Aalto University
School of Engineering

Vili Virkki

The value of open-source river streamflow estimation in Southeast Asia

Thesis submitted for examination for the degree of
Master of Science in Technology.

Espoo 18.12.2018

Thesis supervisor

Prof. Olli Varis

Thesis instructors

MSc. Amy Fallon and MSc. Marko Kallio

Aalto University
School of Engineering

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author BSc. Vili Virkki		
Title of thesis The value of open-source river streamflow estimation in Southeast Asia		
Master programme Master's Programme in Geoinformatics		**Code** ENG22
Thesis supervisor Prof. Olli Varis		
Thesis advisor(s) MSc. Amy Fallon & MSc. Marko Kallio		
Date 18.12.2018	**Number of pages** 53+15	**Language** English

Abstract

Water governance deals with a substance that is essential for sustaining life. Managing a physical substance, which runs in multiple interconnected systems crossing administrative borders of all scales, creates controversies when the interests of different stakeholders collide. Creating policies and making decisions related to water requires efficient science-policy interaction, to which environmental modelling is inarguably an important input that provides salient, credible, and legitimate information to be used. However, modelling the endless complexity of the physical world leaves the modelers facing uncertainties. This work demonstrates the environmental modelling process, conducts an uncertainty quantification, and finally investigates the potential of adopting open-source solutions as a part of environmental modelling and policy support.

This work applies both quantitative and qualitative methods. The main quantitative methods are hydrostreamer, a newly developed open-source tool that can estimate river streamflow, and Monte Carlo simulation, which is applied to quantify a type of uncertainty related to hydrostreamer. Within the qualitative methods, a survey and semi-structured interviews are applied to assess the current state of hydrostreamer and its applicability in the 3S river basin located in Southeast Asia, and also to investigate the potential of open-source based environmental modelling solutions in a more general manner. In the work, one of the contributors of uncertainty related to modelling streamflow with hydrostreamer is quantified and shown to behave with respect to Strahler order. Minor applications for hydrostreamer in its present state are found in the study area, and potential for adopting open-source solutions is shown, primarily due to low costs and through major donor organizations.

In the resulting discussion, it is emphasized that trade-offs between modelling tool applicability and accuracy should be addressed to make the environmental modelling process truly open, instead of limiting the openness to only few of those with the required capabilities. Evaluating uncertainties related to a modelling process helps building confidence on the methods used; however, it should be considered carefully how to present the analysis and results for those who are not familiar with the subject. The work is a technical approach with ultimate goals in incorporating non-technical people and arguing for open data and transparency. It is concluded that open-source tools exhibit potential to be incorporated in complicated policy-making contexts despite the fact that no single tool can provide a panacea for complex issues.

Keywords water governance, water resources management, science-policy interaction, environmental modelling, uncertainty, spatial uncertainty, hydrology, geoinformatics

Aalto-yliopisto
Insinööritieteiden korkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Diplomityön tiivistelmä

Tekijä Tekn. kand. Vili Virkki

Työn nimi Avoimeen dataan perustuvan jokivirtaamatyökalun arvo Kaakkois-Aasiassa

Maisteriohjelma Master's Programme in Geoinformatics **Koodi** ENG22

Työn valvoja Prof. Olli Varis

Työn ohjaaja(t) MSc. Amy Fallon & DI Marko Kallio

Päivämäärä 18.12.2018 **Sivumäärä** 53+15 **Kieli** englanti

Tiivistelmä

Vesi on keskeinen elämää ylläpitävä aine, ja sen merkitys lähes kaikille elinkeinoille on hyvin suuri. Se kiertää luonnon toisiinsa liittyneissä järjestelmissä jatkuvasti, ja vesivarojen hallinnoinnin on kyettävä toimimaan hallinnollisten rajojen yli. Veden käyttöön liittyy usein intressiristiriitoja eri toimijoiden välillä. Tieteellistä tietoa ja ympäristömallintamista hyödynnetään päätöksenteossa vesialalla, jolloin tuotetun tiedon tulisi olla mahdollisimman hyödyllistä ja luotettavaa. Fyysisen maailman loputon monimutkaisuus kuitenkin aiheuttaa epävarmuutta mallintamiseen. Tässä työssä esitellään ympäristömallintamisen prosessi ja arvioidaan erästä epävarmuustekijää numeerisesti. Lopuksi arvioidaan avoimeen lähdekoodiin perustuvien ratkaisujen potentiaalia osana ympäristömallintamista ja päätöksenteon tukea.

Tässä työssä käytetään sekä laskennallisia että laadullisia menetelmiä. Laskennallisia menetelmiä ovat hydrostreamer, uusi avoimeen lähdekoodiin perustuva työkalu jokivirtaaman arviointiin, ja Monte Carlo -simulaatio, jota käytetään epävarmuuden simuloimiseen ja arviointiin. Kyselytutkimusta sekä puolistrukturoituja haastatteluja sovelletaan hydrostreamerin nykytilan ja sovellettavuuden arviointiin käyttäen nk. 3S-jokilaaksoa Kaakkois-Aasiassa esimerkkialueena. Työn tuloksena näytetään arvioidun spatiaalisen epävarmuuden käyttäytyvän käänteisesti Strahlerin luvun suhteen. Pienimuotoisia sovelluskohteita hydrostreamerille osoitetaan olevan tutkimusalueella, ja avointen työkalujen käyttöä ympäristömallintamisessa arvioidaan laadullisten tulosten perusteella.

Työn pohdinnassa korostetaan, että mallintamistyökalujen valinnassa ja käytössä tulisi ottaa huomioon kompromissit sovellettavuuden ja tarkkuuden välillä. Tällöin mallinnusprosessista tulisi aidosti avoin sen sijaan, että avoimuus rajoittuisi vain pienelle joukolle toimijoita, joilla on tarvittava osaaminen ja data käytössään. Mallinnukseen liittyvien epävarmuuksien analysointi auttaa rakentamaan luottamusta kehitettyihin menetelmiin: epävarmuusanalyysin toteutus ja tulosten esittäminen muille voivat kuitenkin olla haastavia ja vaativat huolellista suunnittelua. Työ on lähtökohtaisesti tekninen, mutta sen pohjalta voidaan myös saada ei-teknisiä toimijoita osallistumaan mallintamiseen entistä enemmän. Yhteenvetona voidaan todeta, että potentiaalia sisällyttää avoimet työkalut osaksi monimutkaisia päätöksentekoprosesseja on olemassa, vaikkakaan mikään työkalu ei voi yksinään ratkaista kaikkia ongelmia.

Avainsanat vesihallinto, vesivarojen hallinta, tiede ja päätöksenteko, ympäristömallintaminen, epävarmuus, alueellinen epävarmuus, hydrologia, geoinformatiikka

Foreword

Originally, the idea of writing my Master's Thesis within a university research group was born already in autumn 2017 before my exchange semester in University of Calgary. I discussed with my advisor for this work, Marko Kallio, and professor Kirsi Virrantaus about writing the thesis in the Geoinformatics Research Group. However, when the spring came, the most promising place to get started with the work seemed to be in the Water and Development Research Group, which would have been a semi-deep dive into the uncharted waters for me because I hadn't studied anything water-related since some introductory bachelor's level courses.

After a while, the topic started to form around the research by my advisors, Marko, and Amy Fallon, both doctoral students within WDRG. It became an interesting combination consisting of topics both in technical, far abstracted modelling and in studying social settings and good governance. Certainly, it was a challenge first to grasp the basics of water governance and qualitative research and then try to conduct a study that would move smoothly from technical parts to non-technical parts and still remain somehow focused.

Now that the thesis is finished, I feel glad about taking this opportunity to do a small sidestep to slightly different field of study without forgetting my background in geoinformatics. It has been a pleasant, yet challenging, learning path alongside which I have had to teach myself a lot of completely new things and also reinforce some old things that I was more familiar but not really skilled with. In conclusion, I am happy with the result despite the feeling that the work has invoked more new research questions and topics for future studies than it has been able to answer.

I would like to express my gratitude to the people that have enabled my work and helped me to comprise the thesis – without your sincere help, the work wouldn't be even half done by now. I would like to thank my supervisor, professor Olli Varis, who has provided me with insightful comments on scientific writing and the balance of the study, which is something that I can't really see myself while being bound on the work so tightly. My instructors, MSc. Amy Fallon and MSc. Marko Kallio have given me continuous and inspiring support throughout the work and made me think about the small but important choices made during the research. Dr. Maija Taka has made sure that my working conditions within WDRG have been consistently great and also that I have had the patience to take some days off during the project. The WDRG as a whole has been a great source of peer support during the days of autumn that are as dark as the coffee from old coffee room brewing machine. Financially, I am grateful to Maa- ja vesitekniikan tuki ry. who have provided me with direct and indirect funding, essentially giving me the chance to concentrate only on this work.

Finally, I would like to thank my parents, Arja and Antti, who have been the first and the largest support in my whole life from childhood to graduating from university. Also, friends within Aalto University and elsewhere have helped me throughout the studies as both study support and by coming up with something completely else than studying to relax and unwind. And last but certainly not least, my dear Anna, who has been there for me by listening and helping me through all joys and worries during this time.

Espoo 18.12.2018
Vili Virkki

Table of Contents

List of figures

List of tables

List of abbreviations

ADB	Asian Development Bank
CHP	Corrected Heights Map
DEM	Digital Elevation Model
FOI	Freedom of Information
GFDL	Geophysical Fluid Dynamics Laboratory
GIS	Geographic Information System
GUI	Graphical User Interface
GWP	Global Water Partnership
HEC	Hydrological Engineering Center
HMS	Hydrologic Modelling System
HYMOS	Hydrological Monitoring Station
IUCN	International Union for Conservation of Nature
JPL	Jet Propulsion Laboratory
IWRM	Integrated Water Resources Management
MRC	Mekong River Commission
NASA	National Aeronautics and Space Administration
NGA	National Geospatial-Intelligence Agency
NGO	Non-Governmental Organization
OECD	Organisation for Economic Co-operation and Development
OSM	OpenStreetMap
RMSE	Root Mean Square Error
RTI	Right to Information
SAGA	System for Automated Geoscientific Analyses
SDG	Sustainable Development Goal
SDP	Standard Deviation Map
SRTM	Shuttle Radar Topography Mission
SWAT	Soil & Water Assessment Tool
TAMU	Texas A&M University
UN	United Nations
WDRG	Water and Development Research Group
WWF	World Wildlife Fund

1 Introduction

Scientific information is an essential part of supporting natural resources management and related governance decisions, but it must overcome the science-policy interaction boundary between scientists and stakeholders who use the information (Armitage et al. 2015; Cash et al. 2002). Therefore, it is not only the quality of information that is important but also the evaluation of its contents and delivery. Water governance considers inherently complex and intertwined issues, which require making decisions and creating policies that have significant effects on a substance that is crucial for sustaining life everywhere in the world (e.g. Reed and Kasprzyk 2009). Still, conflicts emerge due to for example difficulties in defining borders of both physical forms of water and the instruments and organizations that manage it (Gupta et al. 2013).

When forming policies and using scientific information in complex contexts, such as transboundary water management, participatory approaches engaging scientists, policy-makers, civil society, and other relevant stakeholders involved are often beneficial (e.g. Cash et al. 2006; Meadow et al. 2015). Ways to enhance communication and integration of science and policy include co-production of information (Meadow et al. 2015), the work of boundary organizations (e.g. Clark et al. 2016; Guston 2001), and knowledge brokering (Meyer 2010). Studies on making the science-policy integration more collaborative through participatory processes are aiming for including end-users in creating and communicating scientific information to achieve better salience[1] for it (Cash et al. 2002; Reed et al. 2009). Production of scientific information to support policy-making would benefit from using solutions that are based on open data and tools: the information would be equal and accessible for all stakeholders involved not depending on each stakeholder's position and interests.

Streamflow in rivers relates closely to various water resources management issues, such as hydropower and irrigation development, which are prominent issues in water-abundant developing countries (e.g. Geheb et al. 2015). Kallio et al. (2018) have developed a tool in R programming language, called *hydrostreamer,* which downscales low-resolution runoff data to a high-resolution river network by assigning the runoff into river segments and accumulating flow from upstream to downstream, finally resulting in discharge estimates. The package can be used with openly available and global data, such as global runoff models, and the source code is freely available for download and modifications. Input data requirements for *hydrostreamer* have been kept minimal and using it does not require expertise in the field of hydrology. With these factors in mind, the developers have aimed the tool for fast, understandable computing of streamflow estimates with enough scientific credibility.

Since *hydrostreamer* is a modelling tool, uncertainty is always present in the modelling process. To increase the credibility of the results produced with *hydrostreamer*, the uncertainties associated with it can be quantified. Though some parts of total uncertainty reach beyond quantifiable factors, one of the main contributors of statistical uncertainty is the uncertainty of the river network used in the downscaling process. While explicit river network data is not always available, the network to be used in streamflow estimation may have to be created from other globally available sources, such as digital elevation models

---

[1] In this work, salience refers to the usefulness and relevance of information for the involved stakeholders, meaning that information is interpreted beneficially in the right place at the right time (Cash et al. 2002).

