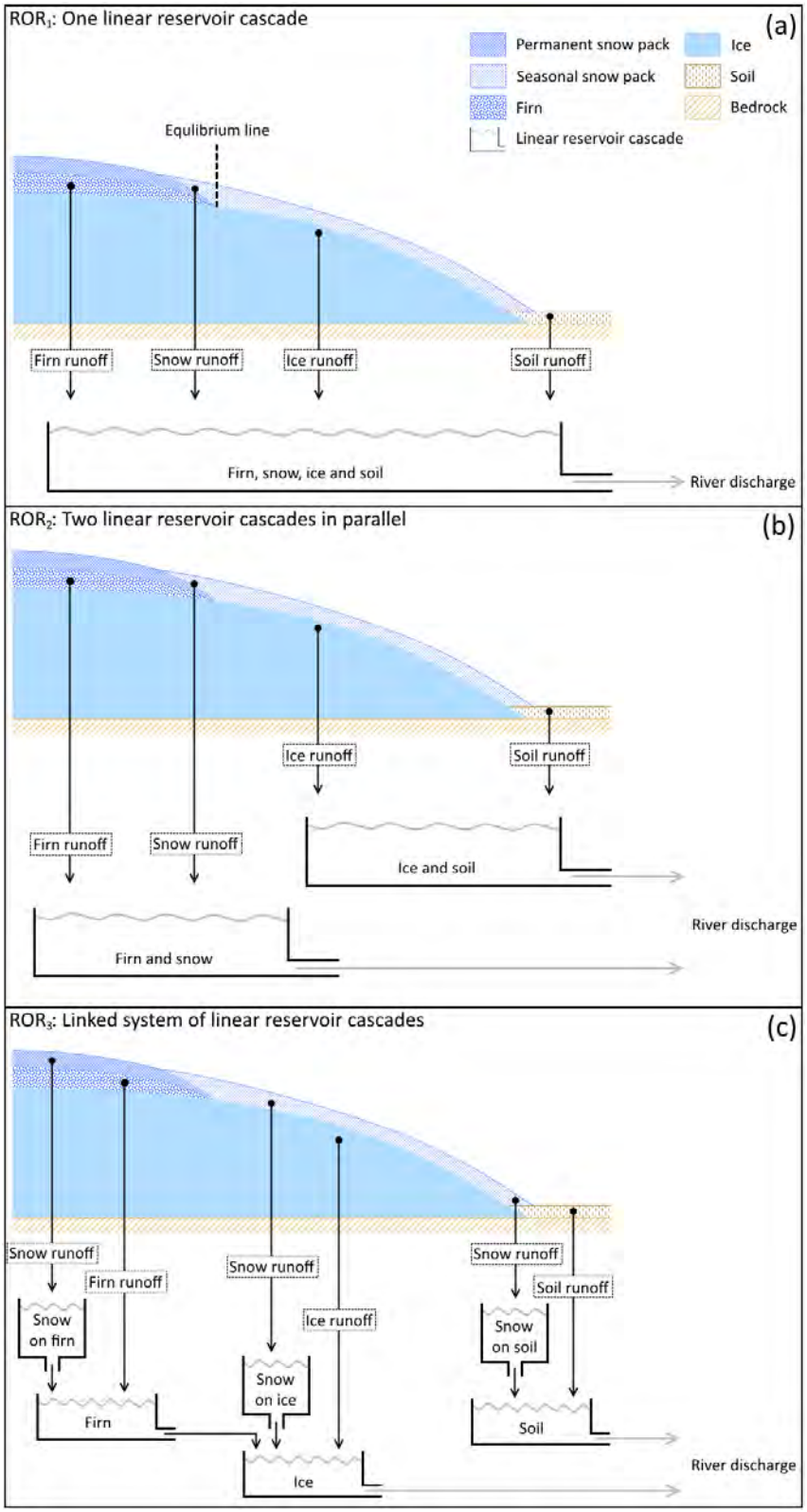

Figure 4.1: Three runoff-routing model structures which relate the linear reservoir cascade configurations to idealised cross-sections of a temperate glacier.

4.4.2 Driving climate data

The GHM was configured to run from the initial ice geometry of 1988 through to the end of 2014[1] It requires continuous measurements of hourly precipitation, near-surface air temperature and incident solar radiation to drive the various model components. The following sections describe how each of these were derived from the available climate data outlined in section 3.

Precipitation

A new gridded precipitation time series was constructed that incorporated the measurements of rainfall from the AWSs in the Virkisá basin and the information on spatial and long-term variations in precipitation from the gridded ICRA data. First, the AWS data were used to bias-correct the ICRA data. Due to the lack of rainfall measurements during the winter months, the AWS4 rainfall data were not used for the bias correction procedure. Furthermore, because the AWS1 and AWS3 gauges overlap the same ICRA data pixel, and because the AWS1 time series is the longest and most complete, it was decided to use the AWS1 data to bias-correct the overlapping ICRA data pixel. Here, the equidistant quantile mapping (EQM) approach (Li et al., 2010; Sachindra et al., 2014; Srivastav et al., 2014) was employed. Quantile mapping is a simple and efficient bias correction procedure that maps the cumulative distribution function (CDF) of the observed 'true' time series onto the biased time series whilst preserving the rank correlation between the two series (Panofsky and Brier, 1968). Generally, it has shown to be superior to other statistical methods for bias correction of precipitation data for hydrological modelling studies (Gudmundsson et al., 2012; Teng et al., 2015). EQM is an adaptation of the original quantile mapping method that accounts for non-stationarity in the statistical properties (moments) of the biased time series and helps to preserve changes in the CDF of the precipitation data that may have occurred over time (Cannon et al., 2015; Switanek

[1]Note, at the time of undertaking this study, there were not sufficient river gauging data to calculate the river discharge beyond 2014. Accordingly, the model was only run up to the end of 2014 for this study.

et al., 2017). To evaluate the effectiveness of the bias correction procedure, a number of statistics were calculated to compare the observed and ICRA precipitation data before and after bias correction (Table 4.1). There were a total of 30,460 hourly measurements of precipitation available for above-freezing days at AWS1 of which the majority were during the autumn months (September, October and November) and the least during the winter months (December, January and February). Overall, the procedure corrects for bias in the mean (Avg) and also improves the spread (SD), relative variability (CV) and skewness of the distribution of precipitation data at hourly, daily and 3-daily time steps. At the seasonal scale, these improvements are notable for spring, summer and autumn. However, the bias correction procedure typically has a slightly negative impact on the winter precipitation statistics, probably because of the limited above-freezing data available for these months. In particular, average hourly winter precipitation is underestimated by 0.11 mm (16%) while the positive bias in relative variability and skewness are amplified after bias correction. Given that EQM preserves the rank correlation of the time series, it has little effect on the R^2 correlation score, with a typical reduction of 0.01-0.02 after bias correction. At an hourly timescale, the bias-corrected data only captured 22% of the observed variance in the AWS1 rainfall record. However, when averaged to a daily timescale the R^2 score increased to 0.49, and for a three-daily timescale the R^2 increased to 0.72. The limited correlation of the ICRA precipitation data at an hourly timescale could hinder the acceptability of the GHM across some of the signatures (e.g. the river discharge signatures related to the timing of flows). However, the AWS1 rainfall record is complete for the years 2013 and 2014 where the GHM is compared against observed river discharge signatures. As such, poor replication of the timing of hourly rainfall events should have minimal influence on the GHM's ability to capture the river discharge signatures. Rather, the role of the bias-corrected ICRA precipitation data was primarily to drive the glacier-mass balance component of the GHM prior to 2009 for which a reliable three-daily temporal correlation with observations was deemed adequate.

To distribute the newly generated continuous time series of precipitation data from AWS1 across the entire study catchment, hourly maps of precipitation anomalies gen-

Table 4.1: Statistics calculated from the observed precipitation data at AWS1 and from the corresponding ICRA precipitation data before and after bias correction. Statistics have been calculated at an hourly, daily and 3-daily time step and include n (total number of above-freezing measurements available at AWS1), Avg (mean), SD (standard deviation), Cv (coefficient of variation), Skewness and R^2 (coefficient of determination).

Time step	Statistic	Overall			Winter (DJF)			Spring (MAM)			Summer (JJA)			Autumn (SON)		
		Obs	Before	After	Obs	Before	After	Obs	Before	After	Obs	Before	After	Obs	Before	After
Hourly	n	30460			4344			6290			8832			10994		
	Avg (mm)	0.33	0.43	0.33	0.65	0.67	0.54	0.20	0.29	0.21	0.17	0.33	0.24	0.41	0.50	0.39
	SD (mm)	1.09	1.24	1.12	1.55	1.63	1.49	0.72	0.83	0.72	0.65	1.01	0.91	1.27	1.39	1.25
	Cv	3.28	2.85	3.37	2.39	2.42	2.78	3.59	2.82	3.46	3.79	3.02	3.74	3.11	2.79	3.23
	Skewness	5.81	5.40	6.01	3.84	4.38	4.83	5.91	4.82	5.43	7.45	6.41	7.51	5.43	4.93	5.35
	R^2		0.24	0.22		0.31	0.29		0.17	0.16		0.15	0.14		0.23	0.22
Daily	n	1264			181			260			368			455		
	Avg (mm)	7.95	10.4	7.93	15.6	16.2	12.9	4.84	7.02	5.00	4.13	7.99	5.86	9.78	11.9	9.31
	SD (mm)	16.2	19.5	17.1	23.2	27.2	24.4	10.0	12.2	10.2	8.80	13.1	11.3	18.9	22.6	20.0
	Cv	2.04	1.87	2.16	1.49	1.68	1.89	2.06	1.73	2.04	2.13	1.64	1.92	1.93	1.89	2.15
	Skewness	4.37	4.67	5.07	2.78	4.50	4.66	3.71	3.38	3.92	4.39	2.64	2.92	4.19	4.02	4.36
	R^2		0.51	0.49		0.64	0.62		0.44	0.43		0.49	0.47		0.45	0.43
3-daily	n	385			47			75			123			140		
	Avg (mm)	23.5	31.0	23.6	48.6	50.1	40.0	13.7	19.8	14.0	12.4	23.8	17.4	30.0	36.8	28.7
	SD (mm)	34.1	41.6	36.4	47.7	67.4	60.9	18.4	23.9	19.5	18.2	27.1	22.8	39.7	45.4	39.7
	Cv	1.45	1.34	1.54	0.98	1.35	1.52	1.35	1.21	1.39	1.46	1.14	1.31	1.32	1.23	1.38
	Skewness	2.94	3.44	3.78	1.86	3.36	3.45	1.96	1.92	2.22	2.82	1.84	1.98	2.58	2.28	2.53
	R^2		0.73	0.72		0.79	0.77		0.60	0.59		0.56	0.54		0.68	0.66

erated from the raw ICRA data were used. As such, it was assumed that the absolute bias correction applied to the ICRA data at AWS1 holds for the rest of the study catchment. While it is difficult to validate this approach because of the scarcity of precipitation observations available at higher elevations in the river basin, Guðmundsson (2000) estimated annual precipitation rates at the summit to be somewhere between 7000 - 8000 mm y^{-1}. The mean bias-corrected ICRA data concur with an estimate of 7600 mm y^{-1} which suggests the assumption a fixed bias correction in space is reasonable.

Note the distinction between precipitation falling as rainfall or snowfall is made within the GHM++ code. Here, following previous modelling studies on Icelandic glaciers (Aðalgeirsdóttir et al., 2006; Aðalgeirsdóttir et al., 2011; Hannesdóttir et al., 2015b; Johannesson et al., 1995), it was assumed that for a given model node, precipitation falls as rainfall when the near-surface air temperature is at least 1 °C.

Near-surface air temperature

To generate a continuous time series of temperature back to 1988, daily measurements of temperature available from the nearby Fagurhólsmýri weather station to the south of the study site were used. A comparison of daily average temperatures showed there to be a good linear relationship between the two stations with an R^2 of 0.92. As such, this linear model was used to bias-correct the daily weather station data so that it could be combined with the AWS1 time series. To downscale the data to an hourly resolution, 24-hour temperature anomalies were randomly sampled from the AWS1 record, thereby ensuring the complete time series had a consistent sub-daily variability. Of course, diurnal cycles in temperature are dependent on the time of year, whereby increased incident solar radiation in the summer enhances sub-daily temperature variability. Therefore, the sampling strategy was employed on a month-by-month basis. The complete hourly time series of temperature at AWS1 is shown in Fig. 4.2b.

The importance of characterising temperature lapse rates for glacio-hydrological modelling is well known because it has a strong control on spatial patterns of melt simulations (Gardner and Sharp, 2009; Heynen et al., 2013; MacDougall et al., 2011). In fact while

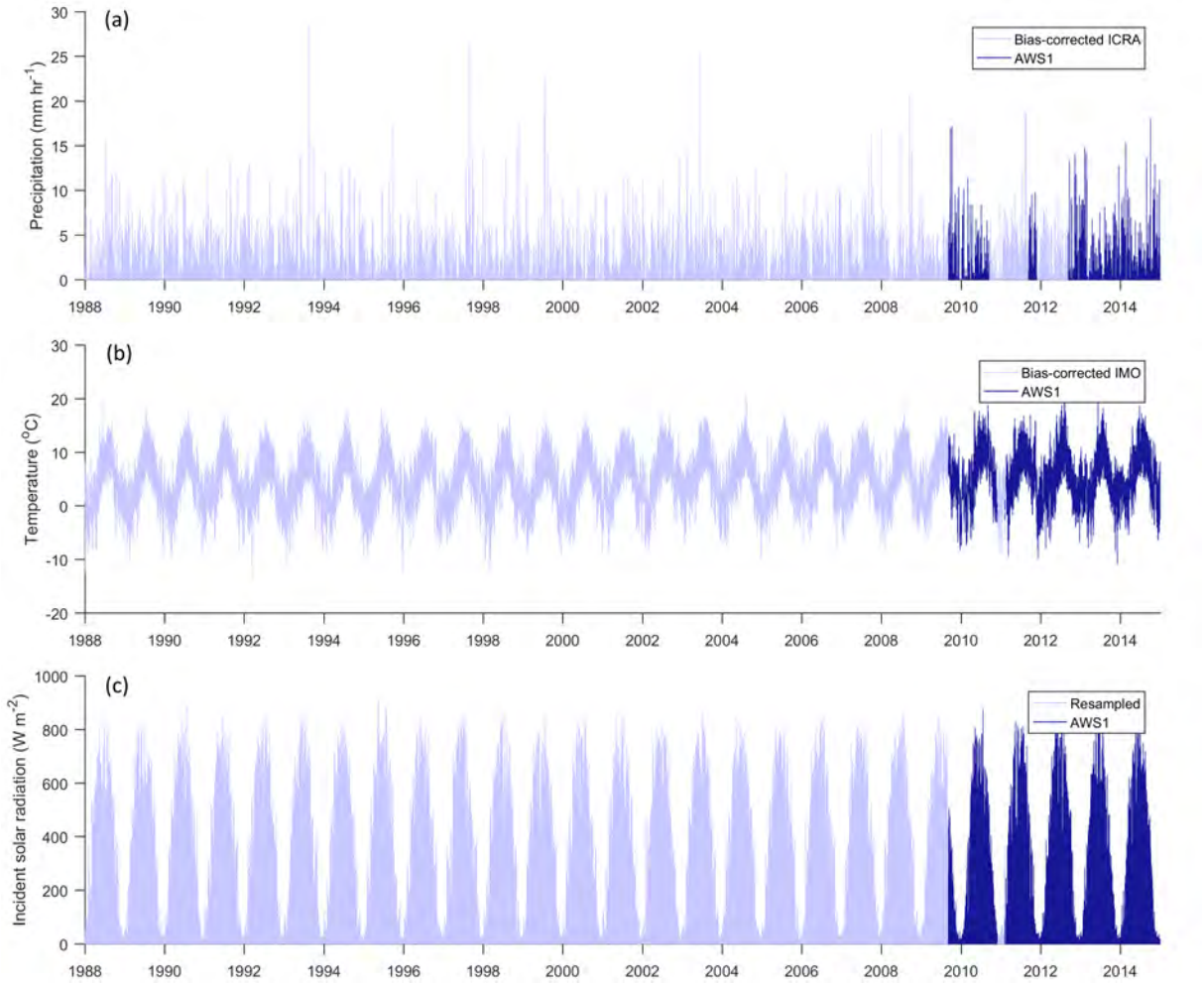

Figure 4.2: Continuous hourly time series of precipitation (a), temperature (b) and incident solar radiation (c) between 1988 and 2015 at AWS1.

many studies employ a fixed temperature lapse rate, in reality seasonal variations in surface characteristics (e.g. albedo and roughness) and atmospheric conditions can bring about strong seasonal and diurnal variations in lapse rates which control melt processes (Gardner et al., 2009; Immerzeel et al., 2014; Minder et al., 2010). Furthermore, local atmospheric phenomena associated with mid-latitude glaciers such as katabatic winds which bring cool dense air over the ice surface can serve to reduce the temperature gradient (Petersen and Pellicciotti, 2011; Ragettli et al., 2014). Having analysed near-surface air temperature variations both on and away from the Virkisjökull glacier, it was deemed most appropriate to extrapolate temperature across the study catchment using a seasonally variable hourly lapse rate in conjunction with an on-ice temperature correction function based on the work of Shea and Moore (2010) (see Appendix D).

Incident solar radiation

The only source of incident solar radiation is the continuous hourly time series from AWS1. To construct a continuous time series back to 1988, a resampling strategy was employed to generate a complete time series that was statistically consistent with the data at AWS1. It was found that during the summer months, the daily range in incident solar radiation and temperature are strongly correlated. Therefore, when generating a continuous time series of hourly incident solar radiation from 1988, it was important to maintain this dependence between intra-day solar radiation and temperature variability. To do this, a coordinated (in time) sampling strategy identical to that used for the near-surface air temperature data was employed. More specifically, for each random 24-hour temperature anomaly sample from the AWS1 record used to build part of the temperature time series, the corresponding 24-hour solar cycle data were extracted and used to build the same part of the incident solar radiation time series. Figure 4.2c shows the complete time series of incident solar radiation used to drive the model.

4.4.3 Signatures and limits of acceptability

Observations of ice melt, snow coverage and river discharge were used to derive 33 unique signatures with LOA to characterise the glacio-hydrological behaviour of the Virkisá River basin over a different spatio-temporal scales and evaluate the acceptability of the different model structures (Table 4.2). For convenience, the signatures have also been subdivided into 11 attributes which encapsulate the main aspects of model behaviour to be assessed.

Ice melt

The average winter (November 2012 - April 2013) and summer (May 2013 - September 2013) melt across the ablation stake network were used to characterise the short-term, seasonal ice melt on the glacier tongue. Of course, point measurements of melt are not directly comparable to simulated melt at the GHM nodes as these simulations represent the average melt over the node area. Therefore, the GHM can only be expected to get as close to the stake measurements as the actual spread in melt over the equivalent model node area. To calculate this spread, the high resolution terrestrial lidar scans taken during the ablation stake campaign (2012-2014) were used. The scans were used to estimate the spread of melt deviations from the mean melt across 50 m square regions (Fig. 4.3). The 95% confidence bounds ($\pm$ 0.78 m y^{-1}) were then used to define the LOA around the winter and summer melt signatures where it was assumed that the spread should be proportional to the total melt. This assumption leads to much narrower LOA around the winter melt signature than the summer melt signature.

A signature to characterise the long-term change in glacier volume was also quantified by differencing the two 3-D models of the ice from 1988 and 2011. DEMs derived from photogrammetric methods are prone to errors due to poor image rectification and stereo image mismatches due to low or poor contrast areas such as cloud or new snow cover (Barrand et al., 2009). Therefore, the 1988 ice DEM was assumed to be the main source of uncertainty in the calculation of the ice volume change signature. A comparison to the

Table 4.2: Summary of signatures used to evaluate model acceptability. Units with asterisk (*) are per section of flow duration curve (FDC).

Group	Attribute	Attribute ID	Signature	Limits of acceptability
Ice melt	Seasonal ice melt on tongue	Seas melt	2013 Summer ice melt	5.22 – 6.44 m w.e.
			2012-2013 Winter ice melt	0.64 – 0.78 m w.e.
	Long-term glacier volume change	Melt vol	Change in ice volume (1988-2011)	-0.36 – -0.28 km^3
Snow coverage	Snow coverage in lower catchment	Low snow	Mean snow coverage in spring	0.32 – 0.45
			Mean snow coverage in early-summer	0.02 – 0.08
			Mean snow coverage in late-summer	0.00 – 0.03
	Snow coverage in mid catchment	Mid snow	Mean snow coverage in spring	0.70 – 0.80
			Mean snow coverage in early-summer	0.17 – 0.27
			Mean snow coverage in late-summer	0.00 – 0.04
	Snow coverage in upper catchment	Upp snow	Mean snow coverage in spring	0.81 – 0.90
			Mean snow coverage in early-summer	0.51 – 0.64
			Mean snow coverage in late-summer	0.02 – 0.09
River discharge	Mean monthly river flow	Mnthly flow	Mean January river flow	1.16 – 1.86 m^3 s^{-1}
			Mean February river flow	1.69 – 2.92 m^3 s^{-1}
			Mean March river flow	0.85 – 1.58 m^3 s^{-1}
			Mean April river flow	0.73 – 1.48 m^3 s^{-1}
			Mean May river flow	1.50 – 2.16 m^3 s^{-1}
			Mean June river flow	4.12 – 6.23 m^3 s^{-1}
			Mean July river flow	6.33 – 10.30 m^3 s^{-1}
			Mean August river flow	5.72 – 9.15 m^3 s^{-1}
			Mean September river flow	4.55 – 7.38 m^3 s^{-1}
			Mean October river flow	3.88 – 7.02 m^3 s^{-1}
			Mean November river flow	3.90 – 7.40 m^3 s^{-1}
	Quick-release high flows	High flows	Volume under highest flow section of FDC	59.4 – 116.0 m^3 s^{-1} *
			Slope of highest flow section of FDC	2.67 – 9.88 m^3 s^{-1} *
			Volume under high flow section of FDC	70.6 – 111.0 m^3 s^{-1} *
			Slope of high flow section of FDC	0.38 – 0.79 m^3 s^{-1} *
	Slow-release low flows	Low flows	Volume under low flow section of FDC	20.9 – 46.1 m^3 s^{-1} *
			Slope of low flow section of FDC	0.03 – 0.05 m^3 s^{-1} *
	Flow variability	Flow var	Coefficient of variation	0.95 – 1.83
	Melt runoff timing	Melt timng	Peak summer flow hour	17:00 – 18:00
	Flashiness	Flow flash	Integral scale	25 – 44 h
			Rising limb density	0.13 – 0.20

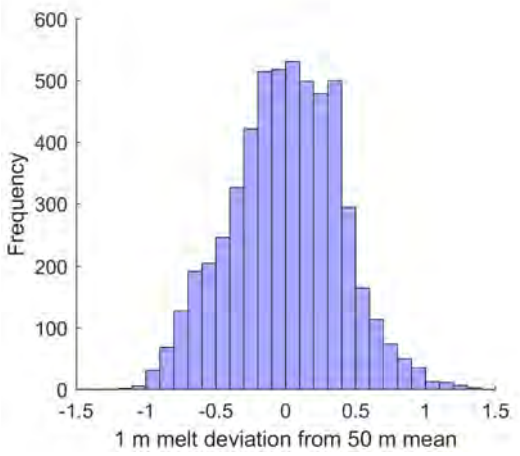

Figure 4.3: Histogram of deviation of 1 m melt from 50 m mean derived from terrestrial lidar scans of static ice front between 2012-2014.

more accurate 2011 DEM shows that the 1988 DEM captures the gridded elevation data across the non-glaciated portion of the study area with reasonable accuracy (Fig. 4.4a). The residuals are approximately normally distributed with a mean error of zero (Fig. 4.4b) and they show to be largest for those parts of the catchment that are steeply sloped (scatter in Fig. 4.4c). To account for these errors in the calculation of the ice volume change signature, 1000 unique DEMs of the 1988 ice surface were generated by randomly perturbing each pixel of the original dataset with perturbations drawn from a normal distribution with mean zero. Given that the spread of the residuals increases for those areas of the catchment that are steepest, the shape parameter of the error distribution (standard deviation) was varied according to the slope of each pixel of the 1988 DEM (see dark blue line in Fig. 4.4c). From these, 1000 equally probable estimates of ice volume change were calculated and the 95% confidence interval was used to define the LOA. The total change in ice volume over 23 years from 1988 was estimated to be between -0.36 and -0.28 km^3.

Snow coverage

No direct observations of snow accumulation or melt exist for the VGO and so instead, the MODIS snow coverage data were used to evaluate snow characteristics of the model.

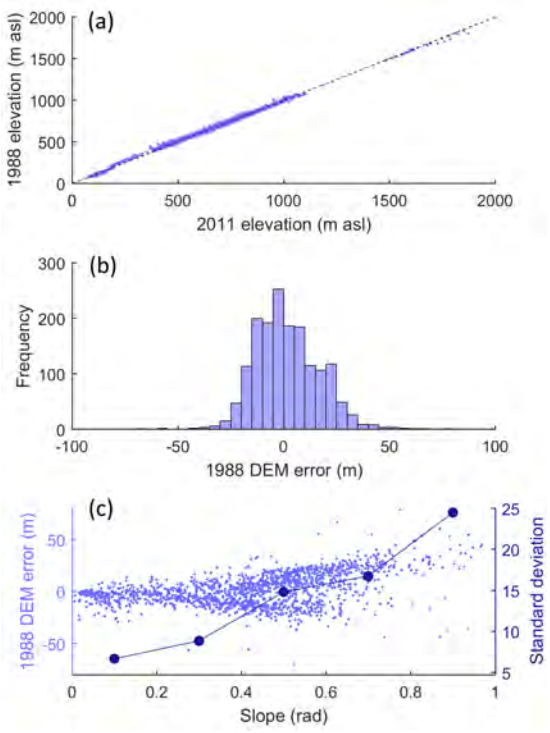

Figure 4.4: Error model for estimating uncertainty in glacier volume change between 1988 to 2011 including: 1988 vs 2011 off-ice DEM elevations (a), distribution of 1988 DEM errors calculated as difference between 1988 and 2011 off-ice elevations (b) and estimation of change in standard deviation of errors with DEM slope (c).

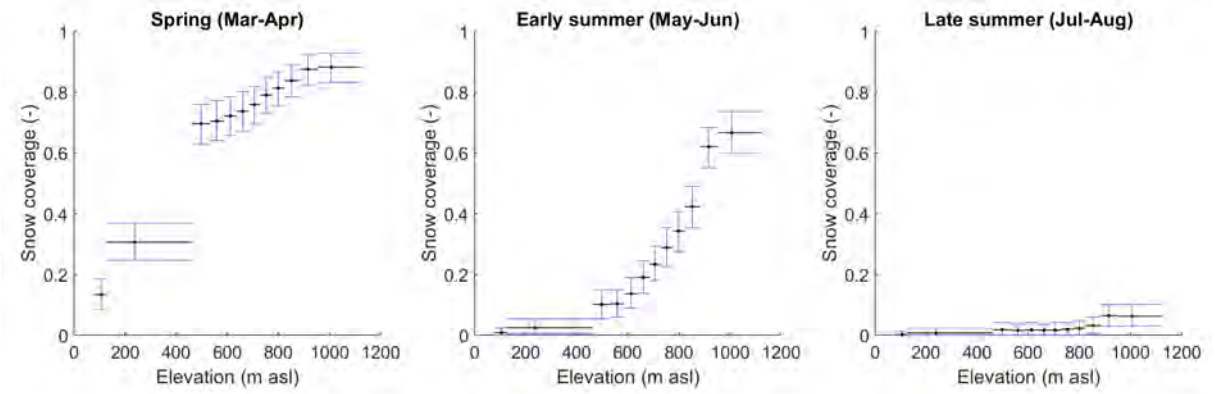

Figure 4.5: Snow coverage curves defined from the MOD10A1 snow cover product from 2000 - 2015 with 95% confidence bounds.

While the MOD10A1 data cannot be used to quantify snow accumulation or melt directly, they can have similar predictive power as ground-based mass-balance observations and, as such, have been widely used in the calibration and evaluation of GHMs (Frey and Holzmann, 2015; Hanzer et al., 2016; Pellicciotti et al., 2012; Valentin et al., 2018).

Given that <5% of the data passed the QA (see section 3.5.1), it was decided that these data should be combined to derive three seasonal average snow coverage maps. From these maps, three snow coverage curves were constructed that define the mean catchment snow coverage at different elevations for spring (March and April), early-summer (May and June) and late-summer (July and August) (Fig. 4.5). The curves provide information on both the spatial and temporal distribution of snowfall in the study catchment. They were constructed by distributing the seasonal average snow distribution maps across the 50 m model grid DEM. For example, for a MODIS pixel value of 0.5, half of the corresponding DEM pixels were assumed to be snow covered. While the data are only available up to an elevation of 1200 m asl, Fig. 4.5 shows that the three curves capture a large amount of variability in seasonal snow cover. From the three snow coverage curves, the mean snow coverage from the lower, middle, and upper terciles of the curves were used as signatures of snow coverage.

There exists no definitive quantification of errors in the MOD10A1 product that can be used to estimate LOA for these signatures. Previous validation of the MODIS data using

satellite imagery has shown the data to be relatively robust (Salomonson and Appel, 2004). Accordingly, it was assumed that as with the ablation stake data, the primary source of uncertainty stems from scale differences between the data and the model simulations. More specifically, because the MODIS data have a coarser resolution (500 m) than the DEM over which the MODIS data were distributed (50 m), a MODIS pixel value of 0.5 only indicates that 50 of the corresponding 100 DEM pixels are snow covered. The construction of a snow distribution curve, therefore necessitates some assumptions about where the snow actually lies which will influence the shape of the snow distribution curve. Accordingly, the LOA were quantified to account for this uncertainty. Here, for each of the seasons, a mean MODIS snow cover map over the study region was derived. Then, for each 500 m pixel, snow was randomly distributed across the corresponding DEM pixels 1000 times. From these, an equal number of snow distribution curves and corresponding snow distribution signatures could be derived, each assumed to be equally probable. The 95% confidence bounds from this distribution of snow cover signatures were used to define the LOA which are indicated by blue error bars in Fig. 4.5.

River discharge

The hourly river discharge data for the years 2013 and 2014 measured at ASG1 (Fig. 4.6a) were used to define 21 different river discharge signatures that cover a range of temporal scales and flow magnitudes. The majority of these signatures were based on previous studies (Andrés-Doménech et al., 2015; Casper et al., 2012; Clausen and Biggs, 2000; Coxon et al., 2014; Euser et al., 2013; Garavaglia et al., 2017; Hrachowitz et al., 2014; Monk et al., 2007; Sawicz et al., 2014; Schaefli, 2016; Shafii and Tolson, 2015; Teutschbein et al., 2015; Viglione et al., 2013; Westerberg et al., 2016; Winsemius et al., 2009; Yadav et al., 2007; Yilmaz et al., 2008).

Mean monthly river flows were calculated to characterise the seasonal river flow regime. Signatures were also derived from sections of the FDC to characterise quick-release high flows and slow-release low flows. These include signatures that quantify the volume under the section (flow magnitude) and the slope of section (flow variability) for the low flow

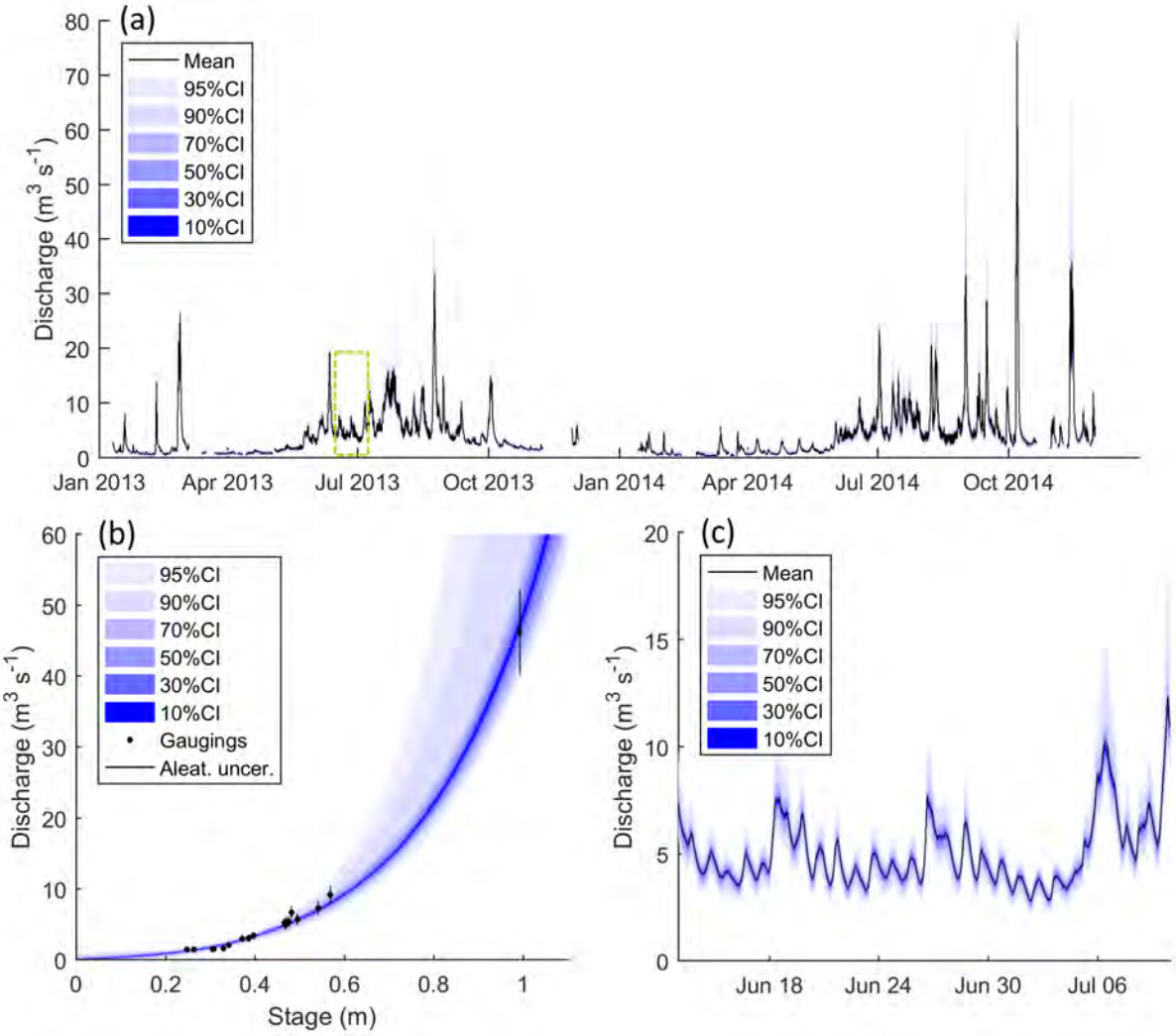

Figure 4.6: River flow time series from ASG1 with quantified confidence intervals (a), rating curve uncertainty used to quantify confidence intervals (b) and zoomed section of river flow time series (see yellow dash box in top plot) with confidence intervals (c).

section (99-66% flow exceedance), high flow section (15-5% flow exceedance) and highest flow section (5-0.5% flow exceedance). An overall estimate of flow variability, the coefficient of variation, was also calculated. Related to this, two further signatures, the rising limb density and integral scale, provide a measure of flashiness. The rising limb density is the ratio of number of flow peaks to the total time to peak where a higher number is more flashy. The integral scale measures the lag time at which the autocorrelation function of the flow time series falls below $\frac{1}{e}$ (diurnal cycles in river flow were removed prior to this using a moving average filter). A higher integral scale therefore indicates a more slowly responding hydrological system. Finally, the peak summer flow hour of the observed discharge time series was calculated to characterise the intra-day river discharge response to melt.

River discharge calculated using the rating curve approach outlined in section 3.5.4 are inherently uncertain (Pappenberger et al., 2006). McMillan and Westerberg (2015) provide a useful definition of two important sources of uncertainty which they distinguish as either aleatory (random) or epistemic (of an unknown character). The first stem from random measurement errors such as those from the instrument used for periodic river gaugings. These cause gauging points to vary around the 'true rating curve', typically according to some formal statistical definition. Epistemic uncertainty stems from the assumptions hydrologists have to make when constructing rating curves such as assuming the river bed profile and horizontal flow velocity distribution is relatively stable over time. These errors make fitting a single rating curve to all of the gauging data invalid. Accordingly, McMillan and Westerberg (2015) propose a method to define rating curve uncertainty which accounts for both sources of error which has been used to estimate uncertainty in river discharge signatures (Westerberg et al., 2016). The random error component was defined from analysis of 27 flow gauging stations in the United Kingdom with stable ratings and without obvious epistemic errors (Coxon et al., 2015). They conclude that this source of error is best approximated by a logistical distribution model. The logistic function with scale parameter σ and zero mean has the following probability

density function (PDF):

$$f_{LG}(x;\sigma) = \frac{e^{-\frac{x}{\sigma}}}{\sigma(1+e^{-\frac{x}{\sigma}})^2} \tag{4.14}$$

and inverse CDF:

$$ICDF_{LG}(p;\sigma) = \sigma ln(\frac{p}{1-p}) \tag{4.15}$$

where p is the probability conditional on σ. Following, McMillan and Westerberg (2015), the scale parameter is calculated for each gauging following:

$$\sigma = 4.18e^{-3.051Q_n} + 3.51 \tag{4.16}$$

where Q_n is he normalised gauged discharge.

To account for epistemic error, they reject the assumption that the rating curve is fixed in time and instead they fit an ensemble of rating curves to all of the gauging data. Each curve is weighted by a 'Voting Point' likelihood function which scores it based on how many points of the periodic gaugings it is able to intersect (and at what location on the logistical PDF of each gauging). The voting point likelihood is calculated as:

$$L_{VP} = w \sum_{g=1}^{n} \begin{cases} \frac{f_{LG}(R_g-1;\sigma_g)}{f_{LG}(0;\sigma_g)} & |R_g - 1| < ICDF_{LG}(0.025;\sigma_g) \\ 0 & otherwise \end{cases} \tag{4.17}$$

where n is the number of gaugings, R_g is the ratio between gauged and rated discharge for point g and w is an additional weighting parameter which rates the curve according to the proportion of the space occupied by the gauging points that the fitted curve captures:

$$w = \left(\frac{max(h_{fit}) - min(h_{fit})}{max(h) - min(h)}\right)\left(\frac{max(q_{fit}) - min(q_{fit})}{max(q) - min(q)}\right) \tag{4.18}$$

where h and q are the gauging stage and flow data and h_{fit} and q_{fit} are those gaugings that are captured ($|R_g - 1| < ICDF_{LG}(0.025;\sigma_g)$) by the rating curve.

This method allows each rating curve to fit a subset of all gaugings only and takes account of the fact that there may be more than one rating curve consistent with the gauge data due to epistemic errors. Following McMillan and Westerberg (2015), Markov chain Monte Carlo sampling was used to define the rating curve ensemble (made up of 667 unique rating curves) which defines the rating curve uncertainty (Fig. 4.6b). From

these an equivalent distribution of each river discharge signature was derived from the ensemble of flow time series (Fig. 4.6c), from which the 95% confidence bounds were used as the LOA. Because the Voting Point method only accounts for uncertainty in the flow magnitude and not the timing, it was not suitable to apply this approach to the three signatures that characterise melt runoff timing and flashiness. For these signatures, Schaefli (2016) proposed that the LOA should be derived by subsampling different periods of the flow time series. For this study a month-by-month subsampling strategy was employed to do this.

4.4.4 Model calibration procedure

The GHM was configured to run from 1988 to 2015 so that simulations could be compared against all observation signatures. The initial ice surface was set to the 1988 DEM of the ice while the bedrock and land surface topography were taken from the Öræfajökull bedrock map (Magnússon et al., 2012). Initial snow coverage, soil moisture, linear reservoir storages and ELA were determined by running the model for three consecutive years prior to the simulation period using climate data from 1985 to 1988.

In total there were nine possible structural configurations of the GHM including all possible combinations of the three melt and ROR structures. For each of the nine configurations, the melt and ROR parameters were calibrated to achieve the closest fit to the observed signatures. To do this, first a set of preliminary runs were undertaken to assess the sensitivity of the simulations to the parameters. Here, it was found that the simulations were insensitive to the firn melt parameters across the range of 33 signatures. Accordingly, these were set to the same values as for snow. Similarly, none of the signatures were sensitive to the threshold above which melt occurs, T^*, and accordingly this was set to 0 °C throughout the model experiments. Finally, it was also decided to fix the albedo parameter for ice in TIM_3 to 0.3. This was because this parameter directly interacts with the b parameter and therefore provides no extra control over model behaviour.

The remainder of parameters were kept for calibration (see Table C.1). For each GHM

configuration, 5000 Monte Carlo simulations with random parameter sets sampled from predefined uniform distributions were undertaken. The prior parameter distributions were defined from a review of previous modelling studies and later refined during the preliminary runs noted above. The quasi-random Sobol sampling strategy (Brately and Fox, 1988) was employed to sample the parameter space as efficiently as possible. The simulated signatures from each model run (parameter set) were then evaluated against the observed signatures using a continuous acceptability score that is analogous to those used in other signature-based hydrological studies (Coxon et al., 2014; Shafii and Tolson, 2015). This objective function explicitly accounts for uncertainty in the observation signatures, so that decisions about model appropriateness can be made within the uncertainties of observation data. The acceptability for signature j is defined as:

$$s_j = \begin{cases} 0 & low_j \leq sim_j \leq upp_j \\ \frac{sim_j - upp_j}{upp_j - obs_j} & sim_j > upp_j \\ \frac{sim_j - low_j}{obs_j - low_j} & sim_j < low_j \end{cases} \tag{4.19}$$

where obs_j and sim_j are the observed and simulated values and upp_j and low_j are the upper and lower LOA. A score of zero indicates that the model captures the observed signature within the LOA. An absolute score greater than 0 is outside of the LOA and therefore unacceptable. The sign of the score indicates the direction of bias while its magnitude indicates the model's performance relative to the LOA. A score of -3 would indicate that the model underestimates the signature by three times the observation uncertainty. This score therefore does not penalise a model if it falls within the observation uncertainty of a signature. It is also tolerant of projections that fall outside of the LOA where observation uncertainty is high which is a desirable attribute given the range of signatures used in the calibration.

Given that there are 33 different signatures to calibrate to simultaneously, it was important to define a weighting scheme to achieve the best overall performance across the range of signatures. It was decided that, for a given GHM configuration, the 5000 runs should be ranked by a weighted average score where each group, each attribute within

each group and each signature within each attribute were given equal weighting so that the scores were not biased to a particular group or attribute. The top 1% of model runs that achieved the smallest weighted average acceptability scores were then taken as the calibrated models for each GHM configuration and the average acceptability scores of these are reported. A bootstrapping with replacement re-sampling scheme was also used to assign 95% confidence intervals around all reported acceptability scores. While not a formal test of statistical significance, these were used to avoid reporting differences between the GHM configurations where issues such as under-sampling of the parameter space would make such conclusions unjustified. Where confidence intervals do not overlap, differences are hereafter referred to as substantial. The different GHM configurations were also compared when calibrated to individual groups of signatures (ice melt, snow coverage and river discharge). In this case the same weighting procedure was applied to a single group only.

4.5 Results

4.5.1 Signature discrimination power

As a first step towards evaluating the LOA framework, the discrimination power of the signatures was investigated to determine their relative usefulness for discriminating between acceptable and unacceptable model structures and parameterisations when used individually. A total of 45,000 calibration runs, each with unique model structures and parameterisations (hereafter referred to as model compositions) were undertaken in this study. The signatures with the highest discrimination power were defined as those that best constrain the range of acceptable model compositions. Here, the total number of acceptable model compositions were calculated for each signature as an indicator of discrimination power (bars in Fig. 4.7a). The results indicate that the ice melt signatures are the best discriminators, where each accepted less than 5000 model compositions. Of these, the winter melt signature from the ablation stake measurements is the best discriminator while the summer melt signature shows the least discrimination power. The snow

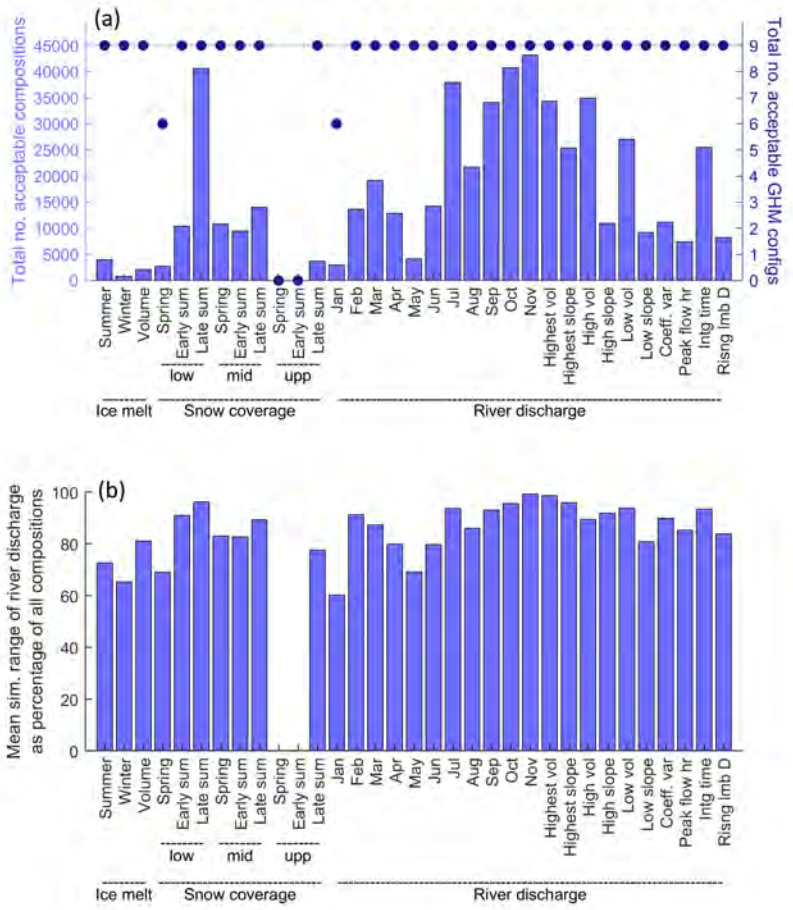

Figure 4.7: Total number of acceptable model compositions (bars) and configurations (dots) for each signature (a) and mean simulated range in river discharge from the population of acceptable models as a percentage of the simulated range using all of the 45000 model compositions (b).

coverage signatures generally show to be inferior discriminators when compared to the ice melt signatures. The late-summer snow coverage signature for the lower catchment shows to be the poorest discriminator, presumably because there is negligible snow cover here at this time of the year; an observation that almost all of the model compositions have no difficulty in replicating. In contrast, no model compositions are deemed acceptable for the signatures of the spring and early-summer snow coverage in the upper catchment.

The discrimination power of the river discharge signatures shows to be highly variable, but there are several discernible patterns. Firstly, the mean monthly flow signatures between January and June, when river discharge is low, show to be better discriminators

than the higher-flow signatures from July to October. The mean monthly January and May flows stand out as being particularly powerful at discriminating between acceptable and unacceptable model compositions suggesting that these are likely to be important focal points for characterising model deficiencies. Those signatures related to the variability of flows such as the coefficient of variation and the FDC slope signatures, as well as peak flow hour (timing) and rising limb density (flashiness) also show to be relatively good discriminators.

To determine the structural discrimination power of each signature, the total number of GHM configurations that returned at least one acceptable simulation has also been calculated for each signature (scatter in Fig. 4.7a). They show that when used individually, most of the discrimination power stems from constraining the parameter space rather than constraining the structural space. Only the lower-catchment spring snow coverage and mean January river flow signatures discriminate between structures where only six of the nine GHM configurations returned acceptable simulations. In both cases it was the GHM configurations that employed the TIM_3 melt model structure that could not capture these signatures within their LOA.

To indicate how each signature helps to reduce river flow prediction uncertainty, a second measure of discrimination power has also been calculated (Fig. 4.7b). Here, the mean simulated range in river discharge from the population of acceptable models has been calculated as a percentage of the simulated range using all of the 45,000 model compositions for each signature. These results show that when used individually, all of the signatures help to constrain the river flow prediction uncertainty, although the effectiveness of each is variable. The mean January and May river flow signatures again exhibit good discrimination power, reducing the mean river discharge uncertainty to 60-70% of that from the full population of model compositions. Similarly, the winter ice melt and spring snow coverage in the lower catchment remain as two of the best discriminators. However, some signatures such as the long-term volumetric change in the glacier, which showed to be a good discriminator of model acceptability, are not as effective at reducing river discharge prediction uncertainty.

4.5.2 Acceptability of melt model structures

While all signatures clearly demonstrate discrimination power when used individually, it remains to be seen how effective the LOA framework is for discriminating between and diagnosing deficiencies in different model structures when using multiple evaluation criteria. Here, the acceptability scores obtained after calibrating the GHM to the different groups of signatures (ice melt, snow coverage and river discharge) using the three different melt model structures have been calculated (Fig. 4.8). The light grey boxes indicate those signatures that have been captured within the LOA and the dark grey boxes and their corresponding acceptability scores indicate those signatures for which the structures were not able to capture within the LOA. So that the river discharge acceptability scores can be compared fairly, they have all been obtained using the ROR_1 runoff-routing structure. When calibrated against the ice melt signatures, the GHM is not able to capture them within their LOA, regardless of the melt model structure used. The different GHM configurations show a tendency to overestimate the measured summer and winter melt from the ablation stake data, yet underestimate the long-term change in total ice volume (note underestimation here refers to the simulated loss in ice volume). This highlights a deficiency in the melt model structures as they are unable to reconcile the three melt signatures simultaneously within the observation uncertainty. The winter melt is by far the most unacceptable simulation, particularly when using the TIM_1 structure where it is overestimated by more than 30 times the observation uncertainty.

Each of the GHM configurations using the three melt model structures have been ranked from 1 to 3 in the top left corner of each box where the acceptability scores are substantially different (Fig. 4.8). While there are clearly differences in the acceptability scores for the summer melt and ice volume signatures, these are not substantially different and therefore it is not possible to say that one structure is more acceptable than another. Indeed, a comparison of the simulated ice thickness change along the Falljökull and Virkisjökull arms of the glacier reveal that all three configurations of the GHM produce almost identical simulations which broadly capture the observed ice thickness change

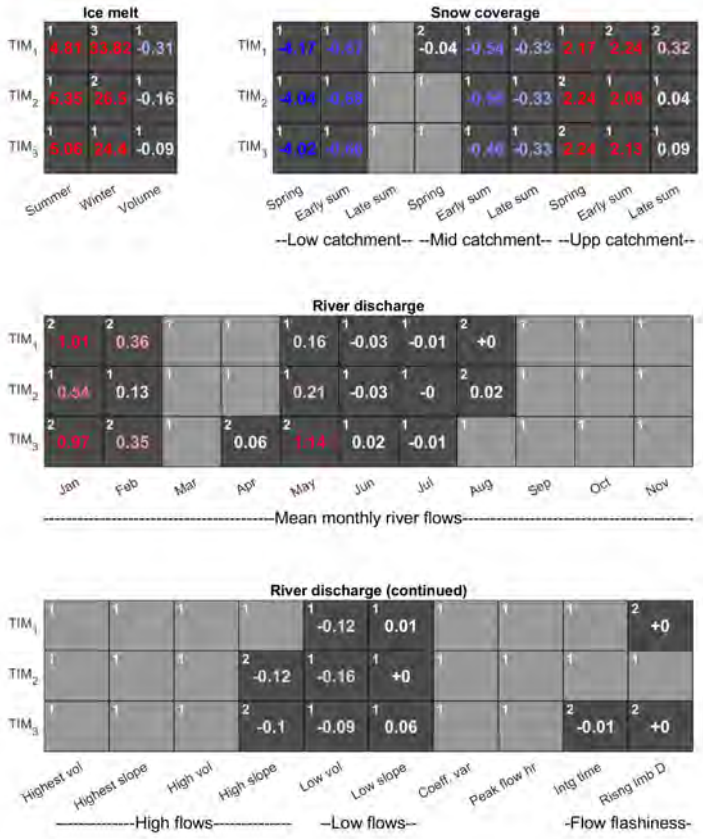

Figure 4.8: Acceptability scores obtained after calibrating the GHM using the three melt model structures in combination with the ROR_1 runoff-routing model structure. The three GHM configurations were calibrated against ice melt, snow coverage and river discharge signatures separately. Light grey boxes indicate acceptable simulations ($s = 0$) and numbered, dark-grey boxes indicate unacceptable simulations coloured blue and red to indicate negative and positive bias respectively. Note, all acceptability scores are rounded to two decimal places. Those non-zero scores that round to zero are accompanied by +/- to indicate sign of score. White numbers in top left of each box indicate relative ranking where acceptability scores are substantially different between the GHM configurations.

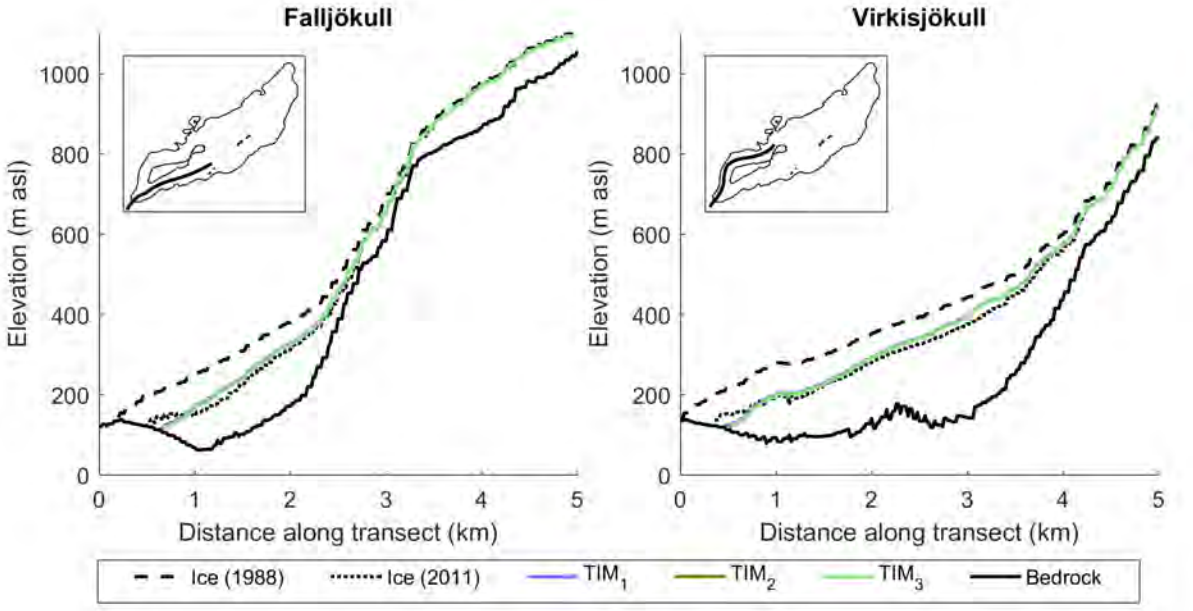

Figure 4.9: Observed and simulated ice thickness change as measured along transects of the Falljökull and Virkisjökull glacier arms. Insets show transect location.

between 1988 to 2011 (Fig. 4.9).

For the winter melt signature, there is a substantial difference in acceptability when using the three melt model structures. Here, the GHM configuration using the TIM_3 structure is the most acceptable while that using the TIM_1 structure is least acceptable, indicating that while all configurations produce simulations outside of the LOA, there is an improvement in ice melt simulations when implementing the most sophisticated TIM_3 melt model structure.