(DEM). One commonly used, near-global river network data set, HydroSHEDS (WWF 2018), is originally derived from a DEM and is therefore uncertain to a degree. Quantifying the uncertainty of *hydrostreamer* estimates with respect to DEM uncertainty is important to evaluate the persistent difference between models and reality: effects of this difference also reflect to the usage and the value of the tool. Various authors, such as Armitage et al. (2015), Beven (2008) and Refsgaard et al. (2007) note that uncertainty estimation should always be addressed and communicated as a part of science-policy interaction.

Investigating the value of scientific information in a governance context can be conducted through different frameworks, such as the salience-credibility-legitimacy framework first proposed by Cash et al. (2002). Other very important and perhaps more practical aspects regarding the value of scientific information include access to data, available software and hardware, and skills to apply available tools and analyze the results. These factors cannot be estimated only by researchers looking at the tools and results produced by themselves – the potential users in their respective context must be identified and interviewed to gather feedback and further develop workflows.

This study presents a combination of quantitative and qualitative research; the complete process of creating scientific information, communicating it to the relevant stakeholders in a water resources management context, and applying the information is outlined and evaluated. To demonstrate the process, a case study is conducted in the 3S river basin, which consists of Sekong, Sesan, and Sre Pok rivers and reaches over the states of Cambodia, Lao PDR, and Vietnam, by applying *hydrostreamer* and contacting stakeholders in the area. The water governance situation in the 3S basin is complex: the transboundary setting, with numerous overlapping policies and desires of economic growth through harnessing natural resources, complicate policy-making and emphasize the existing power inequalities to the expense of those who are in a less powerful position (Geheb et al. 2015; Ham et al. 2015). The results of the study are discussed reflecting on the fact that all methods and tools used are openly available: it is desired that the workflow presented in this study would be applicable globally. The work is not a pure modelling exercise: rather, the environmental model is only a tool that produces knowledge, which is further used in science-policy interaction.

1.1 Research questions

The research aims to determine whether streamflow estimation using *hydrostreamer* would be valuable for stakeholders related to water resources management in Southeast Asia. The key factor increasing the value of *hydrostreamer* is the openness of the process: more generally, the study investigates how open source solutions could be adopted and applied in water resources management and as a supporting instrument for water governance. Prospects of applying open source tools are numerous, aiming to make water resources management processes fairer and more transparent. Another aspect of the work is to quantify a type of uncertainty, which is an inevitable factor in all modelling work. This study will estimate a minor but

Box 1. The focus of the study

What?
Evaluating the value of open source streamflow estimation in water resources management context. Case study in the 3S river basin.
Why?
To investigate the potential of open source solutions in supporting water resources management.
How?
Identifying and interviewing key stakeholders to gather feedback and development proposals based on example data and results provided by the researchers.

significant part of the total uncertainty related to *hydrostreamer* and address the meaning of this estimation.

The key aims of the research are briefly presented in Box 1. To clarify the aims of the work and provide basis for discussion, the following research questions may be formulated. The research questions presented below are explicitly defined with respect to *hydrostreamer*: applicability of this study and the answers to the research questions are further evaluated in the discussion.

Research question 1: *How could hydrostreamer be used in supporting water resources management or water governance?*

Research question 2: *How could open-source and transparent solutions benefit policy-making that is based on environmental modelling?*

Research question 3: *How does spatial uncertainty affect hydrostreamer outputs and use?*

In this study, *hydrostreamer* is used as an example to demonstrate a modelling process using simple open-source tools, but especially research question 2 is an attempt to generalize the study further. For clarification, it should be noted that the study is not involved in or aimed for direct policy-making but is rather an indirect approach through stakeholders that can affect it with technical expertise and guidance. Research question 3 is the most technical one and will be investigated through simulation. However, the concept of uncertainty is also linked to research question 1 as a persistent characteristic of *hydrostreamer* or any other modelling tool.

1.2 Thesis structure

This thesis is divided into six main chapters. This chapter introduced the motivation behind the study and defined the objectives and research questions. Second, the literature is reviewed to create a theoretical framing for the work and to support the research questions. Third, a brief review on the case study area is given, and the quantitative and qualitative methods used in this work are introduced through linking them to literature and justifying the selected data and methodology. Fourth, the results of the empirical parts of the study are presented. Finally, the research questions are answered, the results are discussed, and future study prospects are addressed, after which the work is concluded in brief. Additional material that is not necessary to be included in the main text is appended to the work.

2 Theory

To support the research questions, the theory chapter of this work will review literature related to the context of the study. The broad application area of this work and *hydrostreamer* as a tool may be set in water governance or water resources management, which refer to slightly different concepts. Some authors, such as Eden et al. (2016), Gupta et al. (2013), and Lautze et al. (2011), relate water governance to creating policies with process-oriented targets and goals, while water resources management is more related to implementation of these policies with outcome-oriented targets and goals. Related to the technicality of the approach taken in this study and the interviewee profiles, the empirical parts of this thesis consider more water resources management than water governance. However, since water resources management is very closely related to water governance as a means to influence and implement it, it is worth reviewing some of the most typical issues of water governance.

It is known that water governance is a highly complex field which includes problems approaching the “wicked” state (Reed and Kasprzyk 2009). Thus, it may not be possible to formulate simple or objectively right answers to water resources management or water governance. This must be kept in mind when defining the value of *hydrostreamer* or any other modelling tool or a framework. The conceptual framework describing the theoretical framing of this study is visualized in Figure 1. It should be recognized that the empirical parts of this work represent only a small fraction of the whole conceptual framework, however solutions for complex problems are achieved via breaking down the complex problem into simpler components, and then integrating the solutions of those back together (Lund 2015). This suggests that contributing to one part of the framework will help other parts function better.

In the following sections, the parts of the conceptual framework presented in Figure 1 are described briefly. First, some of the most typical water governance issues are reviewed. Second, good practice of science-policy interaction and some of the main instructive frameworks and concepts are introduced. Third, environmental modelling is reviewed in theory with emphasis on participatory modelling, which is beneficial in solving issues like those in water governance. Finally, the role of uncertainty estimation in the modelling process and characteristics of uncertainty related to geographical information are addressed before moving to the methodology section of this work.

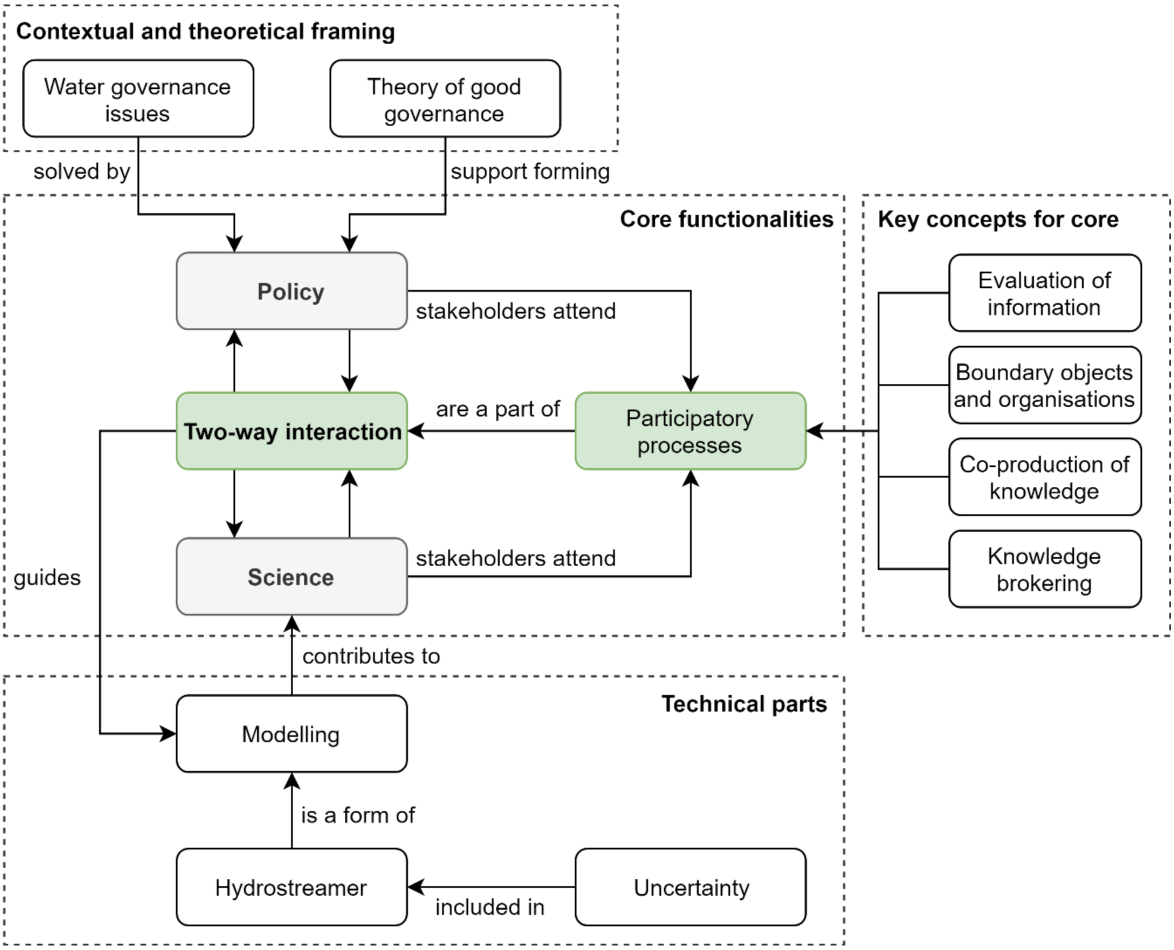


Figure 1. Conceptual framework of this study

2.1 Water governance

In a study conducted for Global Water Partnership, Rogers and Hall (2003) define water governance as *the range of political, social, economic and administrative systems that are in place to develop and manage water resources, and the delivery of water services, at different levels of society*. Highlighting its importance, water governance is closely related to United Nations Sustainable Development Goal (SDG) number 6, aiming for sustainable water management and providing sanitation for all (UN 2018). However, the core definition of water governance is not explicitly defined, as noted by e.g. Lautze et al. (2011). Also, Franks and Cleaver (2007) note that the GWP definition of water governance builds on the assumption that governance consists of linked political processes which result in varying outcomes at different levels of society, some of which are outside of the scope of participating stakeholder organizations. They also add that the definition alone doesn't state anything about how to practice good governance and which outcomes of the governance are good for whom. To decrease the vagueness of definitions, principles of implementing good water governance have been proposed by various researchers and multilateral organizations, such as Organisation for Economic Co-operation and Development (OECD) that has composed 12 widely adopted and cited principles covering the water governance context from information production to policy monitoring (OECD 2015).

However, principles such as those proposed by OECD (2015) are often very general in nature and not directly applicable. Key issues of water governance relate to the large number of political, social, and economic actors in both public and private sectors, multi-national administrative boundaries that do not coincide with hydrological boundaries, the complexity of multi-level governing structures, and differences in regional water laws (Gupta et al. 2013; OECD 2015; Rogers and Hall 2003). A plethora of interconnected water distribution and usage systems and the natural connectedness of hydrological networks further complicate the issues of water governance and force the governance system to adapt to local conditions (Gupta et al. 2013; OECD 2015). Reed et al. (2014) refer to catchment management case studies as a basis for their research, because catchment management often requires interdisciplinary and transdisciplinary work that also combines different forms of knowledge from multiple stakeholders. Pahl-Wostl et al. (2007) note that major regime shifts cannot occur in water governance without taking into account the societal context surrounding the regimes, which is why solving water governance issues expands far beyond technical aspects that often neglect situational and experimental knowledge. Due to water governance being such a complex topic with branches in both the social and physical worlds, not everything about water governance can be reviewed here, but some of the most typical issues are reviewed to provide contextual support for this study.

First, defining the ownership and borders of water is difficult. In their literature review, Gupta et al. (2013), present some concerns also relevant for this study. Water ownership may be claimed by a state or a community, which forces legislative instruments to adapt and may cause overlap with other regimes, especially when water governance laws reach out to governing land areas covered by watersheds. Moreover, state or sub-state boundaries offer convenient borders for institutional infrastructure, but watershed boundaries that don’t coincide with them create multijurisdictional systems for individual watersheds (ibid.). This is supported by Pahl-Wostl et al. (2007), who note that defining the spatial scale of governance by biophysical characteristics, such as river basins, leads to spatial misfits with administration.