For the snow coverage signatures, all three of the GHM configurations capture the late-summer snow coverage in the lower portion of the catchment within the LOA. When using the TIM_2 and TIM_3 structures the mid-catchment spring snow coverage is also captured. The remaining snow coverage signatures are not captured within the LOA where all configurations show a tendency to underestimate snow coverage in the lower and mid parts of the catchment and overestimate snow coverage in the upper part of the catchment. To investigate why this is, Fig. 4.10a shows the simulated early-summer mid-catchment and upper-catchment snow coverage signatures for the 5000 calibration parameter sets (blue dots) used with the TIM_1-ROR_1 GHM configuration. Here it can

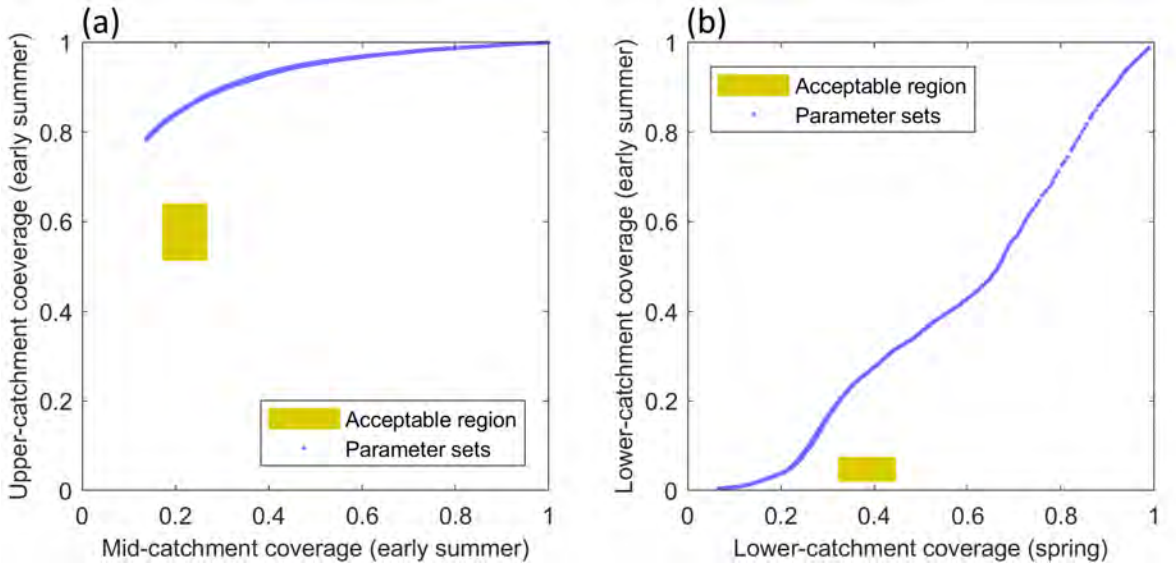

Figure 4.10: Simulated snow coverage signatures from the 5000 calibration runs (blue dots) for the TIM_1-ROR_1 GHM configuration including: early-summer mid-catchment and upper-catchment snow coverage signatures (a), and lower-catchment spring and early-summer snow coverage signatures (b).

be seen that regardless of the choice of melt model parameters, this structure is not able to capture both of these signatures within their LOA simultaneously (indicated by yellow area). A similar inconsistency exists when comparing snow coverage over different seasons where the GHM is not able to capture the lower catchment snow coverage in the early-summer and spring simultaneously (Fig. 4.10b). Indeed, this inconsistency extends across all melt model structures.

A comparison of simulated snow distribution curves from the calibrated models (Fig. 4.11) reveals that all return similar simulations. The simulation using TIM_1 deviates slightly from the curve produced by the GHM when using the TIM_2 and TIM_3 structures, but overall the choice of melt model structure has a limited influence on the simulated seasonal snow coverage.

The acceptability scores for the river discharge signatures in Fig. 4.8 show that regardless of the choice of melt model structure, when used in conjunction with the ROR_1 runoff-routing model structure, all are able to capture a range of the river discharge signatures. The simplest GHM configurations using the TIM_1 and TIM_2 model structures

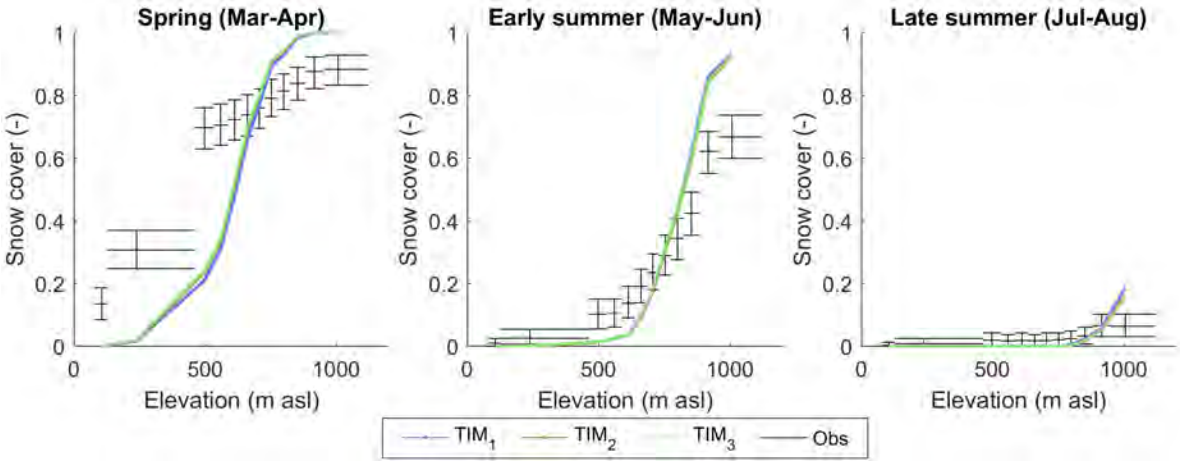

Figure 4.11: Simulated seasonal snow distribution curves when using the three melt model structures.

capture 12 river discharge signatures simultaneously within the LOA while the inclusion of the dynamic snow albedo term and re-arrangement of the melt equation in the TIM_3 melt model actually inhibits the GHM performance where only 10 of the 21 river discharge signatures are captured within the LOA.

The mean monthly flow signatures for January, February and May show some of the highest absolute acceptability scores indicating the models are least efficient at capturing these. For winter flows in January and February, the simulation using the TIM_2 model structure is substantially more acceptable than when using the other melt model structures although it should be noted that, given that flows are very low here, the absolute error is less than 0.2 $m^3 s^{-1}$. A comparison of the simulated ice melt during May 2013 reveals that the TIM_3 structure simulates the highest ice melt of all three melt model structures (Fig. 4.12a) which results in a positively biased river flow time series (see Fig. 4.12b). Note, the full input/output time series over the observation period can be found in Appendix E.

Furthermore, a comparison of the simulated ice melt time series over 2013 with a monthly moving average filter demonstrates that the positive melt bias from TIM_3 extends between April and June (Fig. 4.13b) which corresponds to the period where temperatures are relatively low, but where incoming solar radiation is relatively high (see Fig. 4.13a).

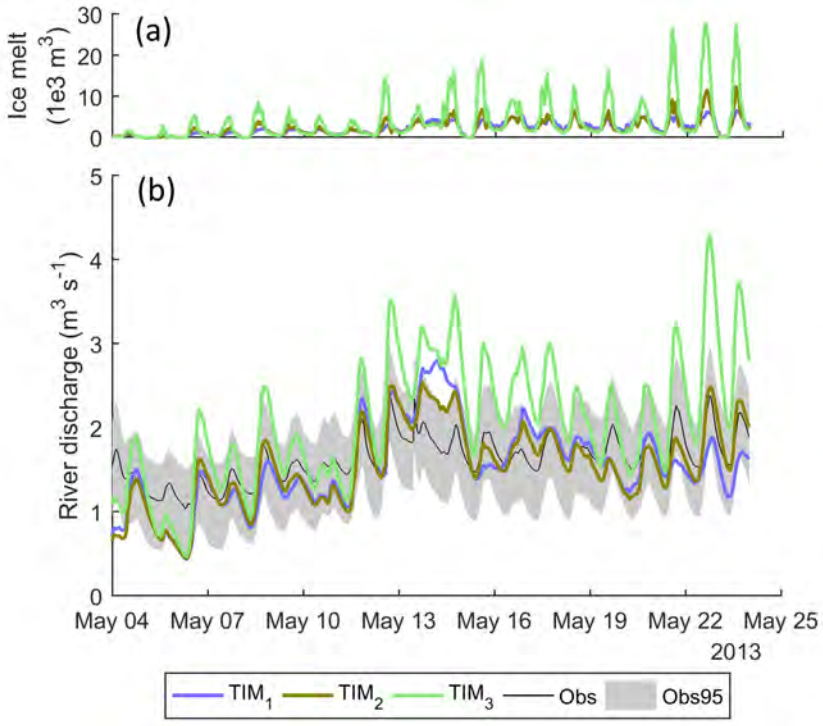

Figure 4.12: Mean simulated hourly ice melt (a) and river discharge (b) during May 2013 using the top 1% of models from the three melt model structures in combination with the ROR_1 runoff routing model structure.

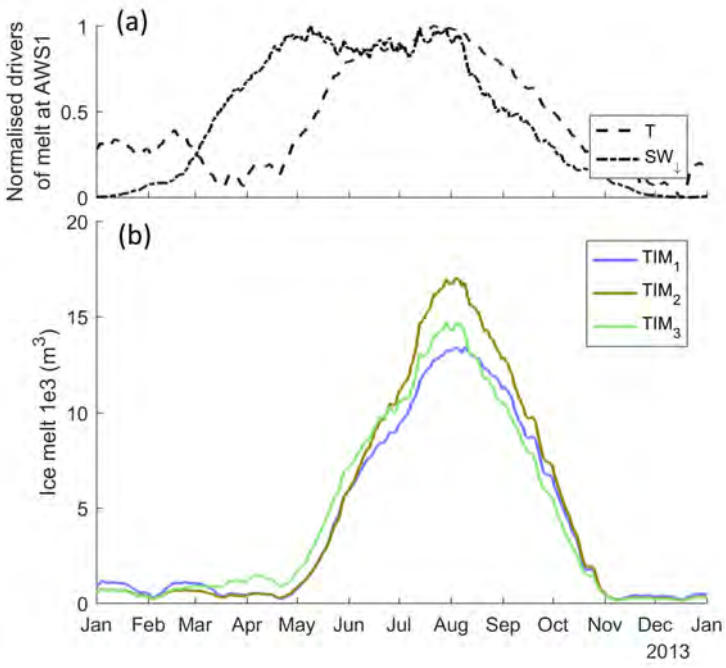

Figure 4.13: Normalised temperature and incident solar radiation (a) and simulated ice melt from the three calibrated ice melt model structures (b) for the year 2013. All time series use a monthly moving average filter.

Of the remaining river discharge signatures, only a handful show any substantial difference when switching between the melt model structures including the mean April and August discharge and the two 'flashiness' signatures: the integral time and the rising limb density. However, the differences here are very small. For the 'high slope' signature, which characterises the variability of high flow river flows, the simulation using the TIM_1 melt model structure is able to capture it within the LOA, while the simulations using the TIM_2 and TIM_3 model structures both show a negative bias suggesting they underestimate high flow variability.

4.5.3 Acceptability of runoff-routing model structures

To evaluate the ROR structures, acceptability scores have been calculated for the river discharge signatures only as these structures do not influence ice melt or snow coverage (Fig. 4.14). To ensure fair comparison between the different structures, all scores have been obtained using the simplest TIM_1 melt model structure in the GHM.

It was noted previously, that all melt model structures used in combination with ROR_1 resulted in positively biased January and February river flows. It could be that including a more complex non-linear ROR structure in the GHM could help to mitigate this bias. Indeed, the calibrated simulations do show a substantial reduction in positive bias for the mean February flows when using ROR_2 and ROR_3, however the simulations are still unacceptable. Furthermore, for the mean January river flow there is no substantial change in acceptability score. This indicates that the runoff-routing representation is also not the reason for this overestimation of flows at the beginning of the year. To investigate this positive bias further, Fig. 4.15c shows the simulated time series from the calibrated models using TIM_1 in combination with ROR_1, ROR_2 and ROR_3 for January and February 2013. Figure 4.15a shows that melt is an insignificant input during these winter months (green line). Rather it is rainfall (black dash) that dominates the runoff input and this results in two pronounced peaks in the simulated river discharge time series. The different behaviour of the simulations using the three ROR structures is much more obvious during the rainfall-runoff events. The simulation using the ROR_1 structure is noticeably more

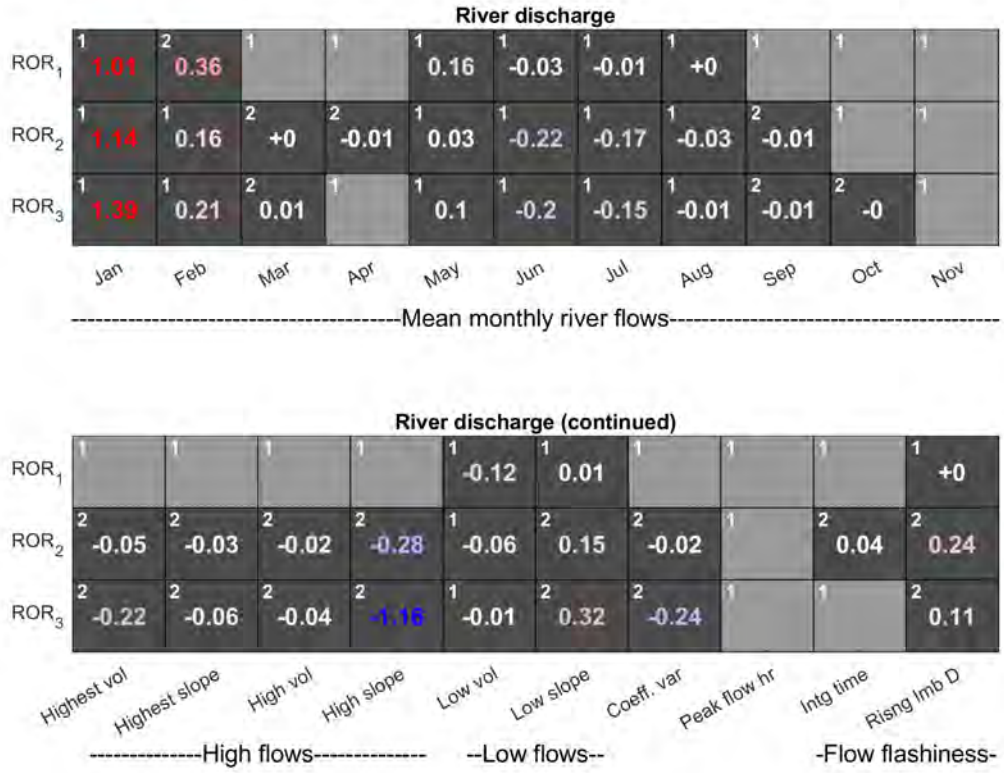

Figure 4.14: Acceptability scores obtained after calibrating the GHM using the three ROR structures in combination with the TIM_1 melt model structure. Light grey boxes indicate acceptable simulations ($s = 0$) and numbered, dark-grey boxes indicate unacceptable simulations coloured blue and red to indicate negative and positive bias respectively. Note, all acceptability scores are rounded to two decimal places. Those non-zero scores that round to zero are accompanied by +/- to indicate sign of score. White numbers in top left of each box indicate relative ranking where acceptability scores are substantially different between the GHM configurations.

flashy in response to the rainfall and overestimates the peak flows while the ROR_2 and ROR_3 simulations, which include additional, more diffusive representations of the flow of water through snow and firn, result in peak flows that are closer to the observed, but with a recession that is too shallow. Regardless of these deficiencies, however, all result in an almost identical positive bias as shown by the cumulative flow in Fig. 4.15b.

There are however differences when assessing other aspects of the river discharge time series, particularly in the signatures relating to high flows. In Fig. 4.14, it can be seen that while the simulation using the ROR_1 routing model structure is able to capture all of the high flow signatures simultaneously, the ROR_2 and ROR_3 structures show an unacceptable negative bias for these signatures indicating underestimation of high flow magnitude and variability. To evaluate this in more detail, Fig. 4.15f shows the simulated time series for the highest recorded river flow event during October 2014. Here, the flashier and more responsive ROR_1 structure achieves the closest fit to the observed peak flow and within the uncertainty bounds while the more diffusive, ROR_2 and ROR_3 structures underestimate the peak flow. Note they also underestimate the overall river flow variability as indicated by the coefficient of variation signature.

4.5.4 Consistency of melt model structures

The results so far have highlighted some inconsistencies in the GHM configurations using the melt and runoff-routing model structures where they are unable to reconcile some combinations of signatures simultaneously. This is important as those inconsistencies could help to further diagnose structural deficiencies in the different model structures. To investigate this, consistency scores have been calculated between pairs of the 33 signatures for each GHM configuration. A model can be deemed consistent across a pair of signatures if it is able to capture both within their LOA simultaneously. The consistency scores are therefore calculated as the minimum sum of the two acceptability scores between a pair of signatures across the 5000 calibration runs for each GHM configuration.

Figure 4.16 shows the average consistency scores calculated across the signatures for each attribute of ice melt, snow coverage and river discharge using the three melt model

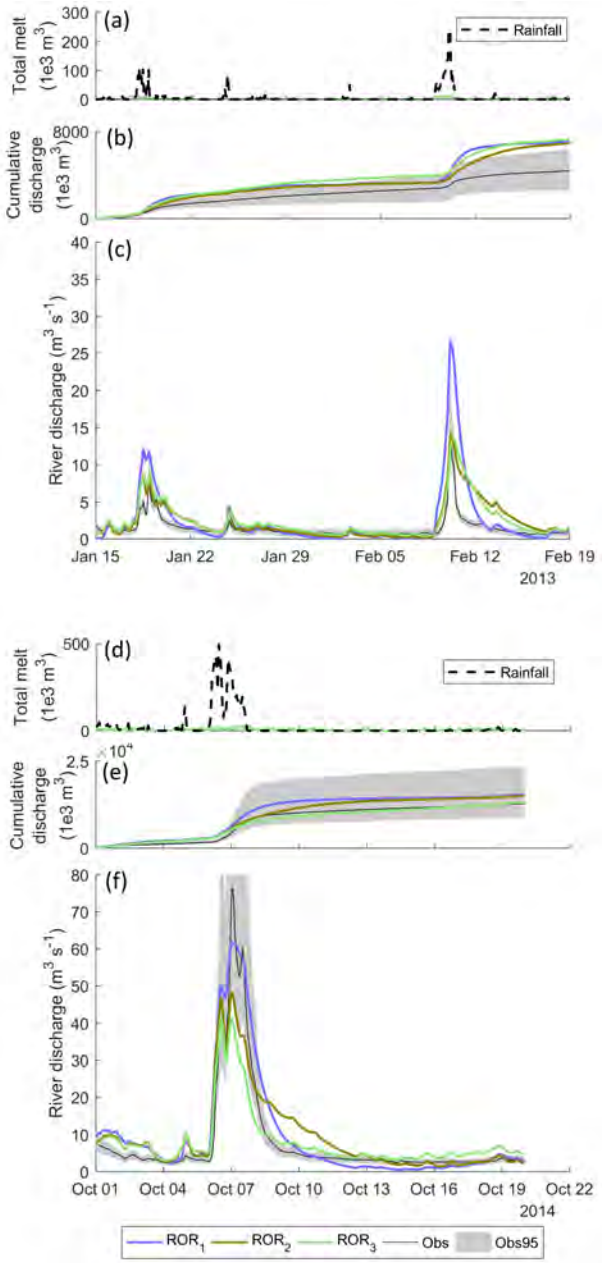

Figure 4.15: Simulation time series using the three different ROR structures in combination with the TIM_1 melt model structure including simulated total melt and rainfall (top), cumulative river discharge (middle) and river discharge time series (bottom) for January and February 2013 (a,b,c) and the October 2014 flood (d,e,f).

structures in combination with the ROR_1 runoff-routing structure. The top panel shows the consistency scores when using the simplest TIM_1 melt model structure. The regions in red highlight the areas where the GHM is inconsistent. The first striking observation is the red band along the upper catchment snow coverage attribute. It has already been demonstrated that the simulations using the TIM_1 structure cannot reconcile the upper-catchment snow coverage with the remaining snow coverage signatures. This further demonstrates that when using the TIM_1 structure, the GHM cannot reconcile the upper-catchment snow coverage with any of the other attributes.

The largest inconsistency score obtained was between the short-term, seasonal melt on the glacier tongue and long-term total glacier volume change. It should be noted that the seasonal melt signatures show a small inconsistency with the lower-catchment snow coverage and a larger inconsistency with the upper-catchment snow coverage. The total glacier volume change signature, however, is also inconsistent with the monthly flow and low flow signatures indicating that it is the long-term glacier wide mass balance that the model is getting wrong.

The use of the TIM_2 model structure which includes topographic effects goes some way to reducing most of the inconsistencies shown using the TIM_1 model structure (Fig. 4.16 middle panel). However, all but one of the inconsistencies (between lower-catchment snow coverage and seasonal melt) remain, indicating that the use of the TIM_2 melt model structure only provides a small improvement in model consistency.

Using the TIM_3 model structure also helps to improve model consistency, particularly those associated with the upper snow coverage, but surprisingly it also introduces new inconsistencies in relation to the lower-catchment snow coverage, where the model is not able to reconcile these signatures with any of the other attributes.

4.5.5 Consistency of runoff-routing model structures

Consistency scores have also been calculated for each pair of river discharge signatures (Fig. 4.17) using the three runoff-routing structures in combination with the TIM_1 melt model structure. The simulations using the ROR_1 structure (top panel) and next simplest

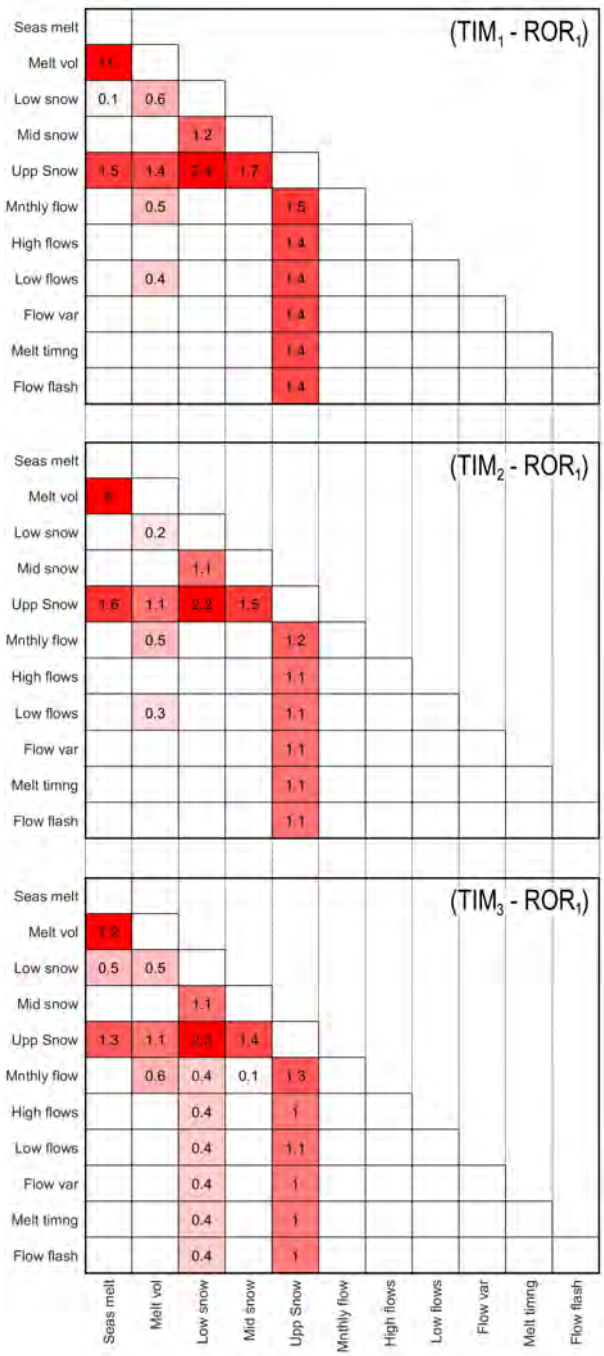

Figure 4.16: Average consistency scores between attributes using the three melt model structures in combination with the ROR_1 runoff-routing structure. Scores of < 0.1 have not been reported.

ROR_2 structure (middle panel) show a very similar pattern of model inconsistencies. Firstly, both sets of simulations do not capture the relatively low flows in February and the relatively high flows in July and August simultaneously. This corroborates the findings from the acceptability analysis which revealed a tendency for the model structures to overestimate low flows in the winter and underestimate high flows in the summer and autumn, particularly with relation to rainfall-induced high flows. Interestingly though, the seasonal flow inconsistency is centred on February and there are not inconsistencies for the other low flow months from January to April. This provides further evidence that it is particularly the rainfall-induced flows that the model is not able to capture effectively. In fact, February has some of the highest flows in the record of winter flows induced by large rainfall events (see average flow signatures in Table 4.2). This suggests this could be the reason that the inconsistencies between winter and summer flows are centred around these months. The inclusion of additional flow pathways in the routing routine only enhances these inconsistencies, particularly when using the ROR_3 model structure where the inconsistencies extend into June (bottom panel).

The ROR_1 simulations show inconsistencies between the February flows and low flow variability as indicated by the low slope signature. The reason for this is not clear, but interestingly, the inclusion of an extra, more diffuse, flow pathway in the ROR_2 model appears to remedy this, suggesting that there is some non-linear behaviour that the ROR_1 model structure cannot capture. However, it comes at the cost of inducing an extra inconsistency between the mean flows in January and the overall flow variability as indicated by the coefficient of variation. This new inconsistency is amplified when using the ROR_3 structure.

Interestingly, the consistency scores when using the ROR_1 and ROR_2 structures are relatively similar, with each configuration demonstrating inconsistencies between four and five pairs of river discharge signatures respectively. In contrast, using the most complex ROR_3 structure introduces a number of new inconsistencies with a total of 12 inconsistent pairs of simulated river discharge signatures. These new inconsistencies are centred around the mean monthly flow signatures as well as the signatures relating to high and low flow

magnitude and variability.

4.6 Discussion

4.6.1 Using the LOA framework for diagnosing GHM structure deficiencies

The first aim of this study was to investigate if a signature-based LOA framework could be used to diagnose deficiencies in the different melt and runoff-routing model structures. The comprehensive set of signatures provided a powerful method to evaluate the model behaviour. Furthermore, when used within a LOA framework, it was straightforward to identify those aspects of the glacio-hydrological system that the GHM configurations could and couldn't capture.

All of the GHM configurations were able to capture 29 of the 33 signatures within their LOA when evaluated against the signatures individually while none of the GHM configurations were able to capture the observed spring and early-summer snow coverage in the upper section of the catchment, highlighting these aspects of glacio-hydrological behaviour as key deficiencies in the tested GHM configurations. A more revealing analysis of structural deficiencies, however, was achieved by evaluating the ability of the models to capture multiple signatures simultaneously. For example, all GHM configurations were able to capture the three signatures of ice melt individually within their LOA, but none of them could capture all of the signatures simultaneously. The challenge here was to reconcile three signatures that characterise glacier melt over different spatial and temporal scales. The fact that none of the total tested GHM configurations could capture these signatures simultaneously is likely because the three TIM structures lump a number of terms from the full energy balance equation into a handful of calibration parameters which may lack robustness in space and time (Gabbi et al., 2014; MacDougall et al., 2011; Matthews et al., 2015). All TIM structures can therefore be deemed 'unacceptable' given they were not able to capture all of the ice melt signatures.

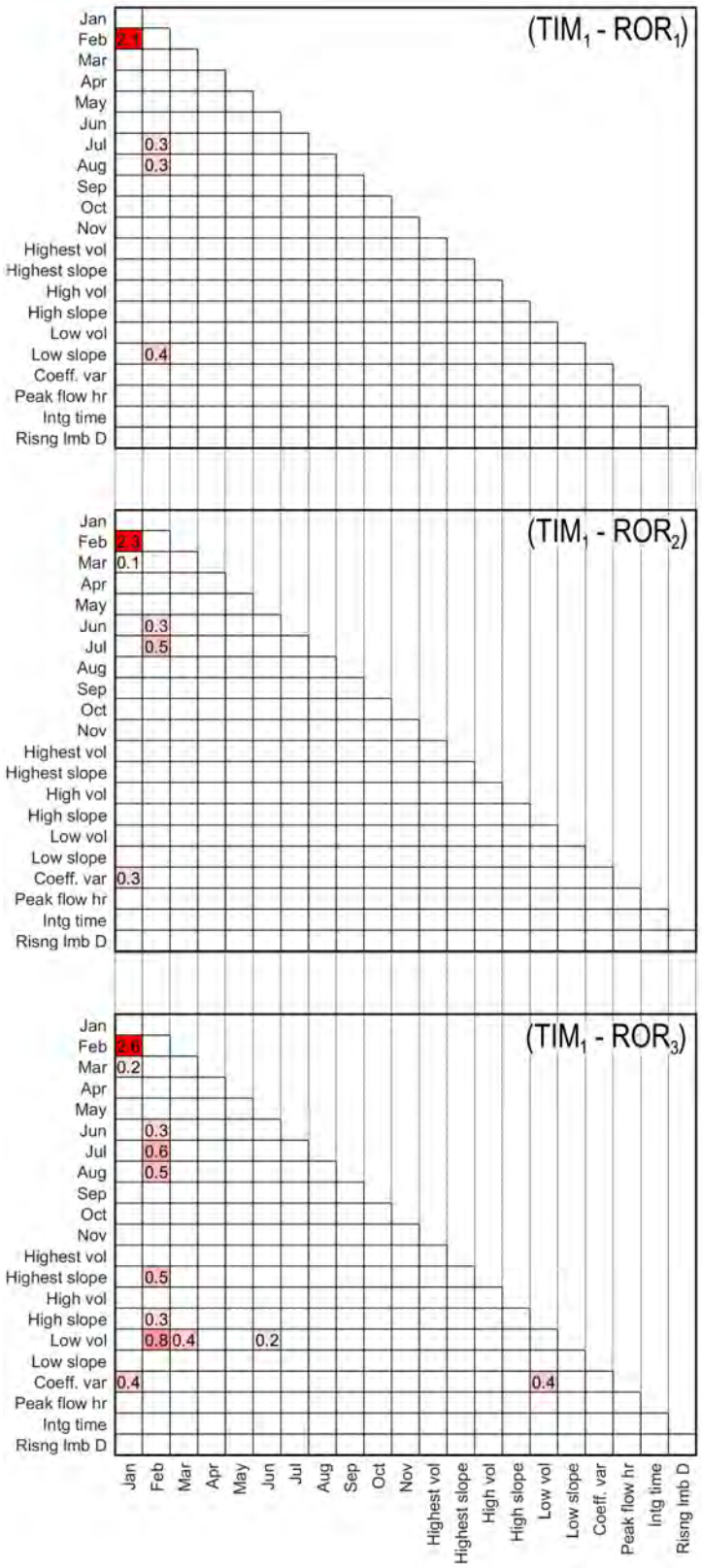

Figure 4.17: Average consistency scores between river discharge signatures using the three ROR structures in combination with the TIM_1 melt model structure.

While none of the TIM structures were technically acceptable, the acceptability objective function adopted in the LOA framework demonstrated that including solar and topographic effects in the TIM_2 and TIM_3 melt model structures improved model efficiency. The inclusion of these processes in conjunction with the dynamic snow albedo parameterisation in TIM_3 returned the most acceptable simulations of the ice melt signatures overall.

The results from this study, however, showed that the inclusion of extra model complexity does not guarantee gains in model acceptability. For example, the most sophisticated TIM_3 melt structure was also the least acceptable structure for the mean May river flow signature where it showed the highest positive bias. This is interesting, as May coincides with the beginning of the main melt season which indicates an inability to capture this initialisation properly. It was shown that the simulated snow coverage signatures were almost identical when using the three melt model structures indicating that this deficiency did not stem from the dynamic snow albedo component of TIM_3. Furthermore, May corresponds to the period where temperatures are relatively low, but where incoming solar radiation is relatively high which suggests that it is the additive form of the TIM_3 melt equation and the subsequent increased influence of solar radiation on melt which induced the positive bias in flow simulations in the early melt season.

Similarly, the ROR_3 structure, originally proposed as the most realistic conceptual representation of water storage and transmission in the river basin, was the least acceptable model overall across the river discharge signatures. It also underestimated peak flows at the end of the melt season and underestimated overall river flow variability. These results highlight the need to exercise caution before introducing complexity to conceptual models of glacio-hydrological processes. They also illustrate the importance of testing prior assumptions about the system against other possible model hypotheses, for which a signature-based LOA framework is ideally suited.

The main deficiencies noted for all of the GHM configurations when compared to the river discharge signatures were an overestimation of the relatively low winter flows in January and February, and the flows at the start of the melt season in May. It was

assumed that the addition of extra 'slow' flow pathways in the ROR_2 and ROR_3 runoff-routing structures would help to correct for any deficiencies in capturing the hydrograph seasonality. Instead, the choice of runoff-routing structure had very little influence on these signatures, indicating that longer term storages of water do not have a major control on the seasonality of the hydrograph. This is probably because of the catchment's small size which leads to an instantaneous seasonal response to melt on a monthly timescale. For larger catchments, the monthly flow signatures are likely to be more sensitive to the choice of runoff-routing structure. Instead, the simulated mean monthly river flow signatures were more sensitive to the choice of melt model structure, particularly in May at the start of the melt season, which is likely due to the high degree of glaciation and thus the high meltwater runoff from the basin.

In contrast to the monthly river flow signatures, the choice of runoff-routing structure had by far the dominant control on those signatures that are controlled by flows operating on much shorter timescales such as the distribution of flows, flow variability and flashiness. This hierarchy of influence between the melt and runoff-routing model structures has important implications for river discharge projection uncertainty in glaciated basins. For example, if one were interested in future seasonal water resource availability, they would be most reliant on projections of mean monthly river flows. The results here indicate that, for this catchment at least, uncertainties in these projections stem primarily from melt model uncertainty. In contrast if one were interested in future changes in flood frequency, the dominant source of model projection uncertainty is the runoff-routing approach. Uncertainties in river flow projections from glacio-hydrological models are therefore likely to be dependent on the river flow characteristic of interest.

4.6.2 Using the LOA framework for GHM structure selection

The second aim of this study was to determine if the signature-based evaluation within a LOA framework could be used to constrain the prior population of model structures and parameter sets (compositions) down to a smaller population of acceptable models. The initial discrimination tests showed that all of the signatures have discrimination

power, although for two of the snow signatures, none of the 45000 model compositions could capture them. The mean January and May river flow signatures were the best discriminators, individually reducing mean river discharge uncertainty to 60 - 70% of that from the full population of model compositions, although it should be noted that the majority of this reduction stemmed from constraining the acceptable parameter sets rather than the model structures. These results indicate that a LOA framework could be used to find a population of acceptable model compositions. However, the fact that none of the prior 45000 compositions were able to capture all of the signatures means that this remains to be seen.

At a fundamental level, this indicates that the structural configurations of the GHM employed in this study are simply not efficient enough to capture the observation data within their observation uncertainty bounds. This could be addressed by including additional process representations within the GHM. For example, one process that is not represented at all in any of the GHM configurations, but which has shown to be important in for Icelandic glaciers, is refreezing of meltwater and rainfall (Johannesson et al., 1995). It is estimated that about 7% of total melt in valley glaciers in Iceland refreezes, and therefore, the inclusion of this process could also help to reduce runoff during the colder winter months where the models showed to overestimate runoff. Another attribute of the study site which was not accounted for was debris cover at the glacier terminus which could be an important control on point scale and overall ice mass balance. Some TIMs that include representations of debris cover do exist (e.g. Carenzo et al., 2016) and the signature-based LOA approach would provide the ideal framework for evaluating the added value of further structural modifications like these.

The results also demonstrated some degree of insensitivity to the different structures tested in this study. For example, none of the prior 45,000 model compositions were able to capture the spring and early-summer snow coverage in the upper catchment and all of the calibrated GHM configurations overestimated snow coverage in the upper catchment whilst underestimating it in the lower catchment. Furthermore, using the most sophisticated TIM_3 structure with the dynamic snow albedo function had almost no effect on

the overall acceptability across these signatures. This of course, could be an indication that the melt model formulation was not the primary source of model deficiency here. Snow coverage simulations are sensitive to other components of the GHM such as the snow redistribution model, which itself, is sensitive to the resolution of the DEM used to parameterise it. A coarser DEM resolution removes peaks and troughs in the land surface which can bring about more complex patterns of snow coverage. Similarly, the glacier volume change signature will be sensitive to the glacier evolution formulation and parameterisation. It is clear, therefore, that while the application of a LOA framework here has highlighted some differences in model behaviour between the different structures, the apparent insensitivity of the snow coverage signatures to structural modifications indicates that further gains may also be made by investigating other components of the GHM structure within this framework.

Beyond the structural nature of a GHM, deficiencies in the boundary conditions may also provide some explanation for the lack of acceptable model compositions. For this study, the driving precipitation data were relatively well constrained by observations within the catchment during the summer and autumn months of recent years, but there were fewer observations during the winter months and none at all before 2009. Furthermore, while the bias-corrected precipitation time series was well correlated over a three-day time step, it was not at an hourly time step. It is also important to note that precipitation observations were all collected at the bottom of the catchment and therefore driving precipitation data at the top of the catchment are less certain. Indeed, one could explain the tendency to overestimate snow coverage higher up in the catchment by a positive bias in the driving precipitation data here. Such a bias could also explain the modelled inconsistencies across signatures that characterise ice melt at different spatio-temporal scales. Furthermore, given the strong coupling between snow, ice and river runoff, deficiencies in capturing the snow and ice signatures could also propagate through to the hydrological representation of the catchment. For example, one could imagine how errors in the spatial distribution of snow could perturb the timing of runoff through the catchment given that snow distribution influences the behaviour of the semi-distributed

runoff-routing routine employed in the GHM. Such perturbations are likely to impact the ability of the GHM to capture the full range of river discharge signatures. Accordingly, it is important to stress the influence that biases in the driving climate data could have on the model acceptability across the different signatures. Of course, for balance, it should be noted that regarding precipitation higher up in the catchment, the limited number of measurements taken at the summit of Öræfajökull indicate that mean annual biases in driving precipitation data are small (Guðmundsson, 2000). Even so, the recent melt model comparison by Reveillet et al. (2017) suggest that uncertainties in driving precipitation data can cloud any differences between melt model behaviour.

Given the fact that none of tested compositions were acceptable across all signatures, one could implement further parameter, structure and boundary condition modifications in an attempt to identify a model, or group of models, that captures all of the signatures within their LOA. While this should be the grand aim of such an approach, it remains beyond the scope of this study. Even so, the LOA framework adopted here has provided a thorough understanding of which aspects of glacio-hydrological behaviour are well represented by the different models and which aren't. This newly-gained knowledge will be fundamental to selecting a justifiable range of GHMs in the subsequent chapters. While this selection process will be detailed in the chapters that follow, the LOA framework has already shown that the inclusion of the most complex ROR_3 runoff-routing structure cannot be justified given that it introduced additional inconsistencies in model behaviour and no clear gains in acceptability across any of the model evaluation signatures. The added complexity and associated uncertainty in the parameterisation of this structure therefore cannot be justified and will not be adopted in the remainder of the research conducted in this thesis.

4.7 Conclusions

The signature-based, LOA framework adopted in this study provided a comprehensive evaluation of different GHM melt and runoff-routing model structures. In contrast to

traditional model evaluation approaches which rely on one or several global summary statistics, the adoption of 33 signatures helped to identify those aspects of the glacio-hydrological system that a particular model could or could not capture and the added value of introducing additional complexities to simplified process models.

When evaluated against individual signatures, the more complex model formulations did improve model simulations in some cases, but were not necessarily more consistent across the full range of signatures. This finding not only emphasises the need to exercise caution before introducing additional model complexities, but it also demonstrates the utility of the LOA framework for justifying such modifications. The often conflicting acceptability scores across the signatures highlights the difficulty and inherent uncertainty in model structure selection. It is clear, therefore, that future glaciological and hydrological projection studies that use simplified model structures should take account of these uncertainties, although to date these have rarely been considered. The results from this study indicate that, for future river flow projections in glaciated basins, the source of model uncertainty will depend on the particular river flow characteristic of interest.

An additional advantage of adopting a LOA framework is that it provides objective criterion for accepting or rejecting particular model structures and parameterisations. While all but two of the signatures demonstrated discrimination power, none of the 45,000 different model compositions tested in this study were able to capture them within their LOA simultaneously. Therefore, it remains to be seen if the framework can be used in this way, although applications that go beyond examining the melt and runoff-routing structural uncertainties and look at other process representations and model boundary conditions may prove more fruitful in obtaining a behavioural population of models.

4.8 Summary

The study undertaken in this chapter aimed to meet the first research objective of this thesis which was to, "implement a novel GHM comparison and selection framework to undertake a rigorous evaluation of multiple GHM structures". To meet this objective

a signature-based LOA framework to compare and diagnose deficiencies in the different snow/ice melt and runoff-routing model structures was implemented. The findings have provided new understanding of deficiencies in the different model structures and this understanding will be used to develop a model selection and calibration approach in the next chapter.

CHAPTER 5:

TWENTY-FIRST CENTURY EVOLUTION AND UNCERTAINTY OF GLACIER-FED RIVER FLOW REGIME

5.1 Introduction

This chapter uses the EURO-CORDEX climate projections to drive an ensemble of GHMs and project twenty-first century changes in different characteristics (magnitude, timing and variability) of river flow regime using a signature-based analysis. By doing so, this chapter will meet the second research objective outlined in chapter 1. This chapter also aims to meet the third research objective by undertaking a decomposition of the projection uncertainties to quantify the relative contributions of five components of the climate-GHM model chain across the signatures. These include the land use and greenhouse gas emission/concentration pathway (RCP), numerical climate model, climate downscaling parameterisation, snow/ice melt model and runoff-routing model. The projections from this study will be used as boundary conditions in the next chapter to drive a distributed groundwater model of the proglacial sandur aquifer.

The material presented in this chapter has been published in Hydrology and Earth System Sciences (Mackay et al., 2019). JM designed and undertook all practical elements of this study, including the downscaling of EURO-CORDEX climate projections, GHM calibration and the decomposition of projection uncertainties. He also led the analysis, interpretation and writing of the manuscript. NB, DH, SK, CJ, JE and GA contributed to the interpretation of the results. AB assisted in the construction of the harmonised river stage time series data used in this study.

5.2 Background

The presence of snow and ice in mountain watersheds profoundly affects characteristics of downstream river flow regime including flow magnitude, timing and variability over a range of timescales (Jansson et al., 2003; Mankin et al., 2015). This is partly due to the periodic (diurnal and seasonal) variations and longer-term (decadal) trends in meltwater inputs brought about by fluctuations in glaciological mass balance. In addition, the dynamic water storage and release properties of snow and ice (runoff-routing) control downstream river flow response to runoff over hourly to seasonal timescales (Willis, 2005). As such, glaciated basins exhibit river flow regimes that differ from their non-glaciated equivalents. Fountain and Tangborn (1985) analysed the effect of temperate glaciers on runoff variations for the North Cascade Mountains in the United states by comparing runoff records from glaciated and non-glaciated catchments and Chen and Ohmura (1990) did the same for the Rhone catchment. They showed that the presence of glaciers: i) increases annual streamflow through melt contribution; ii) delays the seasonal timing maximum runoff due to temporary storage of spring meltwater and peak meltwater production in the summer; and iii) decreases annual and monthly flow variation, particularly in the summer. The last of these effects is often referred to as the 'compensation effect' whereby partially-glaciated catchments demonstrate reduced inter-annual flow variability due to ice melt which compensates for precipitation variability. Indeed, the compensation of runoff from melt inputs can actually serve to increase mean runoff during anomalously dry heatwave events (Zappa and Kan, 2007).

Mountain glaciers are retreating at unprecedented rates (Zemp et al., 2015) while snow coverage is receding (Vaughan et al., 2013) resulting in observable changes to downstream river flow regimes including changes in annual mean flows and inter-annual flow variability (Luce and Holden, 2009; Singh et al., 2016b), changes in flow seasonality (Hernández-Henríquez et al., 2017) and increases in flood frequency (Matti et al., 2017). With near-surface air temperature projected to rise over the coming decades (Collins et al., 2013) future changes in river flow regimes in response to snow and glacier retreat could have

wide-ranging socio-economic and ecological impacts. Long-term reductions in meltwater inputs will disrupt the supply of water available for irrigation (Baraer et al., 2015; Carey et al., 2014; McDowell and Hess, 2012; Nolin et al., 2010) while increased inter-annual flow variability will threaten infrastructure projects such as hydroelectric power stations (Carvajal et al., 2017; Gaudard et al., 2014; Laghari., 2013). The loss of the runoff-regulating effects of snow and ice could result in more frequent short-term very high flows putting downstream populations and infrastructure at risk (Laghari., 2013; Stoffel et al., 2016). Conversely, the loss of melt during drier months could increase frequency and severity of droughts (Van Tiel et al., 2018). Studies of changes in riparian and river species abundance under flow alteration consistently show negative responses to both increases and decreases in flow magnitude (Poff and Zimmerman, 2010). Increased frequency of flow extremes (floods and droughts) typically result in reduced species richness, while shifts in flow seasonality can disrupt fish spawning cues. Increased flashiness and short-term flow variability typically result in decreased riparian germination survival. Indeed, metrics used to define these characteristics of flow regime now underpin decision-support systems for managing these systems (Beamer et al., 2017; Cartwright et al., 2017; Pool et al., 2017).

Computational GHMs driven by numerical climate model projections allow us to assess how future river flow regimes will change in glaciated river basins. Past studies have focussed on projecting changes in decadal, annual and seasonal variations in runoff magnitude. Decadal changes in runoff are inevitable over the coming century (e.g. Bliss et al., 2014; Lutz et al., 2014; Shea and Immerzeel, 2016) where enhanced melt will result in increased river discharge to a point in time termed 'peak water' after which the continued loss of snow and ice will result in an overall decrease in river flow. It has been shown that many basins, particularly those with small glaciers, have already reached peak water and face a future of dwindling water supply (Huss and Hock, 2018). Seasonal flow magnitudes are also projected to change as melt cycles evolve and watersheds shift from glacial-nival to pluvial runoff regimes (Duethmann et al., 2016; Garee et al., 2017; Kobierska et al., 2013; Ragettli et al., 2016).

Some impact studies show robust changes in the magnitude of the highest and lowest river flows including Wijngaard et al. (2017) who projected an increase in the magnitude of the 10% exceedance flow (Q_{10}) for river basins across the Hindu-Kush-Himalayan region. Other studies for the Rhine (Bosshard et al., 2013), upper Indus (Lutz et al., 2016) and upper Yellow river basin (Vetter et al., 2015) show high flow magnitudes will increase. Stewart et al. (2015) projected a decrease in low flow magnitude (Q_{90}) for the snow-covered Sierra Nevada and Upper Colorado river basins due to shifts in the snow melt season and changes in precipitation type from snow to rain. For the Hindu-Kush, Wijngaard et al. (2017) found the opposite impact with an increase in the magnitude of low flow events. The projected trends in Q_{90} for the upper Yellow river basin by (Vetter et al., 2015) were inconclusive as they showed an even spread of positive and negative trends under the warmest climate scenarios.

Of course, one could go beyond projecting changes in seasonal to decadal mean flow magnitudes and quantiles of the FDC. A branch of streamflow analysis that has been widely adopted in hydrology is the calculation of river flow signatures to represent different characteristics of river flow over specific timescales. Signatures have been used in the past to analyse catchment runoff behaviour and similarity (Ali et al., 2012; Yadav et al., 2007). They also offer an opportunity to evaluate past (Sawicz et al., 2014) and future (Casper et al., 2012) river flow regime change. For example, Teutschbein et al. (2015) projected changes in 14 different river flow signatures for 14 snow-covered catchments in Sweden and showed daily to annual river flow magnitude, timing and variability were all sensitive to climate change. An analysis like this is yet to be undertaken for any glaciated river basins.

Projections of river flow regime are inherently uncertain due to assumptions made about the formulation, parameterisation and boundary conditions of the underlying GHM (Huss et al., 2014; Jobst et al., 2018; Ragettli et al., 2013) and climate model, be that a GCM, RCM, or the combination of the two (GCM-RCM) (Giorgi et al., 2009). Uncertainties may also be introduced by intermediary steps employed to link the two sets of models together such as downscaling (DS). Quantifying the propagation of uncertainties from all

sources in the model chain provides a basis for assigning more robust levels of confidence to river flow projections. Additionally, one can assess the relative contributions of model chain components to the total projection uncertainty, providing empirical evidence for future research needs (e.g. Meresa and Romanowicz, 2017). Ensemble-based experiments have been used in the past to provide this understanding. Here, different components of the model chain are perturbed, typically using a one at a time (OAT) approach where the spread in projections for each perturbed component is evaluated. Ragettli et al. (2013) perturbed three components of a model chain applied to the Hunza River Basin, northern Pakistan including the GCM, statistical DS model and parameterisation of the GHM. They showed that all three sources contributed to annual runoff projection uncertainty, but for the heavily glaciated subregions of the catchment, the GHM parameter uncertainty exceeded the effect of other sources. Huss et al. (2014) investigated uncertainty in seasonal river flow projections over the twenty-first century for the Findelengletscher catchment, Switzerland by modifying the GCM-RCM, GHM melt model structure and initial ice volume boundary condition. Of these, they found that the GCM-RCM and initial ice volume were most important while the melt model structure was of secondary importance. Jobst et al. (2018) investigated uncertainties in twenty-first century river flow projections for the Clutha river basin, New Zealand. They evaluated contributions from emission scenario, GCM-RCM, statistical DS approach and melt model structure. Similarly to Huss et al. (2014), they found that uncertainty in the choice of GCM-RCM dominated total projection uncertainty.

The OAT method provides a first-order approximation of the relative contribution of each component to the total projection uncertainty. However, findings are dependent on how the non-perturbed model components are fixed. Furthermore, this approach cannot resolve interactions between model components which may also contribute to projection uncertainty (Pianosi et al., 2016). The analysis of variance (ANOVA) statistical method (Storch and Zwiers, 1999; Tabachnick and Fidell, 2014) addresses these shortcomings and has been adopted in a number of recent regional and global scale hydrological modelling studies (Addor et al., 2014; Bosshard et al., 2013; Giuntoli et al., 2015; Samaniego et al.,

2017; Vetter et al., 2015; Vetter et al., 2017; Yuan et al., 2017) to compare uncertainties stemming from emission scenario/concentration pathway, climate model, hydrological model structure and DS approach. While uncertainties associated with future climate tend to dominate projections of river flow, glacier-fed river flow projections have shown to be highly sensitive to hydrological model structure (Addor et al., 2014; Giuntoli et al., 2015), particularly in relation to high flows (Vetter et al., 2017). Furthermore, the contribution of projection uncertainty from interactions between model chain components can exceed individual components (Addor et al., 2014; Bosshard et al., 2013; Vetter et al., 2015). Several issues not considered in these studies, however, are yet to be addressed. Firstly, none have investigated a full range of characteristic changes in river flow regime covering decadal to sub-daily timescales. Second, all have incorporated hydrological model uncertainty using multiple model codes, each with their own unique set of process representations, resolution, time step and climate interpolation strategies making it difficult to determine which model components contribute most to projection uncertainty. Finally, none included a fully integrated mass-conserving, dynamic glacier evolution model component and therefore could not fully account for atmosphere-cryosphere-hydrosphere feedbacks.

5.3 Aims

This study uses a climate-GHM model chain to simulate the impact of twenty-first century climate change on downstream river flow regime in the Virkisá River basin. Five components of the model chain are perturbed to represent uncertainty of RCP, GCM-RCM, statistical DS parameterisation and structure-parameterisation of two primary controls on river flow regime in the GHM: melt and runoff-routing processes. The study has two principal aims which will address objectives 2 and 3 outlined in chapter 1:

1. Use a signature-based analysis to determine how climate change and consequent cryospheric change will impact different characteristics of downstream river flow regime over the twenty-first century.

2. Use ANOVA to quantify the relative influence of the five model chain components to projection uncertainty across the different characteristics of river flow regime.

This study addresses each of the aforementioned gaps in previous work. Firstly, changes in river flow regime are assessed quantitatively using 25 river discharge signatures which define different characteristics of river flow regime over a range of timescales. Second, a single, consistent, GHM code is used that can incorporate different model structures and parameterisations of melt and runoff-routing processes allowing for uncertainty stemming from these to be localised using ANOVA. Finally, a fully integrated mass-conserving, dynamic glacier evolution routine is included in the GHM code.

5.4 Methodology

5.4.1 Driving climate data

The same historical precipitation, temperature and incident solar radiation data used in Chapter 4 were used in this study, except that the complete time series from 1981 to 2016 inclusive were utilised.

Future climate time series until 2100 were constructed using the CORDEX regional climate projections. It was decided to use the RCP4.5 and RCP8.5 pathways only and omit RCP2.6 from the analysis. This was necessary given that only 8 of 15 GCM-RCMs within the CORDEX archive used this RCP. Furthermore, the probability of achieving the RCP2.6 targets is increasingly unlikely (Fyke and Matthews, 2015; Sanford et al., 2014) and arguably completely infeasible (Mora et al., 2013) given the current global emission trajectory. Of the 15 available GCM-RCMs, one (GCM:CNRM-CM5, RCM:CNRM-ALADIN53) was removed from the ensemble given that it showed an extreme negative winter temperature bias and a consistently low skill when compared to daily observed climate data (see Appendix F). Figure 5.1 shows the seasonal bias of each of the 14 remaining GCM-RCMs when compared to observations between 1981 and 2005. For temperature, the coldest 1% of days (T_1) typically show a negative bias, particularly in winter, spring

and autumn. Biases for warmest 99% of days (T_{99}) are generally positive, but smaller in magnitude. The average absolute bias in mean seasonal temperature (T_{mean}) is 1.4 °C, but the majority of GCM-RCMs show absolute biases <1.2 °C. Biases in seasonal incident solar radiation projections are almost exclusively positive with the largest biases associated with SW_{mean} and SW_{99}, particularly in spring and summer where they can exceed 100 W m^{-2}. Total precipitation biases are typically largest in winter and autumn where proportionally, biases can exceed the magnitude of the observations (see SON for [EC-EARTH]-[HIRHAM5]). The largest biases however are seen in the highest 1% of precipitation values (P_{99}) which range from -86.9 to 77.5 mm d^{-1}. While positive and negative precipitation biases are present throughout the ensemble, the sensitivity of precipitation simulations to the RCM is clear. For example, the CCLM4-8-17 RCM has a systematic negative bias and the HIRHAM5 RCM has a systematic positive bias.

5.4.2 Downscaling regional climate projections

For this study, the statistical delta-change downscaling approach was employed which has been widely applied in hydrological impact studies (Farinotti et al., 2012; Huss et al., 2014; Immerzeel et al., 2013; Kobierska et al., 2013; Lutz et al., 2016). Here, changes (deltas) in the statistical properties of climate variables are calculated from coarse GCM-RCM projection time series and applied to higher-resolution historical observed data. This approach produces future climate time series with finer-scale properties (e.g. sub-daily variability) that are consistent with historical observations, but with broader changes in climate that are consistent with the deltas. While most studies have used monthly mean delta-change values to capture seasonal shifts in climate, several recent investigations have used advanced quantile-based approaches which account for changes in higher-order statistical properties of future climate by evaluating shifts in the empirical cumulative distribution functions (ECDFs) of climate variables. Including these higher-order changes has shown to be important for evaluating shifts in extreme high flows and sub-seasonal metrics of river flow projections (Immerzeel et al., 2013; Jakob Themeßl et al., 2011; Lutz et al., 2016). In addition, shifts in the day-to-day variability of temperature impact

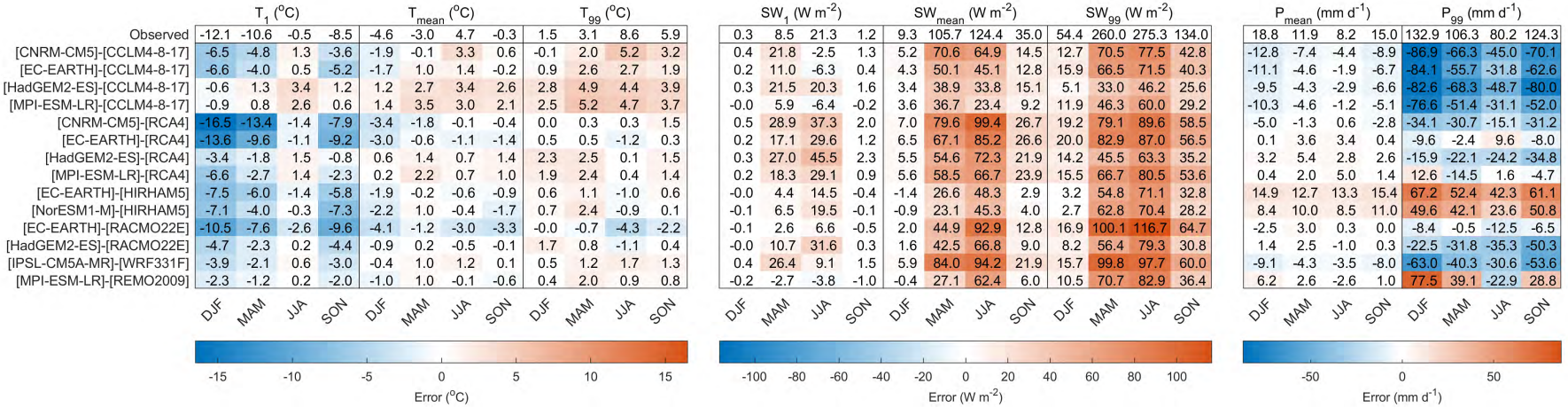

Figure 5.1: Comparison of seasonal catchment-average observed and simulated near-surface air temperature (T), incident solar radiation (SW) and total precipitation (P) between 1981 and 2005 for the 14 GCM-RCM used in this study. The top row shows the observed value and all subsequent rows indicate the GCM-RCM biases. The 1st percentile, mean and 99th are denoted by the subscripts 1, mean and 99 respectively. All statistics are calculated for the recent past (1981-2005) for winter (DJF), spring (MAM), summer (JJA) and autumn (SON).

projections of glacier retreat as these variations control the periodic rising of temperature above the melting point (Beer et al., 2018). Accordingly, the advanced delta-change approach was adopted in this study. The approach is summarised in five steps which were applied to each combination of GCM-RCM, climate variable and RCP separately:

1. The climate variable time series was divided into four 25-year-long periods including the recent past (1981 - 2005) and early (2006 - 2030), mid- (2041-2065) and late (2076 - 2100) twenty-first century.

2. For each of the four periods, all daily data points were further divided into 12 subsamples representing each month of the year. An ECDF was constructed for each month of each period.

3. For each month of each future period, 10 deltas were calculated by taking the mean difference between the recent past and future ECDF for each 10% section (see grey bars in Fig. 5.2a for example).

4. Given the need for transient climate time series to simulate glacier evolution over the twenty-first century, a daily delta time series from 2006 to 2100 was constructed for each ECDF section of each month by linearly interpolating between the calculated deltas of each future period (e.g. as implemented by Farinotti et al., 2012), using the midpoints of the future periods as interpolation points (Fig. 5.2b).

5. The hourly historic observation data for the recent past were randomly sampled (with replacement) on a year-by-year basis to generate an initial unperturbed future climate variable time series (blue dash, Fig. 5.2c). The daily deltas were applied to this time series for each month and ECDF section separately to generate a future perturbed climate time series at an hourly resolution (orange dash Fig. 5.2c). It was noted upon visual inspection, that the inter-annual variability of the future climate time series was very sensitive to the random sampling of the historic climate data. Accordingly, uncertainty associated with this aspect of the DS parameterisation was considered by using 10 different random historic climate samples.

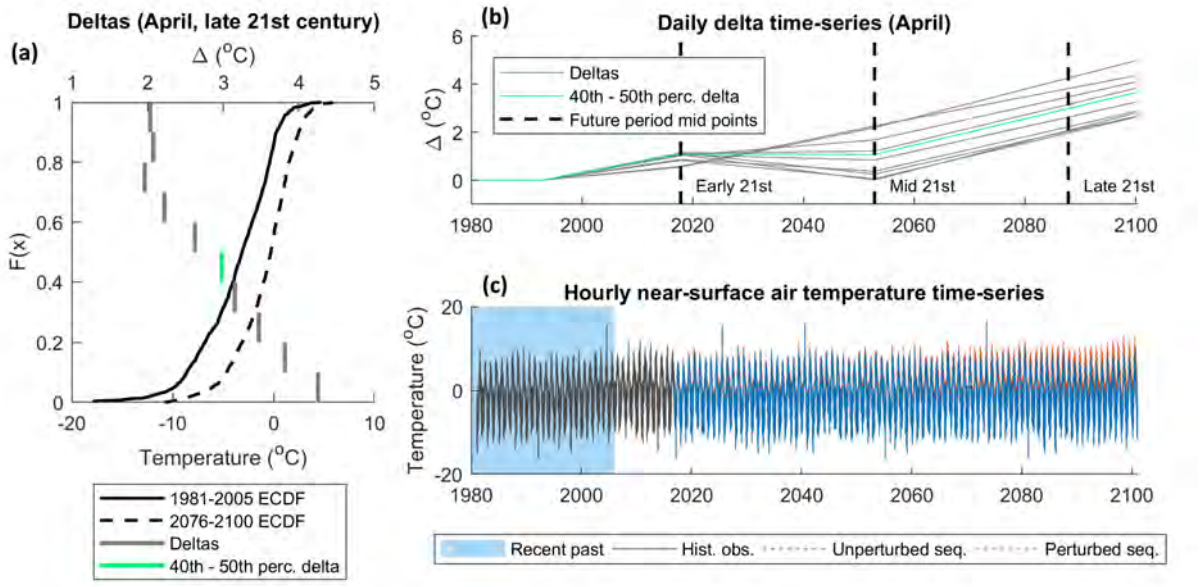

Figure 5.2: Example of advanced delta-change approach when applied to near-surface air temperature data based on the RCP8.5 projections using the CNRM-CM5 GCM and CCLM4-8-17 RCM. Deltas (grey bars) derived from ECDFs (black curves) for April in late twenty-first century (a); Daily delta time series for each section of the April ECDFs (green line represents 40th - 50th percentile section) (b); Initial and perturbed future temperature time series when deltas for all months and ECDF sections are applied (c).

For temperature, catchment-average daily deltas were applied evenly across the catchment and each daily period of the unperturbed time series. Accordingly, diurnal temperature variability and lapse rates were assumed not to change in the future. For incident solar radiation and precipitation, proportional deltas were used to prevent negative values and preserve the sub-daily proportional distribution of these variables in space and time. A total of 2 RCP $\times$ 14 GCM-RCM $\times$ 10 DS parameterisations = 280 future climate time series were generated for this study.

5.4.3 Glacio-hydrological model

The findings from Chapter 4 showed that the most complex runoff-routing structure was consistently the least efficient when compared to the two simpler alternatives, particularly in relation to capturing signatures representing high-river-flow events. Accordingly, the added complexity of this model structure was not deemed justifiable and so only the remaining six combinations of melt and runoff-routing models structures were used in this study made up of every combination of the TIM_1, TIM_2 and TIM_3 melt model structures with the ROR_1 and ROR_2 runoff-routing model structures.

Modification to Δh parameterisation of glacier retreat

The Δh glacier evolution parametrisation is not designed to simulate glacier advance. Under periods of sustained positive mass balance, simulations from the Δh glacier evolution model may result in an unrealistic build up of ice at the glacier tongue without any simulated areal advance. Given the potential for periods of glacier advance under a changing climate, such behaviour is likely to result in significant projection biases. Recently, Seibert et al. (2018) presented an implementation of the original Δh parameterisation that provides more realistic simulations of glacier advance. They propose running the Δh parametrisation a priori outside of the GHM. A small negative mass balance is used to force the Δh model from an initial glacier profile (ideally its maximum observed extent) until the glacier has disappeared completely. At each step, the glacier mass and geometry are stored in the form of a lookup table. On running the GHM, the retreat/advance

of the glacier is derived from the lookup table as a function of the simulated glacier mass. One important drawback of using this static lookup table is that it modifies the behaviour of the Δh formulation during periods of retreat. More specifically, this approach neglects the transient annual sequencing of glacier mass balance which influences simulated glacier geometry due to the non-linear structure of the Δh polynomial that defines the relationship between mass balance and glacier geometry. Accordingly, a modified implementation of the Seibert et al. (2018) approach was used in this study which behaves like the original Δh formulation during periods of glacier retreat and allows for the simulation of glacier advance while accounting for mass balance sequencing effects on the model behaviour. For periods of negative glacier mass balance the original Δh formulation was used. The GHM++ code was then modified so that for each simulation year, the simulated glacier mass and geometry were stored in memory. If a positive glacier mass balance (ΔM) was simulated, GHM++ would log the current glacier mass ($M_{current}$) and then look for the most recent historical simulated glacier mass (M_{hist}) that exceeded $M_{current} + \Delta M$. The Δh model was then run with a negative mass balance (ΔM^*) so that $M_{hist} + \Delta M^* = M_{current} + \Delta M$.

5.4.4 Signatures of river flow regime

Table 5.1 lists the 25 signatures of river discharge used to evaluate future changes in river flow regime. The signatures are grouped into seven different attributes and further categorised by the characteristic(s) of flow regime that they evaluate and their temporal scale. They broadly follow those used in Chapter 4, but have been modified in some cases to provide more relevant indicators of hydrological impacts rather than model behaviour. At the decadal timescale, two signatures were selected. These include the 'peak water', which defines the timing (by year) of maximum flow, as well as the inter-annual flow range which characterises long-term flow variability. Changes in mean annual river flow were also evaluated, while mean monthly flows were used to evaluate changes to the seasonal timing and magnitude of river flow. The range in mean monthly flows was also chosen to evaluate intra-annual flow variability. In addition, eight signatures were selected

Table 5.1: Summary of 25 river discharge signatures used to evaluate future changes in river flow regime. Those with available limits of acceptability were also used as part of the GHM calibration and evaluation procedure.

Attribute	Signature	Limits of acceptability		Regime characteristic	Temporal scale
		Calibration (2013-2014)	Evaluation (2015-2016)		
Inter-annual flow	Peak water (PW)	-	-	Timing and magnitude	Decadal
	Inter-annual flow range (R_{ANN})	-	-	Variability	Decadal
Annual river flows	Mean annual river flow ($\bar{Q}$)	-	-	Magnitude	Annual
Monthly river flows	Mean January river flow ($\bar{Q}_{JAN}$)	1.16 – 1.86 $m^3 s^{-1}$	-	Timing and magnitude	Monthly
	Mean February river flow ($\bar{Q}_{FEB}$)	1.69 – 2.92 $m^3 s^{-1}$	-	Timing and magnitude	Monthly
	Mean March river flow ($\bar{Q}_{MAR}$)	0.85 – 1.58 $m^3 s^{-1}$	1.22 - 2.34 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean April river flow ($\bar{Q}_{APR}$)	0.73 – 1.48 $m^3 s^{-1}$	1.03 - 2.10 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean May river flow ($\bar{Q}_{MAY}$)	1.50 – 2.16 $m^3 s^{-1}$	1.64 - 3.00 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean June river flow ($\bar{Q}_{JUN}$)	4.12 – 6.23 $m^3 s^{-1}$	4.88 - 9.39 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean July river flow ($\bar{Q}_{JUL}$)	6.33 – 10.3 $m^3 s^{-1}$	4.96 - 9.38 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean August river flow ($\bar{Q}_{AUG}$)	5.72 – 9.15 $m^3 s^{-1}$	6.80 - 14.39 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean September river flow ($\bar{Q}_{SEP}$)	4.55 – 7.38 $m^3 s^{-1}$	6.61 - 14.21 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean October river flow ($\bar{Q}_{OCT}$)	3.88 – 7.02 $m^3 s^{-1}$	6.94 - 16.33 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean November river flow ($\bar{Q}_{NOV}$)	3.90 – 7.40 $m^3 s^{-1}$	3.17 - 5.76 $m^3 s^{-1}$	Timing and magnitude	Monthly
	Mean December river flow ($\bar{Q}_{DEC}$)	-	-	Timing and magnitude	Monthly
	Mean monthly flow range (R_{mnth})	-	-	Variability	Seasonal
Slow-release low flows	95% exceedance flow (Q_{95})	0.27 - 1.10 $m^3 s^{-1}$	0.66 - 1.75 $m^3 s^{-1}$	Magnitude	Monthly to seasonal
	99% exceedance flow (Q_{99})	0.12 - 0.88 $m^3 s^{-1}$	0.46 - 1.56 $m^3 s^{-1}$	Magnitude	Monthly to seasonal
	Low flow standard deviation ($\sigma_{99\text{-}95}$)	0.03 - 0.10 $m^3 s^{-1}$	0.02 - 0.09 $m^3 s^{-1}$	Variability	Monthly to seasonal
Moderate flows	50% exceedance flow (Q_{50})	2.38 - 3.70 $m^3 s^{-1}$	3.10 - 5.79 $m^3 s^{-1}$	Magnitude	Daily to monthly
	Moderate flow standard deviation ($\sigma_{52\text{-}48}$)	0.07 - 0.15 $m^3 s^{-1}$	0.08 - 0.18 $m^3 s^{-1}$	Variability	Daily to monthly
Quick-release high flows	1% exceedance flow (Q_{01})	17.71 - 40.31 $m^3 s^{-1}$	21.90 - 61.57 $m^3 s^{-1}$	Magnitude	Hourly to daily
	5% exceedance flow (Q_{05})	9.43 - 15.76 $m^3 s^{-1}$	11.71 - 27.37 $m^3 s^{-1}$	Magnitude	Hourly to daily
	High flow standard deviation ($\sigma_{05\text{-}01}$)	2.08 - 5.68 $m^3 s^{-1}$	2.60 - 8.10 $m^3 s^{-1}$	Variability	Hourly to daily
Flashiness	Integral scale (τ)	25 – 44 h	0 - 54 h	Timing	Hourly to daily

which broadly describe the magnitude and variability of slow-release low flows (99-95% exceedance flows), moderate flows (52-48% exceedance) and quick-release high flows (5-1% exceedance). For these, the quantiles of the FDC were used to assess changes in the magnitude of these flow types. Quantiles were deemed more relevant to hydrological impacts than the volume under the FDC. Additionally, the standard deviation has been used as an indicator of flow variability rather than using the coefficient of variation or slope of the FDC as this provides consistent units with the other flow signatures. Finally, the integral scale, was utilised as an indicator of the response time of the catchment to runoff events (flashiness).

5.4.5 GHM++ calibration

Given the focus on projecting changes in river discharge signatures, these were explicitly included in the GHM calibration procedure as this gives better signature simulations than using traditional global objective functions (Kiesel et al., 2017; Pool et al., 2017). However, calibrating against river flow data alone can lead to unrealistic snow and glacier melt rates, inhibiting model consistency and increasing projection uncertainties (Finger et al., 2011; Hanzer et al., 2016; Konz and Seibert, 2010; Schaefli and Huss, 2011). Accordingly, a novel signature-based calibration of GHM++ was undertaken by evaluating the model against 20 of the river discharge signatures in Table 5.1 for which observation data exists in combination with the 12 signatures of ice melt and snow coverage in Table 4.2.

For each signature, model simulations were compared to observations using the continuous acceptability score outlined in Eq. 4.19 using the LOA in Table 5.1. The aim of the calibration was to extract an ensemble of GHM++ compositions (TIM and ROR structure-parameter combinations) that were most acceptable across the river discharge signatures whilst broadly reproducing the snow coverage and ice melt signatures. This was achieved using a two-stage Monte Carlo procedure which was devised so that the resultant model ensemble reflected the uncertainty in model selection given the known model inconsistencies across the signatures shown in Chapter 4.

Stage 1: TIM calibration

The first stage aimed to extract the optimal TIM compositions (structure-parameter combinations) by calibrating them against the 12 snow coverage and ice melt signatures. Here, for each of the three TIM structures, 5000 TIM parameter sets were drawn from predefined uniform distributions (Table C.1) using the quasi-random Sobol sampling strategy (Brately and Fox, 1988) to sample the parameter space as efficiently as possible. For each parameter set, GHM++ was spun-up for three years from 1985 to 1988 with a static ice geometry fixed to a 1988 ice DEM (Magnússon et al., 2016). The model was then run

from 1988 to the end of 2016 with a freely evolving glacier geometry.

Given the high degree of glaciation in the study catchment, and its recent rapid retreat, an initial emphasis of the calibration was put on the model's ability to capture the long-term glacier volume change signature. Accordingly, only those TIM compositions that captured this signature within the LOA were considered and the rest were discarded. These remaining compositions were then further refined by evaluating them against the remaining 11 snow and ice signatures. First, the TIM compositions were sorted by structure (TIM_1,TIM_2,TIM_3). Then, for a given TIM structure, the following steps were applied:

1. Find the TIM parameter set(s) that capture the signature within the LOA and discard the rest. If more than one parameter set captures the current signature, go to step 2. If none capture the current signature, discard none and go to step 2.

2. Of the remaining models, find that which best captures the 10 remaining snow and ice signatures overall according to the weighted mean scores obtained in Eq. 4.19. The weights were applied to ensure that equal preference was given to ice melt and snow coverage signatures.

24 unique TIM compositions were obtained from this calibration stage made up of eight unique parameterisations of each of the three TIM structures. In some cases the same composition was selected more than once which was accounted for by weighting the simulations in the results presented throughout this study.

Stage 2: ROR calibration

The second calibration stage aimed to extract the optimal ROR compositions when used in combination with the 24 preselected TIM compositions by calibrating them against 20 of the river discharge signatures obtained from observations of river discharge for the years 2013 and 2014 (see signatures with calibration LOA in Table 5.1). Note, the inter-annual flow signatures and the mean December river flow signatures could not be calculated as there were insufficient observation data. Furthermore, the mean annual river flow and mean monthly flow range were not included as this information was already accounted

for in the mean monthly flow signatures. Here, 5000 random ROR parameter sets were drawn for each ROR structure. Each was used in combination with the preselected TIM compositions in GHM++. Then, the two steps outlined in calibration stage 1 were applied using the 20 calibration river discharge signatures with two notable differences. Firstly, for each ROR structure and each river discharge signature, rather than selecting a unique ROR parameter set for each of the 24 TIM compositions, a single parameter set was selected based on its mean performance across the 24 TIM compositions. This was done to satisfy the ANOVA requirements so that the TIM and ROR composition uncertainty could be analysed separately. Furthermore, for step 2, the signatures were weighted so that each of the attributes in Table 5.1 were weighted equally. In total, 14 unique ROR compositions were selected made up of seven unique parameterisations of the ROR_1 and ROR_2 structures, giving a total of $24 \times 14 = 336$ unique GHM++ compositions.

5.4.6 ANOVA uncertainty analysis

For the twenty-first century runs, all 336 GHM++ compositions were run to the end of 2016 using the historic observed climate to capture the evolving ice geometry as accurately as possible. From 2017 to 2100, the 280 downscaled future climate time series were used to drive the models resulting in 94080 individual model runs. For each model run, projections of watershed snow and ice coverage and the 25 river discharge signatures were extracted for six twenty-first century 25-year time slices centred on the 2030s (2023-2047), 2040s (2033-2057), 2050s (2043-2067), 2060s (2053-2077), 2070s (2063-2087) and 2080s (2073-2097). Future changes in these were then calculated relative to a reference 25-year period (1991-2015). This reference period was chosen because ice coverage data used to initialise the GHM is only available from 1988 and historic climate data are available up to the end of 2016.

The ANOVA statistical framework (Storch and Zwiers, 1999; Tabachnick and Fidell, 2014) was used to quantify the effect size of the five components of the model chain, hereafter termed *factors*, on each signature for each twenty-first century time slice. Note, the peak water (PW) signature can only be calculated taking into account the full projection

time series and, as such, it was not possible to apply ANOVA to each time slice for this signature. The five factors include the 2×RCP, 14×GCM-RCM combinations, 10×DS parameterisations, 24×TIM compositions and 14×ROR compositions. ANOVA offers an intuitive approach to estimate the effect size of each factor on each signature by partitioning the total sum of squares (SS_{tot}) in the response variable over all combinations of factor levels:

$$SS_{tot} = SS_a + SS_b + SS_c + SS_d + SS_e + SS_I + SS_\varepsilon \tag{5.1}$$

where:

$$SS_{tot} = \sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\sum_{k=1}^{n_c}\sum_{l=1}^{n_d}\sum_{m=1}^{n_e}(y_{i,j,k,l,m} - \bar{Y})^2 \tag{5.2}$$

where n_a, n_b, n_c, n_d and n_e are the number of levels for each factor, y is the response for a given treatment (i.e. combination of factor levels) and $\bar{Y}$ is the grand mean of the response variable over all treatments. SS_a, SS_b, SS_c, SS_d and SS_e in Eq. 5.1 are the sum of squares due to the main effects, i.e. the variability in the response variable due to varying a given factor on its own. For example:

$$SS_a = n_b n_c n_d n_e \sum_{i=1}^{n_a}(y_{i,\circ,\circ,\circ,\circ} - \bar{Y})^2 \tag{5.3}$$

where $\circ$ indicates averaging over an index. SS_I includes all nonadditive interaction terms where the combined effect of two or more factors is not the sum of their main effects. For a 5-factor ANOVA, one could include all unique n-tuple combinations of factors where $n = (2, 3, 4, 5)$. Given the difficulty in interpreting these higher-order interactions, and computational requirements, it was decided to investigate the nine first-order interactions only, so that:

$$SS_I = SS_{ab} + SS_{ac} + SS_{ad} + SS_{ae} + SS_{bc} + SS_{bd} + SS_{be} + SS_{cd} + SS_{ce} + SS_{de} \tag{5.4}$$

The sum of squares for a first-order interaction are calculated as follows using factors a and b as an example:

$$SS_{ab} = n_c n_d n_e \sum_{i=1}^{n_a}\sum_{j=1}^{n_b}(y_{i,j,\circ,\circ,\circ} - y_{i,\circ,\circ,\circ,\circ} - y_{\circ,j,\circ,\circ,\circ} + \bar{Y})^2 \tag{5.5}$$

Finally, the SS_ε term includes all unexplained variance i.e. error in the ANOVA model.

Having partitioned the sum of squares, the effect size, η^2 for any term in Eq. 5.2 can be taken as the proportion of the total sum of squares:

$$\eta_*^2 = SS_*/SS_{tot} \tag{5.6}$$

where $*$ can be any of the main effects, interactions or error term.

Bosshard et al. (2013) showed that because ANOVA is based on a biased variance estimator that underestimates the variance in small sample sizes, the calculated effect sizes are biased if a different number of levels are used for each factor. Given that the number of factor levels range from 2 to 24, a pure application of ANOVA using all possible treatments would lead to biased results. Bosshard et al. (2013) outlined a method to correct for this which involves subsampling the factor levels down to the smallest number levels across all factors. The procedure is repeated using every possible combination of factor levels with unbiased effect size taken as the mean across all subsamples. However, given that there are $> 10^8$ unique combinations of factor levels when subsampled down to two (and discarding factor level repetitions), it would have been infeasible to account for every possible combination. Instead, it was decided to calculate the effect sizes in this manner using five different subsample sizes ($10^1, 10^2 ... 10^5$). The results were then analysed to see if the effect sizes converged. It was found that 10^3 subsamples were sufficient to converge the effect sizes for all river discharge signatures and projections of snow and ice coverage. Accordingly, this subsampling strategy was adopted in this study.

5.5 Results

5.5.1 Evaluation of calibrated GHM++ compositions

The simulated river discharge time series and signatures using the calibrated GHM++ compositions were evaluated against river discharge observations covering the years 2015 and 2016. Note, no data for mean January and February flows were available for these years. Figures 5.3a and b show the simulated 'capture ratio' (the ratio of the 336 model compositions that capture the observation data within their 95% uncertainty bounds)

time series projected onto the mean observed river discharge for the years 2015 and 2016 respectively. Also shown is the ensemble mean simulated river discharge (black dash) which while not indicative of a single model simulation, does provide an indication of overall projection bias.

56% of the observation time series were captured by at least half of the model compositions, while 41% and 28% of the observations were captured by at least 75% and 90% of the GHM compositions. 12% of the observations were not captured by any of the model compositions. These included some of the low flows observed at the beginning of the year outside of the melt season, particularly in 2015, where the simulations showed consistent negative biases. Some rainfall-induced summer peak flows were also not captured, particularly during the late-summer months of August and September. Furthermore, the sustained summer melt runoff discharge in between rainfall-induced peak flows tended to be overestimated (for example during July and August 2016). Even so, the FDC in Fig. 5.3c shows that almost the entire FDC was captured by all of the GHM++ simulations except for some of the lowest flows on record. Indeed, Fig. 5.3d reveals that models were least efficient at capturing the low flow signatures, particularly the variability signature ($\sigma_{99\text{-}95}$), where simulations were positively biased by almost four times the observation uncertainty. For the remaining signatures though, the ensemble of models were remarkably efficient, with the majority of simulations (and in most cases all of them) capturing these signatures within their LOA.

5.5.2 Future climate projections

Projections of temperature for the late twenty-first century (2076-2100) consistently show an increase relative to the recent past (1981-2005). The largest increases are projected for the coldest days of the year during the winter (Fig. 5.4a), spring (Fig. 5.4d) and autumn (Fig. 5.4j) months as shown by the positive skew in the lower sections of the ECDFs. However, these changes are also typically associated with the greatest projection uncertainty. RCP4.5 projects annual mean near-surface air temperature to rise by between 1.1 and 3.6 °C by the late twenty-first century relative to the recent past with an ensemble

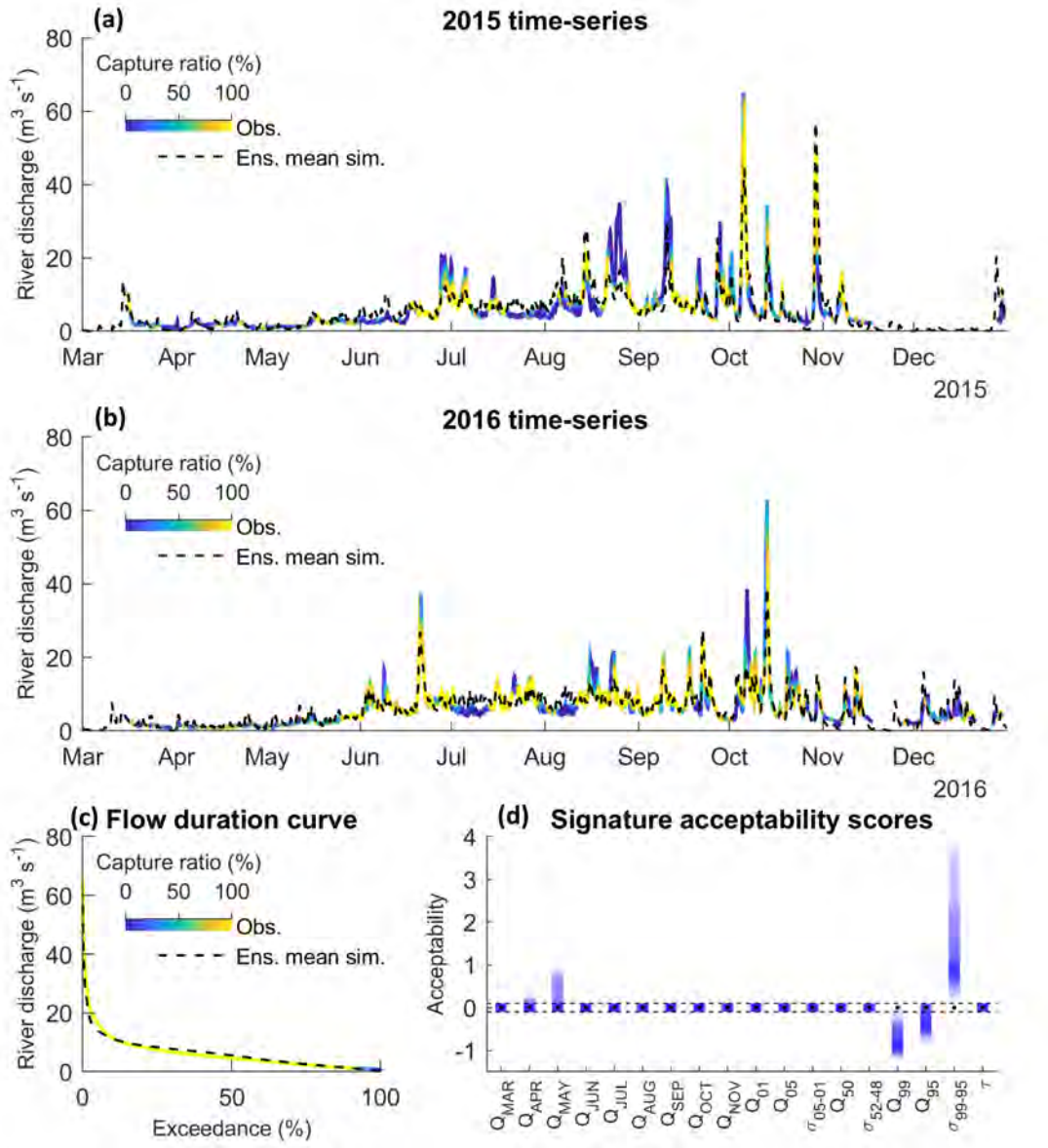

Figure 5.3: Capture ratio projected onto observed river discharge data during evaluation period for 2015 (a); 2016 (b); and over the FDC (c). The weighted ensemble mean simulation is shown as a black dash. Also shown are the range of acceptability scores for each of the available river discharge signatures over the evaluation period (d). Acceptable simulations in (d) are those contained within the black dash lines.

mean projection of +2.0 °C. RCP8.5 projects an equivalent rise of between 2.3 and 4.9 °C with an ensemble mean projection of +3.3 °C.

Projected changes in incident solar radiation span positive and negative values, but the median projections are consistently negative indicating reductions in incident solar radiation are most likely. Uncertainty in the magnitude of change is highest during the spring and summer months (Figs. 5.4e and h) when incident solar radiation peaks. Under RCP4.5 annual mean incident solar radiation is projected to change by between -10.7 to +0.8% by the late twenty-first century with an ensemble mean projection of -4.4%. Under RCP8.5 changes of between -15.3 to 0.4% are projected with an ensemble mean projection of -7.7%.

Projected changes in total precipitation are negligible for the four lowest 10% sections of the precipitation ECDFs, but significant for the two highest sections. In the winter (Fig. 5.4c) and autumn (Fig. 5.4l) months, absolute changes exceed 40 mm d^{-1}. The direction of change is uncertain apart from autumn where median projections are consistently positive for the upper sections of the ECDF. The magnitude of change is also uncertain. RCP4.5 projects annual mean precipitation will change by between -13.5 to +21.6% relative to the recent past by the late twenty-first century with an ensemble mean projection of +1.7%. Under RCP8.5 changes of between -25.7 to 25.1% are projected with an ensemble mean projection of +1.4%.

Figure 5.5 shows the correlation matrix calculated between seasonal average climate variables for the late twenty-first century. For all climate variables, between-season changes (scores within green borders in Fig. 5.5) are positively correlated indicating that an increase in summer temperature typically corresponds with an increase in winter temperature for example. Temperature has the greatest between-season correlation while precipitation is the least well correlated. Within-season, between-variable correlation scores (within purple border in Fig. 5.5) show that precipitation and incident solar radiation are negatively correlated and that the correlation between precipitation and temperature depends on the time of year. For the cooler winter, spring and autumn months, temperature and precipitation are positively correlated, but there is a weak neg-

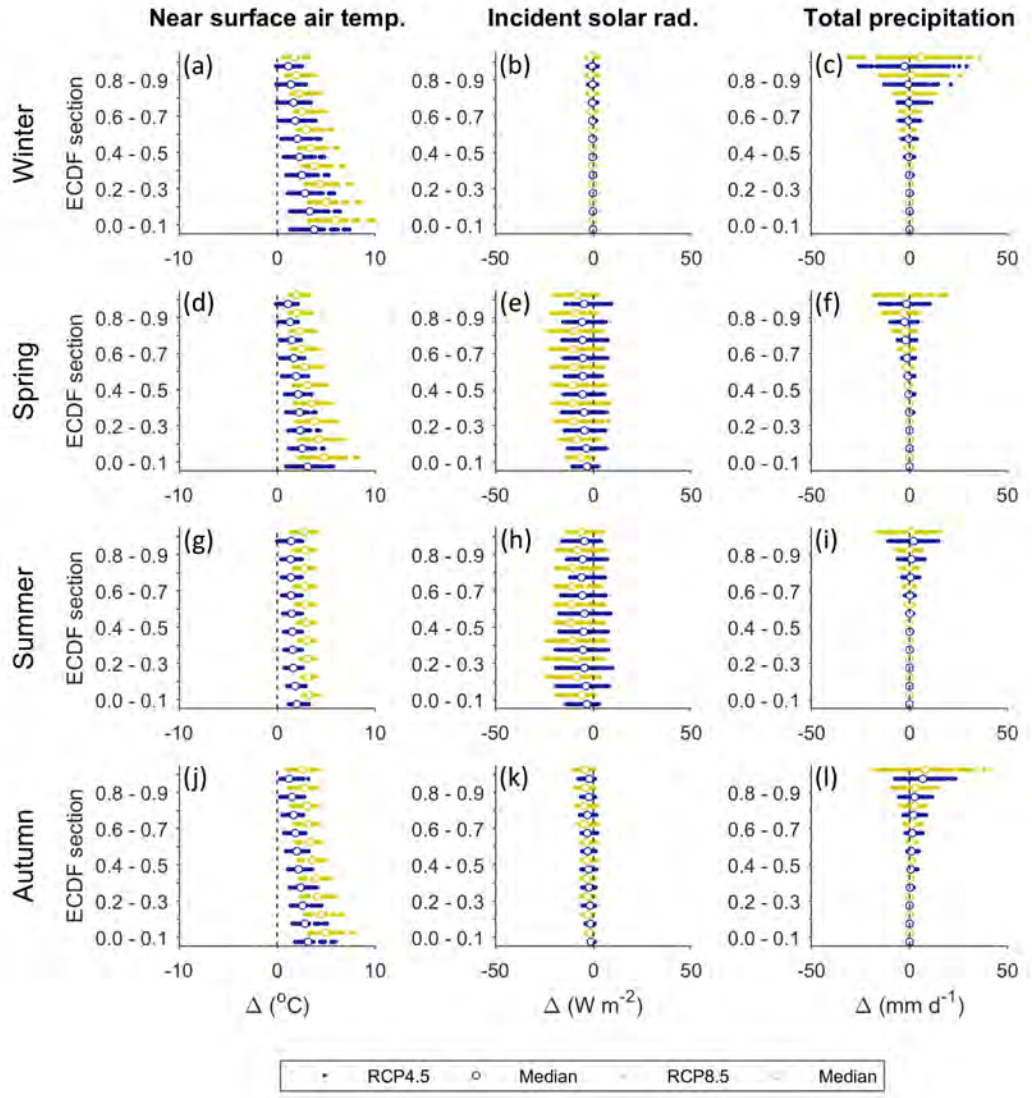

Figure 5.4: Seasonal average projected changes in ECDFs for near-surface air temperature (a,d,g,j), incident solar radiation (b,e,h,k) and total precipitation (c,f,i,l) for the late twenty-first century (2076-2100) relative to the recent past (1981-2005). Changes are plotted for each 10% section of the ECDFs. For each section, blue and yellow dots represent each of the 140 downscaled future climate time series for the RCP4.5 and RCP8.5 pathways respectively (280 in total). Winter: Dec, Jan, Feb; spring: Mar, Apr, May; summer: Jun, Jul, Aug; autumn: Sep, Oct, Nov.

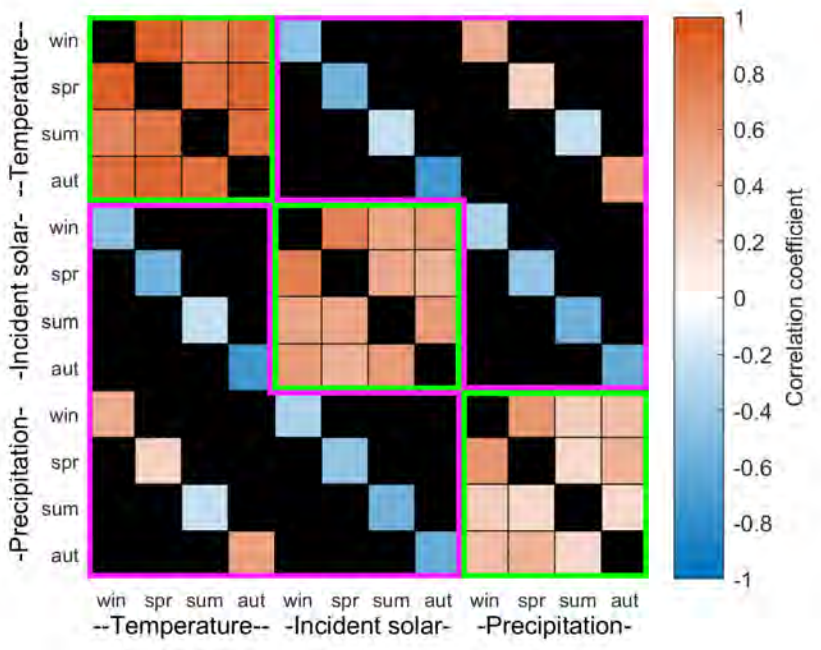

Figure 5.5: Correlation matrix between seasonal average climate variables calculated for late twenty-first century (2076-2100) using the 280 downscaled future climate time series. Within-variable, between-season correlation scores are contained within the green borders and within-season, between-variable correlation scores are contained within the purple borders. Those regions of the correlation matrix that do not cover these two groups are shaded in black.

ative correlation for the summer months. Temperature and incident solar radiation are negatively correlated, most strongly for the cooler winter, spring and autumn months.

5.5.3 Future evolution of snow and ice coverage

The ensemble mean projections of annual mean watershed snow coverage (Fig. 5.6a) show that it will decrease from 12.2 km^2 in 2016 to 9.2 km^2 in 2100 (25% reduction) under RCP4.5 and 6.0 km^2 (51% reduction) under RCP8.5. The 95% projection confidence intervals indicate that by 2100 the watershed could be almost entirely free of snow (2.5 km^2 remaining) under RCP8.5 or could have a coverage exceeding present levels (13.3 km^2) under RCP4.5.

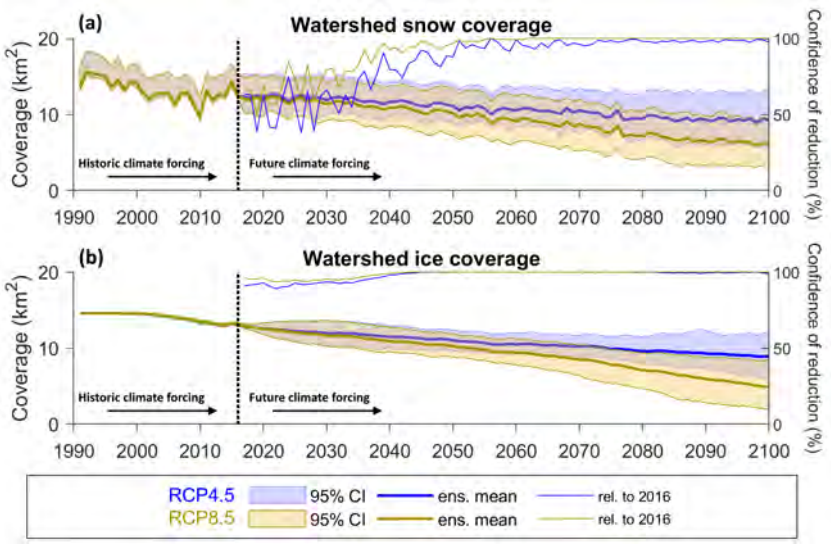

Figure 5.6: Projected annual mean watershed snow coverage (a) and ice coverage (b) including the projection confidence intervals (bands) and ensemble mean projections (thick solid lines) for the RCP4.5 (blue) and RCP8.5 (yellow) projections. Also shown are projection confidence levels for a reduction in coverage relative to 2016 (thin solid lines, right-hand axis).

Beyond 2050, there is high confidence ($\geq$ 95%) that snow coverage will reduce relative to 2016 levels under RCP8.5 (thin yellow line in Fig. 5.6a) and equally high levels of confidence apply to projected reductions in snow coverage beyond 2066 under the cooler RCP4.5 (thin blue line in Fig. 5.6a).

The ensemble mean projection of ice coverage (Fig. 5.6b) projects a 31% reduction relative to 2016 by 2100 under RCP4.5 and a more severe 63% reduction under RCP8.5. There is high confidence ($\geq$95%) that ice coverage will be less than 2016 levels from 2037 onwards under RCP4.5 and from 2030 onwards under RCP8.5 but the magnitude of change is uncertain under both RCPs. By 2100, the 95% confidence intervals for both RCP4.5 and RCP8.5 are 6.5 km^2 wide (more than half the 2016 watershed ice coverage).

The simulation that projected the minimum ice coverage by 2100 under RCP8.5 shows sustained retreat of the glacier between 2000 and 2100 resulting in a watershed that is almost entirely ice free by the end of the century (Fig. 5.7). In contrast, the maximum

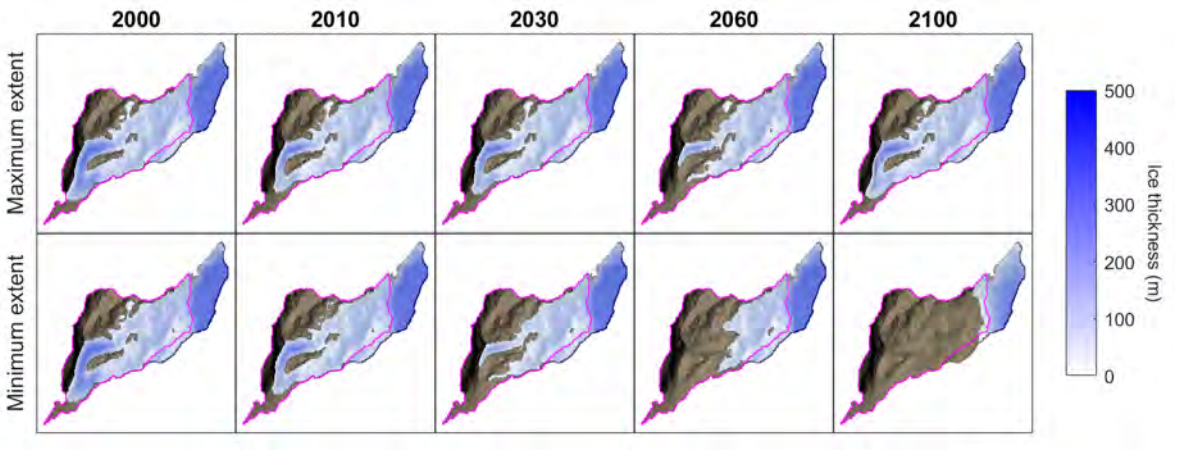

Figure 5.7: Simulated ice thickness between 2000 and 2100 based on simulations that projected the maximum (RCP4.5) and minimum (RCP8.5) ice coverage by 2100. Watershed outline shown in magenta.

ice coverage simulation under RCP4.5 projects two periods of glacier advance between 2010 and 2030 and between 2060 and 2100. By the end of the century, this simulation projects ice coverage will be similar to that in 2000.

Figures 5.8a-c show the climate projection time series that produced the minimum (dotted lines) and maximum (dashed lines) snow (blue lines) and ice (red lines) coverage by 2100. The minimum coverage simulations were forced with some of the highest temperature time series while the maximum coverage simulations were forced with some of the lowest. The maximum coverage simulations show higher-than-average incident solar radiation inputs (Fig. 5.8b) and lower precipitation volumes than the minimum coverage simulations. Indeed, correlation scores calculated between seasonal average climate variables and the simulated snow and ice coverage by 2100 (Fig. 5.8d) show that there is a strong negative correlation between mean temperature and projected snow and ice coverage and a weaker positive correlation between snow and ice coverage and incident solar radiation. An even weaker negative correlation exists between autumn and winter precipitation and snow and ice coverage.

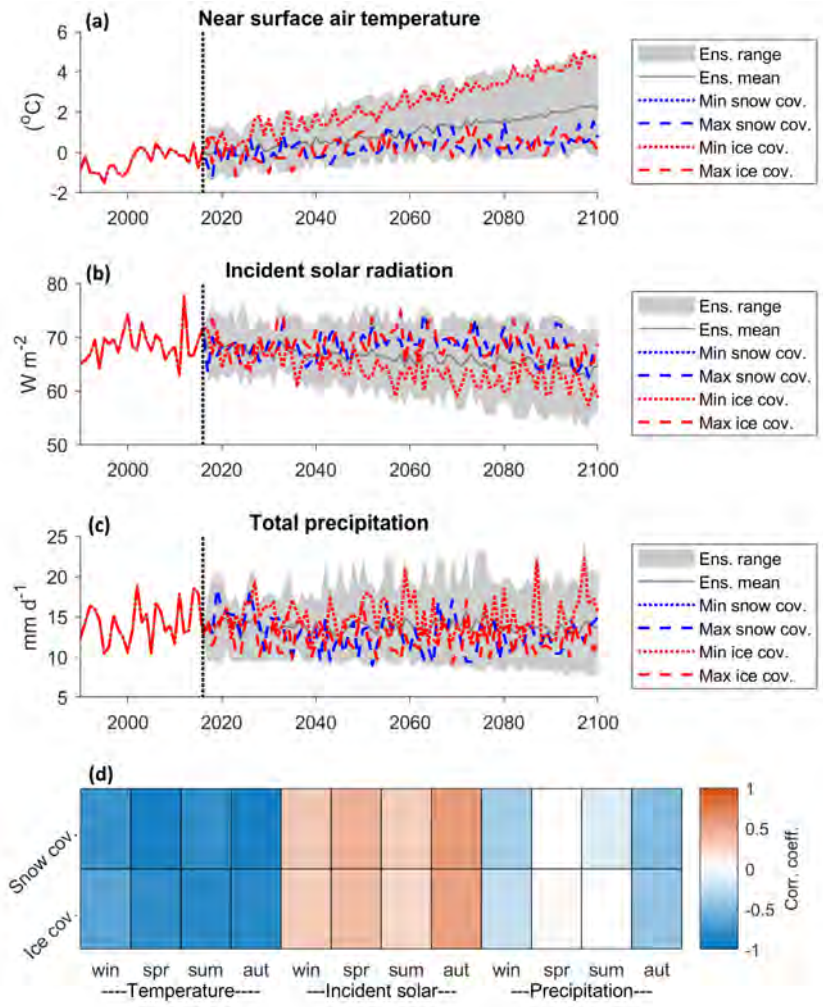

Figure 5.8: Relationship between driving climate data and projected snow and ice coverage including annual mean downscaled climate time series of temperature (a), incident solar radiation (b) and total precipitation (c) with time series that produced the minimum (dotted lines) and maximum (dashed lines) snow and ice coverage by the end of 2100. Also included are correlation scores calculated between seasonal average climate variables over the entire future period (2017-2100) and simulated snow and ice coverage by the end of 2100 (d).

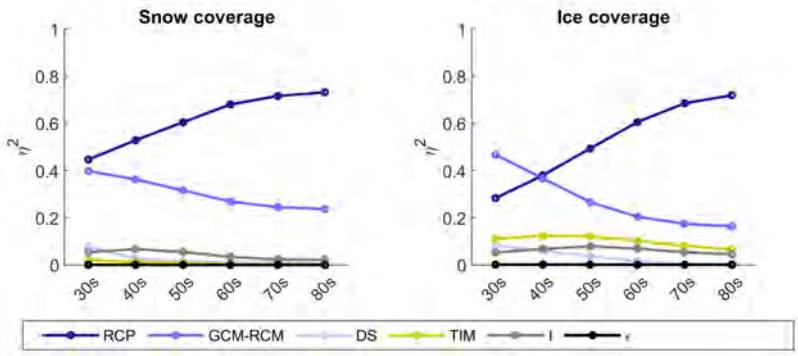

Figure 5.9: Effect size (η^2) of main effects (RCP, GCM-RCM, DS and TIM), interactions (I) and remaining error (ϵ) on projected changes in snow and ice coverage calculated using ANOVA for the six twenty-first century time slices. Note, the ROR main effect is not included here as it does not influence cryospheric processes in GHM++

.

5.5.4 Sources of uncertainty in snow and ice coverage projections

The effect size of the main, interaction and error terms calculated using ANOVA for projected changes in snow and ice coverage is shown in Fig. 5.9. Note, ROR effects are not included here as this model chain component has no influence on cryospheric processes in GHM++. The effect size of each ANOVA term changes through the decades and also varies between snow and ice coverage. Throughout the twenty-first century, TIM uncertainty contributes <3% to the total projection uncertainty of snow coverage. For projections of ice coverage, η^2_{TIM}>0.11 up to and including the 2060s, but then gradually falls to 0.07 by the 2080s. η^2_{DS} and η^2_I never exceed 0.1 for snow and ice coverage and as with η^2_{TIM}, they gradually reduce through the latter half of the twenty-first century. GCM-RCM uncertainty is the largest contributor to ice coverage projection uncertainty in the 2030s with an effect size of 0.47. For snow coverage, RCP and GCM-RCM have similar effect sizes of 0.45 and 0.4 respectively. However, for the mid- and latter half of the twenty-first century RCP uncertainty dominates, contributing 73% and 72% of snow and ice coverage total projection uncertainty by the 2080s respectively.

5.5.5 Future evolution of primary runoff components

As an initial indication of the potential for downstream river flow regime change, Fig. 5.10 shows the twenty-first century evolution of changes in the four primary runoff components relative to the reference period. The ensemble means (solid coloured lines) indicate that by the end of the century rainfall will increase for all months under both RCPs except for August where RCP8.5 shows a small decrease in rainfall on average (Fig. 5.10a). The largest increases are shown during the autumn (SON) and winter (DJF) months under RCP8.5. The confidence in the direction of change by the end of the century is ≥90% for six months under RCP8.5 (as indicated by the coloured bands), but only for two months (March and April) under RCP4.5. However, ≥75% of the projections from both RCPs project an increase in rainfall between October and April (as indicated by the markers in Fig. 5.10a). A comparison of the reference and 2080s monthly ensemble means (inset in Fig. 5.10a) indicates that the peak rainfall month will shift from September to October.

For snow melt, the greatest changes are projected to occur in the summer months of July and August under RCP8.5 where there is ≥90% confidence that melt will reduce relative to the reference period from the 2040s onwards (Fig. 5.10b). RCP4.5 also projects decreases in summer melt, but the magnitude of change is smaller. In the winter months, both RCPs project a small increase in melt on average by the end of the century. The ensemble means project that total summer (JJA) melt will reduce by 19% under RCP4.5 and 37% under RCP8.5 by the 2080s (inset in Fig. 5.10b). Annual melt will decrease by 12% under RCP4.5 and 26% under RCP8.5. A similar pattern of change is projected for ice melt (Fig. 5.10c) where total summer (JJA) melt will reduce by 33% under RCP4.5 and 58% under RCP8.5 by the 2080s. There is high confidence (≥90%) that mean monthly ice melt will reduce for all months except December under RCP8.5. Under RCP4.5 a small increase in winter ice melt is projected for the early and mid twenty-first century, but by the 2080s, winter melt is projected to reduce near to reference levels on average. Under RCP8.5, winter ice melt is projected to reduce relative to reference levels for the latter half of the twenty-first century.

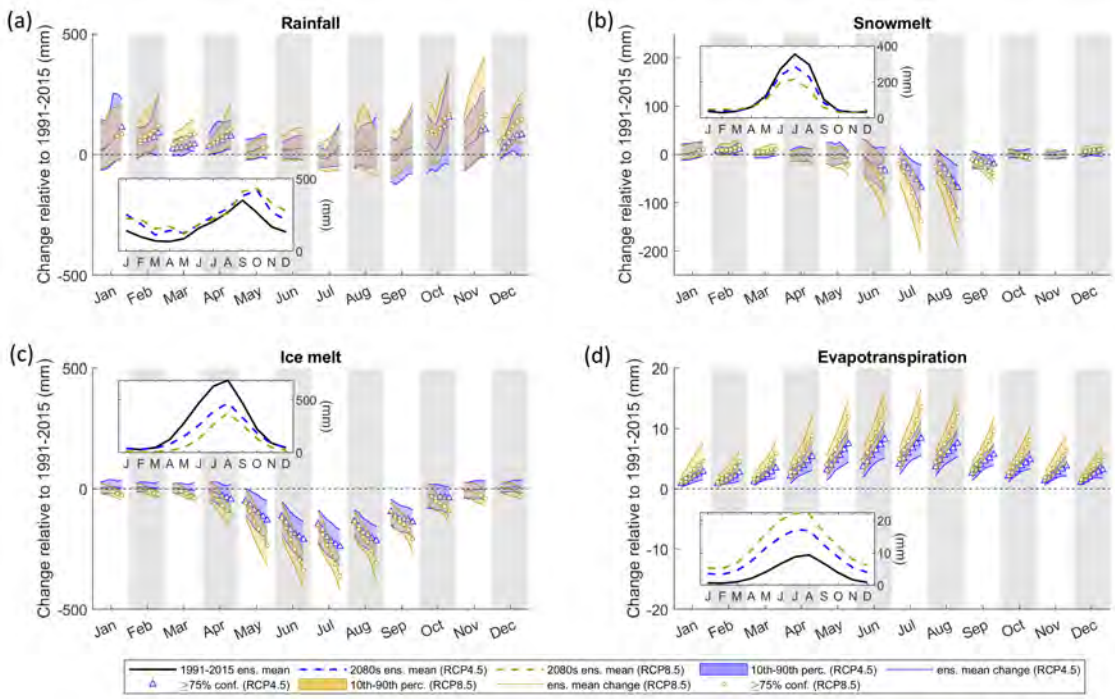

Figure 5.10: Projections of monthly mean runoff components including rainfall (a), snow melt (b), ice melt (c) and evapotranspiration (d) for RCP4.5 (blue) and RCP8.5 (yellow). For each month, the trajectory of the ensemble mean change over the twenty-first century time slices (2030s to 2080s) relative to the reference period (1991-2015) is shown by the solid coloured lines. These lines are marked for each time slice where there is ≥75% confidence in the direction of change. They are bounded by the 10th and 90th percentiles of the projections (bands). Inset in each plot are ensemble mean monthly runoff volumes averaged over the reference period (black solid line) and 2080s (dashed lines).

Projections consistently (≥90%) show an increase in evapotranspiration for all months of the year (Fig. 5.10d) with the largest increases projected under RCP8.5 towards the end of the twenty-first century. However, the volume of evapotranspiration will remain a small component of the overall water balance.

5.5.6 Future evolution of river flow regime

Figure 5.11 shows the projected changes in river discharge signatures relative to the reference period (except peak water for which the raw projections are shown). Under RCP4.5 the ensemble mean projection of peak water is 2045, while under RCP8.5 peak water is projected to occur 17 years earlier in 2028. Indeed, the RCP8.5 projections of the mean annual flow signature ($\bar{Q}$) show a consistent decline through the twenty-first century with ≥90% confidence that flows will reduce by the end of the century by 19% on average. In contrast, under RCP4.5 the magnitude of the decline is smaller (ensemble mean projects 7% decrease for 2080s) and the direction of change is more uncertain. Both RCPs project an increase in inter-annual flow range (R_{ANN}) throughout the twenty-first century (≥75% under RCP8.5). Under RCP4.5 the ensemble mean projects a 47% increase in R_{ANN} by the 2080s while RCP8.5 shows a 71% increase.

Seasonally, monthly winter (DJF) flows are projected to increase while ≥90% of the ensemble project a decrease in summer (JJA) flows by the 2080s under both RCPs. The absolute change in mean monthly flows is larger for summer flows on average, but proportionally, the winter flows are projected to change most, particularly in February where the ensemble mean projects a 60% and 67% increase under RCP4.5 and RCP8.5 respectively by the end of the century. The combined effect of increased winter flows and decreased summer flows results in decreased intra-annual flow variability. Under both RCPs, more than 90% of the ensemble project a decrease in R_{mnth} relative to the reference period from the 2050s onwards. The ensemble mean projections under RCP8.5 show a decade-on-decade reduction in R_{mnth} with time and a 41% reduction by the end of the century.