Since water is in many circumstances considered as a common pool resource with cultural and social properties, informal governance may arise in cases of lacking state-led governance systems, which has been reported in the Mekong region by Lebel et al. (2005). In this informal governance, water is treated as a communal property instead of a state property, and informal organizations act as a replacement for institutionalized organizations (Armitage et al. 2015). Ostrom et al. (1999) address the issue of managing common pool resources: sometimes, state-led governance can benefit the management of common pool resources, but in some cases, local governance using traditional and customary management methods works better. They support this by stating that civilizations have been able to manage the same resources efficiently for thousands of years without strong top-down interventions.

In addition to difficulties in defining the borders and the decision-makers of water governance, the effects of water governance are distributed in a non-consistent manner. Lebel et al. (2005) divide the issues in water governance into three categories: scale, position, and place. Scale issues arise when governance units vary in size: from individual fields and villages to multilateral institutions. Not all scales can be governed with similar methods, and stakeholders at different scales may have varying interests and motivations for participating in the governance processes. While Lund (2015) notes that decentralized governance acts more efficiently than centralized governance, overly fragmented water governance is an issue in the 3S basin and has been addressed in institutional analyses conducted for each of the 3S countries (Nam et al. 2013; Seng et al. 2013; Sisouvanh et al. 2013). Second, position issues defined by Lebel et al. (2005) are simply caused by gravity: water is flowing from upstream to downstream, causing power asymmetries between upstream and downstream states especially when mixed with claims of absolute or relative sovereignty of water inside or on the borders of a state (Gupta et al. 2013; Zeitoun and Warner 2006). Finally, place issues relate to prioritization of some areas to others; for example, providing flood control and hydropower availability to larger cities with the expense of rural areas (Lebel et al. 2005).

Based on early water resources management criticism from the 1970s, Reed and Kasprzyk (2009) summarize the main challenges of water resources management as follows: 1) perfectly optimal solutions in complex systems consisting of humans and water may never be achieved; 2) managing water resources is a “wicked” problem; and 3) future works in water resources management require consensus-based and transparent decision-making. The authors state that optimality achieved via least-cost analysis methods requires both perfect problem statements and perfect information, which are impossible to formulate and gather in complicated water resource systems. Wicked problems, a notion also used by Armitage et al. (2015), can be described as unique, subjectively judged problems that lack definitive formulation but may possess making irreversible decisions and yield unknown or uncertain consequences. Reed and Kasprzyk (2009) conclude their study by remarking that traditional solutions for solving water resources management problems are not enough and closing the information gap between science and management is the key to create alternative solutions for problems as complex as described.

A comprehensive framework for solving water governance issues, Integrated Water Resources Management (IWRM), has been chosen for a practical way and a sub-target of completion of the UN SDG 6 (UN 2018). IWRM was first formally proposed by GWP (2000) as a *process which promotes the coordinated development and management of water, land and related resources, in order to maximize the resultant economic and social welfare in an equitable manner without compromising the sustainability of vital ecosystems.* In a recent report about progress on UN SDG 6, almost 40% of countries have reported practicing IWRM in a level ranging from medium high to very high (UN 2018). However, implementation levels are lower in Southeast Asia; Vietnam reports in medium low level implementation whereas Cambodia and Lao PDR do not even have sufficient data to report (UN 2018). Critique expressed towards implementing IWRM in developing countries include lack of practical instructions how to apply it, dysfunctional or missing institutional frameworks (Gupta et al. 2013), pre-defined but vague IWRM goals that diminish true participation, disregard of local contexts (Lautze et al. 2011), and inability to include social effects (Cook and Spray 2012). Therefore, IWRM may be the best of current comprehensive solutions for water resources management, but its implementation in developing countries is hindered by multiple issues.

Legislation related to freedom of information (FOI) and right to information (RTI) and attempts to open governance have been aiming for more democratic and transparent governance processes (Kosack and Fung 2014). In addition, open data has recently been advocated strongly across scientific disciplines. Important criticisms towards open data have however been expressed by for example Gurstein (2011), who notes only making data open and accessible is not enough, instead, the owners of the data must also ensure that it is possible for the end users to employ the data effectively. Gurstein (2011) argues that the open data should be available for effective use by all of those who may benefit from it instead of only those who have the technical capabilities and appropriate skills to exploit it. He states that this should be already included and considered when planning to distribute open data. Finally, Pedregal et al. (2015) emphasize the need for making open data truly available to everyone while addressing that there's potential to start significant paradigm shifts in water governance through open data. However, they add that opening data requires additional work by environmental managers and implies sharing power with the shared data – which all of those in power may be not willing to do.

In the end, parts of water governance and environmental management issues can be attributed to power inequalities between stakeholders. In the 3S basin, power inequalities exist both between state actors, such as ministries and environmental agencies, and between state and extra-state actors, such as non-governmental organizations and companies investing in natural resources (Pech and Ranamukhaarachchi 2013). Harnessing rivers with hydropower and irrigation schemes by decisions made within a defective governance framework cause controversy in development with the expense of local people, one of the most extreme effects being resettlement from traditional habitats (Ham et al. 2015; Harris 2016). Not everything related to water governance is only exploitation and abuse, though. For example, in all 3S basin countries, access to electricity is increasing both in rural and urban populations, as is the gross domestic product per capita, too (World Bank 2018). Either way, efficient science-policy interaction and transparent governance are in demand, about which literature will be reviewed in the following sections.

2.2 Science-policy interaction

As seen, water governance issues are not solved trivially, and the high number of interdependent actors and boundaries complicates both defining and solving the problems. Literature agrees that the root of good governance is in good science-policy interaction and the integration of social and physical sciences (e.g. Armitage et al. 2015; Lund 2015). In addition, Pahl-Wostl et al. (2007) argue that a democratically legitimate process should give a chance to participate for anyone who is affected by policy decisions. Studies on good governance and frameworks are plenty, but none provide out-of-the-box answers to unique problem settings. Next, some of the most applicable studies are reviewed and reflected on the case study area.

It is widely accepted that bringing stakeholders together and facilitating interaction between scientists and policy-makers will improve governance (Armitage et al. 2015; Cash et al. 2006; Jahn et al. 2012; Kosack and Fung 2014; Meadow et al. 2015; Reed et al. 2014). In addition, a linear process in which science states how policies should be formed has been identified inadequate while co-production of knowledge by science and policy stakeholders together is promoted (Armitage et al. 2015; Gupta et al. 2013; Meadow et al. 2015; Roux et al. 2006; Wehn et al. 2018). Co-producing information related to complex environmental governance issues may also invoke or involve social learning, in which stakeholders learn from each other and achieve better problem definitions and solutions (Wehn et al. 2018). Despite extensive research and promotion of good practices, gaps between scientists and policy-makers exist because of different terminologies and communication preferences, for example (Eden et al. 2016; Lemos et al. 2012). Moreover, interaction between stakeholders is often conducted without systematic and research-based grounding or evaluation (Reed et al. 2014).

First, interaction between science and policy must be facilitated. From the research devoted to the topic, this work closer inspects a few key concepts: evaluation of information, boundary work, co-production of knowledge, and knowledge brokering. These methods may be used to help in creating effective knowledge exchange, which is defined by Reed et al. (2014) as a process of producing, sharing, translating, and transforming information between people that are set in a social context. They propose five principles of effective knowledge exchange: 1) design, 2) represent, 3) engage, 4) impact, and 5) reflect and sustain. These principles aim for facilitating long-term and tailored knowledge exchange between carefully identified stakeholders, providing benefits for everyone involved and sustaining the knowledge exchange process for future use (Reed et al. 2014). The challenge for scientific researchers is to move away from the linear knowledge exchange process and purely technical approaches, and dive into the social context in which their research results are used (Reed et al. 2014).

Meadow et al. (2015) define co-production of knowledge as *the process of producing usable, or actionable, science through collaboration between scientists and those who use science to make policy and management decisions.* Like Reed et al. (2014), also Meadow et al. (2015) argue for conducting collaboration through established protocols based on literature: established practices are easier to guide during and evaluate after the knowledge co-production process. Engaging stakeholders may be divided into four modes in the order of tightness of relationships between scientists and stakeholders: contractual, consultative, collaborative, and collegial modes (Meadow et al. 2015). The authors state that the most legitimate results of knowledge co-production will likely be found through collaborative or collegial modes, but not every science-policy interaction necessarily requires either of the tightest modes.

Regarding the water governance situation in the 3S river basin and the technically focused aim of this study, one potential solution from the study by Meadow et al. (2015) would be a transdisciplinary approach. While highly promoted in research, formal definition of transdisciplinarity had been lacking until Jahn et al. (2012) who define it as a research approach that combines societal and scientific research fields by integrating scientific and extra-scientific knowledge. Again, their research promotes collaborative and inclusive interaction in complex problem settings. The authors address a difference between transdisciplinarity and interdisciplinarity by noting that transdisciplinarity reaches out to extra-scientific knowledge and smooths scientific discipline borders away. However, interdisciplinary and single-disciplinary science are still considered as essential parts of the transdisciplinary process. Finally, Jahn et al. (2012) note that wicked problems may be the “ideal” setting to apply their transdisciplinarity framework.

Despite that transdisciplinarity may appear conceptually sound and desirable, a more concrete and widely adopted solution applicable for a variety of problem settings is boundary work (Guston 2001, Meadow et al. 2015). Conceptual boundaries exist between stakeholders due to factors such as lack of consensus about the problem being addressed (Michaels 2009) or different goals of collaboration (Cash et al. 2002). In addition, Cash et al. (2002) remark that some stakeholders may be interested in retaining the boundaries for personal benefits. Michaels (2009) notes that the boundaries between science and policy are socially constructed while Clark et al. (2016) also add that power relationships may hinder collaboration in developing countries and successful boundary work must carefully give attention to this. In addition, Clark et al. (2016) and Roux et al. (2006) emphasize the need for combining different forms of knowledge from science and society, and conveying technical information in an adequate manner while conducting boundary work.

Often, boundary work concentrates on creating and managing boundary objects and boundary organizations (Guston 2001). Boundary objects enable discussion between stakeholders in different sides of the boundary by providing a common starting point despite that the goals of applying the boundary objects may differ: for example, patents or environmental models can act as a boundary object (Eden et al. 2016; Guston 2001). Boundary organizations are defined as independent groups that are separated from the stakeholders on each side of the boundary, while still having knowledge about both, which can perform tasks that are useful for all involved stakeholders (Scott 2000). One example of a boundary organization in the 3S basin is the Mekong River Commission (Lebel et al. 2005). The challenges of boundary work and creating well-functioning boundary organizations lie in achieving consent of action and trust from all stakeholders related to fairness and reliability (Guston 2001). Successful boundary organizations can facilitate beneficial boundary work for everyone involved (Guston 2001). This is supported by Roux et al. (2006) who note that proper two-way interaction between scientists and managers requires recognizing a common knowledge interface over a subject to enable co-operation.

A slightly analogous concept to boundary work is knowledge brokering: in his commentary, Meyer (2010) defines knowledge brokers as *people or organizations that move knowledge around and create connections between researchers and their various audiences.* In contrary to boundary organizations, the role of knowledge brokers is more active and individual persons may act as one. Meyer (2010) also proposes that the information flowing through knowledge brokers becomes dissembled and reassembled into brokered knowledge, which should serve local audiences better. Michaels (2009) states that the intent of knowledge brokering in environmental policy-making is to create information which wouldn’t otherwise be available to decision makers through intermediaries that act as knowledge brokers. Again, the policy or problem setting closely affects the choices of knowledge brokering method, which can vary from light informing in the form of fact sheets of websites to intensive collaboration and co-production of knowledge between knowledge brokers and stakeholders (Michaels 2009). Furthermore according to Michaels (2009), the role of a scientist as a knowledge broker may vary from simple problem solving to advocacy; it has also been reported by Meadow et al. (2015) and Reed et al. (2014) that scientists and researchers should be encouraged and challenged to start acting as knowledge brokers.

Coordinating science-policy interaction by using numerous concepts and frameworks aims for creating information that is usable and valuable. Research devoted to evaluating this has been conducted by for example Cash et al. (2002) and Cash et al. (2006) who create conceptual metrics to measure information and apply these to real-world phenomena. In Cash et al. (2002), the authors propose a salience-credibility-legitimacy framework: 1) salient information is relevant and timely for a stakeholder and can be adopted beneficially, 2) credible information is scientifically plausible and technically adequate, and 3) legitimate information is perceived as unbiased and fair by the affected stakeholders. However, trade-offs between these factors exist due to different conceptions of each quality by stakeholders in different cultures and negative changes in some qualities when others are improved (Cash et al. 2002). In another study, Cash et al. (2006) give four tasks for organizations between scientists and policy-makers: 1) convening (bringing stakeholders together), 2) translation (literal translation from language to another or metaphorical from a terminology system to another), 3) collaboration (co-production of knowledge), and 4) mediation (reconciling interests so that all parties achieve beneficial results). Completing these tasks as a part of boundary work may help in creating information that complies with the salience-credibility-legitimacy framework and is perceived adequate by the public (Cash et al. 2006).