Of those signatures with units $m^3\ s^{-1}$, the high flow Q_{01} signature shows the largest

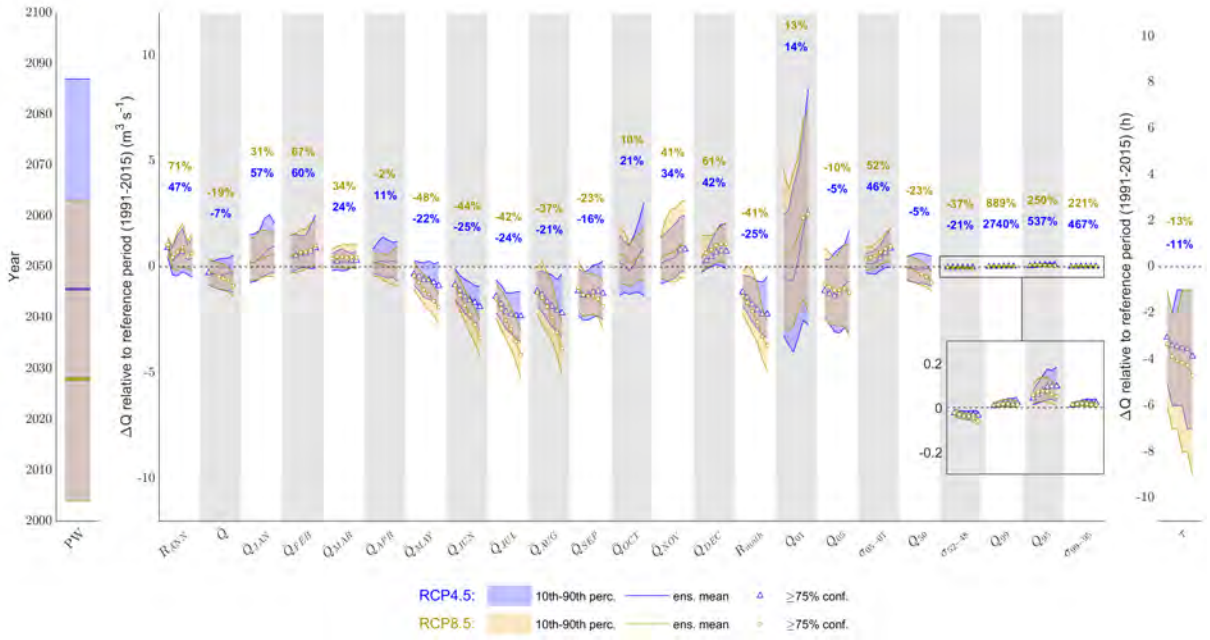

Figure 5.11: Projected changes in river discharge signatures. For each signature, the trajectory of the ensemble mean change over the twenty-first century time slices (2030s to 2080s) relative to the reference period (1991-2015) is shown by the solid coloured lines. These lines are marked for each time slice where there is ≥75% confidence in the direction of change. They are bounded by the 10th and 90th percentiles of the projections (bands). Also shown are 2080s ensemble mean change expressed as a percentage of simulated signatures for the reference period (text). Note, the peak water (PW) signature is not expressed as a change, but as the overall raw projections.

ensemble mean increase of 2.8 m^3 s^{-1} and 2.5 m^3 s^{-1} for RCP4.5 and RCP8.5 respectively by the end of the century. There is high confidence ($\geq$75%) that Q_{01} will increase relative to the reference period under RCP8.5 but the magnitude of change is uncertain under both RCPs. For Q_{05}, the ensemble means from both RCPs both show a reduction throughout the twenty-first century, however the 10th and 90th percentile span positive and negative values of change for all decades. The ensemble mean projections of changes to high flow variability (σ_{05-01}) are positive throughout the twenty-first century under both RCPs. In the latter half of the century, $\geq$75% of the projections under RCP4.5 show an increase in σ_{05-01} while $\geq$90% of the projections under RCP8.5 show an increase.

For moderate flows, the ensemble mean of the RCP4.5 projections show a small reduction in Q_{50} of approximately 0.15 m^3 s^{-1} throughout the twenty-first century while the RCP8.5 ensemble mean projects a decade-on-decade reduction in Q_{50} and by the end of the century there is high confidence ($\geq$90%) that moderate flows will reduce. Moderate flow variability (σ_{52-48}) is projected to reduce with high confidence under both RCPs, albeit by only 0.03 m^3 s^{-1} and 0.06 m^3 s^{-1} by the 2080s under RCP4.5 and RCP8.5 respectively.

For the slow-release low flow signatures, $\geq$90% of the projections are positive throughout the twenty-first century under both RCPs indicating an increase in the magnitude of low flow events (or equivalently a reduction in the frequency of these flow events) and variability of low flows. The absolute changes in the ensemble means never exceed 0.1 m^3 s^{-1} for these signatures, although proportionally, they show the largest degree of change, particularly for Q_{99} where the proportional change exceeds 2000% under RCP4.5.

Finally, the response time to runoff (τ) is projected to decrease throughout the twenty-first century under both RCPs ($\geq$90% confidence) indicating the catchment will likely become more flashy. The magnitude of change is small where the ensemble mean projects a small reduction in τ of 3.9 hours under RCP4.5 and a slightly greater reduction of 4.7 hours under RCP8.5.

5.5.7 Sources of uncertainty in river flow regime projections

Figure 5.12 shows the ANOVA effect sizes calculated for the 2030s and 2080s for each river discharge signature. The error term (η^2_ϵ) never exceeds 0.09 and for 21 of the 25 signatures is < 0.03 indicating that the main effects and first-order interaction terms explain the majority of projection uncertainty. For the 2030s, RCP uncertainty contributes 4-27% of the total projection uncertainty across the signatures. By the 2080s, RCP contributes up to 65% of total projection uncertainty. In fact, for all but four signatures, RCP contributes proportionally more to total projection uncertainty in the 2080s than the 2030s. By the 2080s the five signatures with the highest η^2_{RCP} include the mean monthly flows from May to August and the mean monthly flow range (R_{mnth}) signature (Table 5.2) for which the effect sizes are at least 0.47. GCM-RCM uncertainty is the largest contributor to total projection uncertainty for 21 of the 25 river discharge signatures for the 2030s and it still remains a significant contributor to projection uncertainty by the 2080s with a mean effect size across the signatures of 0.3. Four of the five most sensitive signatures to GCM-RCM uncertainty for the 2030s remain in this top five for the 2080s (Table 5.2) and these include the mean monthly winter flows in January and February and two of the quick-release high flow signatures (Q_{01} and Q_{05}).

On average, the DS parameterisation contributes 18% of the total projection uncertainty across the signatures for the 2030s. In fact, η^2_{DS} is relatively consistent across the signatures, ranging from 0.1-0.2 for 18 of the 25 signatures. For the 2080s, η^2_{DS} reduces for all signatures except mean November and December flows and the inter-annual flow range (R_{ANN}). For R_{ANN}, DS has the largest effect size, contributing 43% of the total projection uncertainty. Autumn and winter monthly mean flows for September, November, December and February make up the remainder of the top five signatures most affected by DS uncertainty for the 2080s. On average TIM uncertainty contributes 9% of the total projection uncertainty across the different signatures for the 2030s. For this period it is the largest contributor to R_{ANN} uncertainty ($\eta^2_{TIM} = 0.35$) and it also shows significant contributions to mean monthly flow projection uncertainty for April ($\eta^2_{TIM} = 0.17$) and

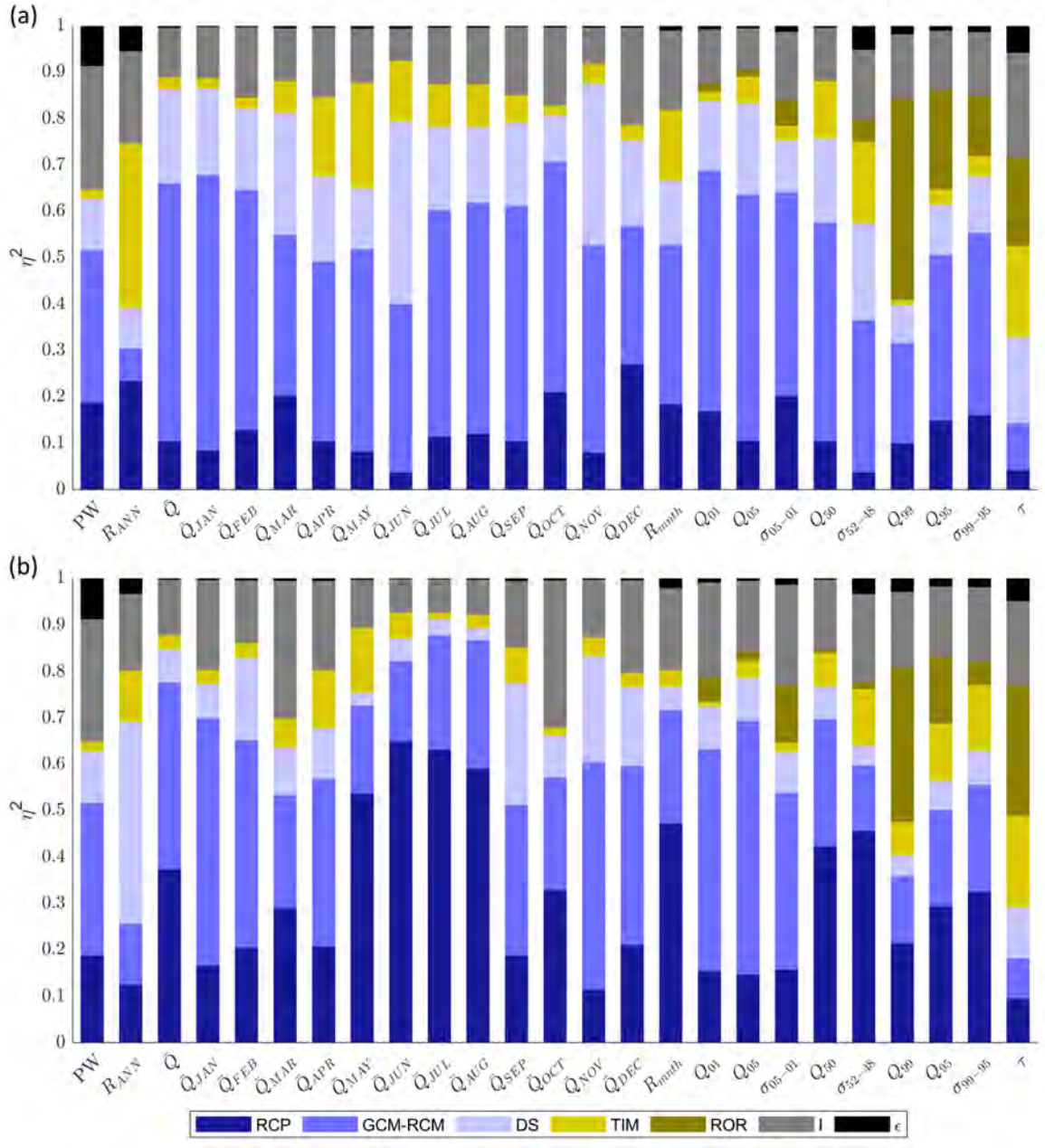

Figure 5.12: Effect size (η^2) of all main effects (RCP, GCM-RCM, DS, TIM and ROR), interactions (I) and remaining error (ϵ) on projected changes in the 25 river discharge signatures at the start (2030s, a) and end (2080s, b) of the twenty-first century.

Table 5.2: Top five river discharge signatures ranked according to the average effect size for each of the main effects, interactions and remaining error on projected changes for the 2030s and 2080s. Effect sizes are included in brackets.

Decade	Rank	RCP (η^2_{RCP})	GCM-RCM ($\eta^2_{GCM-RCM}$)	DS (η^2_{DS})	TIM (η^2_{TIM})	ROR (η^2_{ROR})	I (η^2_I)	ϵ (η^2_ϵ)
2030s	1	$\bar{Q}_{DEC}$ (0.27)	$\bar{Q}_{JAN}$ (0.59)	$\bar{Q}_{JUN}$ (0.39)	R_{ANN} (0.35)	Q_{99} (0.43)	PW (0.27)	PW (0.09)
	2	R_{ANN} (0.23)	$\bar{Q}$ (0.56)	$\bar{Q}_{NOV}$ (0.35)	$\bar{Q}_{MAY}$ (0.23)	Q_{95} (0.22)	τ (0.23)	τ (0.06)
	3	$\bar{Q}_{OCT}$ (0.21)	Q_{05} (0.53)	$\bar{Q}_{MAR}$ (0.26)	τ (0.20)	τ (0.19)	$\bar{Q}_{DEC}$ (0.21)	R_{ANN} (0.05)
	4	$\bar{Q}_{MAR}$ (0.20)	Q_{01} (0.52)	$\sigma_{52\text{-}48}$ (0.21)	$\sigma_{52\text{-}48}$ (0.18)	$\sigma_{99\text{-}95}$ (0.13)	R_{ANN} (0.20)	$\sigma_{52\text{-}48}$ (0.05)
	5	$\sigma_{05\text{-}01}$ (0.20)	$\bar{Q}_{FEB}$ (0.52)	$\bar{Q}$ (0.20)	$\bar{Q}_{APR}$ (0.17)	$\sigma_{05\text{-}01}$ (0.06)	R_{mnth} (0.17)	Q_{99} (0.02)
2080s	1	$\bar{Q}_{JUN}$ (0.65)	Q_{05} (0.55)	R_{ANN} (0.43)	τ (0.20)	Q_{99} (0.33)	$\bar{Q}_{OCT}$ (0.32)	PW (0.09)
	2	$\bar{Q}_{JUL}$ (0.63)	$\bar{Q}_{JAN}$ (0.53)	$\bar{Q}_{SEP}$ (0.26)	$\sigma_{99\text{-}95}$ (0.14)	τ (0.28)	$\bar{Q}_{MAR}$ (0.29)	τ (0.05)
	3	$\bar{Q}_{AUG}$ (0.59)	$\bar{Q}_{NOV}$ (0.49)	$\bar{Q}_{NOV}$ (0.23)	$\bar{Q}_{MAY}$ (0.14)	Q_{95} (0.14)	PW (0.27)	$\sigma_{52\text{-}48}$ (0.03)
	4	$\bar{Q}_{MAY}$ (0.54)	Q_{01} (0.48)	$\bar{Q}_{FEB}$ (0.18)	$\bar{Q}_{APR}$ (0.12)	$\sigma_{05\text{-}01}$ (0.12)	$\sigma_{05\text{-}01}$ (0.21)	R_{ANN} (0.03)
	5	R_{mnth} (0.47)	$\bar{Q}_{FEB}$ (0.45)	$\bar{Q}_{DEC}$ (0.17)	Q_{95} (0.12)	Q_{01} (0.05)	Q_{01} (0.20)	Q_{99} (0.03)

May ($\eta^2_{TIM} = 0.23$) at the beginning of the melt season. It is also the largest contributor to uncertainty of projections of response time to runoff (τ) where $\eta^2_{TIM} = 0.20$. For the 2080s the average η^2_{TIM} falls slightly to 7%, but TIM uncertainty remains an important contributor to total projection uncertainty for τ, April and May flows and two of the low flow signatures (Q_{95} and $\sigma_{99\text{-}95}$) where $\eta^2_{TIM} \geq 0.12$. Uncertainty stemming from the ROR structure-parameterisation has a negligible influence on the decadal signatures (PW and R_{ANN}) and those signature characterising annual and monthly mean flows for the 2030s and 2080s. For the 2030s it is important for projections of low flow magnitude (Q_{99} and Q_{95}, $\eta^2_{ROR} = 0.43$ and 0.20 respectively) and variability ($\sigma_{99\text{-}95}$, $\eta^2_{ROR} = 0.13$). In fact, for Q_{99}, ROR is the single largest contributor to total projection uncertainty. For the 2080s, its influence on low flow quantiles remains significant and it is the single largest contributor to both the Q_{99} and τ projection uncertainty. It also remains a significant contributor to the high flow variability signature, $\sigma_{05\text{-}01}$ where $\eta^2_{ROR} = 0.12$.

Unlike ice and snow coverage, interactions between model components significantly contribute the total projection uncertainty across the signatures where η^2_I ranges between 0.07 and 0.27 for the 2030s and between 0.07 and 0.32 for the 2080s. Figure 5.13 shows the decomposition of the five interaction terms with the largest effect sizes on average for the 2030s (a) and 2080s (b). The interactions between the RCP and GCM-RCM model chain components dominate the contribution to projection uncertainty. However, inter-

actions between the climate model chain components and GHM++ (e.g. DS-TIM) may also contribute to the projection uncertainty. For R_{ANN}, DS-TIM interaction contributes 7% of total projection uncertainty for the 2030s and 2080s. Furthermore interactions between the TIM and ROR contribute some (albeit small) amounts to the total projection uncertainty. For 16 of the signatures, the contribution from interactions between model chain components increases from the 2030s to the 2080s. These include all of the signatures that characterise, high-, moderate- and low-flow magnitude and variability, but the largest increases are shown for March and October mean monthly flows.

5.6 Discussion

5.6.1 Future evolution of river flow regime

There is high confidence that near-surface air temperature will increase by the late twenty-first century (2076-2100) relative to conditions in the recent past (1981-2005). Precipitation and incident solar radiation were projected to slightly increase and decrease respectively on average - a finding that is consistent with other analyses of the EURO-CORDEX projections for northern Europe (Bartók et al., 2017). The primary driver of changes in snow and ice is near-surface air temperature, while precipitation and incident solar radiation are of secondary importance. Because of this, there is high confidence that glacier ice and snow will continue to retreat as near-surface air temperature rises throughout the twenty-first century which would leave the river basin almost free of snow and ice by 2100 under the warmest climate projections.

The signature-based analysis undertaken in this study has revealed how climate change will impact the magnitude, timing and variability of downstream river flows over a range of timescales in the Virkisá River basin. Projected changes in flow seasonality broadly follow those shown for other mid-latitude alpine river basins where the loss of snow and ice will reduce meltwater inputs in the summer months and a phase shift of precipitation from snowfall to rainfall combined with enhanced melt during the colder months will increase winter runoff (Addor et al., 2014; Huss et al., 2014; Jobst et al., 2018; Mandal

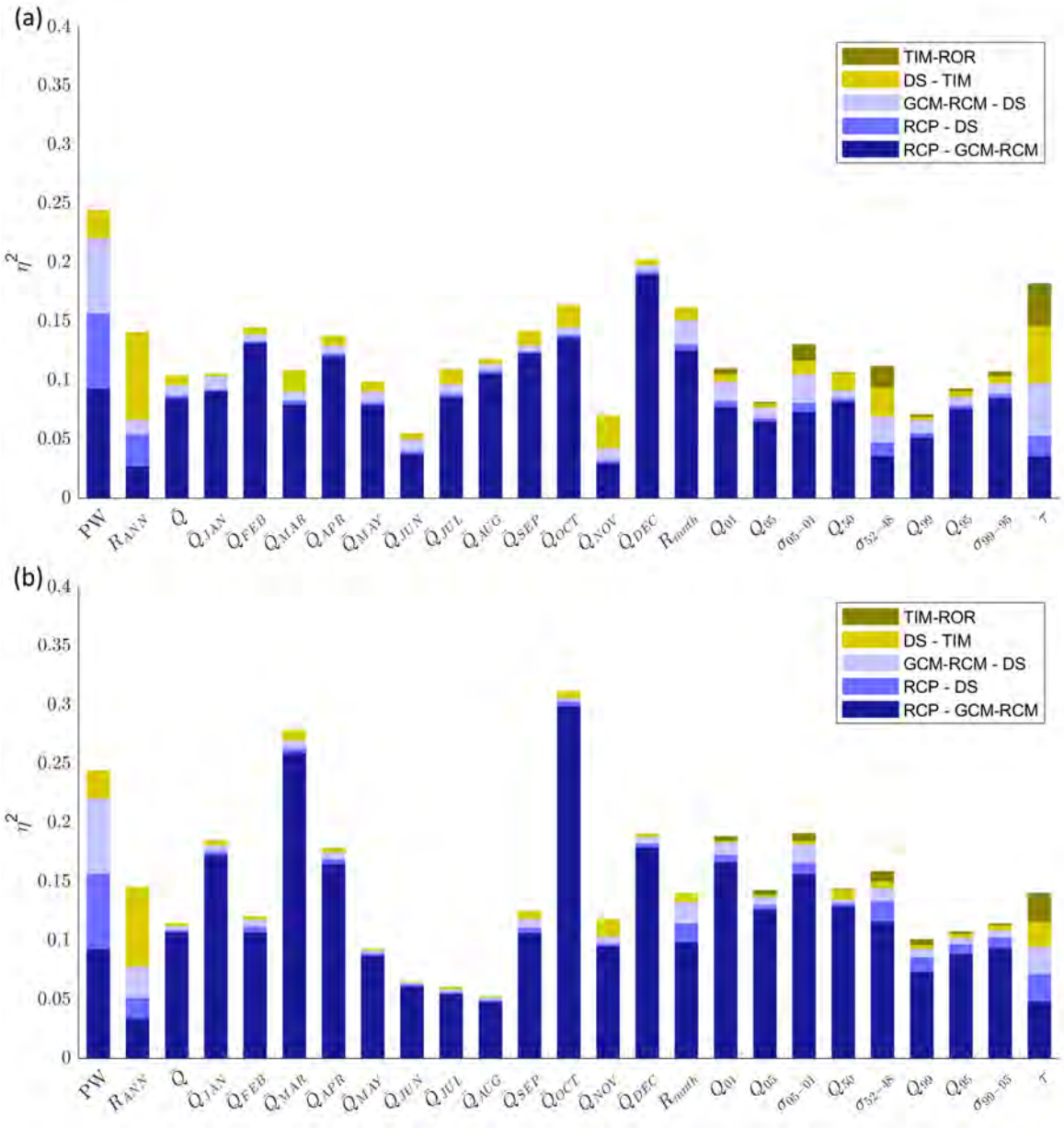

Figure 5.13: Effect size (η^2) of the five most significant interactions on projected changes in the 25 river discharge signatures at the start (2030s, a) and end (2080s, b) of the twenty-first century.

and Simonovic, 2017). Summer runoff is projected to decrease by 24% under RCP4.5 and 40% under RCP8.5 by the 2080s while winter runoff is projected to increase by 59% under RCP4.5 and 57% under RCP8.5 by the 2080s. The consequence of these seasonal shifts in runoff is that intra-annual (monthly) flow variability will reduce by 25% (RCP4.5) and 41% (RCP8.5) by the 2080s. Furthermore, the magnitude of very low flow events (Q_{99}), which typically occur during the winter months, is likely to increase.

On average, the projections indicated that the seasonal redistribution of runoff will have little influence on mean annual flows under RCP4.5 (-7% by the 2080s) as changes in summer and winter flows approximately compensate for one another. Under RCP8.5, however, the more pronounced reduction in summer melt inputs results in a 19% reduction by the 2080s. The loss of a consistent melt input to the river basin and its evolution to a hydrological regime governed by rainfall-runoff processes means inter-annual flow variability (R_{ANN}) will increase by 47% (RCP4.5) and 71% (RCP8.5) by the 2080s. The increase in rainfall inputs, particularly during the storm-prone autumn and winter months, likely explains the projected increased magnitude of very high flow events (Q_{01}), a finding that is in agreement with other studies that have investigated changes in high flow magnitudes in glaciated river basins (Lutz et al., 2016). It is likely that the intensification of peak flow magnitudes will be further exacerbated by the projected decrease in river flow response time to runoff (τ), which is an artefact of losing the runoff-regulating ice and snow water stores. Accordingly, the river basin will become more flashy and flood-prone in the future.

Increased flood frequency has major implications for local infrastructure in the vicinity of the Virkisá River basin. In particular, the southern section of the Route 1 highway which passes over the Skeiðarársandur floodplain navigates a large number of glacier-fed rivers including the Virkisá. Due to the unconsolidated nature of the floodplain lithology, the morphology of these rivers can change rapidly, particularly during high flow events (Marren, 2005) and often at considerable expense to the road authority (Björnsson and Pálsson, 2008). Accordingly, the projected increase in frequency and severity of high flow events will likely incur further expenses to maintain this transport link in the future.

Beyond local implications, one should be cautious in extrapolating the findings from

this study to other glaciated catchments in Iceland or beyond, but it is clear that the timing, magnitude and variability of glacier-fed river flows over a range of timescales are sensitive to climate change. For Iceland, these changes could impact glacier-fed hydroelectric dams, which are a primary source of electricity for the country. Increased frequency and magnitude of high flow events could render current dams unsafe if their designed flood capacity can no longer meet regulation requirements (Thorsteinsson and Björnsson, 2012). The redistribution and levelling out of seasonal flows, however, could actually have a beneficial effect on the running costs and capacity to produce electricity from such projects (Jóhannesson et al., 2007).

5.6.2 Uncertainties in projections of river flow regime

Projections of the direction of change relative to the reference period were well constrained for the majority of river discharge signatures, particularly towards the end of the twenty-first century and for the warmer RCP8.5 pathway. Even so, there was considerable spread in the projected magnitude of these changes due to uncertainties in the driving climate data (RCP,GCM-RCM,DS) and representation of glacio-hydrological processes (TIM,ROR) in the model chain. Uncertainty in future snow and ice coverage primarily stemmed from the RCP due to its control on future near-surface air temperature. In fact, the proportional contribution of the RCP to projection uncertainties increased throughout the twenty-first century and, consequently, the RCP was also found to be the dominant source of uncertainty for projections of mean monthly flows during the melt season by the 2080s. The growing influence of the RCP over time was also shown by Addor et al. (2014) for six alpine catchments in Switzerland and by Duethmann et al. (2016) for two mountain river basins in the Tian Shan. Interestingly though, these studies along with the recent study of Jobst et al. (2018) found that climate model uncertainty was still the dominant source for projections of monthly river flows. Jobst et al. (2018) postulated that this was likely because of the high uncertainty in future precipitation across the climate models. Indeed, others have also attributed future runoff uncertainty in glaciated river basins to variability in precipitation projections (Lutz et al., 2016), a finding which is compounded

by an increasingly warm and thus rainfall-dominated precipitation input. In this study however, the GCM-RCM model chain component only dominated river flow projection uncertainty during the winter months while summer flow uncertainty was dominated by the RCP. There are two key reasons that could explain this. Firstly, precipitation uncertainty across the GCM-RCMs showed to be especially high during winter (Fig. 5.4) which coupled with the fact that rainfall is the primary source of runoff during winter, likely explains the dominant role GCM-RCM plays in projection uncertainty during the winter months. Furthermore, it should be noted that the Virkisá River basin has a much higher proportional glacier coverage (60%) compared to the aforementioned studies (1.8%-22.3%). Therefore, it is postulated that the influence of the RCP in the summer is related to the relatively high proportion of melt runoff that the Virkisá River receives during these months and the fact that the RCP showed to be the dominant contributor to future ice coverage uncertainty. Importantly, this finding also serves to highlight the need to represent atmosphere-cryosphere-hydrosphere feedbacks adequately in future studies, particularly where glacier coverage is high, through the inclusion of a dynamic glacier evolution model in the model chain like that implemented in this study.

For projections of the inter-annual flow range, the DS procedure was the largest contributor to projection uncertainty by the end of the twenty-first century, which should be expected given that the perturbation of this procedure accounted for uncertainty in the random year-by-year sampling of the historic climate data. Uncertainty in the TIM structure-parameterisation was the dominant contributor to the spread in projections of moderate monthly flows during the transition from the cold to melt season, which corroborates the findings in Chapter 4 where the structural representation of melt was important for controlling the initiation of the melt season due to the contrasting sensitivity of the models to temperature and incident solar radiation. It was also found that signatures derived from the flow duration curve as well as those representing flashiness were most sensitive to the configuration of the ROR component. Indeed, in this study, the uncertainty in the ROR structure-parameterisation significantly contributed to the total projection uncertainty of slow-release low flow signatures as well as the response time

(flashiness) of the catchment to runoff. Similar sensitivities in low flow metrics to the choice of hydrological model have been shown for non-glaciated river basins Yuan et al. (2017) and they postulated that these might stem from differences in water storage-release processes in the models. However, a key drawback of this study and other studies that have investigated the role of hydrological model uncertainty in glaciated river basins (e.g. Giuntoli et al., 2015; Vetter et al., 2017) is that they have implemented multiple model codes and therefore cannot make any definite conclusion about the source of the projection uncertainties. For example, Addor et al. (2014) concludes that the sensitivity to the choice of hydrological model could stem from any number of differences between model codes including the structure, climate interpolation method and calibration strategy. In this study, it has been demonstrated that by using a single but flexible model code, it is possible to separate out the sources of projection uncertainties down to the process level. Such insights can be used to help prioritise those aspects of the GHM that require i) additional refinement (e.g. through model development) and ii) adequate representation of their uncertainty to improve projection robustness.

Furthermore, the signature-based analysis undertaken in this study has shown that the importance of these different sources, be it from the GHM or the climate projections is dependent on which signature of river discharge is being evaluated. It is clear, therefore, that signature-based analyses could be used to help prioritise uncertainty sources based on the characteristic of flow one is interested in. For example, the results from this study indicate that for evaluating changes in monthly melt season runoff only, it may be beneficial to ignore ROR uncertainty and focus time and computational resources on quantifying uncertainties stemming from the remaining model components. In this respect, the time frame of the projections should also be considered, given the apparent change in effect sizes with time demonstrated for projections of snow and ice coverage and river flow signatures (see Appendix G).

More broadly, the results from this study emphasise the need for impact studies to represent uncertainties stemming from model chain components that control future climate and glacio-hydrological behaviour, the second of which has been widely neglected.

The need for this is compounded by the fact that interactions between model chain components exceeded individual main effects for some river discharge signatures. Accordingly, an ensemble that includes perturbations of multiple components of the model chain simultaneously will provide the most rigorous quantification of projection uncertainty.

5.6.3 Limitations

While some characteristics of projected river flow regime change are broadly in agreement with other studies in similar mid-latitude alpine settings (e.g. changes in flow seasonality and projected increase in high flow magnitude), it is important to emphasise that the projected river flow regime shifts should not be generalised across glaciated river basins. Indeed, recent regional (Ragettli et al., 2016) and global (Huss and Hock, 2018) studies have shown that local catchment characteristics such as climate and glacier hypsometry largely influence seasonal river flow response to twenty-first century climate change. In this study a small absolute increase in low flow magnitude was projected, indicating climate change and deglaciation could help to limit periods of water scarcity. However, in more arid regions, where rainfall cannot compensate for reductions in melt, the opposite effect has been shown (Stewart et al., 2015). One might also expect to see much greater changes in the river flow response time to runoff as snow and ice retreat in other river basins. For the Virkisá River basin, a relatively small reduction in response time (τ) was projected on average by the end of the twenty-first century. This, perhaps, should not be surprising given the small size of the river basin and the fact that previous investigations have shown that Virkisjökull has a well developed conduit drainage system that routes runoff efficiently year-round (Flett et al., 2017; Phillips et al., 2014). For larger river basins with more expansive cryospheric water stores, changes in the response time to runoff could be much greater, substantially increasing the risk of pluvial flooding.

Similar inter-catchment variability should also be expected with regards to the dominant sources of projection uncertainty. Indeed, as already noted in this discussion, some of the results from this study contrast the limited number of studies that have investigated uncertainty sources in other glaciated basins. Addor et al. (2014) suggests that catchment

elevation influences the importance of the RCP on projection uncertainty whereby runoff from higher elevation catchments with more snow and ice are more sensitive to the RCP. It is therefore vital that signature-based evaluations like the one undertaken in this study are applied to other glaciated river basins in the future so that regional variations in river flow regime change and uncertainty sources can be evaluated.

It is also important to consider potential deficiencies in the calibrated GHMs. In fact, the model evaluation demonstrated that they were able to capture the majority of the observed river discharge signatures within their observation uncertainty bounds. Even so, it should be noted that there are several limitations in the calibration approach that could have hindered the efficiency of the calibrated models. Firstly, given the distributed structure of GHM++ and the fact that it runs on an hourly time step, running it over multiple years required considerable computation time which limited the number of runs that could be undertaken in the Monte Carlo calibration procedure. 5000 runs was adopted as an appropriate compromise, balancing the density of parameter sampling with available computational resources. Even so, it is recognised that particularly for the more complex model structures which employ more calibration parameters, a denser parameter sampling could help to find more efficient model parameterisations. It should also be noted that the models were calibrated and evaluated on four years of river flow data only. This detail is particularly important given the conceptual nature of GHM++ and thus the potential for the calibration parameters to become less applicable when applied to periods outside of the calibration data. Additionally, it is important to highlight possible model deficiencies brought about by the two-step calibration procedure employed in which the TIM and ROR model chain components were calibrated independently. This was necessary so that the main effects (Eq. 5.3) and interaction terms (Eq. 5.5) for both components could be calculated separately (thus achieving the second aim of the study). However, the drawback of implementing this stepwise calibration procedure over one that calibrates both model components simultaneously is that it neglects any interactions between the TIM and ROR models. Of course, its should be noted that the ANOVA results showed that TIM and ROR interactions are negligible except for two of the 25 signatures evaluated.

Chapter 4 highlighted the historic observed precipitation data as a potential source of model deficiencies. This is partly due lack of available precipitation data at higher elevations, making the gridded dataset employed in this study less reliable near the basin summit. Furthermore, the bias correction procedure applied to the precipitation dataset resulted in time series that were well correlated to the AWS data over a 3-day time step, but that this correlation degraded at shorter daily and hourly time steps which could have contributed to the model's inability to capture snow coverage observations higher up in the catchment and river discharge signatures relating to the timing of flows. Biases in baseline climate data also serve to condition the magnitude of simulated impacts of climate change due to the non-linear response climate change (Remesan and Holman, 2015).

Indeed, uncertainties in the historic precipitation data were not included as part of this study, partly because there was almost no information that would have allowed one to quantify these uncertainties (e.g. rain gauge errors), particularly higher up in the catchment where measurements are least reliable. Additionally, though, it would have meant further increasing the size of the model chain ensemble which was already at the very limit of what was computationally feasible. This, however, raises an important broader limitation of the study in that the total projection uncertainties reported are not indicative of the 'true' uncertainty. Further insights could undoubtedly be gained by perturbing other model chain components including the historic climate time series which may in fact contribute more to projection uncertainty than GHM parameter uncertainty (Islam and Déry, 2017). Jobst et al. (2018) calculated that the bias correction of precipitation contributed up to 22% of seasonal streamflow projection uncertainty and other studies have found that the initial ice volume (Gabbi et al., 2012; Huss et al., 2014) and parameterisation of the glacier evolution model (Huss et al., 2014; Linsbauer et al., 2013) are also important sources of projection uncertainty.

Furthermore, the representations of uncertainty in the five components evaluated in this study are themselves not exhaustive. It is well established that uncertainties in climate model ensembles are under-represented (Daron and Stainforth, 2013) and steps

were taken in this study to limit the total ensemble size so that the experiments were computationally feasible. For example, only 10 random DS sequences were generated, and indeed other aspects of the downscaling procedure could have also been modified (e.g. replacing the linear interpolation of change factors with a moving average model). Additionally, the melt and runoff-routing model structures implemented represent a subset of a much larger population of available models. Other model structures that employ more complex physically based energy balance approaches and hydraulic models that simulate discrete flow pathways through the glacier Arnold et al. (1998) could also be implemented to provide a more accurate representation of the true projection uncertainty.

5.7 Conclusions

Twenty-first century climate change is projected to alter the magnitude, timing and variability of river flows over decadal to sub-daily timescales in the Virkisá River basin. Relative to the 1990s reference period, there was high confidence in the direction of change for the majority of the 25 river discharge signatures over the twenty-first century. The magnitude of change, however, was more uncertain. The application of ANOVA demonstrated that the climate model chain components (RCP,GCM-RCM,DS) were the main sources of this uncertainty. However, uncertainty relating to glacio-hydrological process representation in the model chain (TIM,ROR) was the dominant source of projection uncertainty for some river discharge signatures. Furthermore, interactions between model chain components can exceed individual main effects. A rigorous quantification of uncertainty in different characteristics of river flow regime change in glaciated river basins must, therefore, account for uncertainty emanating from these different model components.

5.8 Summary

The study undertaken in this chapter aimed to meet research objectives 2 and 3 of this thesis. Objective 2 was to, "implement a climate-GHM model chain to project twenty-first century changes in different characteristics of river flow regime". This was achieved

by undertaking a signature-based analysis of twenty-first century climate change impacts on river flow regime in Virkisá River basin. The results have shown for the first time that different characteristics of glacier-fed river flow regime including flow magnitude, timing and variability over a range of timescales are sensitive to climate change. Objective 3 was to, "determine the relative contribution of individual model chain components to uncertainty in river flow regime projections". To meet this objective, an ANOVA decomposition of the projection uncertainties was undertaken which demonstrated that all of the perturbed components of the model chain contribute to projection uncertainty in at least some aspects of glacier-fed river flow regime change.

A subsample of the projections from this study will be used as boundary conditions in the next chapter to drive a distributed groundwater model of the proglacial sandur aquifer. The high uncertainty in projected changes in climate, snow/ice coverage and downstream river flow regime indicate that this subsample should aim to capture a range of future glacio-hydrological and climatic behaviours that are likely to drive changes in proglacial groundwater dynamics.

CHAPTER 6:

TWENTY-FIRST CENTURY EVOLUTION OF PROGLACIAL GROUNDWATER LEVEL DYNAMICS AND GROUNDWATER-SURFACE WATER EXCHANGES

6.1 Introduction

This chapter presents the final research study of this thesis which aims to meet the fourth and final research objective by using a subsample of the climate and river flow projections generated in the previous chapter to drive a distributed groundwater model of the proglacial sandur in the VGO. The subsample is selected using a signature-based cluster analysis approach. The groundwater model projections are then used to evaluate impacts of twenty-first century climate change and changes in the Virkisá River flow regime on proglacial groundwater level dynamics and GW-SW exchanges in the sandur.

All aspects of the methodology in this chapter were designed and undertaken by JM including the configuration and calibration of the distributed groundwater model and the signature-based cluster analysis approach. JM also led the analysis and interpretation. NB, DH, SK, CJ, JE and GA provided guidance on to the groundwater model configuration and choice of signatures used in the cluster analysis. They also contributed to the interpretation of the results.

6.2 Background

Groundwater is increasingly being recognised as an important component of water cycling in glaciated mountain catchments that could become strategically more important as water inputs from rainfall and meltwater become less reliable under twenty-first century

climate change (Ghasemizade and Schirmer, 2013; Taylor, 2013; Vincent et al., 2019). Overburden materials such as talus piles and moraine deposits can form hydrologically important groundwater flow pathways that alter catchment water transit times significantly (Baraer et al., 2015; Clow et al., 2003; Hood and Hayashi, 2015; Langston et al., 2013; Roy and Hayashi, 2009), while the consolidated mountain bedrock may also transmit groundwater downstream to the valley bottom along flow pathways that may extend kilometres deep in some cases (Frisbee et al., 2017; Manning and Solomon, 2003; Wilson and Guan, 2004; Yao et al., 2017).

In the foreland areas of glaciated catchments, unconsolidated, high-storage, alluvial aquifers are commonly found (Baraer et al., 2015; Covino and McGlynn, 2007; Ó Dochartaigh et al., 2019; Ward et al., 1999). These aquifers are normally unconfined with a shallow water table offering an easily accessible source of fresh water to downstream communities year-round (e.g. for drinking and irrigation Stefania et al., 2018). Additionally, they provide a steady supply of baseflow to glacier-fed rivers, helping to supplement low flows outside of the melt season (Andermann et al., 2012; Jódar et al., 2017) and meet downstream water demands for drinking, irrigation and industry. Indeed, hydrochemical tracer studies in the tropical Andes (Baraer et al., 2015; Guido et al., 2016; Somers et al., 2016), the Himalayas (Wilson et al., 2016), the Alps (Kobierska et al., 2015; Magnusson et al., 2014; Ravazzani et al., 2016) and Iceland (MacDonald et al., 2016) have shown that even within tens of kilometres of a glacier outlet, groundwater can contribute more than 70% of river flow outside of the melt season. These groundwater inputs provide more favourable habitat conditions for aquatic fauna by regulating ecologically important physiochemical properties such as channel stability, temperature, electrical conductivity and suspended sediment concentration (Blaen et al., 2014; Brown et al., 2007) which strongly influence spatial and temporal aquatic species abundance and diversity (Brown et al., 2009; Hotaling et al., 2017; Khamis et al., 2016). Even in relatively harsh, high-energy glacial outwash floodplain environments, groundwater springs can provide as much as 40% of total riverine habitat by acting as centres of invertebrate diversity (Crossman et al., 2012; Gray et al., 2006; Ward et al., 1999).

Given the socio-economic and ecological significance of groundwater in glaciated catchments, a fundamental uncertainty is how future availability of water in proglacial aquifers and their exchanges with surface water resources will evolve in a changing climate. Allen et al. (2010) found that groundwater level dynamics in mountain aquifers (i.e. those in mountain valleys and forelands) exhibit two main types of behaviour which could provide a model of the likely drivers of change in proglacial mountain aquifers. In their analysis of hydrometric records from nine aquifers in southern British Columbia, Allen et al. (2010) proposed that groundwater level dynamics in mountain aquifers exhibit two main types of behaviour. The first can be classified as recharge-driven. In these aquifers, groundwater level fluctuations are primarily driven by variability in diffuse recharge from rainfall and snow melt. Future changes in groundwater level dynamics in recharge-driven aquifers will be determined by changes in local climate conditions in the vicinity of the aquifer. The second can be classified as stream-driven. Here, groundwater level fluctuations are driven by the bi-directional exchange of water between the aquifer and rivers draining from the neighbouring mountains. In contrast to recharge-driven aquifers, groundwater level dynamics in stream-driven aquifers are closely coupled to river stage variability and, therefore, may exhibit a response to climate change that is more strongly tied to climatic conditions away from the aquifer in the runoff-bearing mountains.

Of course, in reality, groundwater level dynamics are controlled by the combined influence of diffuse recharge inputs and exchanges with hydrologically connected mountain-fed rivers. A full understanding of aquifer evolution in glaciated mountain environments under climate change must, therefore, take into account: i) changes in diffuse recharge inputs driven by climate change in the vicinity of the aquifer; and ii) changes in mountain-fed river runoff driven by climate and glacier changes in the neighbouring highland regions. Numerical glacio-hydrological and groundwater models offer a means to represent the hydrological linkages between climate, mountain glacio-hydrology and proglacial aquifer in order to investigate current and future groundwater storage dynamics in glaciated mountain aquifers. Okkonen and Kløve (2011) implemented a numerical model chain consisting of a hydrological, land surface and groundwater model to investigate future groundwater

level dynamics in a snow-dominated esker aquifer in Finland. The models were forced with climate change projections based on the SRES A1B emission scenario and they showed that a warming climate induced a shift in peak diffuse recharge inputs from spring to winter causing an earlier rise and recession of groundwater levels. Sridhar et al. (2018) projected an increase in average water table elevations of between 2.1 and 2.6 m by 2042 for an unconfined fractured basalt aquifer in the north-west of the United States and attributed this to projected increases in precipitation.

Other modelling studies have shown mountain-fed streams to be the main driver of groundwater level dynamics in mountain aquifers. Allen et al. (2004) ran a series of MODFLOW groundwater model (Harbaugh, 2005) simulations of the alluvial Grand Forks Aquifer in southern British Columbia and conducted a sensitivity analysis to determine the relative sensitivity of groundwater levels to changes in climate-driven diffuse recharge and mountain-driven river stage boundary conditions. Their high and low recharge scenarios perturbed groundwater levels from the control scenario by +0.05 m and -0.025 m respectively while their high and low river stage scenarios perturbed groundwater levels by 3.45 m and -2.10 m respectively indicating that river stage was the dominant control on groundwater level dynamics. Scibek et al. (2007) built on the work of Allen et al. (2004) using more realistic changes in diffuse recharge and river stage. They forced a transient version of the Grand Forks groundwater model with future climate change scenarios up to 2029. They also used Principle Component Analysis to define a statistical mountain-hydrology model which was used to perturb the river stage boundary conditions as a function of the future climate. They found that climate warming led to an earlier initiation of snow melt and peak in river flow which propagated through to the timing of maximum groundwater levels.

More recently, Huntington and Niswonger (2012) investigated the impact of climate change on groundwater level dynamics in three snow-dominated watersheds in California and Nevada using the integrated GW-SW model, GSFLOW forced with an ensemble of future climate projections based on the IPCC A2 and B1 emission scenarios. The results from their study are of particular interest as they show not only how changes in

snow melt seasonality propagates through to groundwater level seasonality, but also the consequent changes in groundwater contributions to the river further downstream. More specifically, their simulations showed that an earlier recession in mountain spring snow melt and subsequent drop in river stage induced earlier drainage of the alluvial aquifer to the stream. The resultant fall in groundwater levels reduced groundwater discharge to the stream in the summer.

A key limitation of the modelling work undertaken to date is that all of the studies have focussed on snow-dominated catchments with no glacier coverage. This is despite observation studies that highlight the sensitivity of proglacial groundwater level dynamics to mountain runoff (Dragon et al., 2015; Ó Dochartaigh et al., 2019; Robinson et al., 2009). Recently, Liljedahl et al. (2017) showed that the glacier-fed Tanana River in Alaska currently loses 46% of its annual flow to a proglacial alluvial aquifer which equates to more than the total recharge from diffuse sources. Accordingly, they conclude that continued glacier retreat would impact groundwater level dynamics and contributions to surface water systems downstream. In addition, Levy et al. (2015) found that an observed reduction in groundwater levels of up to 2 m in the Skeiðarársandur glacier outwash floodplain could be attributed to a reduction in meltwater recharge at the glacier margin. Glacier-fed rivers are particularly relevant for impact studies because of the uniquely pivotal role that groundwater plays in sustaining ecosystem services in otherwise harsh glacier-fed river systems. Furthermore, unlike snow-fed rivers, glacier-fed river flows typically peak during the warmest and driest summer months when diffuse recharge inputs are lowest (Farinotti et al., 2012; Rees and Collins, 2006; Verbunt et al., 2003). Accordingly, proglacial mountain aquifers are likely to be more sensitive to changes in meltwater inputs in a warming climate. Indeed, the need for further research into the links between climate, meltwater and groundwater dynamics in glaciated mountain watersheds has been highlighted by two comprehensive reviews of glaciated catchment hydrology and hydrogeology (La Frenierre and Mark, 2014; Vincent et al., 2019). There is, therefore, a need to investigate how proglacial groundwater level dynamics and GW-SW exchanges will change under a warming climate.

6.3 Aims

This study investigates how twenty-first century climate change and consequent changes in runoff to the glacier-fed Virkisá River will impact on groundwater storage and GW-SW exchanges with the river and springs in the proglacial sandur aquifer. This will be achieved using a subsample of the river runoff and diffuse recharge projections from chapter 5 to drive a transient MODFLOW-NWT coupled GW-SW flow model (Niswonger et al., 2011) of the proglacial sandur aquifer. Specifically, this study aims to:

1. Evaluate changes in seasonal groundwater level dynamics and GW-SW exchanges with the Virkisá River and springs in the proglacial sandur over the twenty-first century.

2. Determine if the aquifer is predominantly recharge- or stream-driven by evaluating the relative influence of changes in diffuse recharge and river runoff on future proglacial groundwater level dynamics.

The first aim is of key relevance to water resource issues and the consequent social and ecological impacts as it addresses the future role of the aquifer as a groundwater resource and contributor to surface water resources. The second aim has a broader scientific relevance as it addresses the importance of hydrological linkages between surface and groundwater in a coupled mountain glacier-alluvial system and follows on from past research in other study regions which have investigated the relative importance of recharge and stream in driving mountain aquifer groundwater dynamics.

6.4 Methodology

6.4.1 MODFLOW-NWT groundwater model

The MODFLOW-NWT code was used to simulate groundwater flow through the sandur as well as water exchanges with the Virkisá River and surface water springs. The model operates over a spatially discretised three-dimensional Cartesian grid of rectangular

model cells and uses a finite-difference approach to solve the three-dimensional saturated groundwater flow equation:

$$\frac{\partial}{\partial x}\left(K_x\frac{\partial h}{\partial x}\right)+\frac{\partial}{\partial y}\left(K_y\frac{\partial h}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial h}{\partial z}\right)=S_s\frac{\partial h}{\partial t}-W \tag{6.1}$$

where K_x, K_y and K_z are the hydraulic conductivity values in the x, y and z directions [L T^{-1}], h is the hydraulic head [L], W represents per unit volume net source (positive) and sink (negative) fluxes [T^{-1}] and S_s is the specific storage [L^{-1}].

The continuity equation for a given model cell can be expressed as:

$$\sum Q_{in}=S_s\frac{\Delta h}{\Delta t}V \tag{6.2}$$

where $\sum Q_{in}$ is the sum of all flows into and out of the cell [L^3 T^{-1}] which includes exchanges with neighbouring aquifer cells and from external stresses such as rivers, drains and diffuse recharge and V is the cell volume. Note, MODFLOW-NWT is specifically designed to simulated unconfined aquifers with a free water table so that groundwater flow is proportional to the saturated thickness of the aquifer. Under such conditions, water released from elastic storage will be small relative to the drainable porosity released under gravity which is defined by the specific yield, S_y [-].

Flow between cells in direction i where $i \in \{x, y, z\}$ is calculated from Darcy's law:

$$q_i=K_iA\frac{\Delta h}{\Delta s_i} \tag{6.3}$$

where s_i is the model cell length in direction i and A is the cross-sectional area of the flow plane normal to the direction of flow.

The streamflow-routing package (SFR1) (Prudic et al., 2004) was used to simulate river-aquifer interactions. The SFR1 package splits the river into a number of interconnected sections so that each section is assigned to a given MODFLOW model cell. For a given section of river with cross-sectional river bed area A_{riv} [L^2], river bed thickness m [L], and river bed hydraulic conductivity K_{riv} [L T^{-1}], flow between the river section and the underlying groundwater model cell (river leakage), q_L [L^3 T^{-1}] is calculated using Darcy's Law:

$$q_L=K_{riv}A_{riv}\frac{(h_{riv}-h^*)}{m} \tag{6.4}$$

where h_{riv} is the river hydraulic head [L] and h^* is the prescribed head in the aquifer which depends on where the simulated aquifer head, h_{aq} lies in relation to the river bed bottom elevation, z_{bot} [L] so that:

$$h^* = \begin{cases} h_{aq} & h_{aq} \geq z_{bot} \\ z_{bot} & h_{aq} < z_{bot} \end{cases} \tag{6.5}$$

Note that in eq. 6.4, the river is losing when q_L is positive and gaining when the leakage is negative.

In addition to simulating river-aquifer interactions, the SFR1 package also routes river flow along the river channel using Manning's equation for uniform open channel flow:

$$Q_{riv} = \frac{A_{chn} R^{2/3} S_0^{1/2}}{n} \tag{6.6}$$

where Q_{riv} is the river discharge [L^3 T^{-1}], n is the Manning's roughness coefficient [-], A_{chn} is the cross-sectional areas of the river channel [L^2], R is the channel hydraulic radius [L] and S_0 is the river bed slope.

The in-built drainage package was also used to simulate groundwater springs. Drains remove water from the aquifer when the aquifer water table, h_{aq} is above the drain elevation, z_{drn}. The drainage rate, q_D [L^3 T^{-1}] is calculated as:

$$q_D = \begin{cases} c_{drn}(h_{aq} - z_{drn}) & h_{aq} > z_{drn} \\ 0 & h_{aq} \leq z_{drn} \end{cases} \tag{6.7}$$

where c_{drn} [L^2 T^{-1}] is the drain conductance.

6.4.2 Groundwater model configuration

The model domain was divided into a Cartesian grid with 50 m horizontal and 10 m vertical resolution. This resolution was selected as a compromise between model complexity and runtime requirements, to ensure consistency with the GHM++ model grid and to ensure that the aquifer geometry could be adequately represented.

The Virkisá groundwater catchment derived by Ó Dochartaigh et al. (2019) was used to define the lateral extent of the sandur aquifer using a combination of no flow and specified flux boundary conditions (Fig. 6.1a). The upstream model boundary (B1 in Fig. 6.1a) coincides with known outcropping of the bedrock approximately 800 m downstream of the lake outlet. Pumping tests conducted by Ó Dochartaigh et al. (2012) have shown that the permeability of the volcanic bedrock material is negligible, and so no flow model nodes were specified at this boundary. No flow conditions were also imposed along the boundaries to the east and west of the river (B2 and B3 in Fig. 6.1a) given that groundwater level contours are perpendicular to the river here (Ó Dochartaigh et al., 2019), and flows across these boundaries are, therefore, negligible. At the downstream boundary (B4 in Fig. 6.1a), specified flux boundary conditions were prescribed by assuming that the hydraulic gradient is equal to the topographic slope. This is a reasonable assumption given that the water table resides at or very close to the ground surface here year-round.

The thickness of the sandur aquifer was set according to a bedrock topography model derived from two Tromino® passive seismic surveys (Fig. 6.1b, see Appendix H for details). No flow boundary conditions were also imposed at the aquifer-bedrock interface.

The Virkisá River was represented by a single rectangular channel running between the upstream and downstream model boundaries. This is somewhat contrary to the known geometry of the channel which braids significantly in the lower (downstream) sandur where sand and fine gravel material dominates (Fig. 6.1a). However, it was decided not to represent the braided channel system explicitly in the model given that diurnal and seasonal variations in meltwater discharges as well as less frequent high-flow-magnitude rainstorm events are known to bring about regular (sub-annual) sediment reworking, causing changes to channel morphology, channel migration and channel deactivation and reactivation (Marren, 2005). As such, any attempt to include a braided channel system (e.g. based on the aerial imagery in Fig. 6.1a) would require the inclusion of hundreds of interconnected channels which would only be valid for a short time frame. Furthermore, the inclusion of any dynamic landscape evolution component in the model was beyond the scope of this study.

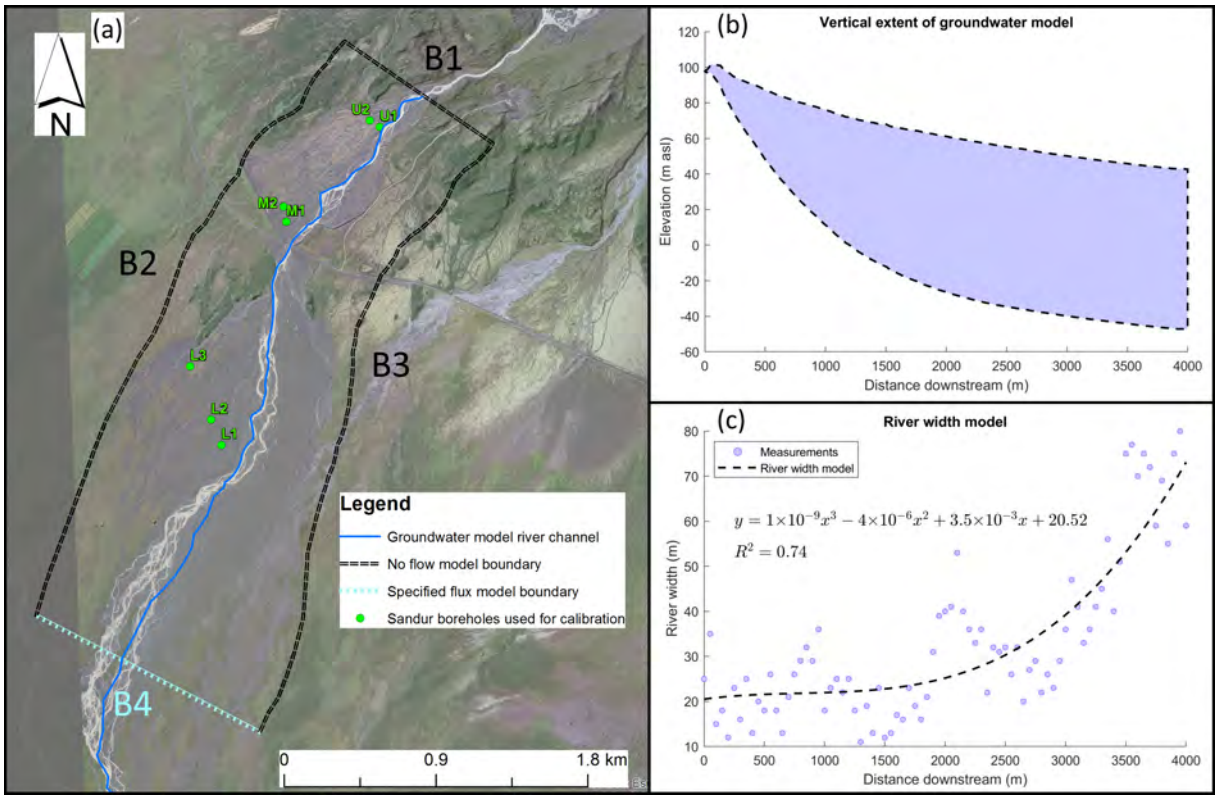

Figure 6.1: Groundwater model lateral (a) and vertical (b) extent and model used to define river width variations (c). Aerial image taken on Aug 7, 2012, source: DigitalGlobe (Vivid - Iceland).

An important implication of river channel braiding is that it promotes water exchanges between the aquifer and the river as the cumulative river bed area (A_{riv} in eq. 6.4) increases with the number of braids. To account for this, the width of the river channel was varied in the groundwater model according to the degree of braiding downstream. More specifically, the total width of all river channels was measured at 50 m intervals along the main river channel (blue dots in Fig. 6.1c). A polynomial was then fit to these data points (black dash line in Fig. 6.1c) to determine the change in river channel width with distance downstream. This was deemed more appropriate than assigning the raw width data to the river channel model as it promotes model stability and convergence during transient simulations.

No direct measurements of the river bed elevation exist except at the ASG1 gauging station. However, numerous sections of the river have been waded into during field excursions where water depth is typically between 0.2 and 1 m with an estimated average water depth of 0.5 m. Accordingly, the river bed bottom elevation was set for each river section by subtracting 0.5 m from the 2011 lidar DEM of the sandur.

In addition to the river, an extensive network of ephemeral and perennial springs exist, particularly in the lower sandur. These springs are thought to form a significant drainage output (MacDonald et al., 2016) and so were included in the model using drains at the land surface elevation at each x and y coordinate covering the lateral extent of the model domain.

6.4.3 Groundwater model driving data

The two major water inputs to the groundwater catchment are mountain runoff from the Virkisá River basin and diffuse rainfall and snow melt recharge over the sandur. Both of these can be specified as transient boundary conditions within MODFLOW-NWT and this was done using the GHM++ simulations from chapter 5. The raw hourly river discharge simulations were used to specify inflows at the upstream end of the river. Hourly diffuse recharge was specified across the groundwater catchment using the simulated excess soil water calculated by the soil water balance model in the GHM++ code. Note that

the GHM++ model domain from chapter 5 does not cover the entirety of the Virkisá groundwater catchment. As such, the domain was extended for this study using the same model parameterisations.

As part of the study in chapter 5, 336 unique GHM compositions (structure-parameter combinations) were calibrated based on simulations using the historical climate data (1980-2016). These simulations are hereafter referred to as the historical ensemble and show variability in river flow and diffuse recharge simulations due to differences in the melt and runoff-routing components of the GHM. Additionally, 94800 twenty-first century (2005-2100) simulations were undertaken, made up of every combination of the GHM compositions and 280 climate projections. These simulations are hereafter referred to as the future ensemble and show variability in river flow and diffuse recharge simulations due to differences in the climate projections and the melt and runoff-routing components of the GHM.

6.4.4 Groundwater model calibration

Within MODFLOW-NWT, the hydraulic parameters can be specified for each model cell individually. Depositional processes on alluvial aquifers can bring about complex distributions of sediment grain size, shape, packing and orientation giving rise to spatial heterogeneity in aquifer permeability and porosity (Chen et al., 2010; Neton et al., 1994). However, the only quantitative data on hydraulic parameter heterogeneity come from the pumping tests undertaken by Ó Dochartaigh et al. (2019). These indicate that transmissivity increases downstream and observations at the site show considerable downstream fining of sandur material which could conceivably bring about changes in hydraulic properties. Even so, the transmissivity data alone do not provide conclusive evidence for changes in hydraulic conductivity given that the aquifer thickness is also known to increase downstream. There is also no information on changes in storage properties of the aquifer. Given the lack of information available to constrain to a heterogeneous parameterisation of the sandur, and the additional model uncertainties that would be introduced by doing so, it was decided to fix the aquifer hydraulic properties across the sandur.

This meant that eight model parameters needed to be specified using available field data or through model calibration (Table 6.1). Prior to undertaking model calibration, a number of preliminary steady-state simulations were used to explore parameter sensitivity and refine their behavioural ranges. Groundwater level simulations showed to be insensitive to the prescribed river bed thickness parameter, m which is likely due to the shallow location of the water table and therefore relatively linear behaviour of eq. 6.4. Accordingly, this was set to an estimated thickness of 0.5 m. It was also decided to fix the specific storage parameter, S_s given that the sandur is unconfined. S_s was set to 10^{-5} m^{-1} based on representative values reported for unconsolidated coarse sand aquifers (Domenico and Schwartz, 1990). In addition, the river bed hydraulic conductivity, K_{riv} was set to 40 m d^{-1} based on the mean of the surface permeability measurements undertaken by Ó Dochartaigh et al. (2019) in abandoned river channels on the sandur. Given the observed similarity between river bed material and the material found on the winder sandur, this parameter was also used to set the drain conductance, c_{drn} (see Table 6.1). Finally, the Manning's roughness coefficient was set to 0.05 based on the 'normal' value for mountain streams with beds made up of cobbles and large boulders (Chow, 1959).

The three remaining model parameters were calibrated against groundwater level time series data from the seven observation boreholes in the sandur. These include: 1) horizontal hydraulic conductivity in the x and y directions, which are assumed equal so that $K_x = K_y = K_h$; 2) the vertical anisotropy, κ which is specified as the ratio between vertical (K_z) and horizontal hydraulic conductivity so that $\kappa = \frac{K_z}{K_h}$; and 3) the specific yield, S_y. Half of the data (up to August 31st 2014) were used for model calibration and the other half (up to December 31st 2016) were used for model evaluation.

The parameters were calibrated using a Monte Carlo approach with 5000 parameter sets drawn randomly from uniform distributions using the Sobol sampling technique adopted in previous chapters (Brately and Fox, 1988). 5000 runs was deemed adequate given that two of the three calibration parameters and all of the main state variables were identifiable (see section 6.5.1). For each run, the model was executed in transient mode and set to run for a simulation period between 2010 and 2017. This period covers

the groundwater level observation time series and also includes an additional two year spin-up period. A prior, steady state simulation was also undertaken for each run using the average diffuse recharge and river inflow fluxes over the simulation period as driving data. The hydraulic head field from the steady state simulation was used as the initial head for the transient simulation.

The model fit was quantified using the average root mean squared error (RMSE) over all observation boreholes. All simulated groundwater levels were extracted using the Observation Process MODFLOW package (Harbaugh, 2005) which interpolates simulated groundwater levels in space and time so that the simulations are spatially and temporally coherent with available observation data.

A key constraint on the model calibration procedure was the computational burden of running the groundwater model thousands of times. Accordingly, two measures were taken to ensure that the calibration could be executed within time constraints. Firstly, the time step of the model was set to daily rather than hourly. This reduced the runtime of a single calibration run from six hours to sixteen minutes. Secondly, a single river discharge and diffuse recharge simulation from the 336 calibrated GHM++ models (historical ensemble) was derived by taking the ensemble mean. The calibrated groundwater model therefore represents the best groundwater model approximation given the average river discharge and diffuse recharge simulations from chapter 5.

6.4.5 Twenty-first century groundwater model simulations

To drive the calibrated groundwater model over the twenty-first century, river discharge and diffuse recharge simulations from the future ensemble were used. Similarly to the calibration runs, all transient twenty-first century simulations were executed with a two year spin-up period and initial hydraulic head field from a prior steady state run.

Given the computational demands of running a distributed groundwater model, it was not feasible to run the entire ensemble of projections from chapter 5, particularly as it was important to run the groundwater model on an hourly time step so that sub-daily characteristics of river flow, which may bring about episodic GW-SW exchange

Table 6.1: Parameters requiring specification or calibration for MODFLOW-NWT.

Parameter	Description	Calibration range	Justification
K_h	Horizontal hydraulic conductivity	1-20 m d^{-1}	Preliminary steady-state runs
κ	Vertical anisotropy	0.01-1.0	Recommended range from Anderson and Woessner (2002)
S_y	Specific yield	0.1-0.3	Estimate from Ó Dochartaigh et al. (2019)
S_s	Specific storage	$1e^{-5}$ m^{-1}	Based on representative values for unconsolidated coarse sand aquifers (Domenico and Schwartz, 1990)
K_{riv}	River bed hydraulic conductivity	40 m d^{-1}	Average of surface permeability measurements in abandoned river channels (Ó Dochartaigh et al., 2019)
m	Rived bed thickness	0.5 m	Not known, but preliminary steady-state runs showed it to be insensitive
n	Manning's roughness coefficient	0.05	Based on 'normal' value for mountain streams with cobble/large boulder beds (Chow, 1959)
c_{drn}	Drain conductance	$2e^5$ m^2 d^{-1}	Set using the river bed parameterisation so that $c_{drn} = \Delta x \Delta y \frac{K_{riv}}{m}$

dynamics, could be simulated. Instead a small subsample of five ensemble members, hereafter referred to as scenarios, were selected to represent a range of potential future outcomes under twenty-first century climate change. While such an approach means that it is not possible to ascribe meaningful confidence intervals to projections like those in chapter 5, using a handful of ensemble members crucially allows one to undertake a more process-oriented analysis of differences between individual simulations and capture the nuanced feedbacks between climate, glacio-hydrology and groundwater that might be lost when averaging simulations over large ensembles (Knutti et al., 2010). This approach is particularly beneficial for evaluating the dynamics of unconfined aquifers and exchanges with surface water bodies given their known nonlinear response to climate stresses (Cayar and Kavvas, 2009).

Given the emphasis on a small number of scenarios, it was important to adopt a robust scenario selection procedure that captured a range of future glacio-hydrological and climatic behaviours from the future ensemble that are likely to drive changes in proglacial groundwater dynamics. Therefore, to aid the scenario selection, eleven signatures were calculated for each member of the future ensemble to characterise key climatic and glacio-hydrological behaviours over the mountain (Virkisá River basin) and alluvial aquifer (sandur) study regions (Table 6.2). It was decided to focus the projections on the

Table 6.2: Eleven signatures used to characterise the glacio-hydrological and climatic drivers of proglacial groundwater dynamics in the sandur aquifer for the 2080s (2073-2097).

Region	Group	Signature	ID	95% CI (2080s)
Mountain (Virkisá River basin)	Ice and snow coverage	Ice coverage	IceCov	4.09 - 11.28 km^2
		Snow coverage	SnowCov	7.99 - 15.7 km^2
	Seasonal river discharge	Winter river discharge	Q_{DJF}	1.52 - 4.53 $m^3 s^{-1}$
		Spring river discharge	Q_{MAM}	1.20 - 3.55 $m^3 s^{-1}$
		Summer river discharge	Q_{JJA}	4.29 - 9.25 $m^3 s^{-1}$
		Autumn river discharge	Q_{SON}	3.43 - 6.84 $m^3 s^{-1}$
	Quick-release high flows	1% exceedance flow	Q_{01}	15.56 - 32.4 $m^3 s^{-1}$
Alluvial aquifer (sandur)	Seasonal total precipitation	Winter total precipitation	P_{DJF}	3.04 - 6.04 mm d^{-1}
		Spring total precipitation	P_{MAM}	2.08 - 4.08 mm d^{-1}
		Summer total precipitation	P_{JJA}	2.26 - 5.09 mm d^{-1}
		Autumn total precipitation	P_{SON}	3.57 - 6.92 mm d^{-1}

time slice centred on the 2080s (2073-2097) given that, in chapter 5, this period showed the largest spread in projected evolution of climate and glacio-hydrological characteristics and therefore is most appropriate for evaluating the range of potential responses to climate change. The signatures include snow and ice coverage as these control runoff volumes and timing over seasonal to hourly timescales. The ice coverage also provides an indicator of long-term climate warming. Seasonal river discharge signatures were also included to characterise changes in runoff inputs throughout the year which are controlled by melt and rainfall-runoff in the Virkisá River basin. The 1% exceedance river flow signature was also included as episodic river recharge during high flow events could be an important recharge mechanism. Finally, total precipitation signatures over the sandur aquifer were calculated to characterise seasonal shifts in the timing and magnitude of diffuse recharge inputs.

In order to define the scenarios, k-means cluster analysis (MacQueen, 1967) was used. The k-means method can be used to find natural groupings in multivariate data that are distinct from one another and has been widely used in the past for climate projection scenario selection (Cannon, 2015; Wilcke and Bärring, 2016) and for undertaking signature-based analyses of watershed hydrological functioning (Jehn et al., 2019; Singh

et al., 2016a) and so was considered ideal for this application. For a given dataset $\mathbf{X}$ with n observations (ensemble members) and p variables (signatures):

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{6.8}$$

The aim is to partition the observations into k clusters $\mathbf{C} = (C_1, C_2, \ldots, C_k)$ so as to minimise the within-cluster variance:

$$\arg\min_{\mathbf{C}} \sum_{i=1}^{k} \sum_{\mathbf{x} \in C_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 \tag{6.9}$$

where $\boldsymbol{\mu}_i$ is the mean (centroid in the p-dimensional Euclidean space) of all observations in C_i.

The MATLAB® *kmeans* function was used to solve eq. 6.9, which uses an iterative approach to determine the optimal location of the centroids. Given the potential to reach local minima in the solution space during the optimisation procedure, 100 replicates were undertaken using random centroid locations initialised using the kmeans++ algorithm.

6.5 Results

The results are split into four sections. The first section covers the model calibration and evaluates the model simulations against the observed groundwater level time series data. The second section then analyses simulations of the main water balance variables from the calibrated model over the historical reference period (1991-2015) to evaluate the current hydrodynamic behaviour of the aquifer. The third section then analyses the driving river discharge and diffuse recharge data for the twenty-first century projections (2080s) before the final section analyses the corresponding projections of groundwater level dynamics and GW-SW exchanges with the Virkisá River and springs and how these relate to shifts in river discharge and diffuse recharge.

Note, the majority of simulated state variables in the results section are presented as a sequence of daily averages for the 365 days of the year to allow for straightforward

analysis of shifts in seasonal timing and magnitude of state variables. These intra-annual distribution plots are calculated after first applying a 30-day moving average (MA) filter to the raw simulation data which smooths out any inter-day variability that can make the plots difficult to interpret, but also allows one to analyse shifts in the timing of peaks and troughs in the simulations at a daily resolution.

6.5.1 Groundwater model calibration

In total, 4398 of the 5000 Monte Carlo calibration runs executed successfully. The remaining 602 runs did not converge. Figure 6.2a shows the frequency histogram and cumulative frequency diagram of the RMSE scores (averaged over the seven observation borehole time series) obtained from the successful runs. The least efficient model run obtained an RMSE of 1.89 m while the most efficient calibrated model obtained an RMSE more than three times smaller of 0.56. More than 90% of the runs returned a RMSE$<$0.9 m. Figures 6.2b-d show the range (light blue lines) and mean (dashed blue line) of the three calibration parameter values for all model runs under each point of the cumulative frequency diagram. It can be seen that as progressively fewer models are included (i.e. the RMSE threshold is reduced), K_h converges towards a value of approximately 4 m d^{-1} (Fig. 6.2b). The yellow line shows the convergence ratio: the proportion by which the calibration parameter range is reduced, for each point on the cumulative frequency diagram. For the the most efficient models (left-most point on the yellow line), the K_h parameter range is reduced by 93% of the original calibration range size indicating high sensitivity and identifiability. The calibrated K_h value was 3.3 m d^{-1}.

The convergence ratio only reaches 57% for κ (Fig. 6.2c) indicating that it is less identifiable than K_h. Even so, the parameter still converges towards a value $>$0.5 and the calibrated κ is 0.98 indicating an aquifer lithology that is approximately isotropic. The calibrated S_y is 0.15, but the convergence ratio only reaches 4% (Fig. 6.2d) indicating that this parameter is not identifiable. This is likely because inter-borehole groundwater level variability is much larger than variations in groundwater levels within individual boreholes as demonstrated by the analysis that follows in this section. Accordingly, the calibration of

hydraulic model parameters to all of the groundwater level time series data simultaneously likely renders S_y insensitive. To investigate this further, an additional side experiment was undertaken whereby the model was calibrated to each borehole individually. It was found that, except for the U1 and U2 boreholes, which are situated close to the river and therefore controlled primarily by river dynamics, S_y showed to be more sensitive with the convergence ratio ranging between 27 and 99% with a mean of 73% and the calibrated S_y ranging between 11 and 25% with a mean of 15%. Given that the S_y obtained when taking the average of the calibration values at all sensitive boreholes is in agreement with the value obtained from the initial Monte Carlo calibration and that this is in the representative porosity range (10 - 20%) for proglacial fans and sandur aquifers (Parriaux and Nicoud, 1990), an S_y of 0.15 was deemed justifiable.

Figures 6.2e-h show equivalent convergence plots for the mean simulated state variables as the RMSE threshold is reduced. All state variables appear to converge towards a calibrated state and indeed, the convergence ratio reaches $\geq$89% for all variables indicating the model behaviour is well constrained by the calibration procedure.

Table 6.3 shows the RMSE scores obtained at each borehole for the most efficient calibrated model over the calibration and evaluation periods. The relative efficiency of the model between the boreholes is consistent across the calibration and evaluation time periods where the model is most efficient at capturing the U1 groundwater level data and least efficient at capturing the U2 groundwater level data. There is a slight decrease in model efficiency between the calibration and evaluation periods where the average RMSE rises from 0.39 to 0.40.

Figure 6.3 shows that the model captures the average spatial distribution of groundwater levels accurately, obtaining an R^2 of 0.997. Temporally, the simulated groundwater level time series (yellow lines in Fig. 6.4) also capture the seasonal timing of peaks in the winter and spring months and troughs in the summer months shown in the observation borehole data. The timing of individual, event-scale peaks are also captured by the model. There are also, however, some deficiencies in the model simulations. For example, the simulated groundwater level dynamics at U2 show a systematic underesti-

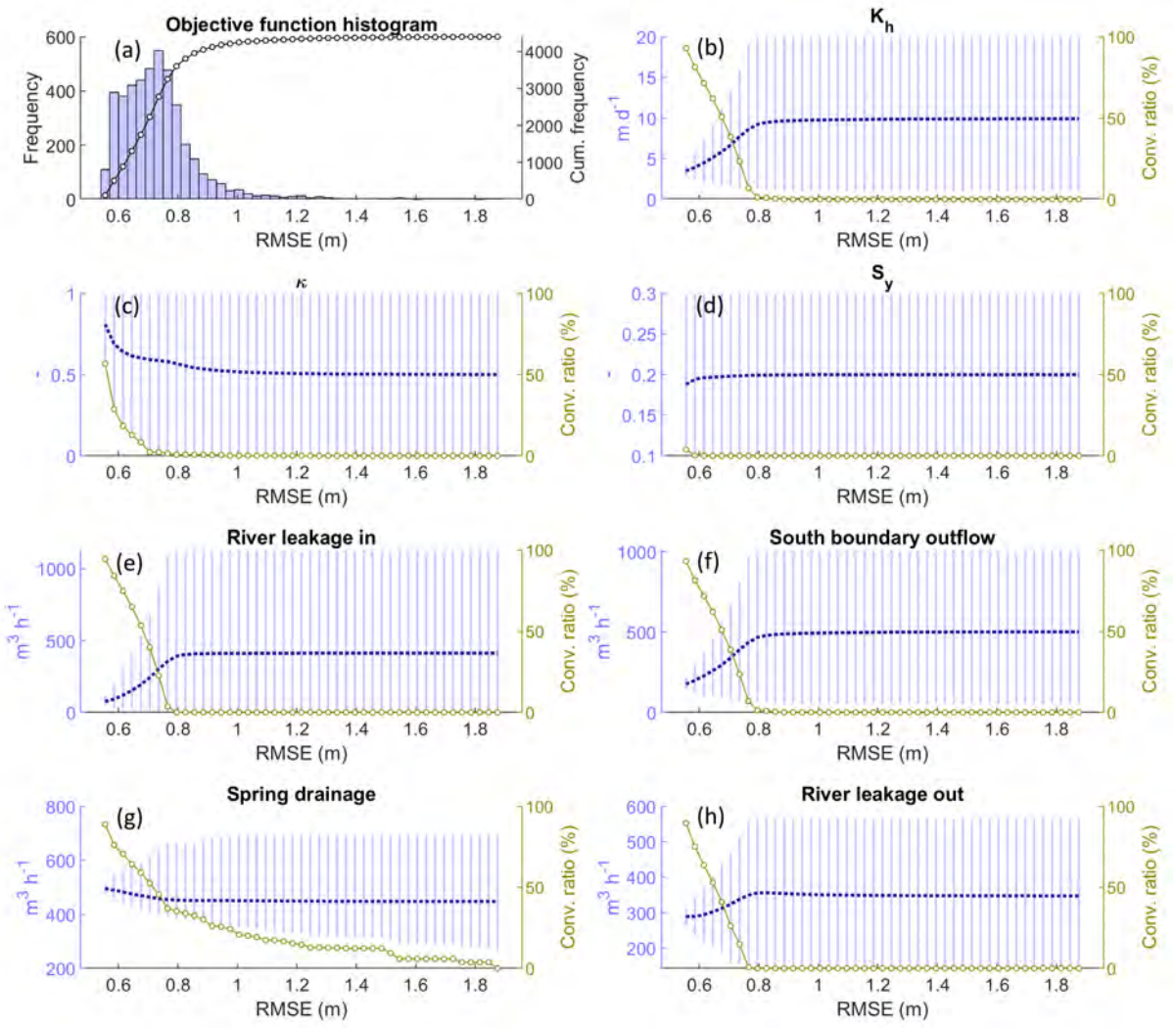

Figure 6.2: Frequency histogram (blue bars) and cumulative frequency diagram (black line) of RMSE scores obtained from successful Monte Carlo calibration runs (a). Also shown is the range (light blue lines), mean (dashed blue line) and convergence ratio (yellow line) of the calibration parameters (b-d) and model state variables (e-h) for all model runs under each point of the cumulative frequency diagram.

Table 6.3: Calibration and evaluation RMSE scores against observed groundwater level time series for the seven observation boreholes used for the model calibration.

Borehole ID	Calibration RMSE (m)	Evaluation RMSE (m)
U1	0.16	0.17
U2	0.65	0.69
M1	0.59	0.65
M2	0.44	0.38
L1	0.29	0.22
L2	0.23	0.22
L3	0.38	0.47

mation of levels throughout the year except during the winter months when levels are highest. This underestimation exceeds 2 m in the summer of 2016. The simulations at M2 show the opposite effect where groundwater levels outside of the winter months are typically overestimated while simulations at M1 and L3 show systematic underestimation and overestimation of groundwater levels respectively throughout the year. Simulated groundwater levels at L1 and L2 show a larger amplitude of response than shown in the observation data where the model overestimates maximum winter groundwater levels and underestimates summer minimum groundwater levels.

6.5.2 Historical reference period simulations

Average simulated annual diffuse recharge rates over the reference period are highest in the upper sandur where rates peak at 1700 mm y^{-1} (Fig. 6.5a). In the lower sandur, recharge rates are significantly lower with a minimum rate of 1100 mm y^{-1} on average. Figure 6.5b shows how the intra-annual distribution of diffuse recharge changes as you move downstream. Here, each line represents the average simulated diffuse recharge for a given 100 m long subdivision of the Virkisá groundwater catchment (Fig. 6.6) where yellow represents the upstream end and blue the downstream end of the groundwater catchment. As noted above, each point on the line is calculated using a 30-day moving average of the raw simulation data, averaged over all years, for each day of the year. In fact, the intra-annual timing of diffuse recharge is approximately identical across the

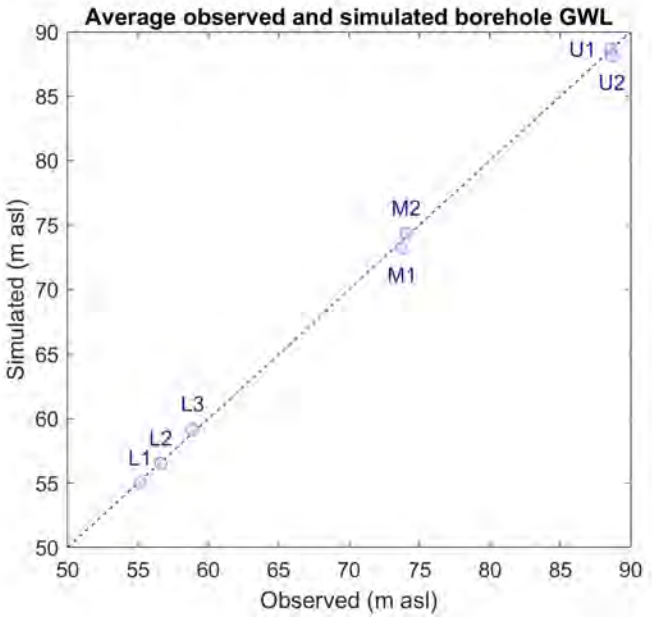

Figure 6.3: Comparison of mean observed and simulated groundwater levels at seven observation boreholes used for model calibration.

entire catchment where it is highest in February and between September and October and lowest in May.

More than 65% of the groundwater catchment has a simulated annual average DTGWL less than 1 m (Fig. 6.5c). The water table is deepest upstream where it can exceed 14 m on average. Within the top 1 km of the catchment (area shown by purple line in Fig. 6.6), the groundwater level fluctuates by more than a metre (yellow lines in Fig. 6.5d). Here, they are at their highest at the end of winter in February and reach a minimum at the end of summer between July and August. Downstream of this (green and blue lines in Fig. 6.5d), the amplitude of fluctuations are typically less than 0.3 m.

The map of river leakage shows that the upper 2 km of the river is net losing on average and the lower 2 km is net gaining on average (Fig. 6.5e) which broadly agrees with the observations made by Ó Dochartaigh et al. (2019). Even so, there is considerable variability in the magnitude and direction of fluxes, as you move downstream (Fig. 6.5f). Seasonally, leakage to groundwater typically peaks in the summer months when groundwater levels are at their lowest. Note, that in the sections furthest downstream (dark blue lines in Fig. 6.5f), the maximum leakage occurs earlier (approximately May) which

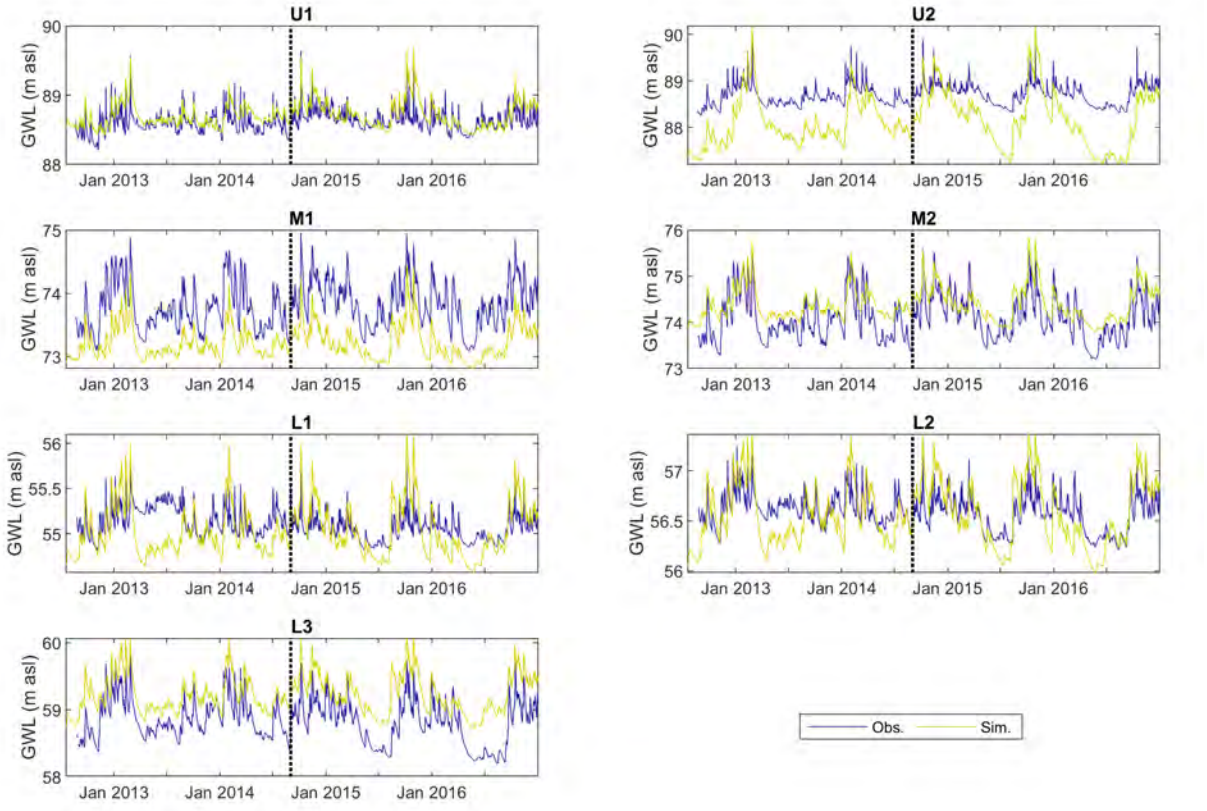

Figure 6.4: Observed (blue) and simulated (yellow) groundwater level time series at the seven observation boreholes. The dashed black line shows the dividing time between the calibration and evaluation periods to the left and right of the line respectively.

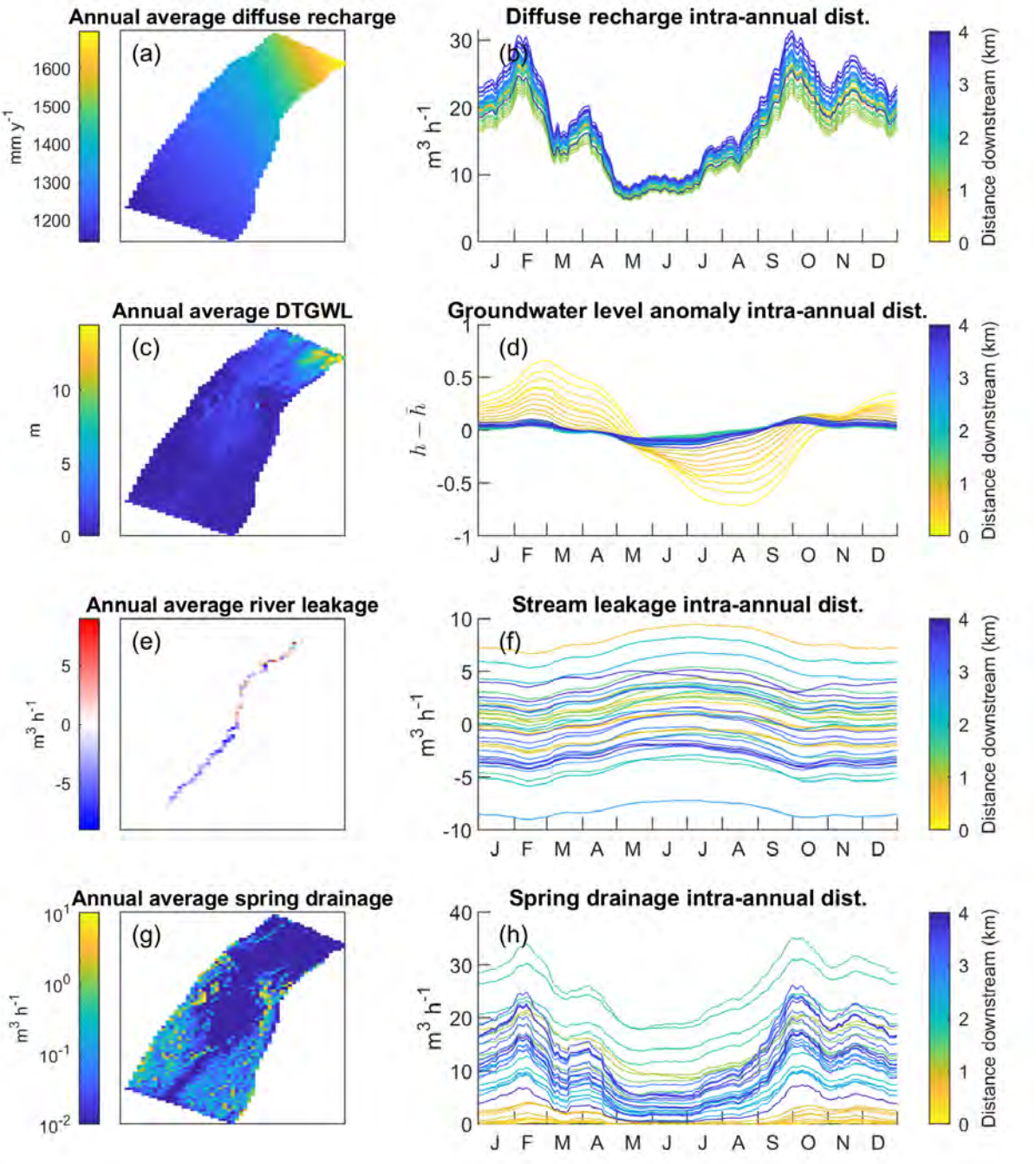

Figure 6.5: Annual average maps (a,c,e,g) and intra-annual distribution plots (b,d,f,h) of diffuse recharge, groundwater level, river leakage and spring drainage over the reference period (1991-2015) using average of simulations from TIM-ROR models in Table 6.4. Note in c, groundwater level is expressed as depth to groundwater level (DTGWL) and in d, groundwater level is expressed as an anomaly, i.e. the deviation from the mean groundwater level. Also note, in g spring drainage is on a log scale.

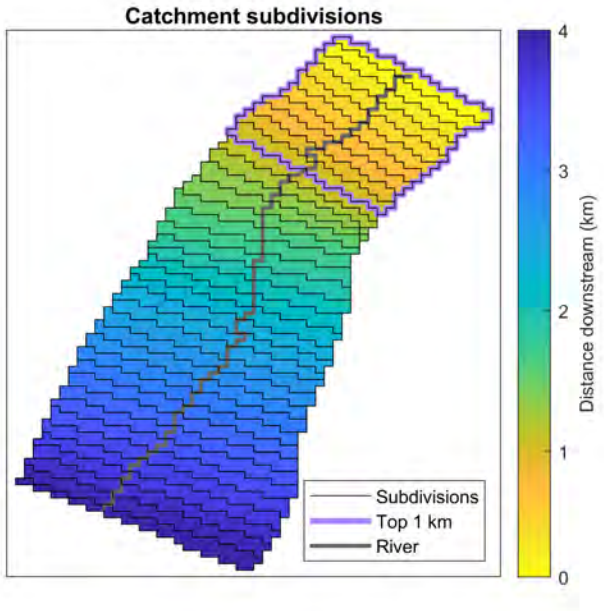

Figure 6.6: 100 m long subdivisions of the Virkisá groundwater catchment (shown by black lines) used to average simulated state variables for a given distance downstream. Yellow represents the upstream end and blue the downstream end. The purple line shows the top 1 km section of the groundwater catchment.

corresponds to an earlier minimum groundwater level (Fig. 6.5d) in the lower sandur.

Spring drainage is most concentrated downstream, particularly on those parts of the sandur that flank either side of the river. The greatest fluxes occur between 1 and 2 km downstream near the moraines (Fig. 6.5g). The seasonal timing of spring drainage closely follows that of diffuse recharge (Fig. 6.5h).

Based on stable isotope data of meltwater and groundwater (see section 2.5), MacDonald et al. (2016) estimated that groundwater contributes 15-20% of surface runoff in the lower sandur and the reference period simulations show close agreement with this where groundwater drainage to the river and springs make up to 14-17% of the 30-day MA sandur runoff.

6.5.3 Twenty-first century projections of river discharge and diffuse recharge

Scenario selection

Figure 6.7 shows the signatures of the five selected scenarios after undertaking k-means cluster analysis (coloured lines) and where they lie on the distribution of signatures from the future ensemble (grey bars). Each is coloured from orange to blue to indicate the relative degree of glacier and snow retreat where orange indicates high retreat and blue indicates low retreat. The selected scenarios cover between 38% (spring river flow) and 76% (summer river flow) of the ensemble with an average coverage of 61%. While the spread of the scenarios across the signatures is generally good, the extremes are not always represented. For example, for ice and snow coverage, the most extreme cases of retreat are not represented by the selected scenarios. Similarly for the Q_{DJF}, Q_{SON} and Q_{01} signatures, the highest flows are not represented, while for the Q_{MAM} signature neither the highest nor lowest ends of the distribution are represented.

Each scenario has been assigned a code with the format G*-Q*-P* where G, Q and P represent glacier coverage, mean annual river flow and mean annual sandur precipitation respectively and * is either 1, 2 or 3 representing low, moderate and high respectively.

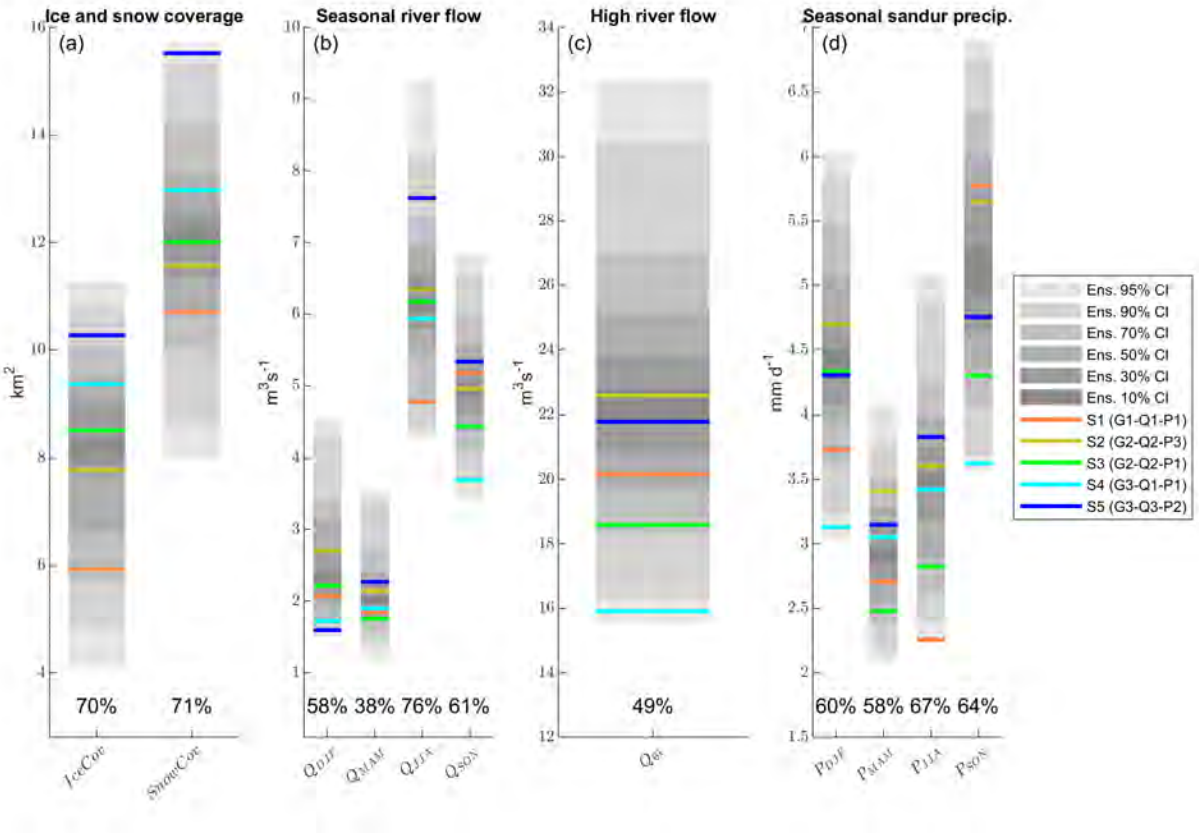

Figure 6.7: Signatures of five scenarios selected using k-means clustering algorithm including ice and snow coverage (a), seasonal river flow (b), high river flow (c) and seasonal sandur precipitation (d) signatures. Grey bars indicate the confidence intervals of the future ensemble from chapter 5. Percentages indicate the proportion of this distribution covered by the selected scenarios.

These bandings were defined based on which tercile of the future ensemble they fall into. Table 6.4 provides a summary of each scenario. Average near-surface air temperature rises between 0.7 and 3.1 °C across the scenarios with the warmest scenarios showing the highest degree of glacier retreat. The change in glacier coverage (ΔG) ranges from -3.7 km^2 for the coolest scenario to -8.1 km^2 for the warmest scenario. All scenarios show a reduction in mean river discharge (ΔQ) between the reference period and the 2080s and all apart from the wettest G2-Q2-P3 scenario show a reduction in total precipitation over the sandur relative to the reference period (ΔP). Both RCPs are included in the scenarios with the warmest scenarios associated with RCP8.5 and the cooler scenarios driven by RCP4.5. Each uses a unique GCM-RCM and all but scenarios 1 and 5 use unique TIM-ROR structure combinations. Each scenario will now be summarised:

G1-Q1-P1: This is one of the warmest scenarios with a mean temperature rise of 3.1 °C and has the lowest glacier and snow coverage of all of the scenarios. Accordingly, upstream river discharge is low with a mean reduction of 1.1 m^3 s^{-1}, particularly in the summer where melt would normally persist. It is also a relatively dry scenario, where sandur precipitation is low for most of the year except autumn.

G2-Q2-P3: This scenario is also one of the warmest, but only shows a moderate glacier retreat, likely due to high snowfall inputs. It is the wettest scenario and the only scenario that shows an increase in total precipitation over the sandur relative to the reference period. River flows are moderate.

G2-Q2-P1: This scenario shows an average warming of 1.9 °C resulting in moderate glacier coverage for the 2080s. River flows are also moderate, but precipitation inputs over the sandur are low throughout the year except in winter. This is a dry scenario.

G3-Q1-P1: This is one of the coldest scenarios with an average warming of only 1.0 °C resulting in high glacier coverage for the 2080s. It is also the driest scenario with

Table 6.4: Scenarios listed in descending order according to their projected glacier retreat for the 2080s. Each have been assigned a code with the format G*-Q*-P* where G, Q and P represent glacier coverage, mean annual river discharge and mean annual sandur precipitation respectively and * is either 1, 2 or 3 representing low, moderate or high respectively. Also shown are the changes in near-surface air temperature (ΔT), glacier coverage (ΔG), upstream river discharge (ΔQ) and total precipitation over the sandur (ΔP).

Scenario	Scenario code	ΔT (°C)	ΔG (km^2)	ΔQ (m^3 s^{-1})	ΔP (mm d^{-1})	RCP	GCM-RCM	TIM-ROR
1	G1-Q1-P1	3.1	-8.1	-1.1	-0.6	8.5	[HadGEM2-ES]-[CCLM4-8-17]	TIM_3-ROR_1
2	G2-Q2-P3	3.1	-6.4	-0.7	0.2	8.5	[EC-EARTH]-[RACMO22E]	TIM_3-ROR_2
3	G2-Q2-P1	1.9	-5.6	-0.9	-0.7	4.5	[HadGEM2-ES]-[RCA4]	TIM_2-ROR_1
4	G3-Q1-P1	1.0	-4.8	-1.2	-0.9	4.5	[NorESM1-M]-[HIRHAM5]	TIM_1-ROR_2
5	G3-Q3-P2	0.7	-3.7	-0.3	-0.2	4.5	[CNRM-CM5]-[CCLM4-8-17]	TIM_3-ROR_1

a mean reduction in total precipitation over the sandur of 0.9 mm d^{-1} and also has the lowest mean upstream river discharge input.

G3-Q3-P2: This is the coldest scenario with an average warming of only 0.7 °C resulting in high glacier coverage for the 2080s. This scenario shows the highest summer river flows and also receives a moderate precipitation input over the sandur.

River discharge projections

Figure 6.8a shows the intra-annual distribution of river discharge out of the Virkisá River basin over the reference period while Figs. 6.8b-d show the contribution from ice melt, snow melt and rainfall. Over the reference period, ice melt is the dominant source of river runoff and this input in conjunction with snow melt has the major control on river flow seasonality where the maximum 30-day MA flow (as indicated by the upwards triangles in Fig. 6.8d) occurs in August. There are small differences in simulated runoff over the reference period. The G1-Q1-P1 (orange) and G3-Q1-P1 (cyan) scenarios show the highest and lowest river discharges respectively over the rising limb of the intra-annual distribution hydrograph between April and July and the opposite over the falling limb

between August and November. These clearly stem from differences in simulated ice and snow melt. Rainfall runoff reaches a maximum between September and October and is the dominant runoff input outside of the melt season.

Projections of river discharge for the 2080s (Fig. 6.8e) and changes relative to the reference period (Fig. 6.8i) show large reductions in melt-season runoff of up to 5 m^3 s^{-1} for the warmest G1-Q1-P1 scenario which is mainly due to reductions in ice melt (Fig. 6.8j). The reduced influence of melt and increased influence of rainfall runoff on the intra-annual distribution of river discharge is demonstrated by the loss in smoothness of the hydrographs (Fig. 6.8e). For the warmest G1-Q1-P1 scenario, a significant increase in October rainfall results in a 51-day shift in the timing of maximum river discharge from August to October. For all of the scenarios, rainfall is projected to increase outside of the melt season between December and May although the two coldest scenarios show a small reduction in December and January rainfall (Figs. 6.8l). These increases are reflected in the projected changes in river discharge (Figs. 6.8i).

Diffuse recharge projections

Projections of month-to-month variability in diffuse recharge over the 2080s (Fig. 6.9e) show significant changes relative to the reference period (Fig. 6.9a). The intra-annual peaks typically become more pronounced in the future (e.g. in February, March/April, and November/December) while the diffuse recharge is consistently projected to decrease for the autumn and winter months across the scenarios. The projected changes in diffuse recharge (Fig. 6.9i) are clearly related to the relative 'wetness' of the scenarios with the G3-Q1-P1 dry scenario (cyan) demonstrating some of the largest reductions in recharge between September and February and the G1-Q1-P1 (orange) scenario showing the greatest reductions in melt-season runoff between May and September. Furthermore, the intra-annual variability of diffuse recharge plots (Figs. 6.9a, e and i) is mirrored in the corresponding rainfall plots (Figs. 6.9b, f and j). Over the reference period, rainfall makes up 94% of the total diffuse recharge and this increases to between 96 and 99% across the scenarios. Diffuse recharge by snow melt, thus, only constitutes a relatively

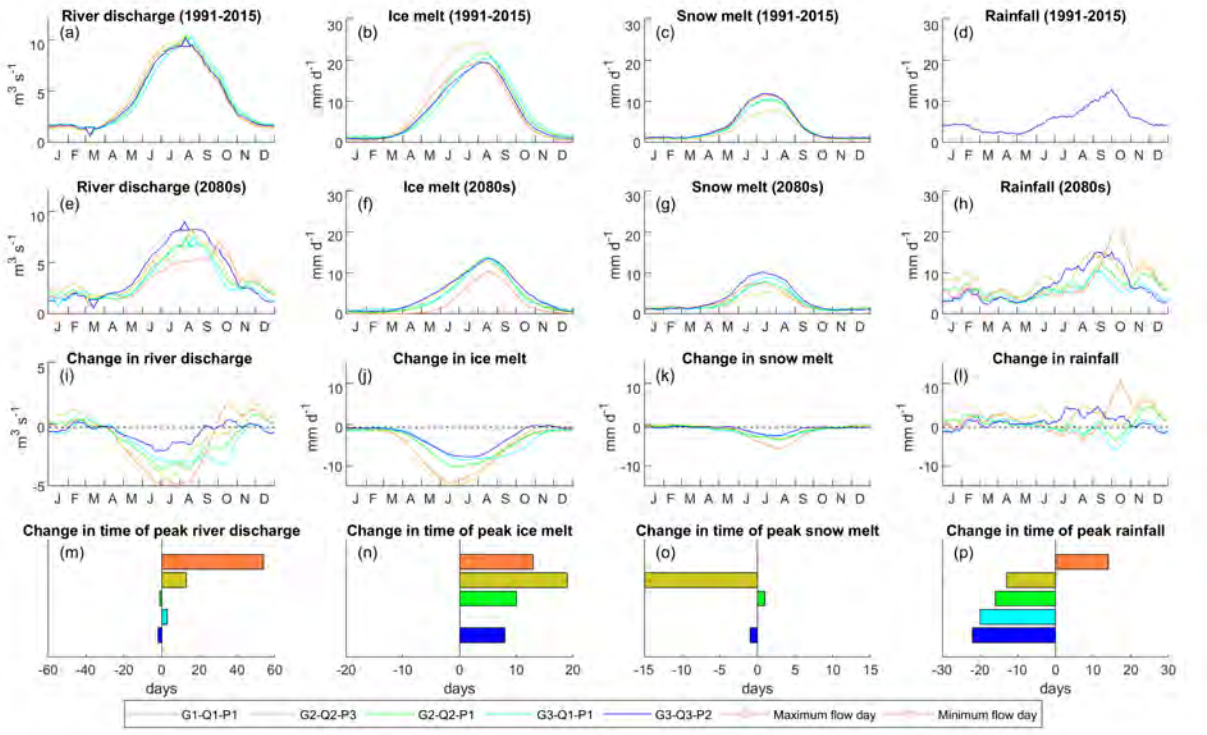

Figure 6.8: Intra-annual distribution plots of river discharge, ice melt, snow melt and rainfall for the Virkisá River basin over the reference (1991-2015) period (a-d); future (2080s) period (e-h) and the difference between the two periods (i-l). Also shown is the change in time of maximum between the reference and future periods (m-p).

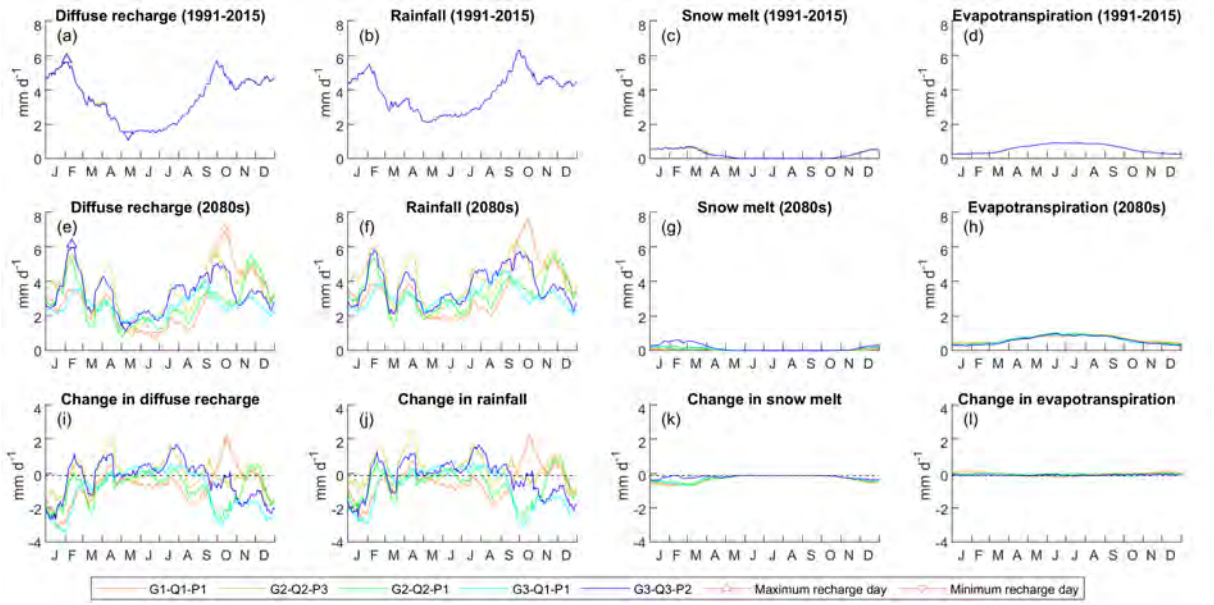

Figure 6.9: Intra-annual distribution plots of diffuse recharge, rainfall, snow melt and evapotranspiration for the Virkisá groundwater catchment over the reference (1991-2015) period (a-d); future (2080s) period (e-h) and the difference between the two periods (i-l).

minor source of groundwater recharge.

During the reference period, evapotranspiration removes as much as 35% of daily rainfall and snow melt inputs to the soil (Fig. 6.9d). However, these fluxes do not change significantly in the future across the different scenarios 6.9l).

6.5.4 Twenty-first century projections of groundwater level dynamics and GW-SW exchanges

Groundwater level dynamics

To evaluate shifts in groundwater level dynamics, average groundwater level, diffuse recharge and river leakage inputs have been calculated for the top 1 km section of sandur (see purple box in Fig. 6.6) where groundwater level dynamics are not constrained by the surface topography (Fig. 6.10). All scenarios except for the wettest G2-Q2-P3 scenario (yellow) project winter and spring groundwater levels to decrease (Fig. 6.10g). The wettest scenarios (yellow and blue) show a small increase in 30-day MA groundwater

level between July and November of up to 0.3 m. The three driest scenarios (orange, green and cyan) show the greatest reductions in groundwater levels. The dry G3-Q1-P1 scenario (cyan) projects 30-day MA groundwater level to reduce by 1 m in February.

Over the reference period, river leakage inputs peak in the summer months (Fig. 6.10c) where 30-day MA contributions to total recharge can reach between 46-50% in the top 1 km of the sandur. Overall river leakage contributes between 21-23% of total recharge in this area. Maximum river leakage inputs for the G2-Q2-P3 scenario (yellow) are approximately 12 m^3 h^{-1} higher than the other scenarios. This difference will be investigated in the next section. All simulations for the 2080s project that river leakage fluxes will increase between November and March (Fig. 6.10i). The wettest scenarios (yellow and blue) project a decrease in spring and summer leakage while the driest scenarios (orange, green and cyan) show an increase in river leakage for the majority of the year. The pattern of river leakage changes closely follows the inverse of changes in groundwater levels across the scenarios.

Significant shifts in the timing of minimum and maximum groundwater levels in the 2080s (Fig. 6.10d) relative to the reference period (Fig. 6.10a) are shown for some scenarios. For the G1-Q1-P1 scenario (orange), a sharp rise in groundwater level in October causes the peak to occur much earlier in December, rather than in February. This corresponds to a significant increase in October diffuse recharge (Fig. 6.10e). The G3-Q1-P1 scenario (cyan) shows an even earlier maximum in groundwater levels in September for the 2080s which can be attributed to a reduction in diffuse recharge during the main recharge season between September and February (Fig. 6.10h). Changes to the timing of minimum groundwater level are much smaller with four of the five scenarios showing an earlier minimum day (up to 23 days earlier), and the G1-Q1-P1 scenario showing a minimum 23 days later in the 2080s.

River leakage dynamics

Figure 6.11 shows intra-annual distribution plots of various aspects of river leakage dynamics over the whole Virkisá groundwater catchment. The seasonal pattern of river

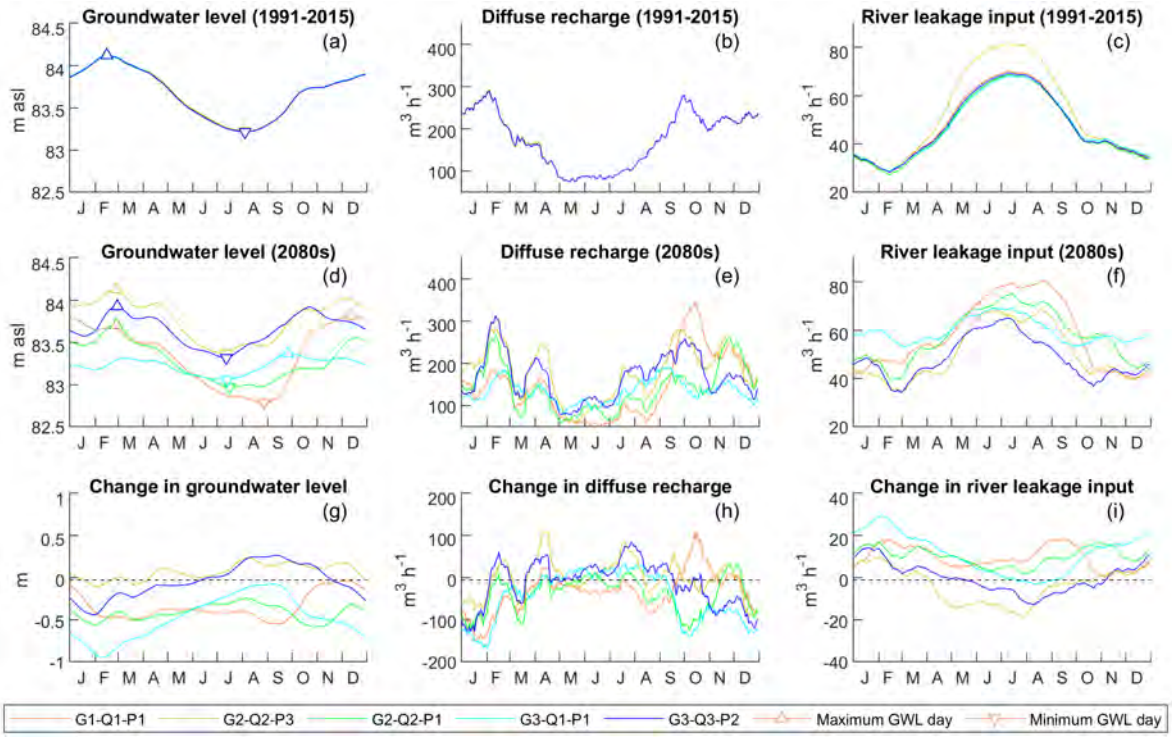

Figure 6.10: Intra-annual distribution plots of groundwater level, diffuse recharge and river leakage input averaged over the top 1 km section of sandur (see purple box in Fig. 6.6) for the reference (1991-2015) period (a-c); future (2080s) period (d-f) and the difference between the two periods (g-i).

leakage over the reference (Fig. 6.11a) and future (6.11e) periods, and the projected change between the two (Fig. 6.11i) broadly follow the patterns observed in the upper sandur (Fig. 6.10c, f and i) where leakage is projected to increase in the winter months while the summer months show a mixture of responses across the scenarios. A comparison of the projected river leakage input plots (Figs. 6.11a, e and i) with the corresponding length of leaking river plots (Fig. 6.11b, f and j) show an almost like-for-like relationship indicating that it is the seasonal evolution of the proportion of river that is losing that drives variability in river leakage. However, some aspects of leakage behaviour do not correlate to the length of leaking river. In particular, the maximum river leakage inputs for the G2-Q2-P3 scenario (yellow) over the reference period (Fig. 6.11a) are approximately 54 m^3 h^{-1} higher than the other scenarios (an observation also made from Fig. 6.10), but this is not shown in the leaking river length data. Instead, it is the rate of leakage at leaking river nodes which is much higher for the G2-Q2-P3 scenario as shown by the plot of specific river leakage i.e. the river leakage rate per unit area of leaking river (Fig. 6.11c).