The final concept in this section is transparency and its benefits in governance; the empirical parts of this work are closely linked to transparency through open-source tools. Kosack and Fung (2014) study the benefits of transparency in governance and divide governance scenarios into five different worlds by 1) evaluating the willingness of societal service providers and policy-makers and 2) short and long routes of accountability[2]. Based on reviewing case studies in social situations classified to each world, they conclude that being context-aware of the transparency intervention conducted and the positive willingness of policy-makers increase the chances of the transparency reform to succeed. However, the authors note that in extreme conditions where everyone in charge of services and policy-making is resistant to conducting reforms, transparency interventions are not always successful. Despite that the study by Kosack and Fung (2014) draws its examples from communal services, such as healthcare, their conclusions agree with studies presented before: context-aware transparency interventions that overcome boundaries between stakeholders are the most favorable approach to pursue. In addition, some water governance related studies (Eden et al. 2016; Lautze et al. 2011; Tortajada 2010) note that transparency is a characteristic of good water governance and enables better science-policy interaction.

Transparency and citizen participation have also been studied before by Fung (2006) who proposes a three-dimensional space to locate the degree of public participation and its effects. The axes represent authority and power (from educating individual participants and benefiting them by this to direct authority over others), participant selection (free to anyone – limited number of expert administrators), and communication mode (listening only – providing technical expertise). As typical, Fung (2006) doesn't propose universal solutions for participatory governance but recommends that participation should be planned carefully to achieve concrete results in form of enhanced governance and inclusion of local citizens' knowledge that wouldn't be included otherwise. Like Kosack and Fung (2014), his study is more concentrated on individual people; however, the benefits of transparent governance should be applicable in the current study too considering especially the complex setting of water governance issues and the life-sustaining role of water and water systems.

2.3 Environmental modelling

As it is clear that scientific information is an essential part of policy-making and good water governance, the production of scientific information in the context of this work should be addressed. Chapter 2.2 concluded that participatory science-policy interaction is beneficial for solving water governance problems, which means that the environmental models used should also be translatable and understood by all involved stakeholders. Problems in this translation and co-operation between expert and local knowledge holders have been reported by Clark et al. (2016) and Eden et al. (2016), and trouble in focusing research efforts and

---

[2] In the long route of accountability, the only way to affect service providers is through politicians in a democratic system (Kosack and Fung 2014). The short route of accountability refers to service providers being directly accountable for their services to the public, and the public may change the service provider in dissatisfaction (Kosack and Fung 2014). In water-related concerns, such as irrigation or drinking water, the "service providers" may not exist, and the only way to influence policy-making is through official government, which in turn might not be listening to individuals' opinions in the 3S river basin.

integrating them with stakeholder interests by Liu et al. (2008). Eden et al. (2016) also note that technical environmental models may be used as boundary objects, which are viable in enabling science-policy interaction as addressed in Chapter 2.2. They remark that models can provide basis for co-production of knowledge, but the models should be a neutral platform, scientifically sound, and defensible to increase confidence and credibility. This is supported by Voinov and Bousquet (2010), who note that stakeholders in dispute may find common ground for discussion through models, and Jakeman et al. (2006), who state that accessible models can increase participation in environmental management. Therefore, it is worth reviewing some theory of environmental modelling with a focus on participatory modelling since the problems being addressed in this work require collaborative solutions.

Beven (2008), complemented by Jakeman et al. (2006), addresses the primary aims of environmental modelling as either gaining and showing understanding of a system driving an environmental process or producing predictions for practical applications. The difference between these two arises from the final aims of the modelling work: practical modelling results for decision-making purposes can settle for imperfect results while the traditional way of modelling for understanding would always require revising the models to the current best understanding of the system being modelled (Beven 2008). There's always a long way from perceiving the physical environmental system to quantified modelling results. According to Beven (2008), formulating the environmental model includes three stages: 1) the perceptual model, which contains knowledge as complete as possible about the complexities of the system being modelled, 2) the formal model, which translates the knowledge from the perceptual model to a mathematical form, and 3) the procedural model, which finally solves the modelling task in form of computer code. Abstractions in this transformation from qualitative understanding to quantitative prediction are inevitable and cause uncertainty in each of these stages, about which will be discussed more in Chapter 2.4.

Including key stakeholders in the modelling process aims for increasing their knowledge of the modelling process and for better assessment of overall impacts of proposed modelling solutions (Jakeman et al. 2006; Voinov and Bousquet 2010). Both studies define a stakeholder as anyone who has interest on the ongoing modelling process; however, it is also considered that in natural resources modelling, the participating stakeholders are typically modelers, managers, and policy-makers instead of members of the public who may not have equally high interest on the subject. On average, increased system knowledge should improve management decisions (Jakeman et al. 2006), which is a good motivation for participation. Voinov and Bousquet (2010) note that the two aims are present together in participatory modelling, and the best results are achieved via a two-way knowledge exchange process between researchers and impacted stakeholders. Forming an explicit conceptual model[3] with stakeholders and practicing transparent modelling are promoted to make the stakeholders familiar with the subject – otherwise unfamiliarity with the models may hinder participation and make the stakeholders feel excluded (Liu et al. 2008; Voinov and Bousquet 2010). Voinov and Gaddis (2008) add that the result of a modelling process is not the only goal; improved understanding of the modelled system and potentially avoided conflict and misunderstanding are as valuable goals as are the results. In addition, they state that the stakeholders should have a genuine interest on the modelling process since otherwise the will to participate and express opinions during the process will not be achieved.

---

[3] In Liu et al. (2008) and Voinov and Bousquet (2010), the conceptual model refers to a non-computational model that explains the general behavior of the phenomenon being modelled including but not limited to inputs, assumptions, simplifications, and outputs.

An important issue regarding hydrological modelling is the uniqueness of place (Beven 2000), which is also one of the key characteristics of problems related with geographical data (Zhang and Goodchild 2002). At catchment scales, problems occur in representativeness: field measurements and modelling results will not represent the complexity of the whole catchment and cause erroneous extrapolation from data points to the rest of the catchment (Beven 2000). In addition, parameterization and calibration of parameters, which are often required in hydrological modelling as parts of creating the formal model from the perceptual model (Beven 2008), suffer from the fact that parameters derived from a given area may not be usable in other areas at all (Beven 2000). The issue of parameterization is also underlined by Jakeman et al. (2006): complicating environmental models too far may lead to overparameterization that will render a model useless beyond the case in which the parameterization and calibration are conducted.

Regarding model complexity, which may decrease the willingness to participate due to lack of understanding, Voinov and Bousquet (2010) and Voinov and Gaddis (2008) note that while simple models may not produce as accurate results as highly complex models, the simple models may be easier understood and better used in participatory environmental modelling. Overly complex models are also considered to contribute negatively to finding the optimal trade-offs between salience, credibility, and legitimacy (Liu et al. 2008). Jakeman et al. (2006) note that a modelling process can be started with a very crude and simplified model that would serve as an explorative tool before more precise models are created.

Participatory and transparent modelling may be used as an instrument to “level the playing field” or “level the players”, i.e. to smooth power relationships between stakeholders and make including all affected stakeholders easier (Campo et al. 2010, Zeitoun et al. 2017). Zeitoun et al. (2017) gives *increasing the technical or negotiations capacity, legitimacy, or authority of the less powerful side* as means to this: building weaker sides’ capacity aims for challenging the existing powers through contestation rather than influencing them through compliance. Zeitoun and Warner's (2006) framework on hydro-hegemony consisting of riparian position, three elements of power, and exploitation potential of an actor (a state in their study) over another is partly related to this study. In particular, the second element of power called “controlling the rules of the game” by Zeitoun and Warner (2006), may be intervened through transparency since equal potential to create new rules and revise old ones would be available. Zeitoun and Warner (2006) note that the power balances between riparian states ultimately define the interactions in between. Therefore, more equal power balance could lead to fair co-operation, and more transparent and accessible environmental modelling could provide means to smooth this power balance.

2.4 Uncertainty in modelling

An important factor closely related to both scientific credibility of modelling and legitimacy of the results perceived by stakeholders in participatory modelling is the evaluation of modelling uncertainties; it has been noted also in water governance studies (e.g. Armitage et al. 2015; Eden et al. 2016; Reed and Kasprzyk 2009). Natural resources managers are often forced into conditions in which they have to make major decisions under considerable uncertainty to achieve results that are equitable for various groups (Beven 2008; Jakeman et al. 2006; Liu et al. 2008). These conditions include multiple spatial and temporal scales and span across various disciplines (Jakeman et al. 2006).

Traditionally, researchers have estimated and evaluated statistical uncertainties with respect to their modelling results, but this may not be the overall largest contributor to total uncertainty as Refsgaard et al. (2007) note. The authors define uncertainty in environmental modelling context with respect to an individual decision maker, who is uncertain if they are lacking confidence about specific outcomes of an event. In their definition, the classical statistical uncertainty is just one of the potential reasons for causing the lack of confidence. Therefore, the definition is an expansion of the classical statistical uncertainty and includes uncertainties in defining the social context of a study and uncertainties related to factors such as model structure and technical implementation (Refsgaard et al. 2007). Again, referring to Lund (2015) noting that small parts create larger ensembles, it is worth to quantify a known uncertainty to increase system knowledge. Also, Beven (2008) argues that uncertainty estimation should be rather a routine than an additional task in environmental modelling, and its goal is to achieve better decision-making in the end, notwithstanding that the optimal ways to assess and communicate uncertainty are not unambiguous. Even in the case of deterministic models and statistical uncertainty, the uncertainty of single-valued outputs must be carefully investigated to understand what the outputs supporting management decisions are representing and how can the related uncertainties affect them (Uusitalo et al. 2015).

The uncertainty estimation within a modelling process should start from determining the acceptable accuracy limits and sources of uncertainty with the stakeholders involved (Refsgaard et al. 2007). The boundaries of acceptability and the methods used to estimate uncertainty will vary case-by-case (Uusitalo et al. 2015). Jakeman et al. (2006) present a drawback on assessing uncertainties related to large, integrated models: the model complexity makes evaluating all uncertain parameters and their combined effects impossible, and always limits the uncertainty analysis to an extent. However, they also state that if the application of the uncertainty analysis is not very critical in terms of accuracy, exact results are not essential.

Stakeholder involvement can also be used to reduce uncertainty: analyzing statistical uncertainties computationally might not be really feasible with stakeholders, but the stakeholders may help notably in assessing conceptual and contextual uncertainties (Refsgaard et al. 2007). However, Beven (2008) note that some uncertainties may not be easily included into formal decision support frameworks, and the communication of uncertainty to the stakeholders should be carefully considered, as should also be the contextual framing to reach out to all affected stakeholders. Uusitalo et al. (2015) remark that developing deterministic models to include uncertainty estimation must be done in a transparent and a justifiable way to avoid creating misleading guidance for the decision makers.

What is common between the notion of uncertainty used by Refsgaard et al. (2007) and the classical statistical uncertainty is that something is not known for sure, and the effects of not knowing something implies effects on model outputs. In this study, the uncertainty conforming to the input uncertainty definition proposed by Refsgaard et al. (2007) is related to elevation data set (DEM), which is a form of geographical information. The main paradigm related to geographical information uncertainty is that it is not possible to express an explicit truth about the endlessly complex geographical reality (Zhang and Goodchild 2002). Uncertainties in individual entities contained within processes of deriving topographic quantities propagate in a complicated way that no analytical means of uncertainty estimation can typically address (Zhang and Goodchild 2002).

Geographical information uncertainty can be related either to deficiencies in capturing the data, in this study by a satellite-borne instrument, or to the necessary abstractions and simplifications in presenting the data in digital form (Zhang and Goodchild 2002). It is easily imagined that all tiny variations in the geographical reality cannot be expressed in digital form, which is restricted by data set spatial resolution and numerical accuracy, for example. Zhang and Goodchild (2002) note that the discrepancy between the modelled reality and the geographical reality always exists and contributes to inaccuracy and uncertainty, which both in turn contribute to decision-making. Therefore, uncertainty is always present in applications related to geographical data.

In this study, not only the simplification of geographical reality into digital data contributes to uncertainty, but also the chain of abstraction from field-based reality to object-based reality (Zhang and Goodchild 2002). In their definition, the field-based reality refers to a continuous function that has a value everywhere (in this study, a DEM grid and a runoff grid), while the object-based reality concerns discrete objects (river lines and discharge values attributed to them). Finally, Zhang and Goodchild (2002) remark that uncertainty estimation is easier when starting from field-based reality and deriving multiple realizations of objects under subjection to uncertainty. Closely related to this, there is an inevitable chain of abstraction from elevation values to discharge computation that induces inaccuracy and uncertainty. This is further addressed in the methodology section of this work.

3 Methodology

This chapter provides a brief introduction to the case study area in the 3S river basin and describes the data and methods used in the empirical parts of this study. Both qualitative and quantitative methods were applied: quantitative methods relate to estimating a type of statistical uncertainty in *hydrostreamer* while qualitative methods are related to defining the value of *hydrostreamer* and open-source solutions in general. In the following sections, the applied methods are justified with references to literature and explained in brief. The computer code used in the quantitative uncertainty estimation is not presented in the appendices of the work but is available in GitHub repository forked from *hydrostreamer* in a folder named uncertainty (Virkki 2018). The materials used in the qualitative parts of the study are presented in the appendices.

3.1 Case study area

The case study area related to this work is in the 3S river basin (Sekong, Sesan and Sre Pok rivers), a major tributary to the Mekong River. It supplies approximately 18 percent of the Mekong's annual discharge and sustains livelihoods for around four million people (Constable 2015). The collaborative water resources management in this transboundary basin across the states of Lao PDR, Cambodia, and Vietnam is threatened by complex issues. While the study area is generally not suffering from water scarcity due to a seasonal monsoon climate (Babel and Wahid 2009), several threats are posed by other factors. The Mekong and its tributaries are invaluably important for the people living in the area, and changes in river streamflow may have drastic consequences in their lifestyle. Particularly, hydropower development has been concluded to alter seasonal streamflow trends significantly (Hoang et al. 2019; Räsänen et al. 2015), and due to this, effects on food security and biodiversity occur (Ziv et al. 2012). Therefore, practicing good water governance is essential in the 3S river basin for sustaining the livelihoods of the local people.

Managing the available water resources in the 3S river basin is complicated due to complex transboundary and national government frameworks and desires to promote economic growth by making use of natural resources. Both Lao PDR and Cambodia have aimed to reach a status of electricity exporter through building additional hydropower despite that electricity access rates in both countries is still relatively low (Constable 2015; Geheb et al. 2015; Ham et al. 2015). In addition, the water governance setting is highly politicized and accused of having a top-down led management system that excludes the opinions and knowledge of people most affected (Ham et al. 2015; Harris 2016).

The complex water governance situation in the 3S river basin has been researched from development aspects by various international organizations, such as Asian Development Bank (ADB 2010), International Rivers (Grimsditch 2012) and International Union for Conservation of Nature (Watt 2015). Hydropower facilities, which inarguably cause the most significant effects on streamflow, have lately been built and planned with strong opposition from the local people and controversial resettlement policies. The Lower Sesan 2 dam in the 3S basin and the Xayaburi dam in mainstream Mekong are prime examples of poor governance practiced by Cambodia and Lao PDR (Baird 2016; Ham et al. 2015). To conclude, water resources management and water governance are clearly not conducted in an adequate manner in the 3S basin and the poor practices persist despite the interest of academic and developmental organizations in the area.

Figure 2. Map of the study area

A map of the study area is presented in Figure 2, and a terrain slope map is presented in Figure 3. As seen, the 3S river basin area is variable both in elevation and slope, which are the driving factors behind the uncertainty estimation conducted in this study. The three rivers begin from mountainous areas and run down through lowlands before uniting into one outlet that connects with the Mekong, further running down to ocean in south. The river network in Figure 2 is presented as provided in HydroSHEDS data (WWF 2018). Regarding rainfall, significantly higher amount of precipitation falls on highlands, leaving especially the lowlands of Sre Pok basin drier and contributing directly to annual runoff, which ranges from 250 millimeters in Sre Pok basin lowlands to nearly 2 000 millimeters in upper Sesan basin mountains (Constable 2015).

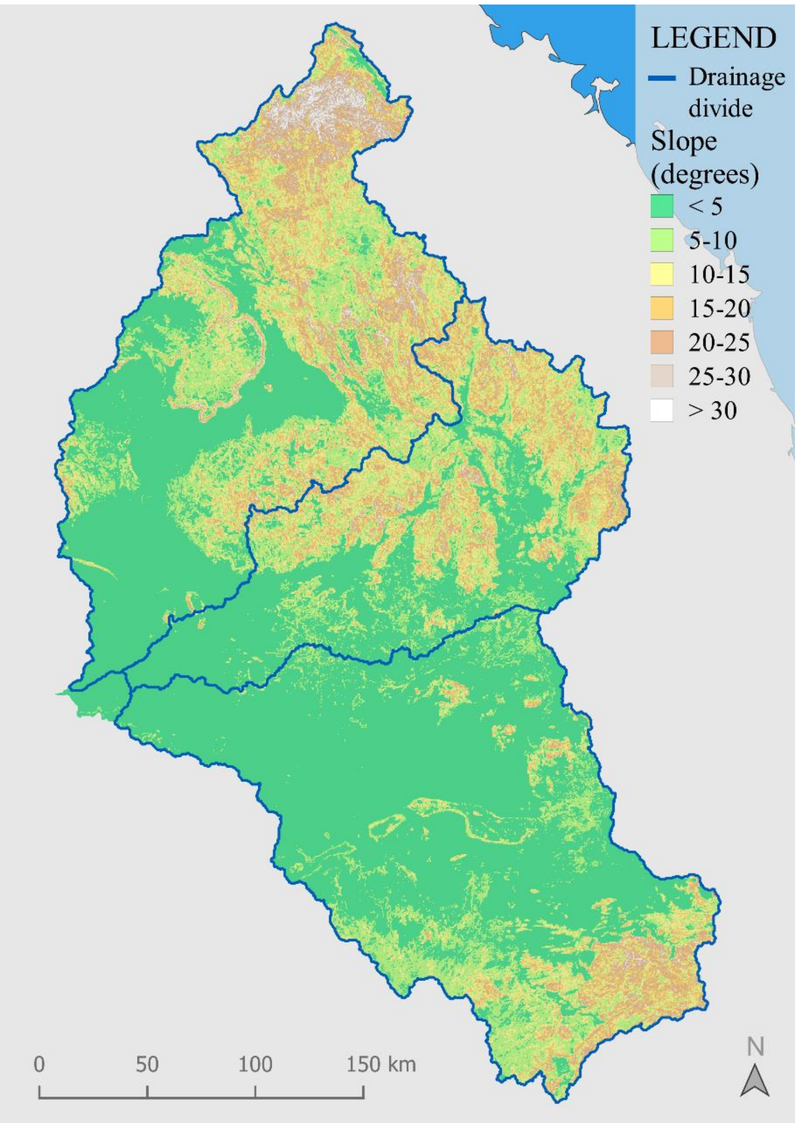


Figure 3. Slope map of the 3S river basin

Agriculture is the primary land use type in upper Sre Pok basin and near the towns of Attapeu and Kon Tum while more mountainous areas are covered in both deciduous and evergreen forest (Constable 2015). Most of the 3S basin inhabitants live in the Vietnamese parts of the basin where the population density is highest, however in population growth rate, Cambodia and Lao PDR surpass Vietnam (Constable 2015). Societal development indicators exhibit lower values in Cambodia and Lao PDR than in Vietnam: access rates to improved water sources, sanitation, and electricity have still been under 50% in many regions until some years ago (Constable 2015). In conclusion, the 3S basin is still a relatively undeveloped area, which contains potential for both population growth and exploitation of the relatively untapped natural resources.

While this study may not provide direct aid for the most critical issues of water governance in the 3S basin, *hydrostreamer* is closely related to one of the underlying factors behind the issues: river streamflow. This is the primary reason why this area was chosen for the study. In other reasons, the seasonal monsoon climate providing an abundant amount of water, variable topography, and relatively unregulated river network make the 3S basin a suitable demonstration area for this work. Finally, the previous works and existing contacts in the Water & Development Research Group in the area provide invaluable help to find interested participants.

3.2 Streamflow uncertainty estimation

The main method to quantify the spatial uncertainty related to *hydrostreamer* is Monte Carlo simulation: the process of deriving river networks and discharge estimates is too complex for analytical methods. Monte Carlo simulation is a general term describing methods that are based on selecting parameters from a random set for consecutive iterations of the model (Beven 2008). It is widely adopted especially in geographic information applications because analytical approaches, in which parameters and their respective uncertainties would be addressed one by one, are impossible due to the complexity of geographic information (Zhang and Goodchild 2002). Using Monte Carlo simulation is popular in research related to uncertainties in digital elevation models (Oksanen and Sarjakoski 2005a, 2005b; Wechsler and Kroll 2006) and has been commonly identified as a method to evaluate uncertainty related to environmental modelling (e.g. Uusitalo et al. 2014, Refsgaard et al. 2007, Matott et al. 2009). This study applies a simple Monte Carlo simulation in which the elevation values of individual cells in the input data are altered by drawing a new elevation value from a normal distribution that is defined by parameters with respect to local terrain slope.

3.2.1 Data used in this study

During the uncertainty estimation, two main data sets were employed to model the streamflow and create discharge estimates, and two auxiliary data sets were used to provide with ground truth data about the rivers and the locations of hydrological monitoring stations (HYMOS). Chapter 3.2.2 will address that *hydrostreamer* requires a runoff data raster and a river network as inputs. In this study, the runoff data raster was applied without modifications while the river network was iteratively created from a digital elevation model (DEM) that was subjected to uncertainty during the study. The two additional data sets were major river lines provided by OpenStreetMap (OSM) and the locations of HYMOS stations, at which the modelling results were observed.

The runoff data used in this study consisted of monthly timesteps from year 1971 to year 2005. The study computed monthly averages for the time period and applied those as the representative runoff. The spatial resolution of the data used was 0.5 geographical degrees, which translates to approximately 50-55 kilometers in the study area[4]. In the data, monthly surface runoff was expressed as the value of each raster cell, which was later assigned to

---

[4] In the study area, the length of a latitude or a longitude degree is approximately 105-110 km (to calculate, see for example NGA (2018)).

river segments and transformed into discharge estimates. The hydrological model originally used to produce the runoff is called PCR-GLOBWB (van Beek and Bierkens 2008), and it applies an earth system model ESM2M as an input climate forcing model (GFDL 2018). In the original runoff model, no social effects (i.e. human influence on streamflow) are included. Since the runoff data is only considered as an input for this study, no further inspection to the data structure, its production process, or uncertainties is provided during the work. Throughout the study, the runoff data is kept constant to generate results that only exhibit the effects of spatial uncertainty in the outputs.

One of the most notable sources for consistent and good quality near-global-scale river networks is HydroSHEDS (WWF 2018) that also provides drainage direction rasters[5] and pre-derived drainage basins that can serve as an area of interest. HydroSHEDS river network has been used by Kallio et al. (2018) during the development of *hydrostreamer.* However, since this study considers the spatial uncertainty related to the river network to be used, HydroSHEDS was applied only as benchmark data to compute discharge values to be compared with the simulation.

The gridded elevation data set, which was subjected to uncertainty and from which the river networks were derived, was originally captured using radar interferometry during Shuttle Radar Topography Mission in 1999, and later void-filled with data from other global-scale elevation models (NASA JPL 2013). The spatial resolution of this data set was 3 arc-seconds[6] or approximately 90 meters for each side of a grid cell. An explicit documentation of the production process of the SRTM DEM is available in NASA (2015). The uncertainty of the SRTM data has been evaluated by numerous case studies, one of the first being conducted by NASA themselves (Rodríguez et al. 2006). A broader review on the related case studies is given in Appendix 1.

Finally, two auxiliary data sets were employed: river lines describing the major rivers (Sekong, Sesan, Sre Pok) extracted from OpenStreetMap (OSM 2018) and the locations of HYMOS stations digitized by Kallio et al. (2018) as a part of *hydrostreamer* development work. The river network data was employed in the study because the uncertainty estimation required a reference that was based on an existing river network and to ensure that the river networks of each iteration will reach the HYMOS stations, in which it is absolutely known that there is a river since discharge measurements are available.

3.2.2 Hydrostreamer

Hydrostreamer, developed by Kallio et al. (2018), is a tool written in R programming language that downscales low-resolution runoff data into a high-resolution river network by dividing the runoff into individual river segments and accumulating flow downstream. Three input data sets are required for using *hydrostreamer* while additional data sets enable more options to use it. Compulsory input data consists of runoff data, a river network, and an area of interest, all of which covering the desired study area. The data sets used in this study were presented in Chapter 3.2.1. Additional data includes a drainage direction raster or pre-defined drainage basins for each river segment contained within the river network. The area of interest may be substituted with the location of a basin outlet if applying estimation

---

[5] Here, drainage direction raster refers to a grid in which each cell contains a value representing the direction to which water flows from that cell, having discrete values in cardinal and half-cardinal directions.

[6] An arc-second represents 1/3600th of a geographical degree. Therefore, 3 arc-seconds translate to approximately 90 meters (see note 4).

options using drainage directions and only for one basin. A noteworthy detail related to *hydrostreamer* is that the package version used during this study did not allow handling any lakes or other water reservoirs as a part of the hydrological network. Therefore, all streamflow effects due to water storages are disregarded.