To investigate this in more detail, Figs. 6.12a and b show the simulated river stage and groundwater level respectively over a three-day period during the melt season in July 2013 at a single river node 1 km downstream for all scenarios. The diurnal melt signal is clear in all time series, but the amplitude of the cycle is highly variable across the scenarios and is highest for the G2-Q2-P3 model run. Furthermore, Fig. 6.12c shows that downward head gradient is proportional to the amplitude of diurnal flow. Accordingly, the high amplitude of the diurnal melt signal for the G2-Q2-P3 model run induces the highest leakage rates from the river to the water table (Fig. 6.12d). The intra-annual distribution of diurnal flow amplitude for each scenario correlates well with the specific leakage rate (Fig. 6.11d). Note, it is the loss of this diurnal flow amplitude for the G2-Q2-P3 scenario during the melt season (Fig. 6.12l) that also explains the large reduction in projected river leakage for the 2080s (Fig. 6.12i).

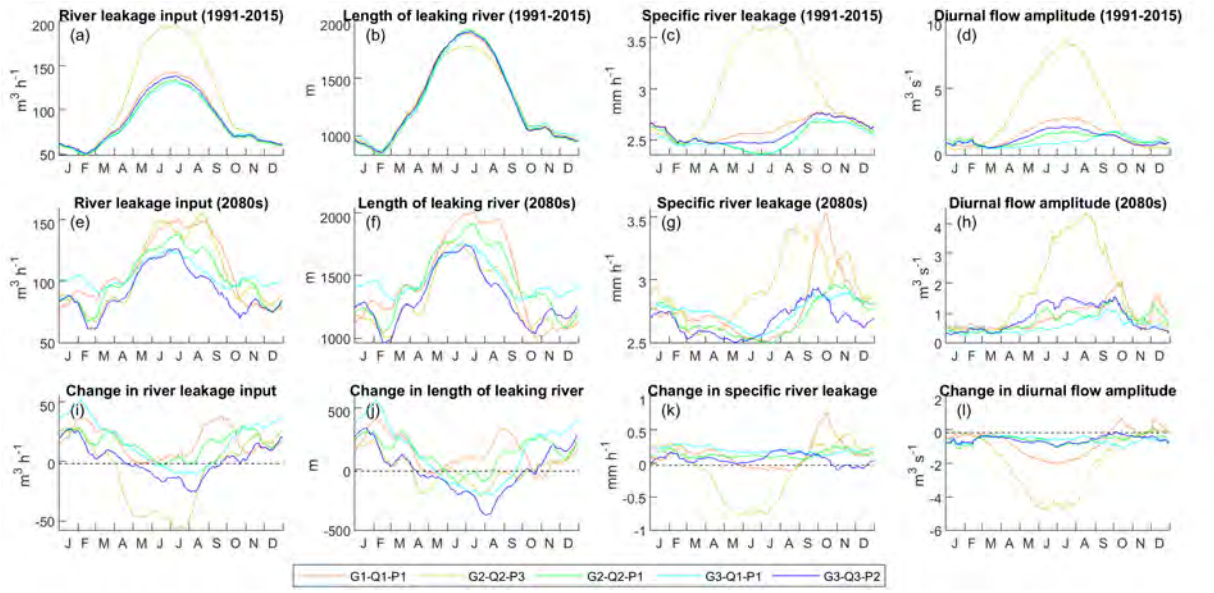

Figure 6.11: Intra-annual distribution plots of river leakage input, length of leaking river section, specific river leakage (river leakage per unit area of leaking river) and the diurnal flow amplitude for all days where ≥90% of runoff is from melt for the reference (1991-2015) period (a-d); future (2080s) period (e-h) and the difference between the two periods (i-l).

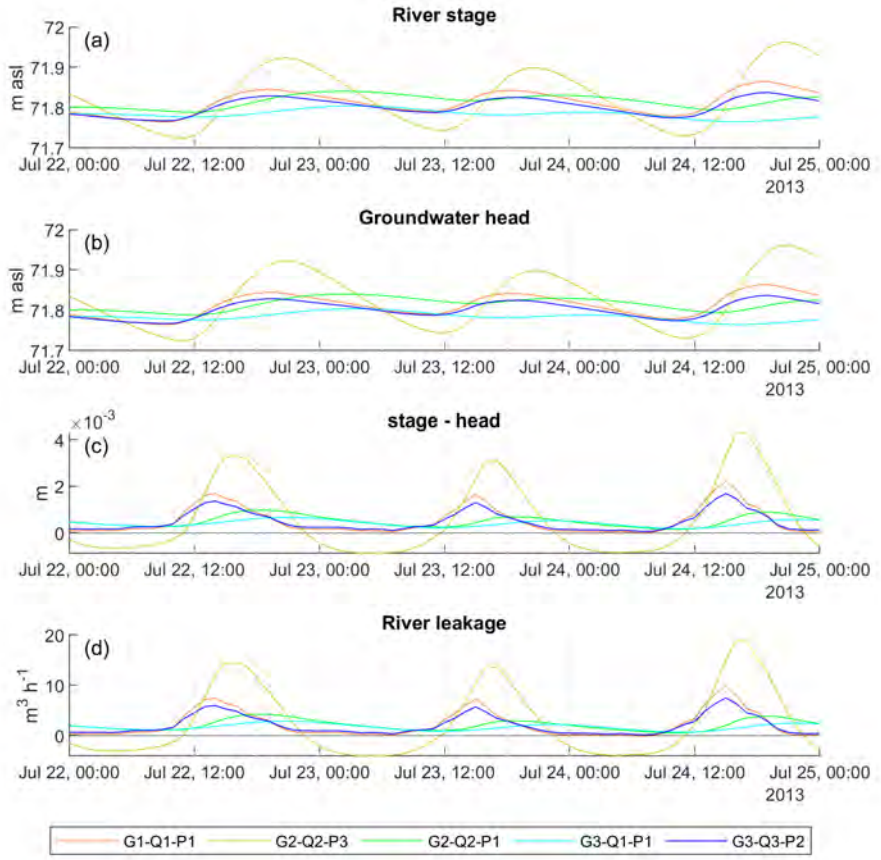

Figure 6.12: Simulated river stage (a); groundwater head (b); stage minus head (c); and river leakage (d) time series for a single river node 1 km downstream for three days during the reference period (July 2013).

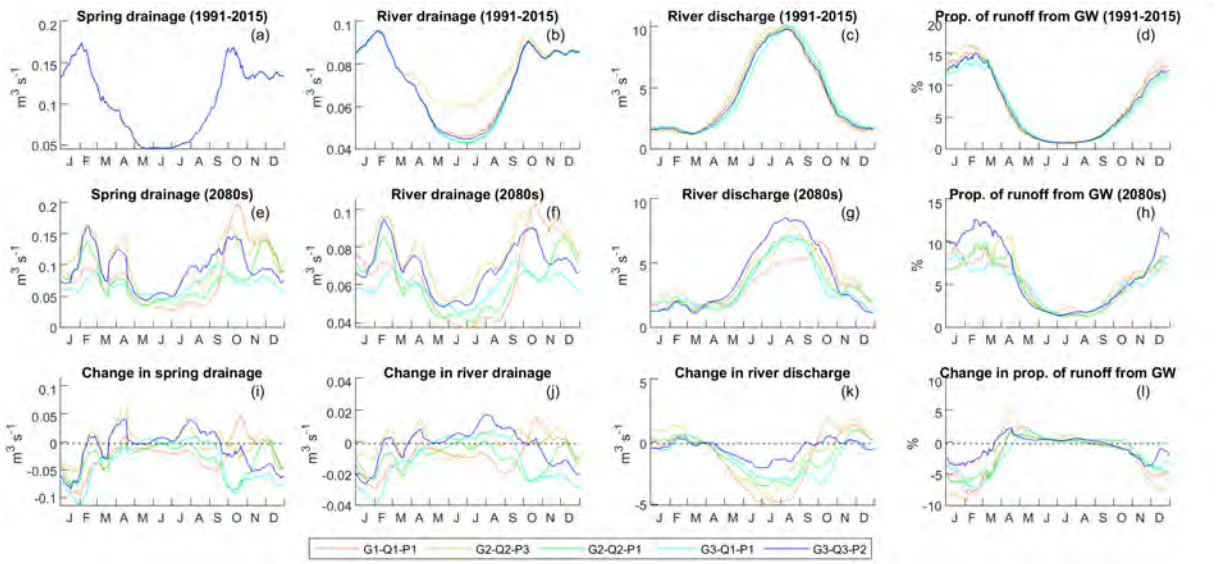

Figure 6.13: Intra-annual distribution plots of surface drainage to springs and the river as well as river discharge from the Virkisá River basin and the proportion of surface runoff from groundwater for the reference (1991-2015) period (a-d); future (2080s) period (e-h) and the difference between the two periods (i-l).

Surface drainage dynamics

Over the reference period, the seasonal patterns of groundwater drainage to springs (Fig. 6.13a) and the river (Fig. 6.13b) are closely aligned with the corresponding pattern of diffuse recharge (Fig. 6.9e) where greatest drainage fluxes occur outside of the melt season between September and March. Note, the G2-Q2-P3 scenario shows elevated drainage to the river over the reference period (Fig. 6.13b) which can be attributed to larger diurnal flow amplitude (Fig. 6.12).

Figure 6.13d shows that over the reference period, spring and river drainage makes up to 14-17% of the total sandur runoff in February and March when drainage is at its highest and river discharge from the Virkisá River basin (Fig. 6.13c) is relatively low. During the melt season, groundwater drainage makes up approximately 1% of surface runoff runoff only.

The projections of spring and river drainage for the 2080s (Figs. 6.13e and f) show seasonal patterns across the scenarios that are closely aligned with the corresponding

projections of diffuse recharge (Fig. 6.9e). There shows to be a consistent reduction in spring and river drainage in the winter months across the scenarios (Figs. 6.13i and j), which, coupled with projected higher winter river discharge (Fig. 6.13k) results in a decreased proportion of winter runoff from groundwater (Fig. 6.13l). For the three warmest scenarios, this proportion falls by up to 9.2%.

Outside of the winter months, significant increases in surface drainage, particularly in April, July and August broadly correspond to those periods where diffuse recharge is projected to increase (Fig. 6.9o), particularly for the wettest G2-Q2-P3 (yellow) G3-Q3-P2 (blue) scenarios. The G1-Q1-P1 scenario (orange) which has the driest projections of summer recharge shows a consistent decrease in summer surface drainage of groundwater. Despite the mixed surface drainage response to climate change outside of winter, all scenarios project an increase in proportion of runoff from groundwater for May through to July (Fig. 6.13l) due to the projected decrease in summer upstream runoff across the scenarios (Fig. 6.13k). It should be noted, however, that the decrease in upstream runoff can be as much as 5 m^3 s^{-1} while the increase from GW spring drainage never exceeds 0.1 m^3 s^{-1} across the scenarios.

6.6 Discussion

6.6.1 Projected changes in proglacial groundwater level dynamics and GW-SW exchanges: are they recharge- or stream-driven?

Changes in groundwater level dynamics over the twenty-first century will be most pronounced in the upper-sandur (top 1 km) region of the Virkisá groundwater catchment where the water table resides up to 14 m below ground level and is therefore free to fluctuate. Four of the five scenarios projected upper-sandur winter (DJF) and spring (MAM) groundwater levels to decrease on average by the 2080s while the three driest scenarios (i.e. lowest sandur precipitation inputs) projected upper-sandur groundwater levels to

fall year-round with a maximum reduction of 30-day MA groundwater levels of 1 m. The two wettest scenarios projected increases in 30-day MA groundwater levels of up to 0.3 m over the summer (JJA) and autumn (SON) months.

The evolution of intra-annual groundwater level in the Virkisá groundwater catchment was dominated by the driving diffuse recharge signal over the sandur. Drier scenarios consistently projected lower groundwater levels in the 2080s while shifts in the timing of maximum groundwater level were controlled by changes in the magnitude of diffuse recharge during the main recharge season. Increases in the early recharge season (SON) and/or decreases in the late recharge season (DJF) shifted the timing of maximum groundwater level forward. The trajectory of future groundwater storage in recharge-driven systems is of course dependent on local trends in climate and the presiding hydrological regime. Sridhar et al. (2018) projected average water table elevations to increase over the twenty-first century in the Snake River basin, United States due to projected increases in precipitation. Okkonen and Kløve (2011) projected winter groundwater levels to increase in an esker aquifer in Finland as a result of increased diffuse recharge from snow melt under a warming climate. For the Virkisá groundwater catchment, enhanced snow melt recharge due to climate warming should be expected to have negligible influence on regional groundwater levels given that snow melt only contributes 6% of total diffuse recharge. A commonality between these recharge-driven aquifers and the Virkisá sandur aquifer is that they are all 'wet' systems. In fact, the sandur receives between 83 and 89% of total recharge from diffuse rainfall and snow melt sources. It is also notable that those studies which have emphasised the sensitivity of future groundwater level dynamics to changes in mountain-fed river runoff (Allen et al., 2004; Huntington and Niswonger, 2012; Scibek et al., 2007) have all focussed on semi-arid regions of North America.

Changes in recharge-driven groundwater level dynamics also showed to control future variations in GW-SW exchanges. The relatively flat topography of the sandur means that even small variations in recharge-driven groundwater level can significantly alter the location at which the water table intersects the land surface and therefore the volume of water draining to the surface and the length of river that is actively leaking to the water

table. Groundwater is projected to contribute proportionally less to winter sandur runoff in the 2080s (3-8% reduction across the scenarios), partly due to projected decreases in groundwater drainage to the surface as a consequence of declining groundwater levels. These results suggest that the role of groundwater in buffering low flow periods outside of the melt season will lessen in the future. During the melt season, the proportional contribution of groundwater drainage to sandur runoff is only projected to rise by 1 or 2% by the 2080s and this is primarily due to decreased melt season mountain runoff to the river indicating that the potential to buffer losses in melt season runoff is negligible.

The simulations from this study indicated that leakage from the Virkisá River currently contributes up to 50% of 30-day MA recharge in the summer and as a consequence of falling groundwater levels in the future and the simultaneous extension of the leaking river length, this contribution could increase. Meltwater-derived river recharge is therefore likely to become more crucial to the sustainability of groundwater storage in the sandur. It should be noted, however, that river leakage only made up for between 7 and 15% of the loss in diffuse recharge across the dry scenarios. Importantly though, the simulations also indicate that the extent to which meltwater-derived runoff might be able to buffer reductions in groundwater levels in proglacial aquifers will depend somewhat on the degree of glacier and snow retreat that has taken place. More specifically, the simulations have shown that the daily rise of river flow driven by the melt of snow and ice induces a downward head gradient at the river-aquifer interface, driving river leakage inputs to the aquifer. A comparison of the simulations over the reference period in this study effectively served as a sensitivity analysis of this aspect of river recharge behaviour (fixed climate and variable melt runoff). Here, the scenario with the largest daily river discharge amplitude due to melting increased river recharge fluxes by up to 53%. To the author's knowledge, the sensitivity of recharge fluxes from mountain-fed rivers to the daily amplitude of flows has not been shown before, but such feedbacks could be important, particularly for aquifers which are more reliant on river recharge e.g. in Iceland (Levy et al., 2015), Alaska (Liljedahl et al., 2017) and Motana, US (Covino and McGlynn, 2007).

6.6.2 Limitations

It is important to highlight potential deficiencies in transferring results from this study to other glaciated catchments. The importance of local climatic factors and hydrological regime have already been noted, but an additional factor to consider when assessing the relevance of the reported magnitude and dynamical behaviour of simulated state variables and fluxes is scale. The section of sandur evaluated in this study is 4 km long (approximately one fifth of the total distance to the ocean). If the groundwater catchment were extended further downstream, the proportion of sandur runoff from groundwater would of course increase. Indeed, other studies have looked at larger scale glaciated groundwater catchments such as Andermann et al. (2012), who studied the contribution of groundwater to glacier-fed rivers over twelve river basins in the Nepal Himalayas covering almost 185,000 km^2. They found much more significant buffering of river discharge from groundwater where snow and ice melt only contributed 10% of river flow. In the studied Virkisá groundwater catchment, groundwater only makes up 7% of total runoff on average while snow and ice melt currently contribute almost 70% of runoff inputs from the Virkisá River basin (refer back to section 5.5.5).

It is also important to consider potential deficiencies in the groundwater model. The model is efficient at capturing the spatial variability of groundwater levels across the seven observation boreholes. However, when analysed against individual borehole time series, a number of model deficiencies were noted including systematic biases in the mean and amplitude of variability of the groundwater level data. Even so, that the RMSE calculated over all observation boreholes was only 0.39 and 0.40 over the calibration and evaluation periods respectively indicating a high degree of accuracy. However, this accuracy should be interpreted with caution given the shallow water table and shallow topographic gradient of the land surface which mean that even small simulation errors could induce relatively large errors in GW-SW flux simulations.

Deficiencies in the simulations partly stem from the assumptions and simplifications employed in the construction of the groundwater model. One key aspect of the model

that was simplified was the surface drainage network. In reality, it is known to consist of a complex interlinking network of groundwater springs and braided meltwater river channels. Furthermore, the morphology of these channels are poorly constrained by the unconsolidated sandur material over which they flow. In fact, a study of proglacial geomorphology on the sandur at the neighbouring Skaftafellsjökull showed major changes to the proglacial drainage network both in terms channel morphology and degree of braiding in response to sustained glacier retreat between 1996 and 2011 (Marren and Toomath, 2013) due to changes in meltwater runoff and proglacial lake development which serve to alter channel incision and aggradation rates. Major channel migrations are also known to occur during and after jökulhlaup events (Marren, 2005) which occur every 20 years on average (Robinson et al., 2008). The inclusion of a landscape evolution component in the modelling of the system was deemed beyond the scope of this study and infeasible given the additional data required to constrain such a model. Accordingly, no attempt was made to explicitly include the connectivity of springs and meltwater channels or the dynamic evolution of these over time, but both of these would undoubtedly influence the flow dynamics in the Virkisá River, and accordingly, the direction and magnitude of GW-SW fluxes significantly.

The parameterisation of the model was also simplified by assuming a homogeneous distribution of aquifer permeability and porosity, the justification for which was given in section 6.4.4. However, when calibrated to individual boreholes, the range of S_y was 0.11 - 0.25 and the range of K_h was 1.1 - 19.2 m d^{-1} which indicated heterogeneity is likely to be high. In fact, the K_h value calibrated for the groundwater model (3.3 m d^{-1}) was an order of magnitude smaller than the surface permeability measurements collected by Ó Dochartaigh et al. (2019). Past seismic analysis of Skeiðarársandur indicates that overburden pressure from past glaciations could result in compaction with depth (Guðmundsson et al., 2002) which would result in a reduction of K_h. However, in their analysis Ó Dochartaigh et al. (2019) suggest that these surface permeability measurements are indicative of the permeability at all depths, although there are no observation data to validate this. If this is the case, it could be an indication that recharge inputs to the aquifer are currently being

underestimated. One potential source of recharge that was not included in this study was recharge via lateral flows draining through the mountain bedrock which has shown to be an important or even dominant source of recharge in other studies (Frisbee et al., 2017; Manning and Solomon, 2003; Voeckler et al., 2014; Wilson and Guan, 2004; Yao et al., 2017). Even so, as noted, pumping tests conducted in the bedrock material suggest the permeability of the bedrock is negligible.

As well as the limitations and potential sources of error from the groundwater model, one must also consider limitations in the driving data. In particular, it should be noted that only five future diffuse recharge and river discharge scenarios were used out of a total ensemble size of 94800. The spread of projections therefore do not represent the full range of uncertainties stemming from the future climate scenarios and GHM compositions implemented in chapter 5. The sensitivity of the simulations to diffuse recharge over the sandur indicates that accounting for the climate uncertainty fully should be a priority for 'wet' catchments like this, however the sensitivity of river leakage to the diurnal river discharge amplitude suggests that the inclusion GHM uncertainty could also be important. It was noted that the five selected scenarios did not cover the extremes of the distributions of the hydro-climatic signatures used to select the scenarios, with a tendency to miss the highest winter and autumn river discharge scenarios. Additionally, none of the scenarios covered the highest precipitation inputs which could explain the tendency for projections of lower winter groundwater levels, leading to increased river leakage and reduced baseflow to the river. It should therefore be stressed that the individual scenarios should not be interpreted as the most likely outcomes for this catchment.

6.7 Conclusions

Across the five analysed scenarios, groundwater levels are consistently projected to decrease in the winter and spring months of the 2080s, while the summer and autumn months show a spread of increase and decreases in groundwater level in proglacial sandur aquifer. The long-term evolution of groundwater level and intra-annual timing of

maximum and minimum groundwater levels is controlled primarily by changes in diffuse recharge over the sandur rather than future river stage dynamics in the Virkisá River. Under the classification of Allen et al. (2010), this aquifer is recharge-driven.

The shallow topography of the Virkisá groundwater catchment makes future changes in GW-SW exchanges extremely sensitive to changes in recharge-driven groundwater levels. Even small variations in groundwater level can drastically alter the volume of spring drainage and the length of river that is actively leaking to the water table. Future changes in GW-SW exchanges are, therefore, also controlled by the evolution of recharge-driven groundwater level variations.

If, as projected by the five scenarios, spring and winter groundwater levels fall in the future, the role of groundwater in buffering low flow periods outside of the melt season will also lessen. Furthermore, the relatively modest contribution of groundwater to summer river discharge means that the potential of groundwater to buffer losses in melt season runoff due to glacier and snow retreat will be negligible. The results also indicate that in response to any future reduction in groundwater level, additional river recharge could serve to buffer up to 15% of reduced diffuse recharge in the driest future scenarios. Importantly, though, the volume of river recharge to groundwater is largest when melt-induced diurnal cycles in runoff are largest. As such, the results from this study indicate that the ability of river recharge to buffer reductions in groundwater levels in proglacial aquifers will depend on the degree of glacier and snow retreat that has taken place. Future retreat of snow and ice and the subsequent changes in diurnal river flow variability could, therefore, have some control over the future evolution of proglacial aquifers in Iceland and in other glaciated catchments around the world.

6.8 Summary

This chapter aimed to meet the fourth and final research objective of this thesis which was to, "incorporate a distributed groundwater model into the climate-GHM model chain to project twenty-first century changes in proglacial groundwater level dynamics and GW-

SW exchanges". This was done by utilising a subsample of the climate and river flow projections generated in chapter 5 to drive a distributed MODFLOW-NWT groundwater model of the proglacial sandur. The groundwater model projections were used to evaluate the impacts of climate change, glacier and snow retreat and river flow regime change on proglacial groundwater level dynamics and GW-SW exchanges in the 2080s and to determine the relative importance of climate-driven diffuse recharge and mountain-driven river runoff in controlling this future evolution. The next chapter will conclude the research undertaken in this thesis by synthesising the findings from this study with those from the previous research chapters and identify areas for future research.

CHAPTER 7:

CONCLUSIONS

7.1 Introduction

Chapter 1 outlined the overarching aims of this thesis which were to utilise a suite of climate, glacio-hydrological and groundwater models and apply them to the glaciated VGO to gain understanding of climate change impacts on glacier-fed river flow regime and downstream groundwater dynamics as well as the uncertainties associated with these projections. Four specific research objectives were identified and addressed in chapters 4, 5 and 6:

1. Implement a novel GHM comparison and selection framework to undertake a rigorous evaluation of multiple GHM structures.

2. Implement a climate-GHM model chain to project twenty-first century changes in different characteristics of river flow regime.

3. Determine the relative contribution of individual model chain components to uncertainty in river flow regime projections.

4. Incorporate a distributed groundwater model into the climate-GHM model chain to project twenty-first century changes in proglacial groundwater level dynamics and GW-SW exchanges.

This final chapter concludes the thesis by synthesising the research and main findings before identifying areas for future work.

7.2 Twenty-first century glacial and hydrological change

7.2.1 Principal research findings

There were three principal research findings related to twenty-first century glacial and hydrological change:

1. The magnitude, timing and variability of river flows in the glacier-fed Virkisá River will change over sub-daily to decadal timescales in response to twenty-first century climate change.

2. Climate change will also result in changes to average groundwater level and intra-annual groundwater level timing and variability in the proglacial sandur aquifer leading to changes in river leakage fluxes to the aquifer and groundwater to surface runoff.

3. The trajectory of twenty-first century groundwater level dynamics and GW-SW exchanges in the proglacial sandur aquifer are primarily controlled by patterns of diffuse recharge to the sandur rather than river stage dynamics in the glacier-fed Virkisá River.

7.2.2 Synthesis of research

Twenty-first century changes in glacier and snow coverage as well as two aspects of hydrological change in the VGO were evaluated. These aspects included the Virkisá River flow regime and proglacial groundwater level dynamics.

The river flow regime projections were simulated and analysed in chapter 5 using an ensemble of GHMs defined using the signature-based LOA framework applied in chapter 4. The suitability of the GHM ensemble for the projection experiments was determined by evaluating the models against historic observations of the river discharge signatures to be simulated. The GHM ensemble was able to capture the majority of the historic

observed river discharge signatures within their LOA indicating that the model ensemble was suitable for the projection experiments.

An ensemble of downscaled EURO-CORDEX climate projections was used to drive the GHM ensemble over the twenty-first century. Projections of glacier and snow coverage as well as 25 different characteristics of river flow regime over decadal to sub-daily timescales were analysed. The study showed for the first time that all characteristics and timescales of glacier-fed river flow regime are sensitive to twenty-first century climate change. A descriptive overview of the projected glacial and hydrological changes at the VGO from chapter 5 will now be given. For a more quantitative summary of the projected changes and their linkages for the end of the twenty-first century (2080s), the reader is referred to the flow diagram above the solid grey line in Fig. 7.1.

The projections showed high confidence that near-surface air temperature will increase over the twenty-first century causing snow and glacier coverage to recede. As the volume of cold-season meltwater runoff increases and more precipitation falls as rainfall, river discharge outside of the melt season is projected to increase. The magnitude of the lowest flow events (Q_{99} and Q_{95}), which typically occur during the colder months, will subsequently increase. However, the loss of snow and ice coverage will reduce melt season runoff volume. Consequently, the seasonality of flows in the Virkisá River will change, resulting in a reduction of intra-annual (month-to-month) flow variability. Inter-annual (year-to-year) flow variability will increase as the runoff-regime of the catchment becomes increasingly dominated by more variable rainfall inputs. The loss of the runoff-regulating effects of snow and ice will result in a more responsive and flashy watershed which, coupled with the increase in rainfall inputs, will lead to an increase in the magnitude of higher than normal flow events (Q_{05} and Q_{01}). Given the clear influence of glacier and snow retreat on the projected shifts in river flow regime at the VGO, it is likely that other glaciated basins around the world will also demonstrate sensitivity across a complex range of river flow regime characteristics. However, it is postulated that the direction and magnitude of these changes will be highly catchment-specific.

Chapter 6 projected changes in proglacial groundwater level dynamics and GW-SW

interactions. Here, a distributed MODFLOW-NWT groundwater model of the proglacial sandur was calibrated against observation borehole groundwater level time series data from the sandur. The model captured the average spatial distribution of observed groundwater levels as well as the seasonal timing of groundwater level peaks and troughs, but also showed some systematic biases in the mean and amplitude of variability of the groundwater level data.

Once calibrated, a subsample of five coupled climate-river discharge scenarios from chapter 5 were used as boundary conditions to drive the groundwater model over the twenty-first century. Cluster analysis was used to select these five scenarios so that a broad range of glacio-hydrological and climatic behaviours were covered.

The groundwater model projections were analysed for change in proglacial groundwater level dynamics and GW-SW exchanges. A descriptive overview of the projected changes in proglacial groundwater dynamics from chapter 6 are given here and a more quantitative summary for the 2080s can be found in the flow diagram beneath the solid grey line in Fig. 7.1.

Average sandur groundwater levels were consistently projected to fall in the spring and winter months due to projected reductions in total precipitation over the main recharge season. River leakage could serve to buffer up to 15% of the loss in diffuse recharge given the inevitable increase in leaking river length in response to a fall in groundwater level. Given the shallow topography of the sandur, the fall in groundwater level reduced the volume of groundwater draining to springs and the Virkisá River indicating that the potential of groundwater to buffer low river flows outside of the melt season will lessen in the future.

The trajectory of future groundwater level dynamics and GW-SW exchanges are primarily controlled by changes in diffuse recharge over the sandur rather than by changes in meltwater runoff to the Virkisá River. Even so, the projections indicated that a reduction in sub-daily river flow variability (the diurnal melt cycle) due to glacier and snow retreat could inhibit future river recharge to the aquifer. This characteristic of glacier-fed river flow regime could, therefore, be key in the future evolution of other proglacial aquifers in

Iceland and elsewhere.

7.3 Uncertainty of twenty-first century glacial and hydrological projections

7.3.1 Principal research findings

There were three principal research findings related to uncertainty of twenty-first century glacial and hydrological projections:

1. The signature-based LOA framework provides a powerful method for identifying deficiencies in different GHM structures and parameterisations. However, its use for objectively selecting a range of acceptable models remains to be seen as no models captured all signatures within their LOA simultaneously.

2. Projections of glacier-fed river flow regime are uncertain due to deficiencies in future climate projections and GHM structural and parameter uncertainty. The relative contribution of uncertainty from climate projections and GHMs depends on the characteristic (signature) of river flow regime and the time frame of the projections.

3. Uncertainty in sandur groundwater level dynamics and GW-SW exchanges primarily stem from uncertainty in future climate. However, structural uncertainty associated with the snow and ice melt routine in the GHM contributes to river leakage projection uncertainty.

7.3.2 Synthesis of research

GHM structure and parameter selection uncertainty contributes to hydrological projection uncertainty in glaciated catchments. Chapter 4 applied a signature-based LOA framework to compare the efficiency of different snow and ice melt model structures and runoff-routing model structures with the aim of determining the framework's utility for diagnosing model deficiencies and constraining a prior population of model hypotheses down to

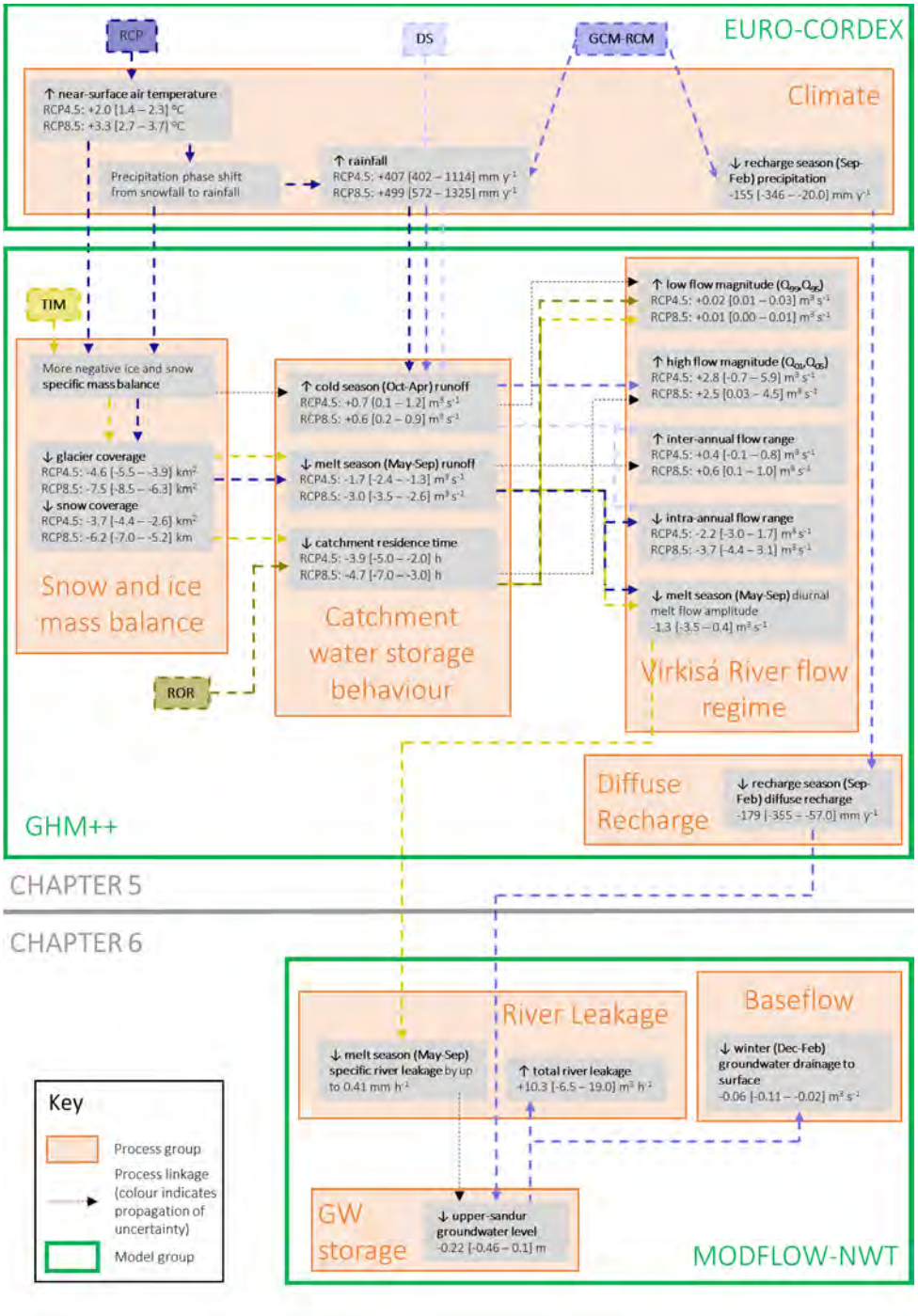

Figure 7.1: Summary of projections for 2080s. Grey boxes show changes in processes/fluxes/stores including mean and interquartile range (square brackets) of projections under each RCP. Values in square brackets from chapter 6 (below grey solid line) show range of the five scenarios. Orange boxes show process groups and green boxes show model groups. Dashed arrows show driver-response linkages (response at arrow head end). Colour of arrows indicates dominant projection uncertainty sources in the model chain.

a smaller population of acceptable ones. 45,000 GHM structure-parameter compositions were applied to the Virkisá River basin and evaluated against 33 signatures of catchment glacio-hydrological behaviour. The study demonstrated that the signatures and their LOA provide an intuitive means to identify aspects of the glacio-hydrological system that a particular GHM configuration can and cannot capture within the uncertainty of the observation data used to calculate those signatures. When evaluated against signatures individually, subsets of the models were able to capture all but two of them. The sensitivity of the acceptability scores to changing the melt and runoff-routing model structures showed to be signature-specific. This provided the first indication that the dominant sources of uncertainty in projections of river discharge could also be signature-specific, emphasising the need to investigate this further in chapter 5.

The acceptability of the models broke down when assessed against two or more signatures simultaneously — an indicator of the model's consistency. None of the GHM compositions were able to capture all of the signatures simultaneously and, therefore, it remains to be seen if the LOA framework can be used to objectively select a population of behavioural GHMs. It was concluded that in order to do this, one would need to expand the model selection framework to include multiple realisations of boundary conditions (particularly the driving precipitation data) and additional model structures.

The inclusion of extra process complexity in the GHM showed to improve model acceptability across specific signatures, but such modifications could also introduce new model inconsistencies. Indeed, the lack of improvement across individual signatures in combination with an increase in model inconsistency provided justification for excluding the most complex ROR_3 runoff-routing model structure from subsequent studies in chapters 5 and 6.

In light of the findings from the LOA framework study, chapter 5 adopted a signature-based LOA model calibration approach to identify a model ensemble that reflected the uncertainty in model selection given the known model inconsistencies. When driven with the downscaled EURO-CORDEX ensemble of climate projections, there was high confidence in the direction of change for each signature, but the magnitude of change was

uncertain. A decomposition of projection uncertainty for snow coverage, ice coverage and river flow regime change was undertaken using ANOVA. All five sources of uncertainty in the model chain contributed to projection uncertainty. For the signatures of river flow regime change, the relative contribution of uncertainty from each source depended on the characteristic of flow and the time frame over which those projections were made. A descriptive overview of the primary sources of projection uncertainty is given here. A summary of the propagation of these sources through the model chain is given in Fig. 7.1 as represented by the coloured arrows.

Uncertainty in projections of snow and ice coverage primarily stemmed from the RCP due to its control on near-surface air temperature. Consequently, the RCP was also the dominant source of uncertainty for mean monthly river flows during the melt season. The DS procedure was the largest contributor to projection uncertainty for inter-annual (year-to-year) variability in river discharge. The TIM structure-parameterisation was the dominant contributor to projections of monthly mean river flows during the transition from the cold to melt season. The GCM-RCM contributed most to uncertainty in future high flow signatures, while the ROR structure-parameterisation significantly contributed to the total projection uncertainty of slow-release low flow signatures and signatures of response time (flashiness) of the catchment.

The translation of large climate and GHM uncertainties quantified in chapter 5 to the proglacial groundwater modelling experiments undertaken in chapter 6 was limited by the computational requirements of running large ensembles through a distributed groundwater model and the desire to undertake a process-oriented analysis of future proglacial groundwater dynamics. Instead cluster analysis was used to select a handful of scenarios that reflected contrasting future glacio-hydrological and climatic conditions for the 2080s. This approach still provided some basis for analysing the sources of uncertainty for the groundwater projections. The results indicate that the GCM-RCM is likely to be the dominant contributor to projection uncertainty for proglacial groundwater dynamics given that the aquifer showed to be recharge-driven and GCM-RCM uncertainty is the main source of future precipitation uncertainty. The results also indicate that the TIM

structure-parameterisation contributes to uncertainty in projections of river leakage given its influence on the amplitude of the melt-induced diurnal river discharge cycle. As such, uncertainty both in future climate and the representation of key glacio-hydrological processes in the GHM have the potential to propagate through to projections of groundwater dynamics in glaciated catchments.

7.4 Recommendations for future research

Based on the findings in this thesis, a number of recommendations for future work have been highlighted and these broadly fall under three themes each of which is detailed in this section.

7.4.1 Application of signature-based methods in glacio-hydrological modelling

This research has demonstrated the advantages of two principal applications of signature-based methodologies within the field of glacio-hydrological modelling and it is recommended that these approaches are adopted in future glacio-hydrological modelling studies where possible.

Firstly, glacio-hydrological modelling studies should seek to diagnose structural model deficiencies through the evaluation diagnostic signatures of system behaviour. Chapter 4 demonstrated that signatures provide a means to interrogate glacio-hydrological model behaviour at the process level by quantifying specific aspects of observed catchment behaviour that can be related to different process representations within the GHM. Additionally, because they are derived directly from observation data, it is relatively easy to quantify their uncertainty, providing a means to evaluate model structure appropriateness within the uncertainties of the observation data and, therefore, diagnose structural deficiencies more objectively.

Of course, a key limitation of the signature-based LOA framework used in chapter 4 was that no acceptable models were found. However, it is hoped that rather than dissuade

others from adopting a similar framework, this finding will encourage them to seek out model configurations (existing or newly-developed) that are able to capture a range of signatures within their observation uncertainty. Model hypothesis testing like this would not only help to deliver more justifiable GHM structures, but could also help to develop process understanding of glacio-hydrological systems (e.g. as has been shown in non-glaciated hydrology Clark et al., 2015). Model processes not investigated in this thesis as well as boundary conditions and driving climate data should be the focus of future work in this area. It should also be emphasised that future applications of a signature-based LOA framework need not adopt the same 33 signatures used in chapter 4. On the contrary, the choice of signatures will depend on the availability of observation data, and information on data uncertainty as well as the complexity (e.g. spatio-temporal resolution) of the model(s) being interrogated. Study sites with good observation data and understanding of data uncertainty would be ideal candidates for future applications.

It is also recommended that signatures are used as a means to identify how different characteristics of glacier-fed river flow regime — which are deemed important for future socio-economic and environmental prosperity — will respond to future climate. Chapter 5 highlighted the complexity of river flow regime shifts and, therefore, the wider adoption of signatures in impact studies would provide a more comprehensive understanding of potential impacts of change which could feed into adaptation strategy planning. Indeed, river flow signatures already underpin decision-support systems for managing hydroecological systems in warmer environments (Beamer et al., 2017; Cartwright et al., 2017; Pool et al., 2017).

Given that the characteristics of river flow regime that have been widely studied before are known to be highly site-specific (e.g. flow seasonality Huss and Hock, 2018; Ragettli et al., 2016), it is hypothesised that changes in specific signatures in other glaciated watersheds will also depend strongly on local climatic, glaciological and hydrological factors. For example, the relatively small reduction in the projected catchment water transit time (τ) in response to glacier and snow retreat is likely due to the fact that, relatively speaking, Virkisjökull is a small glacier and therefore it's ability to slow the flow of water is limited

by the size of its internal drainage network. In addition, the glacier is known to have a well developed conduit drainage system that routes runoff efficiently year-round (Flett et al., 2017; Phillips et al., 2014). Watersheds with larger cryospheric water stores and less efficient internal drainage systems should be expected to show much larger changes in water transit time. Therefore, signature-based analyses of river flow regime change should be encouraged in other glaciated catchments with different geometries, hypsometry and climate to explore physical controls on these characteristic changes.

7.4.2 Investigations of climate change impacts on proglacial aquifers

The study presented in chapter 6 was the first application of a distributed groundwater model driven by a GHM to evaluate the impact of climate change as well as glacier and snow retreat on a proglacial groundwater system. While the sandur in the VGO showed to be relatively insensitive to changes in mountain runoff due to glacier and snow retreat, it is postulated that this is due to the high diffuse recharge inputs that the sandur receives. It is vital, therefore, that future research is undertaken in regions with different climates. For example, it has already been shown that in more arid, snow-dominated mountain catchments, particularly in the semi-arid mountain catchments in the north-west of the United States (Allen et al., 2004; Huntington and Niswonger, 2012; Scibek et al., 2007), changes in mountain-fed river flow regime exhibit a much stronger control on lowland aquifer groundwater dynamics. It is suggested here, therefore, that in more arid glaciated catchments, proglacial groundwater dynamics and GW-SW interactions are likely to be much more sensitive to glacier and snow retreat. Given the potentially important role that groundwater could play in sustaining surface water availability in proglacial regions where snow and ice have retreated, a fuller understanding of climate change impacts on groundwater resources in these regions is paramount. This understanding should be developed through using glacio-hydrological and groundwater models. Based on the findings from chapter 6, a number of recommendations can be made for future modelling applications of this nature. Firstly, they should consider changes in melt-induced sub-daily characteristics of river flow regime given that these could be a significant driver of

river recharge from glacier-fed rivers. As such, it is imperative that such models are run on a sub-daily (ideally hourly) time step so that these characteristics can be captured. These studies should also be underpinned by a well-constrained melt model given the sensitivity of river leakage to the diurnal melt cycle. Some representation of melt model uncertainty would be especially beneficial if the simulations are to be used to inform policy decision making. Finally, the inclusion of a mass-conserving dynamic glacier model and snow routine is vital so that the river flow regime response to climate and cryosphere change and subsequent changes in groundwater dynamics can be captured adequately.

7.4.3 Robust quantification and exploration of projection uncertainties

When considering the different sources of uncertainty in projections of hydrological change in glaciated basins, past studies have typically focussed on including future climate uncertainty only. However, the research undertaken in chapter 5 showed for the first time that the uncertainties stemming from the glacio-hydrological model are high, and dominate projection uncertainty for some characteristic changes in glacier-fed river flow regime. Furthermore, chapter 6 has provided the first model-based evidence that this uncertainty could propagate through to simulations of proglacial groundwater dynamics (although this is subject to further modelling studies recommended above). It is therefore recommended that future glacio-hydrological modelling studies must make efforts to quantify uncertainties stemming both from climate and from incomplete knowledge of glacio-hydrological process complexity. Furthermore, future researchers should be encouraged to include other sources of GHM uncertainty that have not been investigated here. Obvious contenders for these include the snow redistribution routine, the ice evolution model and climate and ice boundary conditions. It follows that if modellers include these different sources of uncertainty in their models (or chain or models), they should be encouraged, not only to quantify projection uncertainty, but to actively explore those uncertainties to direct future best practices (e.g. targeted approaches for most important uncertainty sources) and research directions.

The ANOVA approach applied to projections of river discharge signatures showed that different components of the model chain are more significant sources of projection uncertainty for different signatures. From this, one can conclude that projection studies could be better-designed to prioritise the representation of uncertainty sources that are most important for the characteristics of change they are interested in. Currently, little to no information on this exists for glaciated river basins, yet it must be prioritised given the computational limitations that still limit the number of uncertainty sources that can be realistically investigated at one time. Further applications of approaches like ANOVA would help to address this knowledge gap. For such applications, it is also recommended that a single, but flexible GHM code is used so that sources of projection uncertainty can be easily localised and quantified at the process level.

Finally, it is recognised that the scenario-based analysis undertaken in chapter 6 did not facilitate a thorough analysis of uncertainty in the projections of proglacial groundwater dynamics and GW-SW exchanges. These uncertainties could stem from climate boundary conditions, the GHM and the groundwater model itself. While the computational demands of a large-ensemble experiment that incorporated all of these uncertainty sources is, perhaps, not technically feasible now, such an experiment in the future could deliver potentially important findings to feed into future adaptation strategy planning. This should be treated as a key milestone for future investigations of twenty-first century glacial and hydrological change.

7.5 Final remarks

This thesis has shown how twenty-first century climate and cryosphere change will propagate through the hydrological system at the VGO and manifest as a complex range of shifts in river flow regime and proglacial groundwater dynamics and GW-SW exchanges. It has also shown how different sources of model uncertainty, be it uncertainty in future climate projections or uncertainty due to incomplete understanding of glacio-hydrological processes, map onto these different characteristics of hydrological change. Projections of

change and robust evaluations of their uncertainty like this are fundamental if we wish to increase resilience of populations and ecosystems that will be affected by changes in water cycling due to climate change and glacier and snow line retreat over the coming decades. However, the transferability of the findings from this thesis to other catchments is currently unknown. Accordingly applications of the methods adopted in this thesis to other catchments with different glaciological and climate settings must be undertaken. It is suggested here, that the use of signatures and integrated modelling approaches that include river flow and groundwater would be especially valuable to scientific community and vulnerable communities. Furthermore, through the continued exploration of modelling uncertainties, more robust projections of hydrological changes in glaciated basins can be delivered to the end users that most need them.

Appendices

APPENDIX A:

RAW WEATHER STATION DATA

A.1 Automatic weather stations

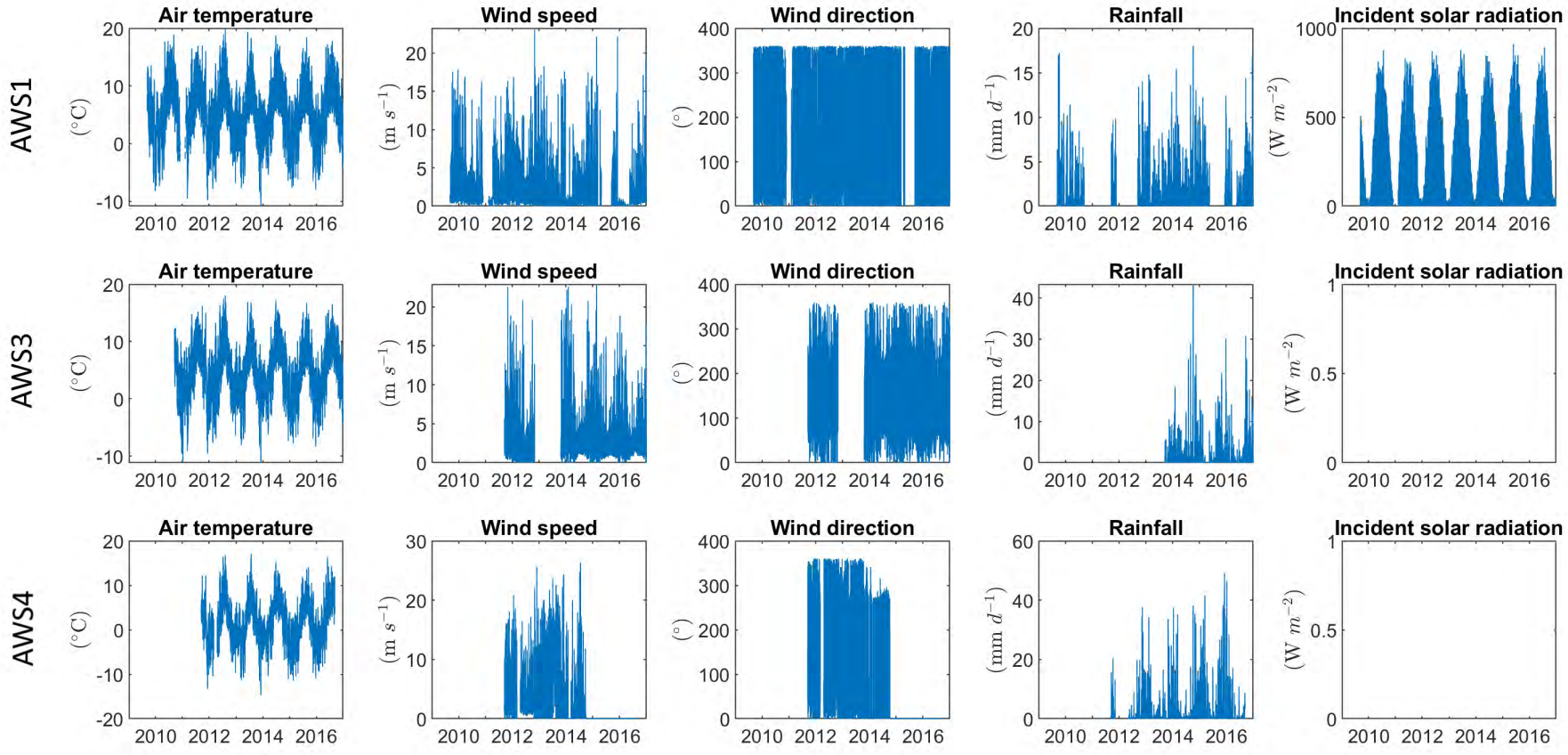

Figure A.1: Raw time series data from three AWSs.

A.2 Fagurhólsmýri weather station

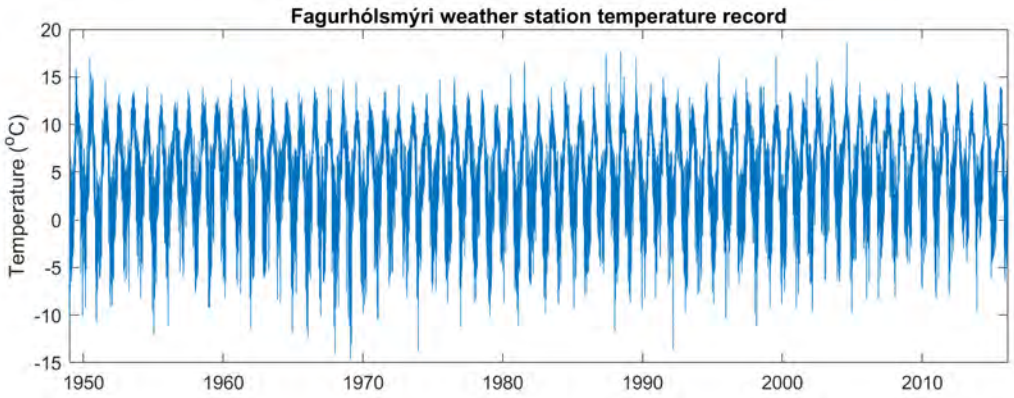

Figure A.2: Raw near-surface air temperature time series data from Fagurhólsmýri weather station.

APPENDIX B:

CORRECTION AND HARMONISATION OF RIVER STAGE TIME SERIES DATA

Prior to combining the logged river stage data into a single harmonised time series, a comparison of water levels logged at the two stilling wells revealed discrepancies between them of up to 40 cm (Fig. B.1a&b). There are a number of distinct flat regions in the difference plot (indicated by green dashed lines in Fig. B.1b) where the dynamics of the two water level records are similar and any difference between the two are systematic. The largest systematic differences occur in 2015 where they jump to -27.5 cm. This corroborates observations made in the field in March 2015 where it was noted that the central stilling well had moved in an upward direction significantly. The reason for this movement is not known, but is thought to have occurred during or shortly after a large storm event in the 2014/2015 winter. As an additional source of evidence, Fig. B.2 shows the difference between logged water levels at both wells and independent readings of stage taken from a stage board permanently fixed to the bridge. Here it can be seen that for 2013 and 2014, logged water levels at both stilling wells lie within the estimated reading error from the stage board ($\pm$5 cm), but in 2015 logged levels from the central stilling are negatively biased with a magnitude equivalent to that shown in Fig. B.1b, while the data from the eastern well is still consistent with the independent measurements of stage. This indicates that the eastern stilling well can be used to correct the systematic bias in the central stilling well data post-2014 which was done according to the derived systematic biases shown in Fig. B.1b. Prior to 2013, there are no obvious flat regions in the difference plot. Instead, there shows to be some longer-term drift between September 2011 and the end of 2012. This could be partly explained by significant bridge maintenance works which were undertaken by the highway authority in October 2012 resulting in a major diversion of the main river channel. Furthermore, comparisons of logged levels to independent readings taken at the stage board show a large scattering of differences in

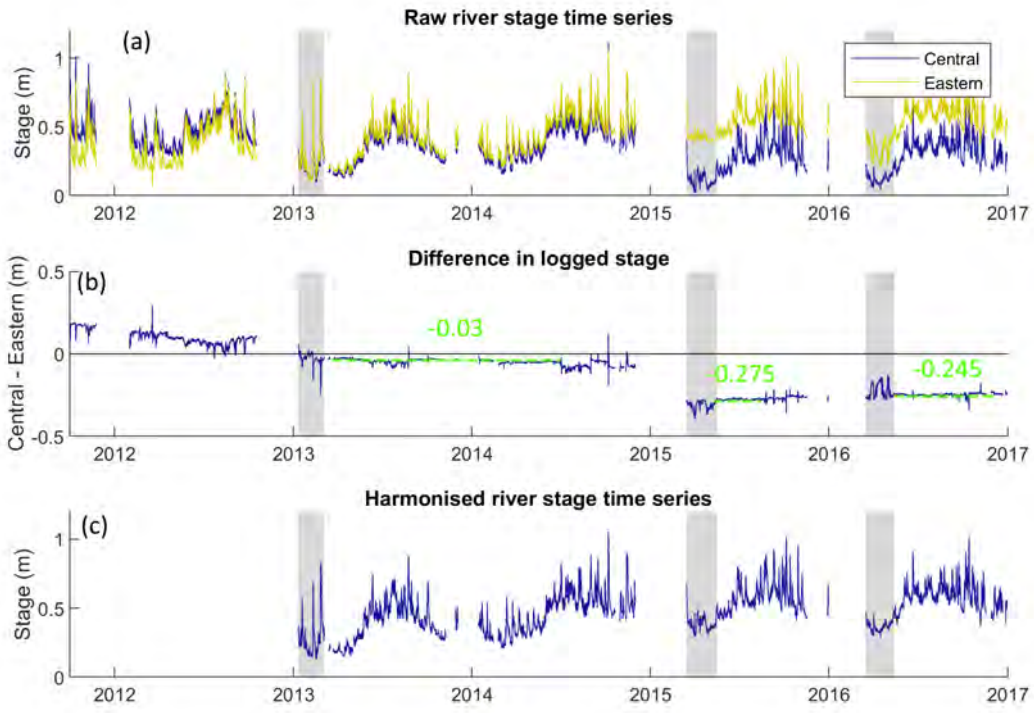

Figure B.1: Raw hourly logged river stage from stilling wells (a); the difference between the two (b) and the harmonised stage time series after correction (c). Note, periods where the river was frozen over have been removed from the time series and shaded areas denote periods where the eastern stilling well data were substituted for corrected central stilling well data.

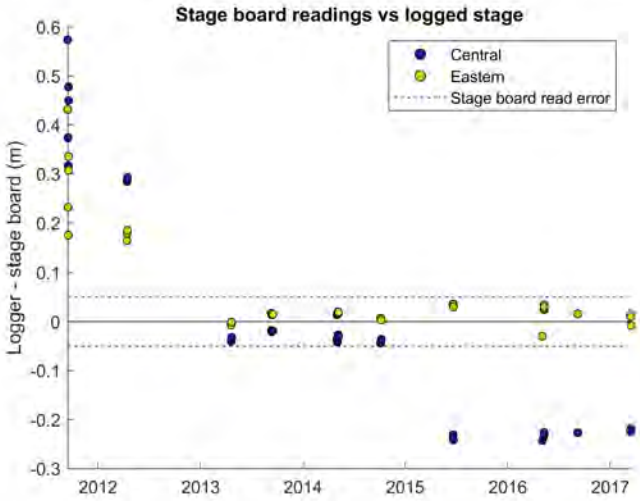

Figure B.2: Difference between logged stage and independent readings from stage board at bridge.