Two key concepts that are closely related to *hydrostreamer* and the uncertainty estimation conducted in this study are Voronoi polygons and Strahler order. Voronoi polygons are parts of a Voronoi diagram, which is a space-partitioning function over a set of features. Starting from a set of features, Voronoi diagram divides a space to parts inside which the nearest source feature is the same. In this study, the initial set of source features is the two-dimensional river network consisting of river segments, and the resulting Voronoi polygons are two-dimensional polygons that represent proximity areas inside which the nearest river segment is always the same. Second, Strahler order is a hierarchical river classification system, which first sets each river without tributaries as order one, and when equal-ordered rivers connect, the resulting downstream river gains an increased Strahler order (Strahler 1957). This hierarchy classification roughly represents the size of a river: downstream rivers consisting of multiple upstream sub-basins gain high Strahler orders while upstream rivers consisting of only small upstream sub-basins gain low Strahler orders. In this study, the largest Strahler order achieved was 7, but only in the main outlet of the 3S basin, which is why it is not shown in the results presentation.

Kallio et al. (2018) have validated the results produced by *hydrostreamer* with respect to observed discharge in the 3S basin. Knowing this, this study doesn't incorporate any measured data into the uncertainty estimation but compares the results with an application using a HydroSHEDS river network. This network was used by Kallio et al. (2018) to validate *hydrostreamer* and was deemed to produce accurate enough results with respect to measured discharge in the 3S river basin. Furthermore, this study takes advantage of inspecting the uncertainty in HYMOS station locations that were the ones used by Kallio et al. (2018) in validation, which means that in those locations, *hydrostreamer* outputs are acceptable with respect to observed flows. Below, the phases of deriving discharge estimates with *hydrostreamer* are addressed briefly.

(1) **Transform runoff raster into polygons with respect to the area of interest**. In the first phase, the runoff data is clipped to cover only the area given by the area of interest and then polygonised. As a result, the study area is covered with runoff units (polygons) created from runoff raster cells, each containing monthly runoff values for arbitrarily long time. It is assumed that the runoff falls uniformly inside one raster cell, and if cells need to be clipped, the remaining runoff is proportional to the area of the clipped part. In addition, *hydrostreamer* contains a function that can average all timesteps of one runoff data product into monthly average values. This method was applied in this study to decrease computation times.

(2) **Compute weights for each river segment to represent the amount of runoff caught.** The weights to be computed for each river segment represent how big of a fraction of the runoff is caught by each river segment. All weights inside one runoff unit (polygon) will sum up to one, meaning that all runoff will be caught by rivers inside the runoff polygon. Three methods were used in this study to compute river segment weights:

a. Weighting by length: the weights are directly proportional to the lengths of the river segments within each runoff unit.
b. Weighting by Voronoi diagram: Voronoi polygons of each river segment represent catchment areas, and the weights are proportional to the areas of these polygons within each runoff unit.
c. Weighting by drainage basin size: drainage basins are delineated for each river segment, and weights are proportional to the areas of the drainage basin parts within each runoff unit.

Weighting by length and Voronoi diagram do not require a drainage direction raster since the weights are based only on geometry of the river segments themselves. Weighting by drainage basins requires a drainage direction raster in order to delineate individual river segment drainage basins.

(3) **Compute runoff caught by each river segment by using the weights derived.** After the weights for each river segments are computed, the runoff from each runoff polygon is divided into river segments using the derived weights. Now, each river segments contains information about how much runoff it will catch in each timestep in the data.

(4) **Apply river routing to accumulate flow from all river segments.** When the runoff caught by each river segment in each timestep is finally known, the total discharge may be computed. At the time of conducting this study, only instantaneous flow accumulation was implemented in *hydrostreamer*. It assumes that contributions to upstream flow in transfer immediately (during the duration of one timestep) to downstream. It is known that when either computing estimates for larger (continental-scale) areas or more frequent than monthly timesteps, the speed of streamflow is too slow for instantaneous flow accumulation to work properly.

3.2.3 Uncertainty estimation process

The uncertainty estimation consists of two main processes: first, the uncertain river networks are created by randomizing a DEM, and then discharge estimates are computed with *hydrostreamer* for each of them. Despite that *hydrostreamer* is a deterministic tool producing identical outputs from identical inputs, the estimation process results in differing discharge estimates between consecutive Monte Carlo simulation runs due to changes in the river network. The complete uncertainty estimation process is presented in Figure 4, from which the process of creating river networks is presented in more detail in Figure 5. Creation of the river networks and simulating the uncertainty effects were programmed completely during this study while *hydrostreamer* code was taken as-is from Kallio (2018) using version 0.2.2. In addition to the base R package, the work takes advantage of various open-source R packages, which are presented with the versions used in Table 1.

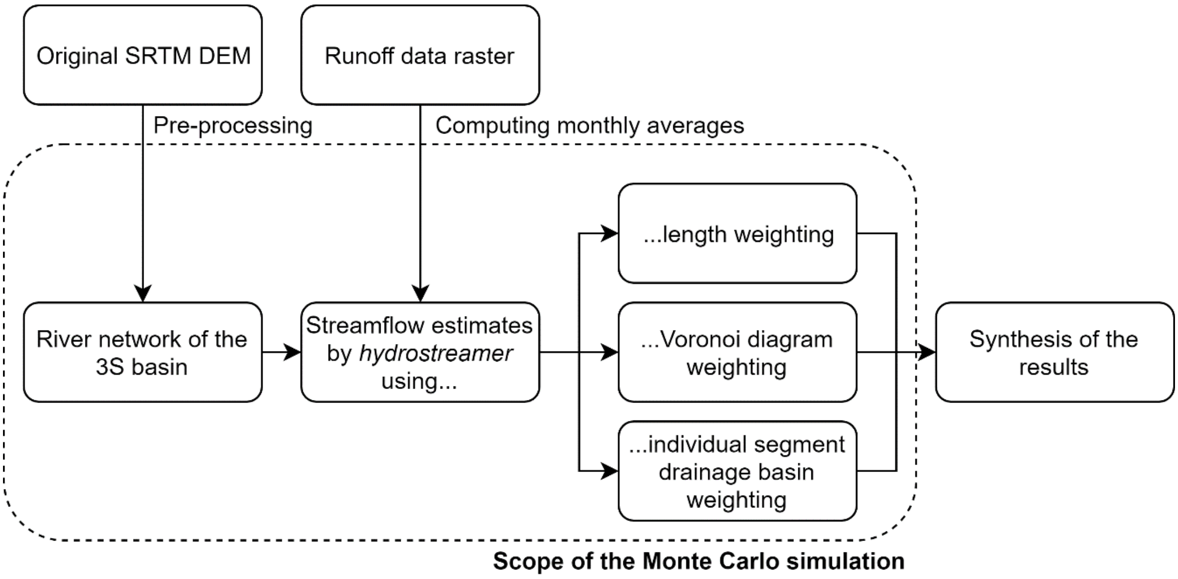


Figure 4. Uncertainty estimation

Table 1. R packages and versions used in this study

Package name	Version used	Utility
DataCombine	0.2.21	graphing
dplyr	0.7.6	analysis
ggplot2	3.1.0	graphing
ggpubr	0.1.8	graphing
gridExtra	2.3	graphing
hydroGOF	0.3.10	graphing
hydrostreamer	0.2.2	analysis
magrittr	1.5	analysis
ncdf4	1.16	analysis
raster	2.6.7	analysis
RColorBrewer	1.1.2	graphing
reshape2	1.4.3	graphing
rgeos	0.3.28	analysis
rgrass7	0.1.10	analysis
RSAGA	1.2.0	analysis
sf	0.6.3	analysis
spatial	7.3.11	analysis

Pre-processing steps before starting derivation of river network realizations[7] include computing a slope map, a mean error map, and a standard deviation map covering the study area, all of which hold the same extent and resolution as the input DEM. The slope map was computed using 8-neighbors algorithm proposed by Horn (1981), contained in the R raster package. Mean error and standard deviation maps were computed with linear functions based on SRTM accuracy case studies (mainly Mouratidis et al. 2010; Rexer and Hirt 2014; Shortridge and Messina 2011; Szabó et al. 2015). A closer review on SRTM accuracy studies

---

[7] A realization may refer to a potential state of a DEM that has been altered from the original by adding a random error field (Wechsler 2007). In this study, realization refers to a unique river network structure generated in each iteration due to changes in the elevation values of the input DEM.

is provided in Appendix 1. The functions for determining the shape of normal distribution from which to draw new elevation values for each cell were defined as

$$ME_{cell} = 0.2 * slope_{cell} \qquad (1)$$
$$STD_{cell} = 1.2 * slope_{cell} \qquad (2)$$

where ME stands for mean error, and STD stands for standard deviation. These definitions imply that both mean error and standard deviation increase linearly with slope, which has consistently been observed in case studies evaluating SRTM accuracy (Appendix 1). Furthermore, the properties of the error field should not be over-assumed when the absence of ground control points effectively eliminates methods based on deriving an error field from punctual observed errors, such as sequential Gaussian simulation (Wechsler 2007). The primary arguments for choosing this uncertainty estimation method were that no ground truth elevation data can be captured from the study area within the scope of this study and no studies about SRTM errors have been conducted in the nearby area to provide more detailed error descriptions. In addition, it has been observed in the literature that neither mean error nor standard deviation are constant in SRTM data, gaining magnitude with respect to slope. Satellite orbit as a driving factor for errors varying with respect to aspect has been reviewed in Szabó et al. (2015) and Shortridge and Messina (2011) but was not considered in this work.

After pre-processing, each cell of the DEM was related to three properties: measured height value, estimated mean error of the height value, and estimated standard deviation of the height value. To further simplify computing, the mean error map was subtracted[8] from the original DEM map to get a corrected height map (CHP). Another map from where parameters for normal distributions were drawn is the standard deviation map (SDP) From this point, the Monte Carlo simulation started, and the following phases of the process were repeated for a total of 500 times. The iteration steps are presented in Figure 5.

---

[8] The mean errors reported in literature have been mainly positive (Appendix 1): hence, SRTM data is overestimating the surface height and needs to be corrected down to represent heights closer to reality.

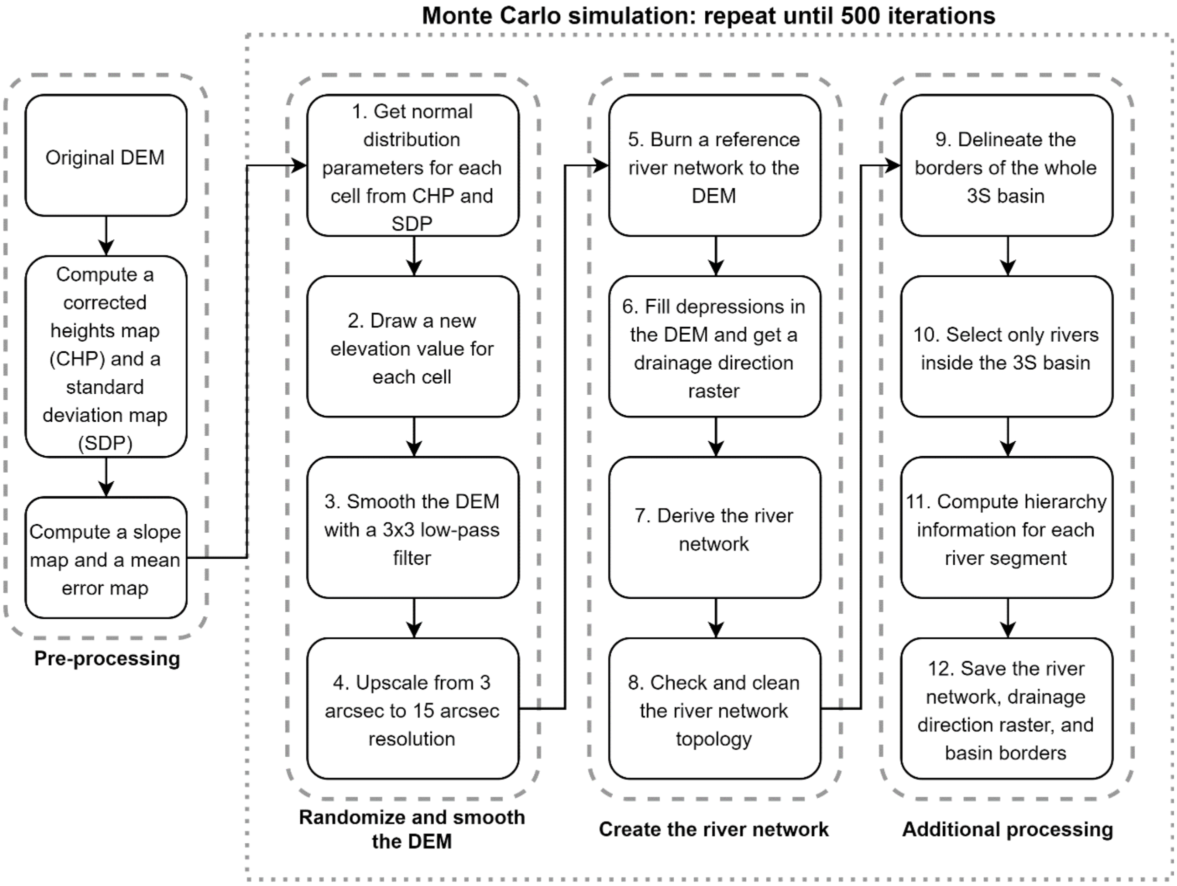


Figure 5. Creation of river networks

Phase 1. For each cell, draw parameters for normal distributions from the previously defined maps of mean errors and standard deviations.