2011 indicating the water level data from both stilling wells are likely to be unreliable. In 2012, both stilling wells show a positive bias relative to the stage board. Given the apparent drift shown in Fig. B.1b, the difficulty in relating the data to independent stage board readings and the known modifications made to the channel in October 2012, it was decided not to use measurements of river stage prior to 2013.

In addition to the systematic differences between the two river stage datasets indicated by the flat regions of the difference plot, there are localised periods where the differences diverge, often increasing in magnitude. These occur in the spring months of 2013, 2015 and 2016 as shown by the shaded regions in Fig. B.1. After discussions with the team at the University of Dundee who set up the gauging station, it was determined that for these periods, given flow is typically lower than average which can lead to drying and/or pooling effects, the central stilling well data is most reliable and should be used. Accordingly, the corrected central stilling well data was utilised here. Outside of these periods, where flow is typically higher and hydraulic effects bias the central stilling well data, only the logged stage from the eastern stilling well have been used.

APPENDIX C:

GLACIO-HYDROLOGICAL MODEL

C.1 Soil infiltration and evapotranspiration

The semi-vegetated nature of the VGO coupled with the relatively cool temperatures year-round mean that evapotranspiration is generally low (Einarsson, 1972). Even so, to satisfy the water balance, an explicit representation of the soil zone for model nodes that are not ice or snow-covered was included using the method developed by Griffiths et al. (2008) based on the well-established FAO56 soil moisture accounting procedure (Allen et al., 1998). This model has been applied extensively (Jackson et al., 2016; Mackay et al., 2014; Mackay et al., 2015; Mansour et al., 2018) and has shown to compare favourably to physically based models at the field scale where interception losses are small (Sorensen et al., 2014). For each bare ground node, the soil is represented as a finite storage reservoir with a soil water capacity, termed the total available water, TAW [L], which defines the maximum volume of water available to plants for evapotranspiration after the soil has drained to its field capacity. This is calculated as:

$$TAW = Z_r(FC - WP) \tag{C.1}$$

where Zr [L] is the maximum root depth of the vegetation and FC [L] and WP [L] are the field capacity and wilting point of the soil respectively, all of which can be defined from lookup tables with basic information on vegetation and soil information (Allen et al., 1998). Parameters defined for the 'Talus' soil class and 'semi-vegetated' land surface class derived by Flett (2016) were used across the model domain, giving an average TAW value of 7 mm. Soil storage is replenished by infiltration from rainfall and melting of residual snow overlying the bare ground and is depleted by evapotranspiration giving a soil water balance:

$$\frac{\Delta S_{soil}}{\Delta t} = I - ET \tag{C.2}$$

where S_{soil} [L] is the soil water storage, t is time, I [LT^{-1}] is the infiltration rate and ET [LT^{-1}] is the evapotranspiration rate. Because measured ET is rarely available, Griffiths et al. (2006) propose using the potential evapotranspiration rate, ET_0, instead which defines the evapotranspiration rate from a reference grass-covered wet soil (see Appendix C.2 for calculation of ET_0). Using ET_0 as the maximum possible evapotranspiration rate, they define a separate function which accounts for the fact that as the soil becomes drier, plants find it more difficult to extract moisture from the soil matrix, and therefore ET is typically less than ET_0. While this is conceptually sound, it was decided not to include this function and instead assume that $ET = ET_0$. There are three reasons for doing this. Firstly, because the inclusion of this function requires an additional parameter which is uncertain and must be calibrated. Secondly because ET is a relatively small component of the overall water balance in this catchment and it was not the aim of this study to investigate this aspect of the catchment hydrology. Thirdly, because previous studies have shown that this parameter (and therefore the behaviour of this function) is relatively insensitive and unidentifiable (Mackay et al., 2014).

In the original formulation by Griffiths et al. (2006), any excess soil water (i.e. when $S_{soil} > TAW$) is distributed between overland flow and groundwater recharge pathways. They use a fixed baseflow index (BFI) parameter which defines the proportion of soil water excess that recharges the groundwater. Given the nature of the Virkisá River basin (thin soils overlying impermeable bedrock), it was assumed that soil water migrates to the river outlet via relatively fast, overland flow pathways only and so the BFI parameter was set to zero.

C.2 Potential evapotranspiration

Potential evapotranspiration can be calculated from measured meteorological data, most simply as a linear function of measured temperature (e.g. Blaney and Morin, 1942), or where measurements of windspeed, air pressure and solar radiation exist, the full Penmen-Monteith combination equation can be solved. Given that these additional variables are

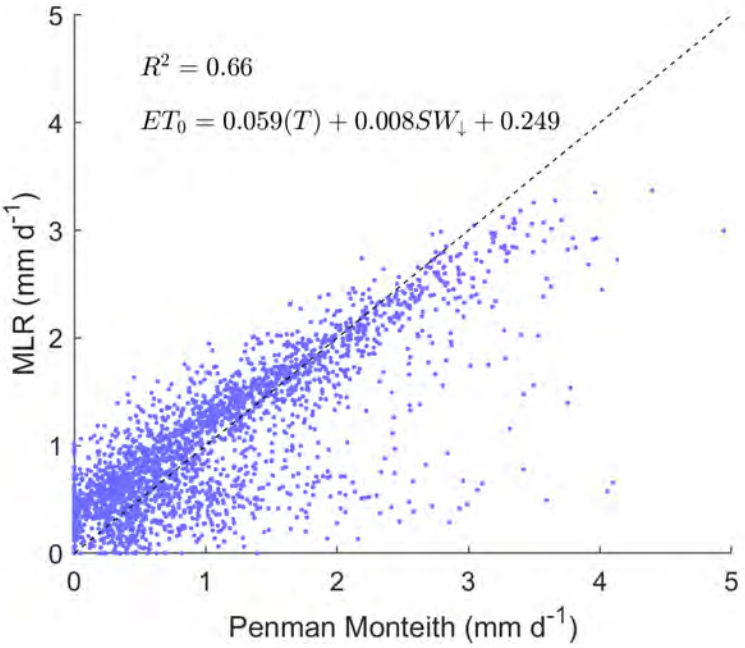

Figure C.1: Multiple linear regression model used to convert ambient air temperature and incoming solar radiation into potential evapotranspiration.

measured at AWS1 from 2009, the combination equation as defined by Allen et al. (1998) was used to calculate hourly potential evapotranspiration over this period:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{T+273}u(e_s - e_a)}{\Delta + \gamma(1 + 0.34u)}h \tag{C.3}$$

where ET_0 is the daily average potential evapotranspiration rate (mm d^{-1}), R_n is the net radiation (MJ m^{-2} d^{-1}), G is the soil heat flux (MJ m^{-2} d^{-1}), e_s and e_a are the saturation and actual vapour pressure respectively (kPa), Δ is the rate of change of the saturation vapour pressure with temperature (kPa °C^{-1}), γ is the psychrometric constant (kPa °C^{-1}), u is the wind speed (m s^{-1}) and T is the mean daily ambient air temperature (°C).

Prior to 2009, the viability of using T as a proxy for ET_0 in a linear regression model framework like Blaney and Morin (1942) was investigated. Similarly, incident solar radiation was also used as the independent variable for this model. In fact, the best fit was achieved using both variables in a multiple linear regression model which was able to explain 66% of the ET_0 variance (Fig. C.1). This model was used to distribute ET_0 in space and time using the driving temperature and incident solar radiation data.

C.3 Glacier geometry evolution

The empirical Δh parametrisation (Huss et al., 2010) requires the availability of at least two digital elevation models of the glacier separated in time. The difference between the two is used to define the Δh polynomial which has the form:

$$\Delta h = (h_r + a)^\gamma + b(h_r + a) + c \tag{C.4}$$

where Δh is the normalised surface elevation, h_r is the normalised elevation range and a, b, γ and c are fitted parameters. The two digital elevation models from 1988 and 2011 were used to define this relationship. Figure C.2a shows the raw change data against the 1988 ice elevation. It was decided that the data at the very front of the glacier should not be used as here the ice has completely melted and as such the bedrock beneath skews the raw change data. Figure C.2b shows the fitted Δh model to the normalised mean elevation change curve. Following Huss et al. (2010), the glacier geometry is updated each year by distributing the net glacier mass balance across the glacier according to this relationship.

C.4 Calibration parameters

Table C.1 lists all of the calibration parameters for the melt and runoff-routing model structures which were randomly perturbed during the GHM calibration procedure.

Table C.1: Calibration parameters for the melt and runoff-routing model structures.

Structure	Parameter	Calibration range	Units
TIM_1	a_{ice}	2.0e-4 - 7.0e-4	m w.e. $°C^{-1}$ h^{-1}
	$a_{snow/firn}$	4.0e-7 - 2.0e-4	m w.e. $°C^{-1}$ h^{-1}
TIM_2	a_{ice}	2.0e-4 - 7.0e-4	m w.e. $°C^{-1}$ h^{-1}
	$a_{snow/firn}$	4.0e-7 - 2.0e-4	m w.e. $°C^{-1}$ h^{-1}
	b_{ice}	4.0e-7 - 2.0e-6	m^3 w.e. W^{-1} $°C^{-1}$ h^{-1}
	$b_{snow/firn}$	4.0e-8 - 4.0e-7	m^3 w.e. W^{-1} $°C^{-1}$ h^{-1}
TIM_3	a_{ice}	1.5e-4 - 3.0e-4	m w.e. $°C^{-1}$ h^{-1}
	$a_{snow/firn}$	6.0e-5 - 2.0e-4	m w.e. $°C^{-1}$ h^{-1}
	b_{ice}	1.0e-5 - 8.0e-5	m^3 w.e. W^{-1} h^{-1}
	$b_{snow/firn}$	2.0e-7 - 4.0e-6	m^3 w.e. W^{-1} h^{-1}
	p_2	0.01 - 0.4	
ROR_1	k	1 - 30	h
	n	1 - 5	
ROR_2	$k_{ice/soil}$	0.1 - 5	h
	$k_{snow/firn}$	20 - 100	h
	$n_{ice/soil}$	1 - 5	
	$n_{ice/snow}$	1 - 5	
ROR_3	k_{soil}	0.1 - 5	h
	k_{ice}	0.1 - 5	h
	k_{snow}	10 - 50	h
	k_{firn}	50 - 300	h
	n_{soil}	1 - 5	
	n_{ice}	1 - 5	
	n_{snow}	1 - 5	
	n_{soil}	1 - 5	

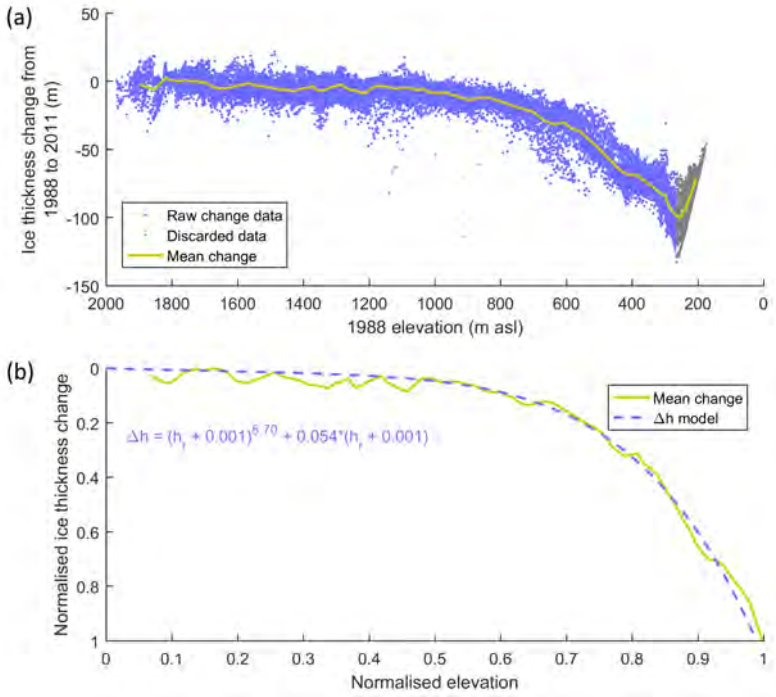

Figure C.2: Raw elevation change data from 1988 and 2011 ice DEMs (a) and fitted Δh model to normalised mean elevation change curve following Huss et al. (2010) (b).

APPENDIX D:

TEMPERATURE LAPSE RATES

In order to investigate seasonal variations in lapse rate, the temperature gradient between the lowest (AWS1) and highest (AWS4) weather stations were analysed. The results showed a remarkable degree of variation in hourly average lapse rate between the months of the year (white lines in Fig. D.1). During the winter months between November and February, the lapse rate is a relatively stable -5 °C km^{-1} throughout the day. In contrast, between March and October there is a pronounced diurnal variation in lapse rate where it is strongest in the late afternoon/early evening. The heat maps in Fig. D.1 represent the frequency distribution of wind direction for each month and show that the development of the strongest lapse rates in the afternoon correspond with a break-up of the prevailing north-east winds that flow down from the summit of Öræfajökull and a switch to winds from the south-west. Petersen and Pellicciotti (2011) found a similar phenomenon on the Juncal Norte Glacier in the semi-arid Chilean Andes. They attributed the shallow temperature gradient in the morning with katabatic winds flowing down-glacier which serve to cool the air over the glacier and weaken the lapse rate. In the afternoon, they showed that a breaking up of this layer by valley winds served to increase the temperature gradient by warming the air over the lower glacier. This suggests that winds flowing down from the Öræfajökull summit in the warmer months could serve to cool near-surface air temperatures over the ice, thereby retarding ice melt. To account for this phenomenon, Petersen and Pellicciotti (2011) suggest adopting the Shea and Moore (2010) model to correct on-ice temperatures relative to ambient off-ice weather station measurements. Shea and Moore (2010) found that for three glaciers on the southern coast mountains of British Columbia, Canada, there was a threshold in ambient off-ice air temperature, above which the winds flowing over the glacier served to cool the near-surface on-ice air. They suggest this temperature lies somewhere between 4 and 8 °C, but is likely to be site specific.

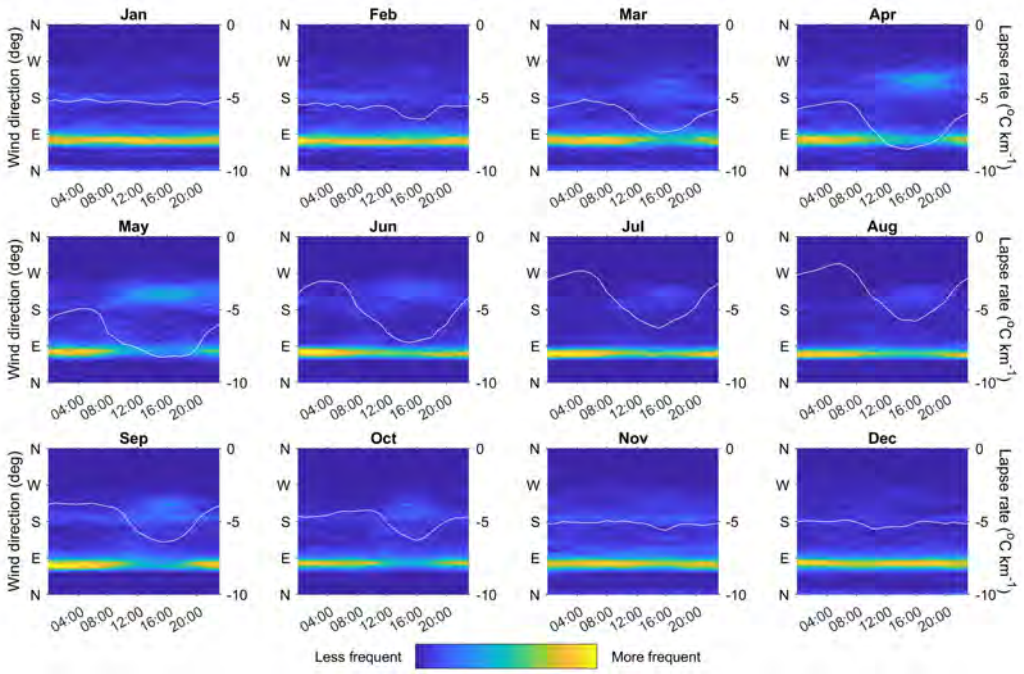

Figure D.1: Monthly average hourly temperature lapse rates (white lines, right hand axis) derived from AWS1 and AWS4 temperature time series overlying heat maps which represent the frequency distribution of hourly wind direction data from AWS4 (left hand axis).

To investigate if such a threshold exists on the Virkisjökull glacier, five Gemini Tinytag Aquatic 2 temperature loggers were deployed across the glacier at elevations ranging from 150 - 400 m asl. Each logger was secured at 1.5 m above the ice in a white PVC radiation shield attached to a tripod (Fig. D.2). The sensors were deployed for 7 days in late August 2016 and then for a further 7 days in early March 2017 to represent summer and winter on-ice temperatures respectively. The loggers were synchronised in time with the AWSs and set to measure temperature every 15 minutes. This allowed for the direct comparison of on and off-ice near-surface temperatures.

Figure D.3 shows the synchronised on and off-ice temperatures from all of the measurements taken in winter (blue dots) and summer (yellow dots). The off-ice temperatures were derived assuming a linear lapse rate between AWS1 and AWS3 as these are situated at elevations similar to the Tinytag temperature loggers. The results show that there is a temperature threshold above which on-ice temperature falls below off-ice temperature which was estimated to be 5.27 °C. Following Petersen and Pellicciotti (2011), Ragettli et al. (2014), and Shea and Moore (2010), this cooling effect was interpreted as being due to north-east winds which bring cooler air from above over the tongue of the glacier, thereby cooling the on-ice air temperature and the piecewise function derived from Fig. D.3 was employed to correct temperatures on the ice during the warmer months when ambient air temperatures exceed this threshold:

$$T_{on} = \begin{cases} T_{off} & T_{off} \leq 5.27 \\ 0.74 \cdot T_{off} + 1.38 & T_{off} > 5.27 \end{cases} \tag{D.1}$$

where T_{on} and T_{off} are the on and off-ice near-surface air temperature (°C).

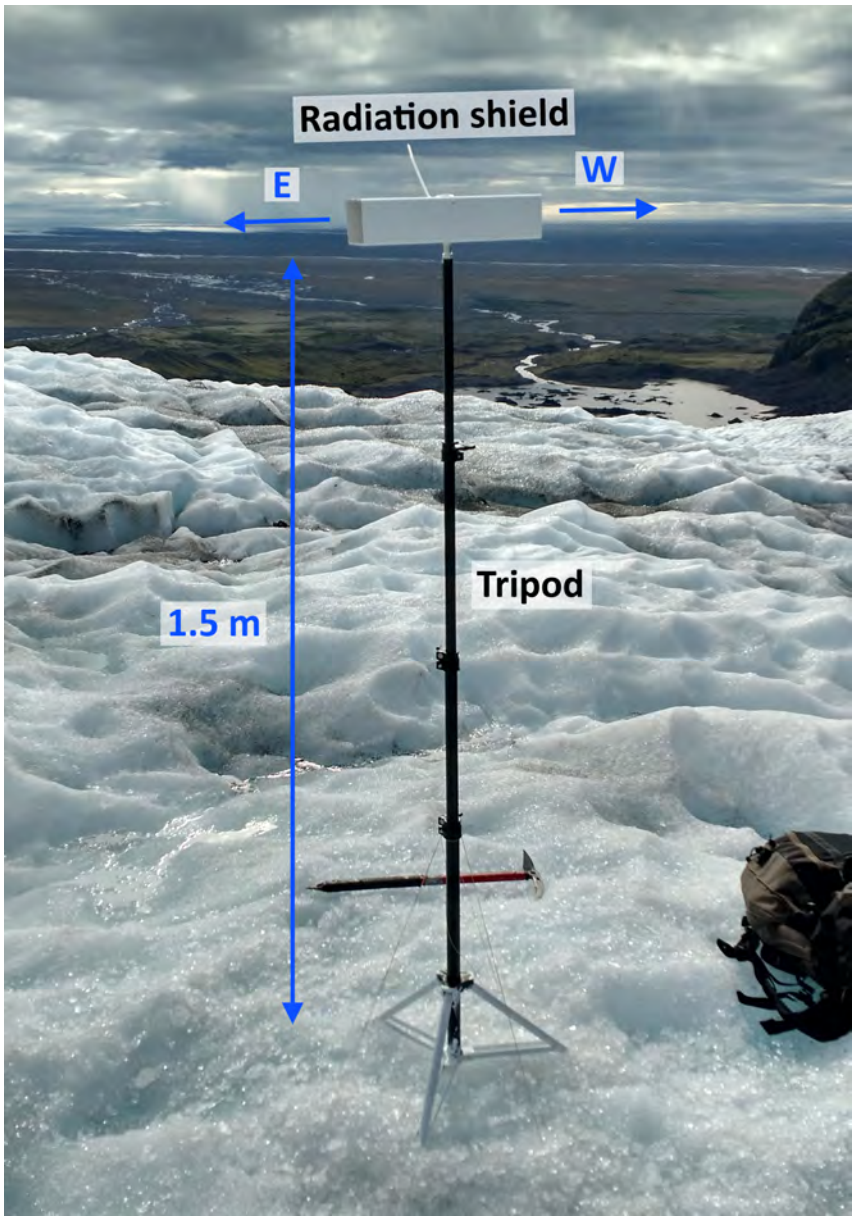

Figure D.2: Example of Gemini TinyTag housing used for measuring on-ice temperature at one location on ice.

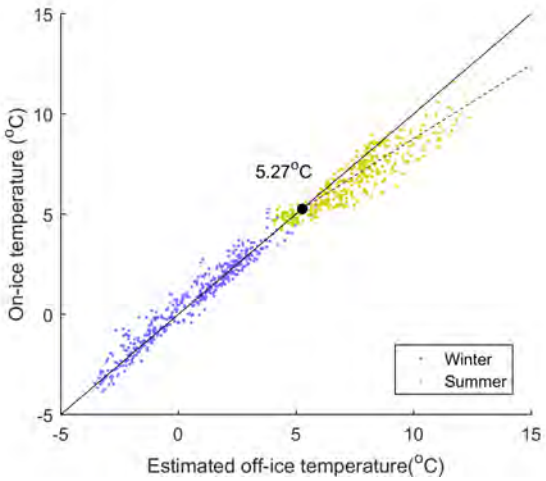

Figure D.3: Derived temperature threshold where on-ice temperature is cooler than the ambient off-ice temperature using Shea and Moore (2010) model.

APPENDIX E:

GHM++ INPUT AND OUTPUT TIME SERIES

Figure E.1 shows the complete GHM++ input and output time series. These include the watershed total precipitation, watershed average temperature and incident solar radiation data used to drive GHM++ as well as the simulated watershed total snow melt, ice melt and river discharge using the TIM_1, TIM_2 and TIM_3 melt model structures in conjunction with the simplest ROR_1 runoff-routing structure. Figure E.2 shows the same set of plots when using the ROR_1, ROR_2 and ROR_3 runoff-routing model structures in conjunction with the simplest TIM_1 melt model structure.

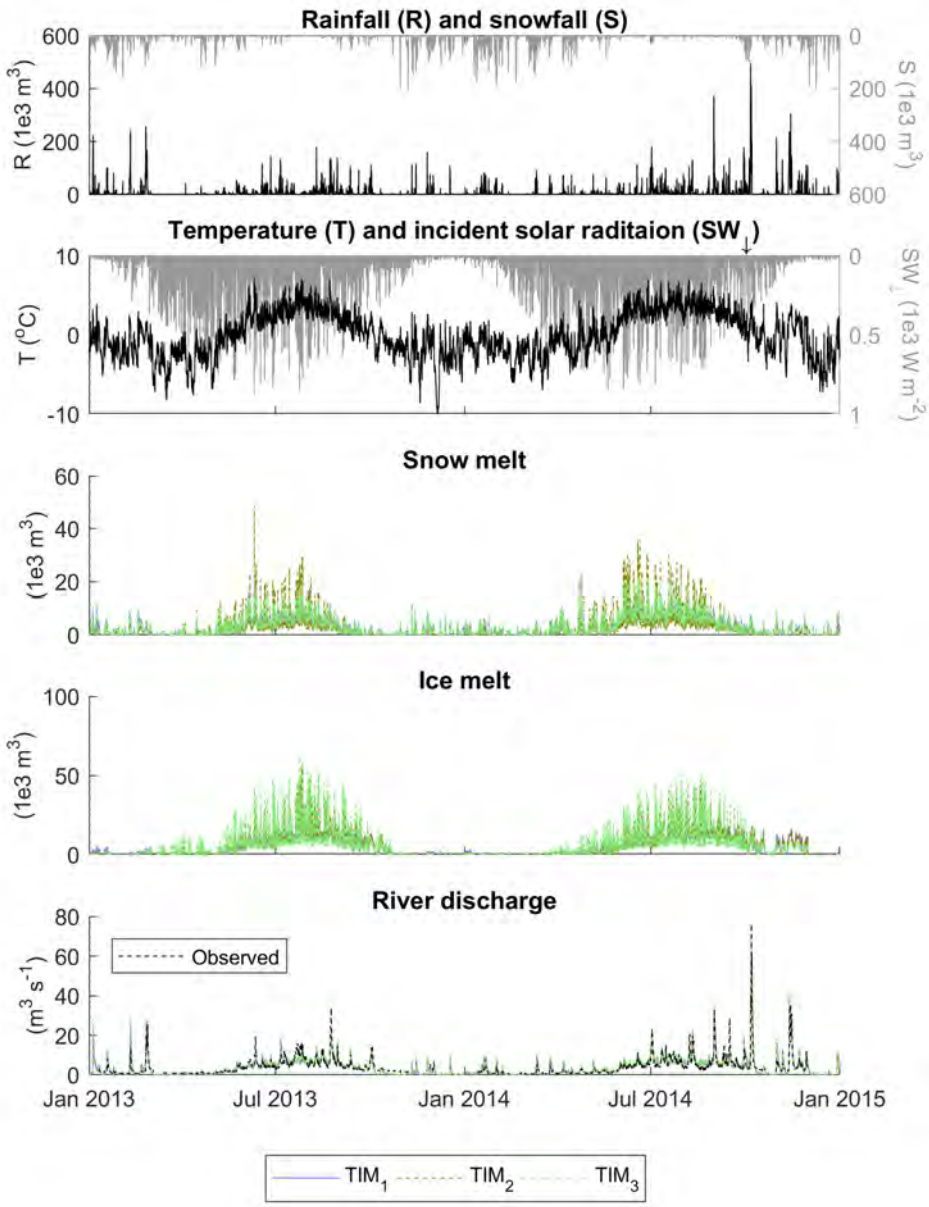

Figure E.1: Time series of driving precipitation, temperature and incident solar radiation data and simulated snow melt, ice melt and river discharge using the TIM_1, TIM_2 and TIM_3 melt model structures in conjunction with the simplest ROR_1 runoff-routing structure. Note, the proportion of rainfall and snowfall is an output from the GHM which is approximately equal across the different configurations. Ice melt includes melt of bare ice and the firn.

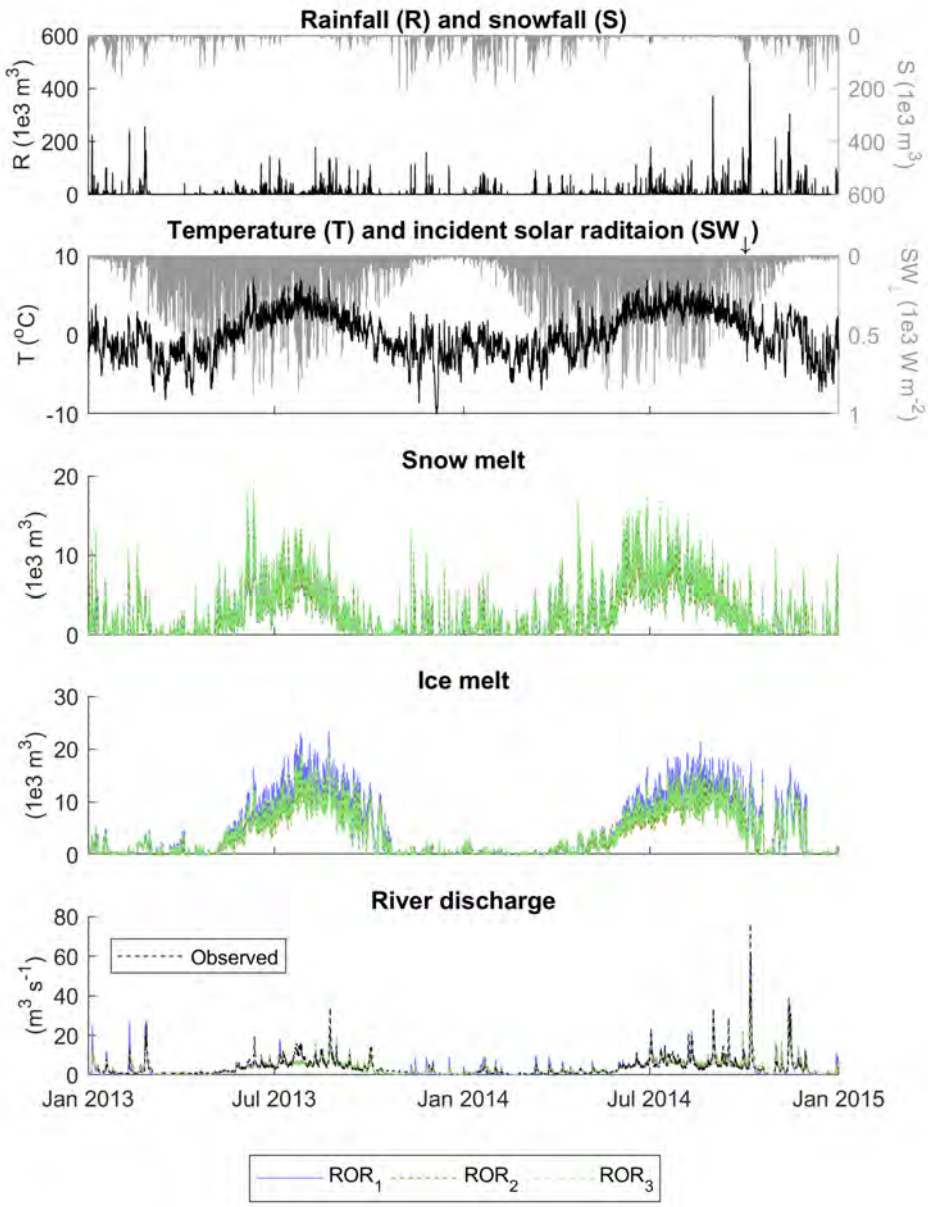

Figure E.2: Time series of driving precipitation, temperature and incident solar radiation data and simulated snow melt and ice melt and river discharge using the ROR_1, ROR_2 and ROR_3 runoff-routing model structures in conjunction with the simplest TIM_1 melt model structure. Note, the proportion of rainfall and snowfall is an output from the GHM which is approximately equal across the different configurations. Also note, ice melt includes melt of bare ice and the firn.

APPENDIX F:

REMOVAL OF GCM-RCM ENSEMBLE MEMBER

After comparing monthly average simulations from each GCM-RCM over the recent past (1981-2005) against the observed climate data, it was found that the [CNRM-CM5]-[ALADIN53] GCM-RCM had anomalously large negative temperature biases, particularly during the winter months of the year (see red line in Fig. F.1d). In addition to this, a RMSE score was calculated for each climate variable by comparing monthly observed and simulated empirical distribution functions constructed from catchment-average daily climate data (Fig. F.1a-c). When ranked according to their RMSE scores, the [CNRM-CM5]-[ALADIN53] GCM-RCM ranked 14, 13 and 15 out of 15. Given the anomalously high biases in temperature and the importance of temperature for driving hydrological change in the catchment (both in terms of melt rate and the proportion of precipitation falling as rainfall), coupled with the fact that the model was relatively poor across all three climate variables, this model was not deemed to be a reliable representation of future climate trends and was removed from the ensemble.

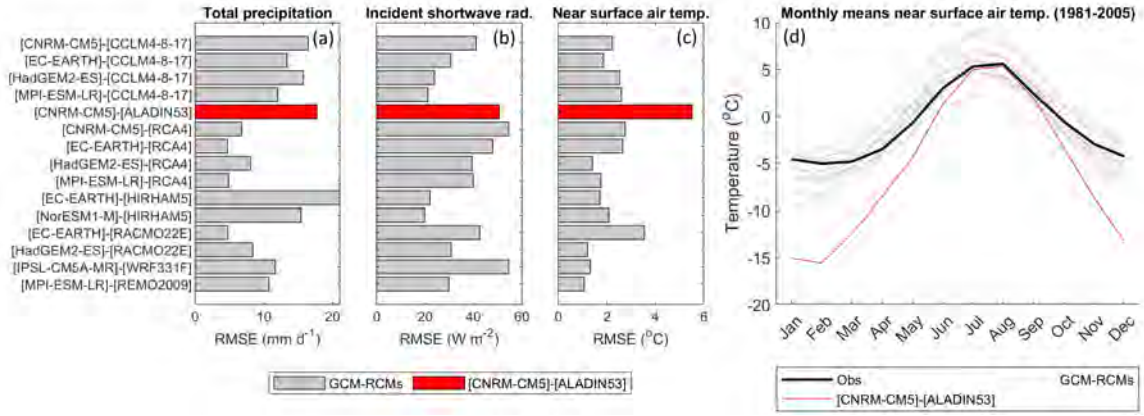

Figure F.1: RMSE scores calculated by comparing monthly empirical distribution functions constructed from catchment-average daily observed and simulated (GCM-RCM) total precipitation (a), incident solar radiation (b) and near-surface air temperature (c) data over the recent past (1981-2005). Also shown are the observed and simulated monthly mean near-surface air temperatures for the recent past (d).

APPENDIX G:

DECADAL CHANGES IN EFFECT SIZE FOR RIVER DISCHARGE SIGNATURES

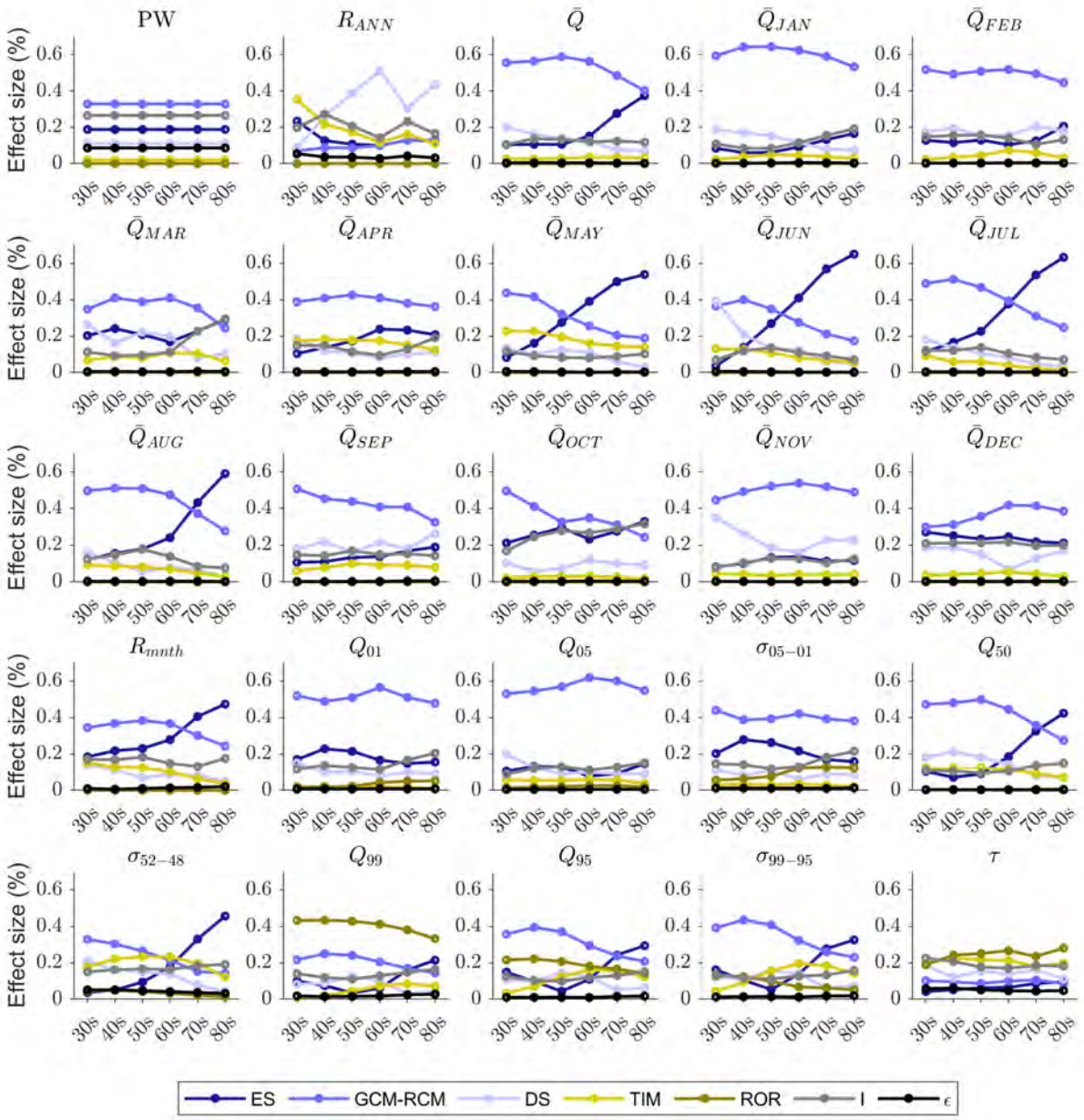

Figure G.1: Effect size of all main effects, interactions and remaining error on projected decadal changes in the 25 river discharge signatures for all future time slices centred on the 2030s to the 2080s.

APPENDIX H:

DELINEATION OF SANDUR BEDROCK TOPOGRAPHY

Using available observations of bedrock depth on the sandur, a map of bedrock elevation (Fig. H.1a) was derived to define the geometry of the sandur groundwater model. Observations of bedrock outcroppings at the north-east groundwater catchment boundary indicate that depth to bedrock here is at, or close to zero. Additionally, they also indicate the orientation (strike) of the bedrock is approximately perpendicular to the river channel. Two Tromino® passive seismic surveys conducted 1.1 and 2.4 km downstream of the north-east boundary (Fig. H.1b&c) provide additional observations of bedrock elevation. The lower survey shows a distinct high spectral ratio (Log H/V) zone indicating the bedrock resides approximately 90 m bgl. The orientation of this survey indicates the strike of the bedrock is beyond perpendicular to the river. The upper survey shows three distinct high spectral ratio zones, but interpretation of the bedrock topography is much more subjective here. Based on discussions with the authors of Ó Dochartaigh et al. (2019), the high spectral ratio zone in the centre of the survey is the most probable location of the bedrock giving a depth to bedrock of 70 m. However, variations in bedrock elevation along the survey are uncertain.

The observations therefore indicate the depth to bedrock increases downstream and that the strike of the bedrock is near or beyond perpendicular to the river which is in agreement with bedrock surveys undertaken in neighbouring catchments on Skeiðarársandur (Guðmundsson et al., 2002). To incorporate these observations into the MODFLOW-NWT groundwater model, a digital 3-D model of the bedrock geometry was constructed. Here, a transect was drawn through the catchment just beyond perpendicular to the river (red dash in Fig. H.1a). The bedrock elevation data were then imposed on that transect (red circles in Fig. H.1d) and a third-order polynomial was then fit to the bedrock elevation data to produce a continuous bedrock depth profile along the transect (blue line in

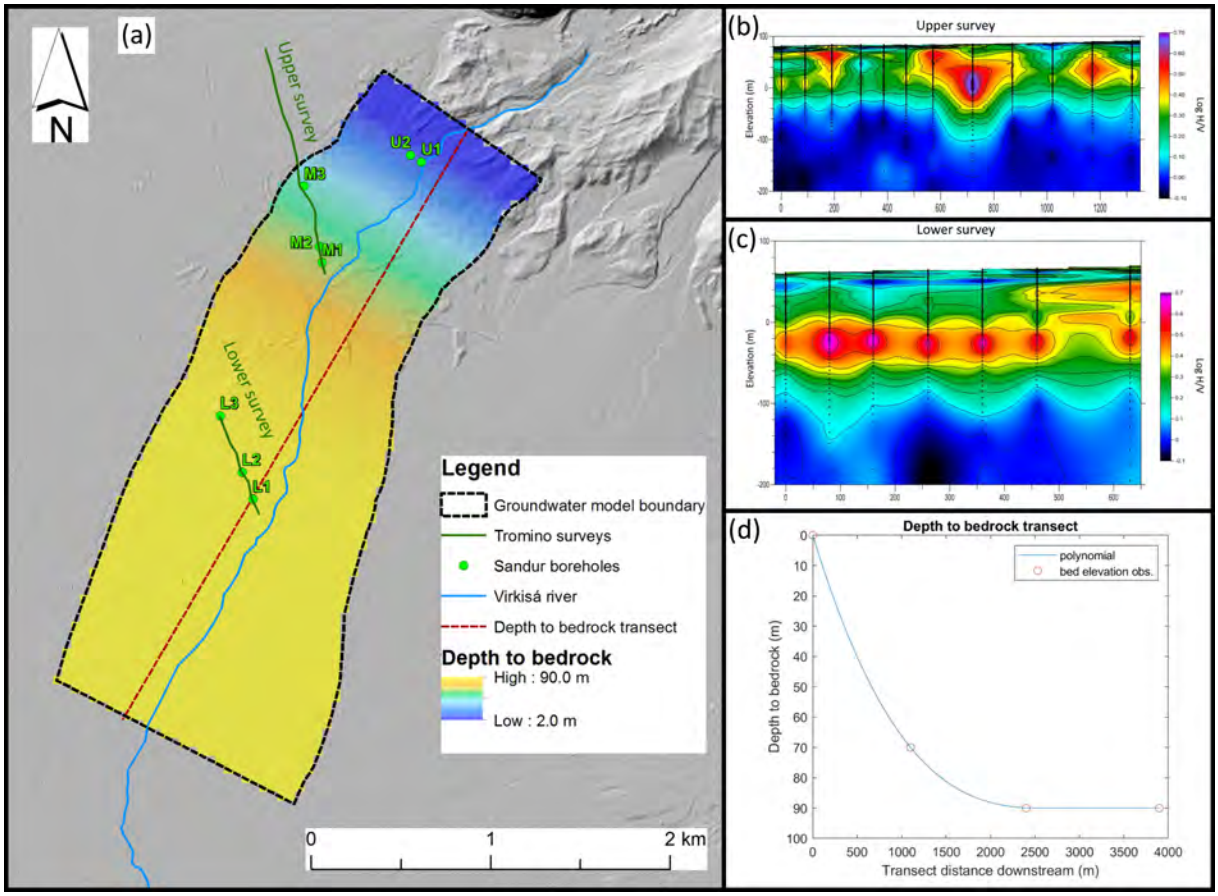

Figure H.1: Depth to bedrock map for sandur groundwater model (a), upper (b) and lower (c) tromino survey data and depth to bedrock transect (d).

Fig. H.1d). Note, a third-order polynomial was chosen based on trial and error experiments using orders ranging from one to five. Also note that at the lower boundary, the depth to bedrock was assumed to be close to that recorded from the lower transect. This was based on the conceptual model of Ó Dochartaigh et al. (2019), but no observation data exist to verify this. The resultant bedrock model used in the groundwater model is shown by the coloured Depth to bedrock map in Fig. H.1.

REFERENCES

Aðalgeirsdóttir, G., T. Jóhannesson, H. Björnsson, F. Pálsson, and O. Sigurðsson (2006). Response of Hofsjökull and southern Vatnajökull, Iceland, to climate change. In: *Journal of Geophysical Research* 111(F3), F03001.

Aðalgeirsdóttir, G., S. Guðmundsson, H. Björnsson, F. Pálsson, T. Jóhannesson, H. Hannesdóttir, S. Sigurðsson, and E. Berthier (2011). Modelling the 20th and 21st century evolution of Hoffellsjökull glacier, SE-Vatnajökull, Iceland. In: *The Cryosphere* 5(4), pp. 961–975.

Addor, N., O. Rössler, N. Köplin, M. Huss, R. Weingartner, and J. Seibert (2014). Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. In: *Water Resources Research* 50(10), pp. 7541–7562.

Ali, G., D. Tetzlaff, C. Soulsby, J. J. McDonnell, and R. Capell (2012). A comparison of similarity indices for catchment classification using a cross-regional dataset. In: *Advances in Water Resources* 40, pp. 11–22.

Allen, D. M., D. C. Mackie, and M. Wei (2004). Groundwater and climate change: A sensitivity analysis for the Grand Forks aquifer, southern British Columbia, Canada. In: *Hydrogeology Journal* 12(3), pp. 270–290.

Allen, D. M., P. H. Whitfield, and A. Werner (2010). Groundwater level responses in temperate mountainous terrain: Regime classification, and linkages to climate and streamflow. In: *Hydrological Processes* 24(23), pp. 3392–3412.

Allen, R., L. Pereira, D. Raes, and M. Smith (1998). *Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56.* Tech. rep. Rome, Italy: Food and Agriculture Organization of the United Nations.

Andermann, C., L. Longuevergne, S. Bonnet, A. Crave, P. Davy, and R. Gloaguen (2012). Impact of transient groundwater storage on the discharge of Himalayan rivers. In: *Nature Geoscience* 5(2), pp. 127–132.

Anderson, M. and W. Woessner (2002). *Applied groundwater modeling.* Elsevier: San Diego, USA, p. 70.

Andrés-Doménech, I., R. García-Bartual, A. Montanari, and J. B. Marco (2015). Climate and hydrological variability: The catchment filtering role. In: *Hydrology and Earth System Sciences* 19(1), pp. 379–387.

Arnold, N, K Richards, I Willis, and M. J. Sharp (1998). Initial results from a distributed , physically based model of glacier hydrology. In: *Hydrological Processes* 12, pp. 191–219.

Arnold, N. S., W. G. Rees, A. J. Hodson, and J. Kohler (2006). Topographic controls on the surface energy balance of a high Arctic valley glacier. In: *Journal of Geophysical Research: Earth Surface* 111(2), F02011.

Baraer, M., J. Mckenzie, B. G. Mark, R. Gordon, J. Bury, T. Condom, J. Gomez, S. Knox, and S. K. Fortner (2015). Contribution of groundwater to the outflow from ungauged glacierized catchments: A multi-site study in the tropical Cordillera Blanca, Peru. In: *Hydrological Processes* 29(11), pp. 2561–2581.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. In: *Nature* 438(7066), pp. 303–309.

Barrand, N. E., T. Murray, T. D. James, S. L. Barr, and J. P. Mills (2009). Optimizing photogrammetric DEMs for glacier volume change assessment using laser-scanning derived ground-control points. In: *Journal of Glaciology* 55(189), pp. 106–116.

Bartók, B., M. Wild, D. Folini, D. Lüthi, S. Kotlarski, C. Schär, R. Vautard, S. Jerez, and Z. Imecs (2017). Projected changes in surface solar radiation in CMIP5 global climate models and in EURO-CORDEX regional climate models for Europe. In: *Climate Dynamics* 49(7-8), pp. 2665–2683.

Beamer, J. P., D. F. Hill, D. McGrath, A. Arendt, and C. Kienholz (2017). Hydrologic impacts of changes in climate and glacier extent in the Gulf of Alaska watershed. In: *Water Resources Research* 53(9), pp. 7502–7520.

Beer, C., P. Porada, A. Ekici, and M. Brakebusch (2018). Effects of short-term variability of meteorological variables on soil temperature in permafrost regions. In: *The Cryosphere* 12, pp. 741–757.

Bengtsson, L. et al. (2017). The HARMONIE-AROME model configuration in the ALADIN-HIRLAM NWP system. In: *Monthly Weather Review* 145, pp. 1919–1935.

Beniston, M. et al. (2017). The European mountain cryosphere: A review of past, current and future issues. In: *The Cryosphere Discussions*(January), pp. 1–60. URL: http://www.the-cryosphere-discuss.net/tc-2016-290/.

Beven, K. (2006). A manifesto for the equifinality thesis. In: *Journal of Hydrology* 320(1-2), pp. 18–36.

Beven, K. (2016). Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. In: *Hydrological Sciences Journal* 61(9), pp. 1652–1665.

Björnsson, H, F Pálsson, M. T. Guðmundsson, and H. H. Haraldsson (1998). Mass balance of western and northern Vatnajökull, Iceland, 1991-1995. In: *Jökull* 45(45), pp. 35–58.

Bjornsson, H. (1982). Drainage Basins on Vatnajokull Mapped by Radio Echo Soundings. In: *Nordic Hydrology* 13(4), pp. 213–232.

Björnsson, H. and F. Pálsson (2008). Icelandic glaciers. In: *Jökull* 58(58), pp. 365–386.

Björnsson, H., F. Pálsson, S. Gudmundsson, E. Magnússon, G. Adalgeirsdõttir, T. Jõhannesson, E. Berthier, O. Sigurdsson, and T. Thorsteinsson (2013). Contribution of Icelandic ice caps to sea level rise: Trends and variability since the Little Ice Age. In: *Geophysical Research Letters* 40(8), pp. 1546–1550.

Blaen, P. J., D. M. Hannah, L. E. Brown, and A. M. Milner (2014). Water temperature dynamics of high Arctic river basins. In: *Hydrological Processes* 28, pp. 3521–3538.

Blaney, H. F. and K. V. Morin (1942). Evaporation and consumptive use of water empirical formulas. In: *Eos, Transactions American Geophysical Union* 23(1), pp. 76–83.

Blazkova, S. and K. Beven (2009). A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic. In: *Water Resources Research* 45(12), W00B16.

Bliss, A., R. Hock, and V. Radić (2014). Global response of glacier runoff to twenty-first century climate change. In: *Journal of Geophysical Research* 119, pp. 717–730.

Blöschl, G. et al. (2019). Twenty-three unsolved problems in hydrology (UPH)–a community perspective. In: *Hydrological Sciences Journal* 64(10), pp. 1141–1158.

Boscarello, L., G. Ravazzani, D. Rabuffetti, and M. Mancini (2014). Integrating glaciers raster-based modelling in large catchments hydrological balance: The Rhone case study. In: *Hydrological Processes* 28(3), pp. 496–508.

Bosshard, T., M. Carambia, K. Goergen, S. Kotlarski, P. Krahe, M. Zappa, and C. Schär (2013). Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. In: *Water Resources Research* 49(3), pp. 1523–1536.

Bradwell, T., O. Sigurdsson, and J. Everest (2013). Recent, very rapid retreat of a temperate glacier in SE Iceland. In: *Boreas* 42(4), pp. 959–973.

Braithwaite, R. J. (1995). Positive degree-day factors for ablation on the Greenland Ice-sheet studied by energy balance modeling. In: *Journal of Glaciology* 41(137), pp. 153–160.

Brately, P. and B. L. Fox (1988). Algorithm 659: Implementing Sobol's quasirandom sequence generator. In: *ACM Transactions on Mathematical Software* 14(1), pp. 88–100.

Brock, B. W., I. C. Willis, and M. J. Sharp (2000). Measurement and parameterisation of albedo variations at Haut Glacier d'Arolla , Switzerland. In: *Journal of Glaciology* 46(155), pp. 675–688.

Brown, L. E. and D. Hannah (2008). Spatial heterogeneity of water temperature across an alpine river basin. In: *Hydrological Processes* 22, pp. 954–967.

Brown, L. E., D. M. Hannah, A. M. Milner, C. Soulsby, A. J. Hodson, and M. J. Brewer (2006). Water source dynamics in a glacierized alpine river basin (Taillon-Gabiétous, French Pyrénées). In: *Water Resources Research* 42(8), W08404.

Brown, L. E., A. M. Milner, and D. M. Hannah (2007). Groundwater influence on alpine stream ecosystems. In: *Freshwater Biology* 52(5), pp. 878–890.

Brown, L. E., D. M. Hannah, and A. M. Milner (2009). ARISE: A classification tool for Alpine RIver and Stream Ecosystems. In: *Freshwater Biology* 54(6), pp. 1357–1369.

Bunn, S. E. and A. H. Arthington (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. In: *Environmental Management* 30(4), pp. 492–507.

Buytaert, W. and S. Domzalski (2015). *Climate change impacts on water resources in the Tropical Andes: Prioritizing scientific research for developing adaptation policies.* Tech. rep. Paris, France: UNESCO International Hydrological Programme (IHP).

Buytaert, W., S. Moulds, L. Acosta, B. D. Bièvre, C. Olmos, M. Villacis, C. Tovar, and K. M. J. Verbist (2017). Glacial melt content of water use in the tropical Andes. In: *Environmental Research Letters* 12(11), p. 114014.

Cannon, A. J. (2015). Selecting GCM scenarios that span the range of changes in a multimodel ensemble: Application to CMIP5 climate extremes indices. In: *Journal of Climate* 28(3), pp. 1260–1267.

Cannon, A. J., S. R. Sobie, and T. Q. Murdock (2015). Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? In: *Journal of Climate* 28(17), pp. 6938–6959.

Carenzo, M, F Pellicciotti, J Mabillard, T Reid, and B. W. Brock (2016). An enhanced temperature index model for debris-covered glaciers accounting for thickness effect. In: *Advances in Water Resources* 94, pp. 457–469.

Carey, M., M. Baraer, B. G. Mark, A. French, J. Bury, K. R. Young, and J. M. McKenzie (2014). Toward hydro-social modeling: Merging human variables and the social sciences with climate-glacier runoff models (Santa River, Peru). In: *Journal of Hydrology* 518(PA), pp. 60–70.

Carey, M., O. C. Molden, M. B. Rasmussen, M Jackson, A. W. Nolin, and B. G. Mark (2016). Impacts of Glacier Recession and Declining Meltwater on Mountain Societies. In: *Annals of the American Association of Geographers* 107(2), pp. 350–359.

Cartwright, J., C. Caldwell, S. Nebiker, and R. Knight (2017). Putting flow-ecology relationships into practice: A decision-support system to assess fish community response towater-management scenarios. In: *Water (Switzerland)* 9(3), pp. 13–16.

Carvajal, P. E., G. Anandarajah, Y. Mulugetta, and O. Dessens (2017). Assessing uncertainty of climate change impacts on long-term hydropower generation using the CMIP5 ensemble—the case of Ecuador. In: *Climatic Change* 144, pp. 611–624.

Casper, M. C., G. Grigoryan, O. Gronz, O. Gutjahr, G. Heinemann, R. Ley, and A. Rock (2012). Analysis of projected hydrological behavior of catchments based on signature indices. In: *Hydrology and Earth System Sciences* 16(2), pp. 409–421.

Cayar, M. and M. L. Kavvas (2009). Ensemble average and ensemble variance behavior of unsteady, one-dimensional groundwater flow in unconfined, heterogeneous aquifers: An exact second-order model. In: *Stochastic Environmental Research and Risk Assessment* 23(7), pp. 947–956.

Chen, J. and A. Ohmura (1990). On the influence of Alpine glaciers on runoff. In: *IAHS Publication (Hydrology in Mountainous Regions. I-Hydrological Measurements; the Water Cycle)* 193, pp. 117–126.

Chen, X., J. Song, and W. Wang (2010). Spatial variability of specific yield and vertical hydraulic conductivity in a highly permeable alluvial aquifer. In: *Journal of Hydrology* 388(3-4), pp. 379–388.

Chow, V. T. (1959). *Open-channel hydraulics*. McGraw-Hill International Editions: Singapore, pp. 110–113.

Ciarapica, L. and E. Todini (2002). TOPKAPI: A model for the representation of the rainfall-runoff process at different scales. In: *Hydrological Processes* 16(2), pp. 207–229.

Clark, M. P. et al. (2015). A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. In: pp. 1–28.

Clausen, B. and B. J. F. Biggs (2000). Flow variables for ecological studies in temperate streams: Groupings based on covariance. In: *Journal of Hydrology* 237(3-4), pp. 184–197.

Clow, D. W., L Schrott, R Webb, D. H. Campbell, A Torizzo, and M Dornblaser (2003). Ground water occurrence and contributions to streamflow in an alpine catchment, Colorado Front Range. In: *Ground Water - Watershed Issue* 41(7), pp. 937–950.

Collins, M. et al. (2013). "Long-term Climate Change: Projections, Commitments and Irreversibility". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley. Cambridge University Press: Cambridge, United Kingdom and New York, USA, pp. 1029–1136.

Covino, T. P. and B. L. McGlynn (2007). Stream gains and losses across a mountain-to-valley transition: Impacts on watershed hydrology and stream water chemistry. In: *Water Resources Research* 43(10).

Coxon, G., J. Freer, T. Wagener, N. A. Odoni, and M. Clark (2014). Diagnostic evaluation of multiple hypotheses of hydrological behaviour in a limits-of-acceptability framework for 24 UK catchments. In: *Hydrological Processes* 28(25), pp. 6135–6150.