Phase 2. Draw a new elevation value for each cell from normal distributions defined in Phase 1. By the definition of the standard deviation map, cells with higher slope values are more probable of gaining larger deviations from the original SRTM DEM elevation value.

Phase 3. Smooth the randomized DEM from Phase 2 by applying a low-pass filter with 3 x 3 cell kernel for each cell. This will increase the spatial autocorrelation of height values, which should be included in the simulation because both elevation values and SRTM errors are spatially autocorrelated (Oksanen and Sarjakoski 2005a; Shortridge and Messina 2011; Wechsler 2007).

Phase 4. Upscale the randomized and smoothed DEM from Phase 3 to 15 arcsec resolution (approx. 450 meters) by computing the mean elevation values in each 5 x 5 cell neighborhood. This upscaling phase is necessary to decrease the number of cells and further computation time. It should be noted that this phase again increases the spatial autocorrelation of the elevation values and also that every smoothing and upscaling step performed simplifies the representation of topography and decreases accuracy.

Phase 5. Burn[9] a reference river network to the DEM from Phase 4 using OpenStreetMap rivers. This will ensure that there will be rivers leading to observation stations that possess actual discharge data. It should be noted that the OSM river network is the only ground truth data used in the whole process and is thus assumed to represent the rivers correctly.

Phase 6. Fill depressions in the DEM from Phase 5 by using an algorithm proposed by Wang and Liu (2006) implemented in SAGA-GIS. Filling the depressions is necessary for deriving the river network in a way that there's only one outlet from the study area basin and there exist a continuous flow path from every basin cell to this outlet. While filling the DEM, a one-degree minimum slope between adjacent cells is preserved to avoid creating flat areas. After the filling, a drainage direction raster is created by setting the drainage direction to the direction of steepest descent (Wang and Liu 2006). The drainage direction raster is later used in deriving drainage basins for individual river segments when applying one of the three weighting methods of *hydrostreamer* and in deriving the borders of the whole 3S basin to be used as an area of interest.

Phase 7. Derive a river network based on the filled DEM from Phase 6. SAGA-GIS tool Channel Network is used to derive the river network from a drainage direction raster by following the steepest slope path (Olaya 2004). No upstream basin size restrictions are applied. The tool results in rivers that flow only in combinations of eight directions, cardinal and half-cardinal. A minimum length of 15 cells (approx. 6750 m – 9750 m) is set for the river to decrease the count of small rivers being formed.

Phase 8. Check and clean the river network topology since *hydrostreamer* accepts only polylines that are connected to each other only on endpoints. If the river network doesn't fulfill this requirement, break and reform the lines. This will ensure that all runoff will flow out of the basin through one outlet.

Phase 9. Delineate the complete study area drainage basin by delineating the drainage basin of the river segment from which the whole basin runs out. This is guaranteed to happen in each iteration because the outlet segment location is known with enough precision due to reference river network being burnt into the DEM in Phase 5 and because the DEM filling in Phase 6 ensures that continuous flow paths from every cell inside the basin to the basin outlet are created.

Phase 10. Select only those river segments that are inside the basin delineated in Phase 9 to decrease river segment count in later computations.

Phase 11. Compute hierarchy information from the river network selected in Phase 10. Strahler order and upstream segment count are attached to each river segment as attributes.

Phase 12. Save the river network, drainage direction raster, and study area basin boundary for later use in flow computations with *hydrostreamer* and fall back to Phase 1 again until a desired number of iterations is completed.

Once the Monte Carlo simulation was complete, the results were analyzed to remove gross outliers and to prepare for presenting the results. In total, only one iteration was rejected

---

[9] Here, burning refers to decreasing elevation values along each river enough to force river network derivation to end up creating a river in the place defined by the river data that was burnt.

after visual inspection of the results due to expressing extremely low discharge values. This rejection was acceptable since it was caused by the whole 3S basin breaking into two basins near the midpoint of the basin, which is not possible knowing the ground truth rivers from OSM. In addition, an unknown software issue caused 12 iterations using Voronoi weighting in *hydrostreamer* to produce invalid results, after which the same 12 iterations were rejected from other weighting methods too[10]. Finally, 487 of 500 iterations were deemed acceptable.

The results of the uncertainty estimation are presented in Chapter 4.1. Here, it is important to address the two main ways of creating the graphs presented: the graphs related to HYMOS observation points and the general assessment of uncertainty in all river segments of the basin area. First, the locations of the HYMOS stations were drawn as a separate layer, after which each river network with computed discharge was compared against the HYMOS stations and the river segment nearest to the crossing point of the station was chosen to represent flow related to the HYMOS station in question[11]. This was required because individual river segments couldn't be queried through common identifiers in different iterations due to the changing number of river segments. Further, the distribution of flow values in the selected river segments was analyzed and is presented in Chapter 4.1.

The general assessment of spatial uncertainty in the river segments inside the 3S river basin was a slightly more challenging task to implement. To approach this, rivers that lie in a roughly similar position in the study area and that have an equal Strahler order were sought after to evaluate the effect of uncertainty present in the DEM. Again, no common identifier between iterations could be used, so the query resorted to geometrical operations and is briefly addressed by the following steps:

1) Create a buffer area around each river segment in one of the iterations (any iteration).
2) From all other iterations, find all river segments completely within this buffer that also possess the same Strahler order than the initial river segment.
3) Compare the endpoints of the initial river segments and the river segments found in 2).
4) If the starting and ending points of the river segments are equal, consider that the rivers are representing a similar kind of a river segment concerning geometry (roughly equal length, significant especially in low Strahler order rivers) and to some extent, also upstream configuration (Strahler order).

Finally, those river segments from the initial river network realization that were found in 50 or more iterations (slightly over 10% chance to exist) were chosen as representative river segments for the uncertainty estimation. In total, 3038 river segments, which is approximately 82% of the roughly 3700 river segments typically found in the iterations, fulfilled this condition. By doing this, the uncertainty estimation was expanded from punctual samples to a more generalized solution.

---

[10] As Figure 4 shows, all weighting methods were applied for each different river network, making comparison of weighting methods possible.

[11] In the technical implementation, the HYMOS station was represented as a line. Considering that the stream burning from Phase 5 of the river network derivation, which reached out to all HYMOS stations, guaranteed that there always was a river line in the presence of each HYMOS station, the geometric intersection between the rivers of each iteration and the HYMOS station line always returned one and only one river segment from which the estimated discharge was read.

Finally, since the absolute discharge differs greatly with respect to the position in the river network, a relative uncertainty metric was adopted to make the results comparable with each other. In this study, the relative uncertainty of a river segment is defined as the ratio of standard deviation of the 488 discharge estimates and the mean of the same estimates. The definitions above will be recalled and discussed in the results section of this study.

3.3 Qualitative approaches

In this study, qualitative methods were applied to gather insights from the local context, to attempt evaluating the value of *hydrostreamer,* and to gain topics for discussion. Two key methods are addressed in further sections: an internet-based survey and semi-structured interviews. The primary ways to search for participants for the research were existing contacts provided by the members of Water and Development Research Group and institutional analyses of the 3S river basin countries conducted by Nam et al. (2013), Pech and Ranamukhaarachchi (2013), Seng et al. (2013), and Sisouvanh et al. (2013). During the research, a booklet describing the study and the quantitative results was composed, and it was sent to the potential participants via e-mail. The booklet is presented in Appendix 2. Approximately 60 organizations were contacted via e-mail, consisting of universities, research institutes, ministries, and non-governmental organizations (NGO). This approach was not successful; not a single response was achieved. However, the existing contacts inside WDRG proved better, and the interviewees forwarded their interest on the research to other people acting in the 3S river basin area.

Therefore, it may be concluded that the primary sampling technique to gather participants for this study was snowball sampling, in which a small group of people are contacted first, after which the contacted people distribute the word in their own networks and provide the researchers with more potential participants (Bryman 2008; Reed et al. 2009). In total, five participants responded to the survey and four participants were interviewed. One interviewed participant didn’t reply to the survey and two participants who replied to the survey weren’t interviewed due to lack of time when compiling this text. Furthermore, interest on the study was expressed by a few other persons and organizations in the 3S basin area and elsewhere, but those people didn’t participate in the study due to limited time resources.

3.3.1 Survey

Bryman (2008) defines a self-completion questionnaire as a set of questions that the respondent will answer to by themselves without the presence of an interviewer, which is what this study refers to as a survey. Compared to a structured interview, which would consist of the same questions only adding an interviewer to ask them, the key strengths of a survey relate to the convenience of answering for the respondents and the relative ease of analyzing the results (Bryman 2008). The main drawbacks of a survey include difficulties in forming the right set of questions to provide enough data and attaining the respondents interest on answering the survey (Bryman 2008). In this study, the survey was mainly a probe to search for topics of discussion in later interviews than an instrument that would provide data for statistical analysis. However, it should be noted that a well-formed survey will attract more respondents.

Bryman (2008) suggests means to make a survey more inviting to answer: keeping the presentation short and clear, providing clear instructions on how to respond, and forming closed question answer patterns in a coherent and unambiguous way. He concludes that including more closed than open questions is the preferred way to go in planning a survey, and the questions should be planned carefully to avoid leading questions and questions asking two things in once, for example. This study applied Bryman's guidelines in forming the survey: primary aims of designing the survey were to create a simple enough set of questions that can be answered in an adequate time without getting deep into technicalities beyond the booklet provided with the survey.

When the contacts to potential stakeholders were formed, a link to the internet-based (on Webropol platform) survey was sent with the message. The survey was chosen for a method of this study to collect preliminary data from a large number of participants and query their interest on participating in an interview. The survey outline is presented in Appendix 3 with wordings corresponding to the actual survey sent to the potential stakeholders. The survey questions were tested by WDRG personnel before opening the survey for stakeholders to ensure that the questions are valid and according to best practice guidelines introduced above.

3.3.2 Semi-structured interviews

Since the topic of this study revolves around defining the value of a tool and being rather abstract in nature, semi-structured interviews were chosen as a follow-up method for the survey to gather insights from expert stakeholders in the study area. In addition, it was not expected to gather a large number of participants, which disables precise statistical analysis. Other arguments for semi-structured interviewing in this study include the freedom to dive deeper into the topics emerged during the interview and the potentially variable background of the participants. Semi-structured interviews have been previously used in research concerning social aspects in water governance (e.g. Cash et al. 2006; Felipe-Lucia et al. 2015; Lienert et al. 2013; Stein et al. 2011) and as a method of conducting systematic stakeholder analysis (Reed et al. 2009).

Bryman (2008) lists two main strengths of qualitative interviewing, of which semi-structured interviewing is a form: setting more emphasis on the interviewees' opinions rather than on quick answers to strict questions set by the researchers, and the flexibility of the interviewing structure that can be used in adjusting the interview during the research. However, he also notes that the interpretation of the results achieved via qualitative interviewing are harder and more time-consuming to process than those from fully structured interviews. For a semi-structured interview, it is required to prepare an interview guide that will provide the interviewer with preliminary questions that should be asked in some order during the interview, but the interviewee has a great freedom to answer while also the interviewer may pick on the topics emerged (Bryman 2008).

For the interviews, an interview guide presented in Appendix 4 guide was prepared. However, not all interviews consisted of the exact questions in exact wordings, but all interviewees were presented the topics listed in the interview guide. Therefore, it can even be argued whether the approach was closer to a non-structured interview in which the basis of the interview is a very simplified list of key topics (Bryman 2008). Notwithstanding the terminology of the approach taken, the aims of interviewing in this study were to gather feedback to develop *hydrostreamer,* evaluate its value, and further search for topics to base discussion on.

In this study, three of the four interviews were conducted over an internet voice call while one interview was conducted in person in WDRG premises during the interviewee's research visit. In the first interview, three interviewers (the author and the instructors of this work) were present, but other interviews were conducted by the author alone. In the voice call interviews that the author conducted alone, the interview was recorded and the recording was later listened to make notes. No word-to-word transcripts were made from the interviews; rather, short notes were made about the key observations during the interview, which were expanded after the interview with optional support of recordings. Analysis of the interview results was done by combining all extended notes into one and grouping observations into a number of themes. Since the interviewee count was small, no statistical analysis was conducted.