Coxon, G, J Freer, I. K. Westerberg, T Wagener, R Woods, and P. J. Smith (2015). A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. In: *Water Resources Research* 51(7), pp. 5531–5546.

Crochet, P. and T. Jóhannesson (2011). A data set of gridded daily temperature in Iceland, 1949-2010. In: *Jökull* 61, pp. 1–17.

Crochet, P., T. Jóhannesson, T. Jónsson, O. Sigurðsson, H. Björnsson, F. Pálsson, and I. Barstad (2007). Estimating the Spatial Distribution of Precipitation in Iceland Using a Linear Model of Orographic Precipitation. In: *Journal of Hydrometeorology* 8(6), pp. 1285–1306.

Crossman, J., C. Bradley, J. N. W. David, and a. M. Milner (2012). Use of remote sensing to identify areas of groundwater upwelling on active glacial floodplains: Their frequency, extent and significance on a landscape scale. In: *Remote Sensing of Environment* 123, pp. 116–126.

Cuffey, K. and W. Paterson (2010). *The physics of glaciers: Fourth edition*. Elsevier: Oxford, UK.

Daron, J. D. and D. A. Stainforth (2013). On predicting climate under climate change. In: *Environmental Research Letters* 8(3), p. 034021.

Domenico, P. and F. Schwartz (1990). *Physical and Chemical Hydrogeology*. John Wiley & Sons, Ltd: New York, p. 824.

Dragon, K., M. Marciniak, J. Szpikowski, G. Szpikowska, and T. Wawrzyniak (2015). The hydrochemistry of glacial Ebba River (Petunia Bay, Central Spitsbergen): Groundwater influence on surface water chemistry. In: *Journal of Hydrology* 529, pp. 1499–1510.

Duethmann, D., C. Menz, T. Jiang, and S. Vorogushyn (2016). Projections for headwater catchments of the Tarim River reveal glacier retreat and decreasing surface water availability but uncertainties are large. In: *Environmental Research Letters* 11(5), p. 054024.

Dyurgerov, M., M. Meier, and R. L. Armstrong (2002). *Glacier mass balance and regime: data of measurements and analysis*. Institute of Arctic and Alpine Research, University of Colorado Boulder, USA.

Einarsson, M. Á. (1972). *Evaporation and potential evapotranspiration in Iceland*. Veðurstofa Íslands: Reykjavík, pp. 1–22.

Einarsson, M. Á. (1984). "Climate of Iceland". In: *Climates of the Oceans*. Ed. by H van Loon. Vol. 15. World Survey of Climatology. Elsevier, pp. 673–697.

Euser, T., H. C. Winsemius, M. Hrachowitz, F. Fenicia, S. Uhlenbrook, and H. H. G. Savenije (2013). A framework to assess the realism of model structures using hydrological signatures. In: *Hydrology and Earth System Sciences* 17(5), pp. 1893–1912.

Everest, J., T. Bradwell, L. Jones, and L. Hughes (2017). The geomorphology of svínafellsjökull and virkisjökull-falljökull glacier forelands, Southeast Iceland. In: *Journal of Maps* 13(2), pp. 936–945.

Farinotti, D., S. Usselmann, M. Huss, A. Bauder, and M. Funk (2012). Runoff evolution in the Swiss Alps: projections for selected high-alpine catchments based on ENSEMBLES scenarios. In: *Hydrological Processes* 26(13), pp. 1909–1924.

Finger, D., F. Pellicciotti, M. Konz, S. Rimkus, and P. Burlando (2011). The value of glacier mass balance, satellite snow cover images, and hourly discharge for improving the performance of a physically based distributed hydrological model. In: *Water Resources Research* 47(7), W07519.

Finger, D., M. Vis, M. Huss, and J. Seibert (2015). The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments. In: *Water Resources Research* 51(4), pp. 1939–1958.

Flett, V., L. Maurice, A. Finlayson, A. R. Black, A. M. MacDonald, J. Everest, and M. P. Kirkbride (2017). Meltwater flow through a rapidly deglaciating glacier and foreland catchment system: Virkisjökull, SE Iceland. In: *Hydrology Research* 48(6), pp. 1666–1681.

Flett, V. T. (2016). "Glacier retreat and projected river regime changes in the hydrologically highly-coupled Virkisjökull catchment , SE Iceland". Doctor of Philosophy. University of Dundee.

Fountain, A. G. and W. V. Tangborn (1985). The Effect of Glaciers on Streamflow Variations. In: *Water Resources Research* 21(4), pp. 579–586.

Frey, S. and H. Holzmann (2015). A conceptual, distributed snow redistribution model. In: *Hydrology and Earth System Sciences* 19(11), pp. 4517–4530.

Frisbee, M. D., D. G. Tolley, and J. L. Wilson (2017). Field estimates of groundwater circulation depths in two mountainous watersheds in the western U.S. and the effect of deep circulation on solute concentrations in streamflow. In: *Water Resources Research* 53, WR019553.

Fyke, J. and H. D. Matthews (2015). A probabilistic analysis of cumulative carbon emissions and long-term planetary warming. In: *Environmental Research Letters* 10(11), p. 115007.

Gabbi, J., D. Farinotti, A. Bauder, and H. Maurer (2012). Ice volume distribution and implications on runoff projections in a glacierized catchment. In: *Hydrology and Earth System Sciences* 16(12), pp. 4543–4556.

Gabbi, J., M. Carenzo, F. Pellicciotti, A. Bauder, and M. Funk (2014). A comparison of empirical and physically based glacier surface melt models for long-term simulations of glacier response. In: *Journal of Glaciology* 60(224), pp. 1140–1154.

Gao, H., Y. Ding, Q. Zhao, M. Hrachowitz, and H. H. Savenije (2017). The importance of aspect for modelling the hydrological response in a glacier catchment in Central Asia. In: *Hydrological Processes* 31(16), pp. 2842–2859.

Garavaglia, F., M. Le Lay, F. Gottardi, R. Garçon, J. Gailhard, E. Paquet, and T. Mathevet (2017). Impact of model structure on flow simulation and hydrological realism: from lumped to semi-distributed approach. In: *Hydrology and Earth System Sciences* 21, pp. 3937–3952.

Gardner, A. S. and M. Sharp (2009). Sensitivity of net mass-balance estimates to near-surface temperature lapse rates when employing the degree-day method to estimate glacier melt. In: *Annals of Glaciology* 50(50), pp. 80–86.

Gardner, A. S., M. J. Sharp, R. M. Koerner, C. Labine, S. Boon, S. J. Marshall, D. O. Burgess, and D. Lewis (2009). Near-surface temperature lapse rates over arctic glaciers and their implications for temperature downscaling. In: *Journal of Climate* 22(16), pp. 4281–4298.

Garee, K., X. Chen, A. Bao, Y. Wang, and F. Meng (2017). Hydrological Modeling of the Upper Indus Basin: A Case Study from a High-Altitude Glacierized Catchment Hunza. In: *Water* 9(17), pp. 1–20.

Gaudard, L., F. Romerio, F. Dalla Valle, R. Gorret, S. Maran, G. Ravazzani, M. Stoffel, and M. Volonterio (2014). Climate change impacts on hydropower in the Swiss and Italian Alps. In: *Science of the Total Environment* 493, pp. 1211–1221.

Ghasemizade, M. and M. Schirmer (2013). Subsurface flow contribution in the hydrological cycle: Lessons learned and challenges ahead-a review. In: *Environmental Earth Sciences* 69(2), pp. 707–718.

Giorgi, F, C Jones, and G. R. Asrar (2009). Addressing climate information needs at the regional level: The CORDEX framework. In: *World Meteorological Organization Bulletin* 58(3), pp. 175–183.

Giuntoli, I., J. P. Vidal, C. Prudhomme, and D. M. Hannah (2015). Future hydrological extremes: The uncertainty from multiple global climate and global hydrological models. In: *Earth System Dynamics* 6(1), pp. 267–285.

Gosseling, M. (2017). *CORDEX climate trends for Iceland in the 21st century*. Tech. rep. Reykjavik, Iceland: Icelandic Meteorological Office.

Gray, D., M. R. Scarsbrook, and J. S. Harding (2006). Spatial biodiversity patterns in a large New Zealand braided river. In: *New Zealand Journal of Marine and Freshwater Research* 40(4), pp. 631–642.

Griffiths, J., V. Keller, D. Morris, and A. Young (2006). *Continuous Estimation of River Flows (CERF) : Model Scheme for Representing Rainfall Interception and Soil Moisture. Environment Agency R & D Project W6- 101*. Tech. rep. Wallingford, UK: Centre for Ecology and Hydrology.

Griffiths, J, V Keller, D Morris, and a. R. Young (2008). *Continuous Estimation of River Flows (CERF)*. Tech. rep. Bristol, UK: Environment Agency, pp. 15–16.

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen (2012). Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods. In: *Hydrology and Earth System Sciences* 16(9), pp. 3383–3390.

Gudmundsson, S., H. Björnsson, E. Magnússon, E. Berthier, F. Pálsson, M. T. Gudmundsson, T. Högnadóttir, and J. Dall (2011). Response of Eyjafjallajökull, Torfajökull and Tindfjallajökull ice caps in Iceland to regional warming, deduced by remote sensing. In: *Polar Research* 30, p. 7282.

Guido, Z., J. C. McIntosh, S. A. Papuga, and T. Meixner (2016). Seasonal glacial meltwater contributions to surface water in the Bolivian Andes: A case study using environmental tracers. In: *Journal of Hydrology: Regional Studies* 8, pp. 260–273.

Guðmundsson, M. T. (2000). Mass balance and precipitation on the summit plateau of Öræfajökull, SE-Iceland. In: *Jökull* 48(48), pp. 49–54.

Guðmundsson, M. T., A Bonnel, and K Gunnarsson (2002). Seismic soundings of sediment thickness on Skeidararsandur, SE-Iceland. In: *Jökull* 51(51), pp. 53–64.

Guðmundsson, S., H. Björnsson, T. Jóhannesson, G. Aðalgeirsdóttir, F. Pálsson, and O. Sigurðsson (2009). Similarities and differences in the response to climate warming of two ice caps in Iceland. In: *Hydrology Research* 40(5), pp. 495–502.

Gupta, H. V., T. Wagener, and Y. Liu (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. In: *Hydrological Processes* 22(18), pp. 3802–3813.

Hannah, D. M. and A. M. Gurnell (2001). A conceptual, linear reservoir runoff model to investigate melt season changes in cirque glacier hydrology. In: *Journal of Hydrology* 246(1-4), pp. 123–141.

Hannesdóttir, H., H. Björnsson, F. Pálsson, G. Aðalgeirsdóttir, and S. Guðmundsson (2015a). Changes in the southeast Vatnajökull ice cap, Iceland, between ˜ 1890 and 2010. In: *The Cryosphere* 9(2), pp. 565–585.

Hannesdóttir, H., G. Aðalgeirsdóttir, T. Jóhannesson, S. Guðmundsson, P. Crochet, H. Ágústsson, F. Pálsson, E. Magnússon, S. Sigurðsson, and H. Björnsson (2015b). Downscaled precipitation applied in modelling of mass balance and the evolution of southeast Vatnajökull, Iceland. In: *Journal of Glaciology* 61(228), pp. 799–813.

Hanzer, F., K. Helfricht, T. Marke, and U. Strasser (2016). Multilevel spatiotemporal validation of snow/ice mass balance and runoff modeling in glacierized catchments. In: *Cryosphere* 10(4), pp. 1859–1881.

Harbaugh, A. (2005). *MODFLOW-2005, The U.S. Geological Survey modular groundwater model—the Ground-Water Flow Process: U.S. Geological Survey Techniques and Methods 6-A16*. Tech. rep. Reston, Virginia: U.S. Geological Survey.

Hariharan, V. and M. Uma Shankar (2017). A review of visual modflow applications in groundwater modelling. In: *IOP Conference Series: Materials Science and Engineering* 263, p. 032025.

Hernández-Henríquez, M. A., A. R. Sharma, and S. J. Déry (2017). Variability and trends in runoff in the rivers of British Columbia's Coast and Insular Mountains. In: *Hydrological Processes* 31(December 2016), pp. 3269–3282.

Herschy, R. W. (1999). *Hydrometry: Principles and Practice, 2nd Edition*. Ed. by R. W. Herschy. John Wiley & Sons, Ltd: Chichester, UK, pp. 9–83.

Heynen, M., F. Pellicciotti, and M. Carenzo (2013). Parameter sensitivity of a distributed enhanced temperature-index melt model. In: *Annals of Glaciology* 54(63), pp. 311–321.

Hock, R. (1999). A distributed temperature-index ice- and snowmelt model including potential direct solar radiation. In: *Journal of Glaciology* 45(149), pp. 101–111.

Hock, R. and P. Jansson (2005). "Modeling Glacier Hydrology". In: *Encyclopedia of Hydrological Sciences 4*. Ed. by M. G. Anderson and J. J. McDonnell. John Wiley & Sons, Ltd: Chichester, UK, pp. 2647–2655.

Hood, J. L. and M. Hayashi (2015). Characterization of snowmelt flux and groundwater storage in an alpine headwater basin. In: *Journal of Hydrology* 521, pp. 482–497.

Hopkinson, C., L. Chasmer, S. Munro, and M. N. Demuth (2010). The influence of DEM resolution on simulated solar radiation-induced glacier melt. In: *Hydrological Processes* 24(6), pp. 775–788.

Hotaling, S., D. S. Finn, J. Joseph Giersch, D. W. Weisrock, and D. Jacobsen (2017). Climate change and alpine stream biology: Progress, challenges, and opportunities for the future. In: *Biological Reviews* 2033, pp. 2024–2045.

Hrachowitz, M., O. Fovet, L. Ruiz, T. Euser, S. Gharari, R. Nijzink, J. Freer, H. Savenije, and C. Gascuel-Odoux (2014). Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. In: *Water Resources Research* 50, pp. 7445–7469.

Huntington, J. L. and R. G. Niswonger (2012). Role of surface-water and groundwater interactions on projected summertime streamflow in snow dominated regions: An integrated modeling approach. In: *Water Resources Research* 48(11), pp. 1–20.

Huss, M., G. Jouvet, D. Farinotti, and A. Bauder (2010). Future high-mountain hydrology: a new parameterization of glacier retreat. In: *Hydrology and Earth System Sciences* 14(5), pp. 815–829.

Huss, M. and R. Hock (2015). A new model for global glacier change and sea-level rise. In: *Frontiers in Earth Science* 3(September), pp. 1–22.

Huss, M. and R. Hock (2018). Global-scale hydrological response to future glacier mass loss. In: *Nature Climate Change* 8(February), pp. 135–140.

Huss, M., A. Bauder, M. Funk, and R. Hock (2008a). Determination of the seasonal mass balance of four Alpine glaciers since 1865. In: *Journal of Geophysical Research: Earth Surface* 113(F1), F01015.

Huss, M., D. Farinotti, A. Bauder, and M. Funk (2008b). Modelling runoff from highly glacierized alpine drainage basins in a changing climate. In: *Hydrological Processes* 22(19), pp. 3888–3902.

Huss, M., M. Zemp, P. C. Joerg, and N. Salzmann (2014). High uncertainty in 21st century runoff projections from glacierized basins. In: *Journal of Hydrology* 510, pp. 35–48.

ICIMOD (2010). *Islamic Republic of Pakistan: Glacial Melt and Downstream Impacts on Indus Dependent Water Resources and Energy*. Tech. rep. July. Kathmandu: International Centre for Integrated Mountain Development.

IGS (2017). *Icelandic Glaciological Society Terminus monitoring*. URL: http://spordakost.jorfi.is (visited on 06/16/2017).

IHA (2019). *Hydropower Sector: Climate Resilience Guide*. Tech. rep. London, UK: International Hydropower Association Limited.

Immerzeel, W. W., F. Pellicciotti, and M. F. P. Bierkens (2013). Rising river flows throughout the twenty-first century in two Himalayan glacierized watersheds. In: *Nature Geoscience* 6(9), pp. 742–745.

Immerzeel, W. W., L. Petersen, S. Ragettli, and F Pellicciotti (2014). The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. In: *Water Resources Research* 50(3), pp. 2212–2226.

Immerzeel, W. W., L. P. H. van Beek, and M. F. P. Bierkens (2010). Climate Change Will Affect the Asian Water Towers. In: *Science* 328(5984), pp. 1382–1385.

IPCC (2018). *Summary for Policymakers*. Tech. rep.

Irvine-Fynn, T. D. L., E Hanna, N. E. Barrand, P. R. Porter, J Kohler, and A. J. Hodson (2014). Examination of a physically based , high-resolution , distributed Arctic

temperature-index melt model , on Midtre Lovénbreen , Svalbard. In: *Hydrological Processes* 28, pp. 134–149.

Islam, S. U. and S. J. Déry (2017). Evaluating uncertainties in modelling the snow hydrology of the Fraser River Basin, British Columbia, Canada. In: *Hydrology and Earth System Sciences* 21, pp. 1827–1847.

Jackson, C., L. Wang, M. Pachocka, J. Mackay, and J. Bloomfield (2016). Reconstruction of multi-decadal groundwater level time-series using a lumped conceptual model. In: *Hydrological Processes* 30(18), pp. 3107–3125.

Jacob, D. et al. (2014). EURO-CORDEX: New high-resolution climate change projections for European impact research. In: *Regional Environmental Change* 14(2), pp. 563–578.

Jacobsen, D., A. M. Milner, L. E. Brown, and O. Dangles (2012). Biodiversity under threat in glacier-fed river systems. In: *Nature Climate Change* 2(5), pp. 361–364.

Jakob Themeßl, M., A. Gobiet, and A. Leuprecht (2011). Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. In: *International Journal of Climatology* 31(10), pp. 1530–1544.

Jansson, P., R. Hock, and T. Schneider (2003). The concept of glacier storage: a review. In: *Journal of Hydrology* 282(1-4), pp. 116–129.

Jehn, F. U., K. Bestian, L. Breuer, P. Kraft, and T. Houska (2019). Clustering CAMELS using hydrological signatures with high spatial predictability. In: *Hydrology and Earth System Sciences Discussions*.

Jobst, A. M., D. G. Kingston, N. J. Cullen, and J. Schmid (2018). Intercomparison of different uncertainty sources in hydrological climate change projections for an alpine catchment (upper Clutha River, New Zealand). In: *Hydrology and Earth System Sciences* 22(6), pp. 3125–3142.

Jódar, J., J. A. Cabrera, S. Martos-Rosillo, A. Ruiz-Constán, A. González-Ramón, L. J. Lambán, C. Herrera, and E. Custodio (2017). Groundwater discharge in high-mountain watersheds: A valuable resource for downstream semi-arid zones. The case of the Bérchules River in Sierra Nevada (Southern Spain). In: *Science of the Total Environment* 593-594, pp. 760–772.

Jóhannesson, G. T. and E. H. Huijbens (2010). Tourism in times of crisis: exploring the discourse of tourism development in Iceland. In: *Current Issues in Tourism* 13(5), pp. 419–434.

Johannesson, T., O. Sigurdsson, T. Laumann, and M. Kennett (1995). Degree-day glacier mass-balance modelling with applications to glaciers in Iceland, Norway and Greenland. In: *Journal of Glaciology* 41(138), pp. 345–358.

Jóhannesson, T. et al. (2006). The impact of climate change on glaciers and glacial runoff in the Nordic countries. In: *European Conference on Impacts of Climate Change on Renewable Energy Sources*, pp. 1–7.

Jóhannesson, T., G. Aðalgeirsdóttir, H. Björnsson, and P. Crochet (2007). *Effect of climate change on hydrology and hydro-resources in Iceland*. Tech. rep. Orkugarður: National Energy Authority, pp. 1–91.

Jóhannesson, T., H. Björnsson, E. Magnússon, S. Guðmundsson, F. Pálsson, O. Sigurðsson, T. Thorsteinsson, and E. Berthier (2013). Ice-volume changes, bias estimation of mass-balance measurements and changes in subglacial lakes derived by lidar mapping of the surface of Icelandic glaciers. In: *Annals of Glaciology* 54(63), pp. 63–74.

Jost, G., R. D. Moore, B. Menounos, and R. Wheate (2012). Quantifying the contribution of glacier runoff to streamflow in the upper Columbia River Basin, Canada. In: *Hydrology and Earth System Sciences* 16(3), pp. 849–860.

Khamis, K., L. E. Brown, D. M. Hannah, and A. M. Milner (2016). Glacier–groundwater stress gradients control alpine river biodiversity. In: *Ecohydrology* 9(7), pp. 1263–1275.

Kiesel, J., B. Guse, M. Pfannerstill, K. Kakouei, S. C. Jähnig, and N. Fohrer (2017). Improving hydrological model optimization for riverine species. In: *Ecological Indicators* 80(April), pp. 376–385.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010). Challenges in combining projections from multiple climate models. In: *Journal of Climate* 23(10), pp. 2739–2758.

Kobierska, F., T. Jonas, J. W. Kirchner, and S. M. Bernasconi (2015). Linking baseflow separation and groundwater storage dynamics in an alpine basin (Dammagletscher, Switzerland). In: *Hydrology and Earth System Sciences* 19(8), pp. 3681–3693.

Kobierska, F., T. Jonas, M. Zappa, M. Bavay, J. Magnusson, and S. M. Bernasconi (2013). Future runoff from a partly glacierized watershed in Central Switzerland: A two-model approach. In: *Advances in Water Resources* 55, pp. 204–214.

Konya, K., T. Matsumoto, and R. Naruse (2004). Surface heat balance and spatially distributed ablation modelling at Koryto Glacier, Kamchatka peninsula, Russia. In: *Geografiska Annaler* 86 A(4), pp. 337–348.

Konz, M. and J. Seibert (2010). On the value of glacier mass balances for hydrological model calibration. In: *Journal of Hydrology* 385(1-4), pp. 238–246.

La Frenierre, J. and B. G. Mark (2014). A review of methods for estimating the contribution of glacial meltwater to total watershed discharge. In: *Progress in Physical Geography* 38(2), pp. 173–200.

Laghari., J. R. (2013). Melting glaciers bring energy uncertainty. In: *Nature* 502, pp. 617–618.

Langston, G., M. Hayashi, and J. W. Roy (2013). Quantifying groundwater-surface water interactions in a proglacial moraine using heat and solute tracers. In: *Water Resources Research* 49(October 2012), pp. 5411–5426.

Levy, A., Z. Robinson, S. Krause, R. Waller, and J. Weatherill (2015). Long-term variability of proglacial groundwater-fed hydrological systems in an area of glacier retreat, Skeiðarársandur, Iceland. In: *Earth Surface Processes and Landforms* 40, pp. 981–994.

Li, H., J. Sheffield, and E. F. Wood (2010). Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. In: *Journal of Geophysical Research Atmospheres* 115(D10), p. D10101.

Liljedahl, A. K., A Gädeke, S O'Neel, T. A. Gatesman, and T. A. Douglas (2017). Glacierized headwater streams as aquifer recharge corridors, subarctic Alaska. In: *Geophysical Research Letters* 44(13), pp. 6876–6885.

Linsbauer, A., F. Paul, H. Machguth, and W. Haeberli (2013). Comparing three different methods to model scenarios of future glacier change in the Swiss Alps. In: *Annals of Glaciology* 54(63), pp. 241–253.

Luce, C. H. and Z. A. Holden (2009). Declining annual streamflow distributions in the Pacific Northwest United States, 1948-2006. In: *Geophysical Research Letters* 36, p. L16401.

Lutz, a. F., W. W. Immerzeel, a. B. Shrestha, and M. F. P. Bierkens (2014). Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. In: *Nature Climate Change* 4(7), pp. 587–592.

Lutz, A. F., W. W. Immerzeel, P. D. A. Kraaijenbrink, A. B. Shrestha, and M. F. P. Bierkens (2016). Climate change impacts on the upper indus hydrology: Sources, shifts and extremes. In: *PLoS ONE* 11(11).

MacDonald, A. M., A. R. Black, B. É. Ó. Dochartaigh, J Everest, W. G. Darling, V Flett, and D. W. Peach (2016). Using stable isotopes and continuous meltwater river monitoring to investigate the hydrology of a rapidly retreating Icelandic outlet glacier. In: *Annals of Glaciology* 57(72), pp. 1–8.

MacDonald, A. M. (2019). *Personal communication, 28th February.*

MacDougall, A. H., B. A. Wheler, and G. E. Flowers (2011). A preliminary assessment of glacier melt-model parameter sensitivity and transferability in a dry subarctic environment. In: *Cryosphere* 5(4), pp. 1011–1028.

Machguth, H., J. E. Box, R. S. Fausto, and W. T. Pfeffer (2018). Editorial: Melt Water Retention Processes in Snow and Firn on Ice Sheets and Glaciers: Observations and Modeling. In: *Frontiers in Earth Science* 6(July), pp. 1–3.

Mackay, J., C. Jackson, and L. Wang (2014). A lumped conceptual model to simulate groundwater level time-series. In: *Environmental Modelling and Software* 61, pp. 229–245.

Mackay, J., C. Jackson, A. Brookshaw, A. Scaife, J. Cook, and R. Ward (2015). Seasonal forecasting of groundwater levels in principal aquifers of the United Kingdom. In: *Journal of Hydrology* 530, pp. 815–828.

Mackay, J. D., N. E. Barrand, D. M. Hannah, S. Krause, C. R. Jackson, J. Everest, and G. Aðalgeirsdóttir (2018). Glacio-hydrological melt and run-off modelling: application of a limits of acceptability framework for model comparison and selection. In: *The Cryosphere* 12, pp. 2175–2210.

Mackay, J. D., N. E. Barrand, D. M. Hannah, S. Krause, C. R. Jackson, J. Everest, G. Aoalgeirsdóttir, and A. Black (2019). Future evolution and uncertainty of river flow regime change in a deglaciating river basin. In: *Hydrology and Earth System Sciences* 23(4), pp. 1833–1865.

MacQueen, J. B (1967). Some Methods for classification and Analysis of Multivariate Observations. In: *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability* 1, Univers, pp. 281–297.

Magnússon, E, F Pálsson, H Björnsson, and S Guðmundsson (2012). Removing the ice cap of Öraefajokull central volcano, SE-Iceland: Mapping and interpretation of bedrock topography, ice volumes, subglacial troughs and implications for hazards assessments. In: *Jökull* 62, pp. 131–150.

Magnússon, E, J. M.-c. Belart, F Pálsson, H Ágústsson, and P Crochet (2016). Geodetic mass balance record with rigorous uncertainty estimates deduced from aerial photographs and lidar data – Case study from Drangajökull ice cap , NW Iceland. In: *The Cryosphere* 10(1), pp. 159–177.

Magnusson, J., F. Kobierska, S. Huxol, M. Hayashi, T. Jonas, and J. W. Kirchner (2014). Melt water driven stream and groundwater stage fluctuations on a glacier forefield (Dammagletscher, Switzerland). In: *Hydrological Processes* 28(3), pp. 823–836.

Mandal, S. and S. P. Simonovic (2017). Quantification of uncertainty in the assessment of future streamflow under changing climate conditions. In: *Hydrological Processes* 31(11), pp. 2076–2094.

Mankin, J. S., D. Viviroli, D. Singh, A. Y. Hoekstra, and N. S. Diffenbaugh (2015). The potential for snow to supply human water demand in the present and future. In: *Environmental Research Letters* 10(11), p. 114016.

Manning, A. H. and D. K. Solomon (2003). Using noble gases to investigate mountain-front recharge. In: *Journal of Hydrology* 275(3-4), pp. 194–207.

Mansour, M. M., L Wang, M. Whiteman, and A. G. Hughes (2018). Estimation of spatially distributed groundwater potential recharge for the United Kingdom. In: *Quarterly Journal of Engineering Geology and Hydrogeology* 51(2), pp. 247–263.

Markovich, K. H., R. M. Maxwell, and G. E. Fogg (2016). Hydrogeological response to climate change in alpine hillslopes. In: *Hydrological Processes* 30(18), pp. 3126–3138.

Marren, P. M. (2005). Magnitude and frequency in proglacial rivers: a geomorphological and sedimentological perspective. In: *Earth-Science Reviews* 70(3-4), pp. 203–251.

Marren, P. M. and S. C. Toomath (2013). Fluvial adjustments in response to glacier retreat: Skaftafellsjökull, Iceland. In: *Boreas* 42(1), pp. 57–70.

Matthews, T., R. Hodgkins, R. L. Wilby, S. Gumundsson, F. Pálsson, H. Björnsson, and S. Carr (2015). Conditioning temperature-index model parameters on synoptic weather types for glacier melt simulations. In: *Hydrological Processes* 29(6), pp. 1027–1045.

Matthews, T. O. M. and R. Hodgkins (2016). Interdecadal variability of degree-day factors on Vestari Hagafellsjökull (Langjökull, Iceland) and the importance of threshold air temperatures. In: *Journal of Glaciology* 62(232), pp. 310–322.

Matti, B., H. E. Dahlke, B. Dieppois, D. M. Lawler, and S. W. Lyon (2017). Flood seasonality across Scandinavia - Evidence of a shifting hydrograph? In: *Hydrological Processes* 31(24), pp. 4354–4370.

Mayr, E., W. Hagg, C. Mayer, and L. Braun (2013). Calibrating a spatially distributed conceptual hydrological model using runoff, annual mass balance and winter mass balance. In: *Journal of Hydrology* 478, pp. 40–49.

McDonald, M. and A. Harbaugh (2003). The History of MODFLOW. In: *Groundwater* 41(2), pp. 280–283.

McDowell, J. Z. and J. J. Hess (2012). Accessing adaptation: Multiple stressors on livelihoods in the Bolivian highlands under a changing climate. In: *Global Environmental Change* 22(2), pp. 342–352.

McMillan, H. K. and I. K. Westerberg (2015). Rating curve estimation under epistemic uncertainty. In: *Hydrological Processes* 29(7), pp. 1873–1882.

Meresa, H. K. and R. J. Romanowicz (2017). The critical role of uncertainty in projections of hydrological Extremes. In: *Hydrology and Earth System Sciences Discussions* 21, pp. 4245–4258.

Milner, A. M. et al. (2017). Glacier shrinkage driving global changes in downstream systems. In: *Proceedings of the National Academy of Sciences* 114(37), pp. 9770–9778.

Minder, J. R., P. W. Mote, and J. D. Lundquist (2010). Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. In: *Journal of Geophysical Research Atmospheres* 115(14), p. D14122.

Monk, W. A., P. J. Wood, D. M. Hannah, and D. A. Wilson (2007). Selection of river flow indices for the assessment of hydroecological change. In: *River Research and Applications* 23(1), pp. 113–122.

Mora, C. et al. (2013). The projected timing of climate departure from recent variability. In: *Nature* 502, pp. 183–187.

Mosier, T. M., D. F. Hill, and K. V. Sharp (2016). How much cryosphere model complexity is just right? Exploration using the conceptual cryosphere hydrology framework. In: *Cryosphere* 10(5), pp. 2147–2171.

Naiman, R. J., J. J. Latterell, N. E. Pettit, and J. D. Olden (2008). Flow variability and the biophysical vitality of river systems. In: *Comptes Rendus - Geoscience* 340(9-10), pp. 629–643.

Nawri, N., B. Pálmason, G. N. Petersen, H. Björnsson, and S. Þorsteinsson (2017). *The ICRA atmospheric reanalysis project for Iceland*. Tech. rep. Reykjavík, Iceland: Icelandic Meteorological Office.

Nepal, S., W.-A. Flügel, P. Krause, M. Fink, and C. Fischer (2017). Assessment of Spatial Transferability of Process-Based Hydrological Model Parameters in Two Neighboring Catchments in the Himalayan Region. In: *Hydrological Processes* 31(April), pp. 2812–2826.

Neton, M., J. Dorsch, C. Olson, and S. Young (1994). Architecture and directional scales of heterogeneity in alluvial-fan aquifers. In: *Journal of Sedimentary Research* B64(2), pp. 245–257.

Niswonger, R., S. Panday, and M. Ibaraki (2011). *MODFLOW-NWT, A Newton formulation for MODFLOW-NWT, A Newton formulation for MODFLOW-2005: U.S. Geological Survey Techniques and Methods 6–A37*. Tech. rep. Reston, Virginia: U.S. Geological Survey, p. 44.

Nolin, A. W., J. Phillippe, A. Jefferson, and S. L. Lewis (2010). Present-day and future contributions of glacier runoff to summertime flows in a Pacific Northwest watershed: Implications for water resources. In: *Water Resources Research* 46(12), W12509.

Ó Dochartaigh, B. É., A. M. MacDonald, P Wilson, and H Bonsor (2012). *Groundwater investigations at Virkisjőkull, Iceland: Data Report 2012*. Tech. rep. Keyworth, Nottingham: British Geological Survey, pp. 1–48.

Ó Dochartaigh, B. É., A. M. MacDonald, A. R. Black, J. Everest, P. Wilson, W. G. Darling, L. Jones, and M. Raines (2019). Groundwater / meltwater interaction in proglacial aquifers. In: *Hydrology and Earth System Sciences Discussions*.

Oerlemans, J (2001). *Glaciers and Climate Change*. A. A. Balkema Publishers: Rotterdam, Netherlands.

Ohmura, A. (2001). Physical Basis for the Temperature-Based Melt-Index Method. In: *Journal of Applied Meteorology* 40, pp. 753–761.

Okkonen, J. and B. Kløve (2011). A sequential modelling approach to assess groundwater-surface water resources in a snow dominated region of Finland. In: *Journal of Hydrology* 411(1-2), pp. 91–107.

Orlove, B. (2009). "The Past, the Present, and Some Possible Futures of Adaptation". In: *Adapting to Climate Change: Thresholds, Values, Governance*. Ed. by W. N. Adger, I Lorenzoni, and K. O'Brien. Cambridge University Press: Cambridge, pp. 22–35.

Panofsky, H. and G. Brier (1968). *Some Applications of Statistics to Meteorology*. The Pennsylvania State University Press: Philadelphia, US.

Pappenberger, F., P. Matgen, K. J. Beven, J. B. Henry, L. Pfister, and P. Fraipont (2006). Influence of uncertain boundary conditions and model structure on flood inundation predictions. In: *Advances in Water Resources* 29(10), pp. 1430–1449.

Parriaux, A and G. Nicoud (1990). Hydrological behaviour of glacial deposits in mountainous areas. In: *Hydrology ofMountainous Areas (Proceedings of the Strbské Pleso Workshop, Czechoslovakia, June 1988)* 1(190).

Pellicciotti, F., B. Brock, U. Strasser, P. Burlando, M. Funk, and J. Corripio (2005). An enhanced temperature-index glacier melt model including the shortwave radiation balance : development and testing for Haut Glacier d ' Arolla , Switzerland. In: 51(175), pp. 573–587.

Pellicciotti, F., J. Helbing, A. Rivera, V. Favier, J. Corripio, J. Araos, J.-E. Sicart, and M. Carenzo (2008). A study of the energy balance and melt regime on Juncal Norte Glacier, semi-arid Andes of central Chile, using melt models of different complexity. In: *Hydrological Processes* 22, pp. 3980–3997.

Pellicciotti, F., C. Buergi, W. W. Immerzeel, M. Konz, and A. B. Shrestha (2012). Challenges and Uncertainties in Hydrological Modeling of Remote Hindu Kush–Karakoram–Himalayan (HKH) Basins: Suggestions for Calibration Strategies. In: *Mountain Research and Development* 32, pp. 39–50.

Petersen, L. and F. Pellicciotti (2011). Spatial and temporal variability of air temperature on a melting glacier: Atmospheric controls, extrapolation methods and their effect on melt modeling, Juncal Norte Glacier, Chile. In: *Journal of Geophysical Research Atmospheres* 116(23), p. D23109.

Phillips, E., A. Finlayson, and L. Jones (2013). Fracturing, block faulting, and moulin development associated with progressive collapse and retreat of a maritime glacier: Falljökull, SE Iceland. In: *Journal of Geophysical Research: Earth Surface* 118(3), pp. 1545–1561.

Phillips, E., A. Finlayson, T. Bradwell, J. Everest, and L. Jones (2014). Structural evolution triggers a dynamic reduction in active glacier length during rapid retreat: Evi-

dence from Falljökull, SE Iceland. In: *Journal of Geophysical Research F: Earth Surface* 119(10), pp. 2194–2208.

Pianosi, F., K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. In: *Environmental Modelling & Software* 79, pp. 214–232.

Poff, N. L. and J. K. Zimmerman (2010). Ecological responses to altered flow regimes: A literature review to inform the science and management of environmental flows. In: *Freshwater Biology* 55(1), pp. 194–205.

Ponce, V. M. (1989). *Engineering hydrology: Principles and practices*. Prentice-Hall: Englewood Cliffs, New Jersey, pp. 313–317.

Pool, S., M. J. P. Vis, R. R. Knight, and J. Seibert (2017). Streamflow characteristics from modeled runoff time series – importance of calibration criteria selection. In: *Hydrology and Earth System Sciences* 21, pp. 5443–5457.

Prudic, D., L. Konikow, and E. Banta (2004). *A new streamflow-routing (SFR1) package to simulate stream-aquifer interaction with MODFLOW-2000*. Tech. rep. Carson City, Nevada: U.S. Geological Survey, p. 96.

Puckridge, J. T., F. Sheldon, K. F. Walker, and A. J. Boulton (1998). Flow variability and the ecology of large rivers. In: *Marine Freshwater Research* 49, pp. 55–72.

Rada, C. and C. Schoof (2018). Channelized, distributed, and disconnected: Subglacial drainage under a valley glacier in the Yukon. In: *Cryosphere* 12(8), pp. 2609–2636.

Radić, V. and R. Hock (2014). Glaciers in the Earth's Hydrological Cycle: Assessments of Glacier Mass and Runoff Changes on Global and Regional Scales. In: *Surveys in Geophysics* 35(3), pp. 813–837.

Radić, V., A. Bliss, A. C. Beedlow, R. Hock, E. Miles, and J. G. Cogley (2014). Regional and global projections of twenty-first century glacier mass changes in response to climate scenarios from global climate models. In: *Climate Dynamics* 42(1-2), pp. 37–58.

Ragettli, S., F. Pellicciotti, R. Bordoy, and W. W. Immerzeel (2013). Sources of uncertainty in modeling the glaciohydrological response of a Karakoram watershed to climate change. In: *Water Resources Research* 49(9), pp. 6048–6066. arXiv: Ragettli2013.

Ragettli, S., G. Cortés, J. Mcphee, and F. Pellicciotti (2014). An evaluation of approaches for modelling hydrological processes in high-elevation, glacierized Andean watersheds. In: *Hydrological Processes* 28(23), pp. 5674–5695.

Ragettli, S., W. W. Immerzeel, and F. Pellicciotti (2016). Contrasting climate change impact on river flows from high-altitude catchments in the Himalayan and Andes Mountains. In: *Proceedings of the National Academy of Sciences of the United States of America* 113(33), pp. 9222–9227.

Ravazzani, G, D Curti, P Gattinoni, S Della Valentina, A Fiorucci, and R Rosso (2016). Assessing Groundwater Contribution to Streamflow of a Large Alpine River with Heat Tracer Methods and Hydrological Modelling. In: *River research and applications* 32(5), pp. 871–884.

Reda, I. and A. Andreas (2008). *Solar Position Algorithm for Solar Radiation Applications*. Tech. rep. NREL/TP-560-34302. Colorado, USA: National Renewable Energy Laboratory.

Rees, G. and D. N. Collins (2006). Regional differences in response of flow in glacier-fed Himalayan rivers to climatic warming. In: *Hydrological Processes* 20, pp. 2157–2169.

Remesan, R. and I. P. Holman (2015). Effect of baseline meteorological data selection on hydrological modelling of climate change scenarios. In: *Journal of Hydrology* 528, pp. 631–642.

Reveillet, M., C. Vincent, D. Six, and A. Rabatel (2017). Which empirical model is best suited to simulate glacier mass balances? In: *Journal of Glaciology* 63(237), pp. 39–54.

Riggs, G. and D. Hall (2015). *MODIS Snow Products Collection 6 User Guide*. Tech. rep. URL: https://nsidc.org/sites/nsidc.org/files/files/MODIS-snow-user-guide-C6.pdf.

Roberts, M. J. and M. T. Gudmundsson (2015). "Öræfajökull Volcano: Geology and Historical Floods". In: *Volcanogenic floods in Iceland: Anassessment of hazards and risks at Öræfajökull and on the Markafljot outwash plain*. Ed. by E. Pagneu, M. T. Gud-

mundsson, S. Karlsdóttir, and M. J. Roberts. IMO, IES-UI, NCIP-DCPEM: Reykjavik, Iceland, pp. 17–43.

Robinson, Z. P., I. J. Fairchild, and A. J. Russell (2008). Hydrogeological implications of glacial landscape evolution at Skeiðarársandur, SE Iceland. In: *Geomorphology* 97(1-2), pp. 218–236.

Robinson, Z. P., I. J. Fairchild, and C. Arrowsmith (2009). "Stable isotope tracers of shallow groundwater recharge dynamics and mixing within an Icelandic sandur, Skeiðarársandur". In: *Hydrology in mountain regions : observations, processes and dynamics*. International Association of Hydrological Sciences, pp. 119–125.

Roy, J. W. and M. Hayashi (2009). Multiple, distinct groundwater flow systems of a single moraine-talus feature in an alpine watershed. In: *Journal of Hydrology* 373(1-2), pp. 139–150.

Rye, C. J., I. C. Willis, N. S. Arnold, and J. Kohler (2012). On the need for automated multiobjective optimization and uncertainty estimation of glacier mass balance models. In: *Journal of Geophysical Research: Earth Surface* 117(2), pp. 1–21.

Sachindra, D. A., F. Huang, A. Barton, and B. J. C. Perera (2014). Statistical downscaling of general circulation model outputs to precipitation-part 2: Bias-correction and future projections. In: *International Journal of Climatology* 34(11), pp. 3282–3303.

Salomonson, V. V. and I. Appel (2004). Estimating fractional snow cover from MODIS using the normalized difference snow index. In: *Remote Sensing of Environment* 89(3), pp. 351–360.

Samaniego, L., R. Kumar, L. Breuer, A. Chamorro, M. Flörke, I. G. Pechlivanidis, D. Schäfer, H. Shah, T. Vetter, M. Wortmann, and X. Zeng (2017). Propagation of forcing and model uncertainties on to hydrological drought characteristics in a multi-model century-long experiment in large river basins. In: *Climatic Change* 141(3), pp. 435–449.

Sanford, T., P. C. Frumhoff, A. Luers, and J. Gulledge (2014). The climate policy narrative for a dangerously warming world. In: *Nature Climate Change* 4, pp. 164–166.

Sawicz, K. A., C. Kelleher, T. Wagener, P. Troch, M. Sivapalan, and G. Carrillo (2014). Characterizing hydrologic change through catchment classification. In: *Hydrology and Earth System Sciences* 18(1), pp. 273–285.

Schaefli, B. and M. Huss (2011). Integrating point glacier mass balance observations into hydrologic model identification. In: *Hydrology and Earth System Sciences* 15(4), pp. 1227–1241.

Schaefli, B., L. Nicótina, C. Imfeld, P. Da Ronco, E. Bertuzzo, and A. Rinaldo (2014). SEHR-ECHO v1.0: A spatially explicit hydrologic response model for ecohydrologic applications. In: *Geoscientific Model Development* 7(6), pp. 2733–2746.

Schaefli, B. (2016). Snow hydrology signatures for model identification within a limits-of-acceptability approach. In: *Hydrological Processes* 30(22), pp. 4019–4035.

Schaner, N., N. Voisin, B. Nijssen, and D. P. Lettenmaier (2012). The contribution of glacier melt to streamflow. In: *Environmental Research Letters* 7(3), p. 034029.

Schoen, E. R. et al. (2017). Future of Pacific Salmon in the Face of Environmental Change: Lessons from One of the World's Remaining Productive Salmon Regions. In: *Fisheries* 42(10), pp. 538–553.

Schulla, J. (2015). *Model Description WaSiM. Technical report.* Tech. rep. Zürich: Hydrology Software Consulting, 332p.

Scibek, J., D. M. Allen, A. J. Cannon, and P. H. Whitfield (2007). Groundwater-surface water interaction under scenarios of climate change using a high-resolution transient groundwater model. In: *Journal of Hydrology* 333(2-4), pp. 165–181.

Seibert, J., M. Vis, I. Kohn, M. Weiler, and K. Stahl (2018). Technical note: Representing glacier geometry changes in a semi-distributed hydrological model. In: *Hydrology and Earth System Sciences* 22(4), pp. 2211–2224.

Shafii, M. and B. A. Tolson (2015). Optimizing hydrological consistency by incorporating hydrological signatures intomodel calibration objectives. In: *Water Resources Research* 51(5), pp. 3796–3814.

Shannon, S., R. Smith, A. Wiltshire, T. Payne, M. Huss, R. Betts, J. Caesar, A. Koutroulis, D. Jones, and S. Harrison (2019). Global glacier volume projections under high-end climate change scenarios. In: *Cryosphere* 13(1), pp. 325–350.

Shea, J. M. and R. D. Moore (2010). Prediction of spatially distributed regional-scale fields of air temperature and vapor pressure over mountain glaciers. In: *Journal of Geophysical Research Atmospheres* 115(23), p. D23107.

Shea, J. M. and W. W. Immerzeel (2016). An assessment of basin-scale glaciological and hydrological sensitivities in the Hindu Kush-Himalaya. In: *Annals of Glaciology* 57(71), pp. 308–318.

Singh, S. K., H. McMillan, A. Bárdossy, and C. Fateh (2016a). Nonparametric catchment clustering using the data depth function. In: *Hydrological Sciences Journal* 61(15), pp. 2649–2667.

Singh, S., R. Kumar, A. Bhardwaj, L. Sam, M. Shekhar, A. Singh, R. Kumar, and A. Gupta (2016b). Changing climate and glacio-hydrology in Indian Himalayan Region: A review. In: *Wiley Interdisciplinary Reviews: Climate Change* 7(3), pp. 393–410.

Somers, L. D., R. P. Gordon, J. M. McKenzie, L. K. Lautz, O. Wigmore, A. M. Glose, R. Glas, C. Aubry-Wake, B. Mark, M. Baraer, and T. Condom (2016). Quantifying groundwater–surface water interactions in a proglacial valley, Cordillera Blanca, Peru. In: *Hydrological Processes* 30(17), pp. 2915–2929.

Sorensen, J. P. R., J. W. Finch, A. M. Ireson, and C. R. Jackson (2014). Comparison of varied complexity models simulating recharge at the field scale. In: *Hydrological Processes* 28(4), pp. 2091–2102.

Soruco, A., C. Vincent, A. Rabatel, B. Francou, E. Thibert, J. E. Sicart, and T. Condom (2015). Contribution of glacier runoff to water resources of La Paz city, Bolivia. In: *Annals of Glaciology* 56(70), pp. 147–154.

Sridhar, V., M. M. Billah, and J. W. Hildreth (2018). Coupled Surface and Groundwater Hydrological Modeling in a Changing Climate. In: *Groundwater* 56(4), pp. 618–635.

Srivastav, R. K., A. Schardong, and S. P. Simonovic (2014). Equidistance Quantile Matching Method for Updating IDF Curves under Climate Change. In: *Water Resources Management* 28(9), pp. 2539–2562.

Stefania, G. A., M. Rotiroti, L. Fumagalli, F. Simonetto, P. Capodaglio, C. Zanotti, and T. Bonomi (2018). Modeling groundwater/surface-water interactions in an Alpine valley (the Aosta Plain, NW Italy): the effect of groundwater abstraction on surface-water resources. In: *Hydrogeology Journal* 26(1), pp. 147–162.

Stewart, I. T., D. L. Ficklin, C. A. Carrillo, and R. McIntosh (2015). 21st century increases in the likelihood of extreme hydrologic conditions for the mountainous basins of the Southwestern United States. In: *Journal of Hydrology* 529, pp. 340–353.

Stocker, T. et al. (2013). "Technical Summary". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley. Cambridge University Press: Cambridge, UK and New York, USA.

Stoffel, M., B. Wyżga, and R. A. Marston (2016). Floods in mountain environments: A synthesis. In: *Geomorphology* 272(Supplement C), pp. 1–9.

Storch, H. von and F. W. Zwiers (1999). *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge, United Kingdom, pp. 171–190.

Switanek, M. B., P. A. Troch, C. L. Castro, A. Leuprecht, H.-I. Chang, R. Mukherjee, and E. M. C. Demaria (2017). Scaled distribution mapping: a bias correction method that preserves raw climate model projected changes. In: *Hydrology and Earth System Sciences* 21, pp. 2649–2017.

Tabachnick, B. G. and L. S. Fidell (2014). *Using Multivariate Statistics Sixth Edition*. Sixth. Pearson Education Limited: Essex, United Kingdom, pp. 65–91.

Taylor, K. E., R. J. Stouffer, and G. a. Meehl (2012). An Overview of CMIP5 and the Experiment Design. In: *Bulletin of the American Meteorological Society* 93(4), pp. 485–498.

Taylor, R. G. (2013). Ground water and climate change. In: *Nature Climate Change* 3(April), pp. 322–329.

Teng, J., N. J. Potter, F. H. S. Chiew, L. Zhang, B. Wang, J. Vaze, and J. P. Evans (2015). How does bias correction of regional climate model precipitation affect modelled runoff? In: *Hydrology and Earth System Sciences* 19(2), pp. 711–728.

Teutschbein, C, T Grabs, R. H. Karlsen, H Laudon, and K Bishop (2015). Hydrological response to changing climate conditions: Spatial streamflow variability in the boreal region. In: *Water Resources Research* 51(12), pp. 9425–9446.

Thompson, A (1988). Historical development of the proglacial landforms of Svínafellsjökull and Skaftafellsjökull, Southeast Iceland. In: *Jökull* 38, pp. 17–30.

Thorsteinsson, T. and H. Björnsson (2012). *Climate Change and Energy Systems: Impacts, Risks and Adaptation in the Nordic and Baltic countries*. Tech. rep. Copenhagen: Nordic Council of Ministers.

UNESCO (2017). *The impact of glacier retreat in the Andes: International Multidisciplinary Network for Adaptation Strategies*. Tech. rep. Paris, France: UNESCO International Hydrological Programme (IHP).

Valentin, M. M., T. S. Hogue, and L. E. Hay (2018). Hydrologic regime changes in a high-latitude glacierizedwatershed under future climate conditions. In: *Water* 10(128), pp. 1–24.

Van Tiel, M., A. J. Teuling, N. Wanders, M. J. Vis, K. Stahl, and A. F. Van Loon (2018). The role of glacier changes and threshold definition in the characterisation of future streamflow droughts in glacierised catchments. In: *Hydrology and Earth System Sciences* 22(1), pp. 463–485.

Vaughan, D. et al. (2013). "Observations: Cryosphere". In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley. Cambridge University Press: Cambridge, United Kingdom and New York, NY, USA, pp. 358–359.

Verbunt, M., J. Gurtz, K. Jasper, H. Lang, P. Warmerdam, and M. Zappa (2003). The hydrological role of snow and glaciers in alpine river basins and their distributed modeling. In: *Journal of Hydrology* 282(1-4), pp. 36–55.

Vetter, T., S. Huang, V. Aich, T. Yang, X. Wang, V. Krysanova, and F. Hattermann (2015). Multi-model climate impact assessment and intercomparison for three large-scale river basins on three continents. In: *Earth System Dynamics* 6(1), pp. 17–43.

Vetter, T. et al. (2017). Evaluation of sources of uncertainty in projected hydrological changes under climate change in 12 large-scale river basins. In: *Climatic Change* 141(3), pp. 419–433.

Viglione, A., J. Parajka, M. Rogger, J. L. Salinas, G. Laaha, M. Sivapalan, and G. Blöschl (2013). Comparative assessment of predictions in ungauged basins - Part 3: Runoff signatures in Austria. In: *Hydrology and Earth System Sciences* 17(6), pp. 2263–2279.

Vincent, A., S. Violette, and G. Aðalgeirsdóttir (2019). Groundwater in catchments headed by temperate glaciers: A review. In: *Earth-Science Reviews* 188(2019), pp. 59–76.

Viviroli, D. and R. Weingartner (2004). The hydrological significance of mountains: from regional to global scale. In: *Hydrology and Earth System Sciences* 8(6), pp. 1016–1029.

Viviroli, D., H. H. Dürr, B. Messerli, M. Meybeck, and R. Weingartner (2007). Mountains of the world, water towers for humanity: Typology, mapping, and global significance. In: *Water Resources Research* 43(7), pp. 1–13.

Voeckler, H. M., D. M. Allen, and Y. Alila (2014). Modeling coupled surface water - Groundwater processes in a small mountainous headwater catchment. In: *Journal of Hydrology* 517, pp. 1089–1106.

Vuille, M. (2015). *Challenges in Sustainable Water Supply in the Tropical Andes due to Climate Change*. Tech. rep. Paris, France: UNESCO International Hydrological Programme (IHP).

Vuuren, D. P. van et al. (2011). The representative concentration pathways: An overview. In: *Climatic Change* 109(1), pp. 5–31.

Ward, J., F. Malard, K. Tockner, and U. Uehlinger (1999). Influence of ground water on surface water conditions in a glacial flood plain of the Swiss Alps. In: *Hydrological Processes* 13(3), pp. 277–293.

Welling, J. T., o. Árnason, and R. Ólafsdottír (2015). Glacier tourism: a scoping review. In: *Tourism Geographies* 17(5), pp. 635–662.

Westerberg, I. K., T. Wagener, G. Coxon, H. K. McMillan, A. Castellarin, A. Montanari, and J. Freer (2016). Uncertainty in hydrological signatures for gauged and ungauged catchments. In: *Water Resources Research* 52(3), pp. 1847–1865.

Wijngaard, R. R., A. F. Lutz, S. Nepal, S. Khanal, S. Pradhananga, A. B. Shrestha, and W. W. Immerzeel (2017). Future changes in hydro-climatic extremes in the Upper Indus, Ganges, and Brahmaputra River basins. In: *PLOS ONE* 12(12), e0190224.

Wilcke, R. A. and L. Bärring (2016). Selecting regional climate scenarios for impact modelling studies. In: *Environmental Modelling and Software* 78, pp. 191–201.

Willis, I. (2005). "168: Hydrology of Glacierized Basins". In: *Encyclopedia of Hydrological Sciences: Part 14. Snow and Glacier Hydrology.* Ed. by M. G. Anderson and J. J. McDonnell. John Wiley & Sons, Ltd: Chichester, UK, pp. 2601–2631.

Wilson, A. M., M. W. Williams, R. B. Kayastha, and A. Racoviteanu (2016). Use of a hydrologic mixing model to examine the roles of meltwater, precipitation and groundwater in the Langtang River basin, Nepal. In: *Annals of Glaciology* 57(71), pp. 155–168.

Wilson, J. L. and H. Guan (2004). Mountain-Block Hydrology and Mountain-Front Recharge. In: *Groundwater Recharge in a Desert Environment: The Southwestern United States* 9, pp. 113–137.

Winsemius, H. C., B Schaefli, A Montanari, and H. H. G. Savenije (2009). On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. In: *Water Resources Research* 45(12), W12422.

Woul, M. de, R. Hock, M. Braun, T. Thorsteinsson, T. Jóhannesson, and S. Halldórsdóttir (2006). Firn layer impact on glacial runoff: A case study at Hofsjökull, Iceland. In: *Hydrological Processes* 20(10), pp. 2171–2185.

Yadav, M., T. Wagener, and H. Gupta (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. In: *Advances in Water Resources* 30(8), pp. 1756–1774.

Yao, Y., C. Zheng, C. Andrews, Y. Zheng, A. Zhang, and J. Liu (2017). What controls the partitioning between baseflow and mountain block recharge in the Qinghai-Tibet Plateau? In: *Geophysical Research Letters* 44(16), pp. 8352–8358.

Yilmaz, K. K., H. V. Gupta, and T. Wagener (2008). A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. In: *Water Resources Research* 44(9), W09417.

Yuan, F., C. Zhao, Y. Jiang, L. Ren, H. Shan, L. Zhang, Y. Zhu, T. Chen, S. Jiang, X. Yang, and H. Shen (2017). Evaluation on uncertainty sources in projecting hydrological changes over the Xijiang River basin in South China. In: *Journal of Hydrology* 554, pp. 434–450.

Zappa, M and C Kan (2007). Extreme heat and runoff extremes in the Swiss Alps Natural Hazards and Earth System Sciences Extreme heat and runoff extremes in the Swiss Alps. In: *European Geosciences Union* 7(3), pp. 375–389.

Zemp, M. et al. (2019). Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. In: *Nature* 568(7752), pp. 382–386.

Zemp, M. et al. (2015). Historically unprecedented global glacier decline in the early 21st century. In: *Journal of Glaciology* 61(228), pp. 745–762.

Zhang, Y., Y. Hirabayashi, Q. Liu, and S. Liu (2015). Glacier runoff and its impact in a highly glacierized catchment in the southeastern Tibetan Plateau: past and future trends. In: *Journal of Glaciology* 61(228), pp. 713–730.