After each interview, the extended notes were sent to the interviewee to ensure that there were no misconceptions about their views. None of the interviewees provided with any corrections to the extended notes. The transcripts were not distributed to anyone else than the interviewees themselves for a review and they were kept in a safe place. In addition, the interviewees were given the option to stay completely anonymous during their participation and an opportunity to withdraw consent to the study at any time.

4 Results

In this chapter, the results from both quantitative and qualitative parts of this work are presented. First, the uncertainty estimation results are shown starting from the probabilities of rivers laying across the study area and finishing with a general outlook on DEM-induced uncertainty in discharge estimates computed with *hydrostreamer*. Second, the answers from the survey and the extended notes from the semi-structured interviews are summarized to present the most prominent topics that emerge from the local stakeholders' opinions. The results are further discussed with respect to theory and research questions of this study in Chapter 0.

4.1 Uncertainty quantification

Presenting the results of the uncertainty quantification starts from displaying the uncertainty of the river networks produced during the estimation. Figure 6 and Figure 7 present the uncertainty of the river networks: since the input DEM consists of raster cells and the river network is expressed as stating for each cell whether it's a part of a river or not, it is possible to compute probabilities of belonging to a river for each cell. In other words, all iterations (487 of them) were inspected for each cell, and probabilities of a river existing in individual cells were computed by dividing the amount of iterations in which the cell is classified as a river cell by the total number of iterations. Figure 6 presents the results of this computation. It is seen that variation happens mostly in headwaters located in lowlands and other low-slope areas (refer back to Figure 2 and Figure 3): larger streams and small streams in mountainous high-slope areas are more probable to follow consistent routes, also beyond the reach of rivers that were set in a fixed place by stream burning.

In addition, the river networks derived during this study are denser than the HydroSHEDS river network. This can be attributed to the fact that HydroSHEDS has a threshold of 100 upstream basin cells (in 15 arcsec, or 450 meters, resolution) before initiating a river (Lehner 2013), but this limit was not applied in the SAGA-GIS tool that was used in this study (Olaya 2004). Hence, the upper reaches of the river networks produced in this study may represent significantly smaller rivers than in HydroSHEDS despite that the utmost smallest rivers are left out by defining a minimum stream length value as explained in Chapter 3.2.3. Altogether, the densification of river networks applies mainly in headwaters only; larger streams are corresponding with each other in HydroSHEDS and this study.

To further investigate differences in the uncertainty of the river network, a continuous surface showing the nearest river cell probability is presented in Figure 7. For this figure, a continuous grid was formed over the study area, and the nearest cell, which had a probability greater than zero in Figure 6, was sought for each cell in the newly created grid. Therefore, the figure extends the uncertainty estimation: from estimating the probability of individual cells being a part of a river or not, to estimating the uncertainty in areas through showing which areas are near to rivers that are uncertain and which areas are near to rivers that are more certain. Comparing Figure 7 with the study area map in Figure 2 and the slope map in Figure 3, it is observed that the areas where the nearest rivers are the most certain are located in highlands and the most uncertain areas lie in lowlands. This behavior occurs despite that according to the error model adopted in this study, areas with high slope gain more uncertainty and areas with low slope gain less uncertainty in the DEM. Hence, it can be concluded that even though the DEM is more uncertain in high-slope areas, the river network is less uncertain in those areas than in lowlands.

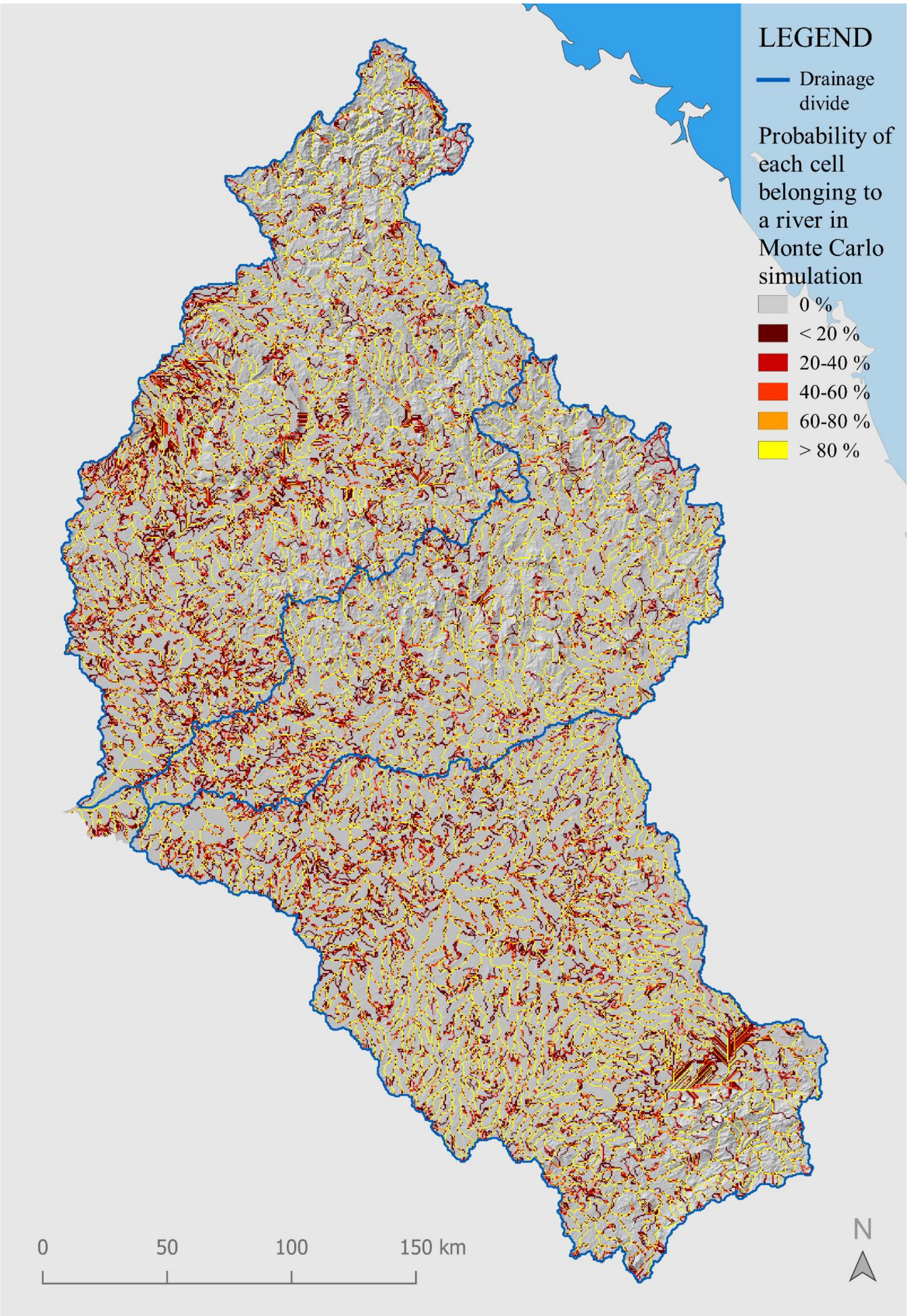


Figure 6. Probabilities of all cells being a part of a river

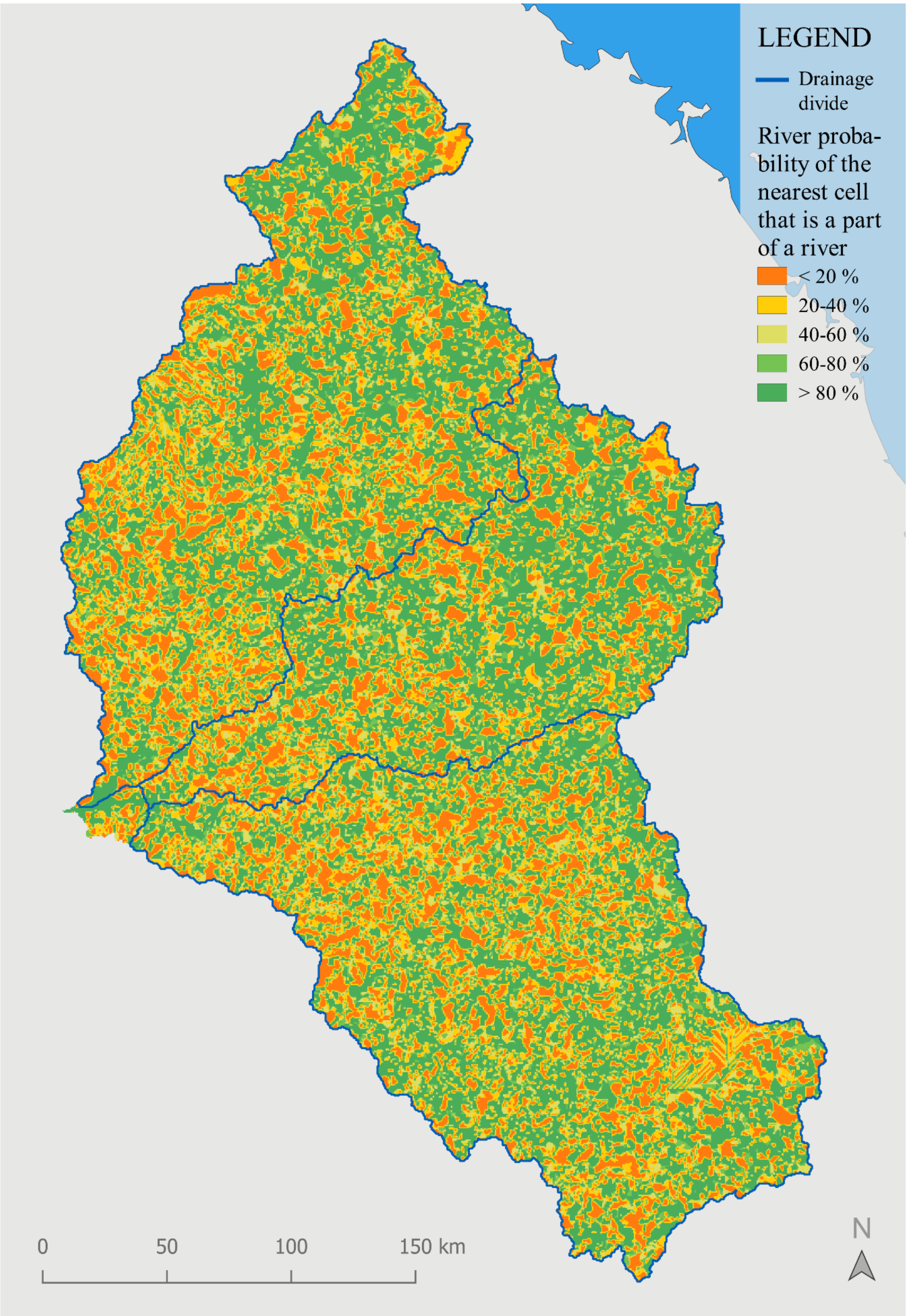


Figure 7. The probability of the nearest cell, which is a part of a river at least once, being a part of a river

After the river network uncertainty was assessed, discharge estimates were computed with *hydrostreamer* as outlined in Figure 4. Three different weighting methods for dividing runoff into river segments were applied and the discharge estimates were inspected in HYMOS station locations. The geographical locations of the HYMOS stations are marked on the map in Figure 2. Figure 8, Figure 9, and Figure 10 present the distribution of discharge estimates computed with the three weighting methods in three HYMOS stations along each of the 3S basin rivers. The violin plots represent the distribution of discharge estimates, and the median estimate is drawn to be compared with a constant reference estimate computed using HydroSHEDS river network and the respective weighting method. In each graph, the HYMOS station number is given with a root mean square error metric (RMSE) computed between the two lines in the graph (median of simulated estimates and reference HydroSHEDS estimate).

It is observed that the simulated estimates follow the reference estimate[12] well, indicating that the effect of spatial uncertainty is not very significant. From the distributions, it is also observed that streams with less total discharge exhibit relatively larger uncertainty in the simulated discharge. Furthermore, some HYMOS stations, such as 440201 and 430105, tend to have a clustered distribution of discharge estimates. This may be attributed to parts of the upstream basin being sensitive to small changes in the river network: those changes cause sub-basins to fall outside of the total basin above a HYMOS station and decrease the total discharge. Finally, the effect of spatial uncertainty in the most downstream HYMOS stations starts to fade out completely since the RMSE values and the width of distribution of discharge estimates are small when compared to total discharge. This is caused by the process of *hydrostreamer* dividing the runoff and accumulating flow: when a basin grows, a growingly large part of the total runoff is caught by it and the relative changes in the basin size have less and less effect, which decreases the changes in how much *hydrostreamer* will divide runoff into those river segments and directly decrease the variation in discharge.

---

[12] While not addressing correspondence to observed discharge values here, it is known that estimates using HydroSHEDS network are rather close to reality (Kallio et al. 2018). Therefore, comparisons between the median simulated estimate and HydroSHEDS network estimate are viable.

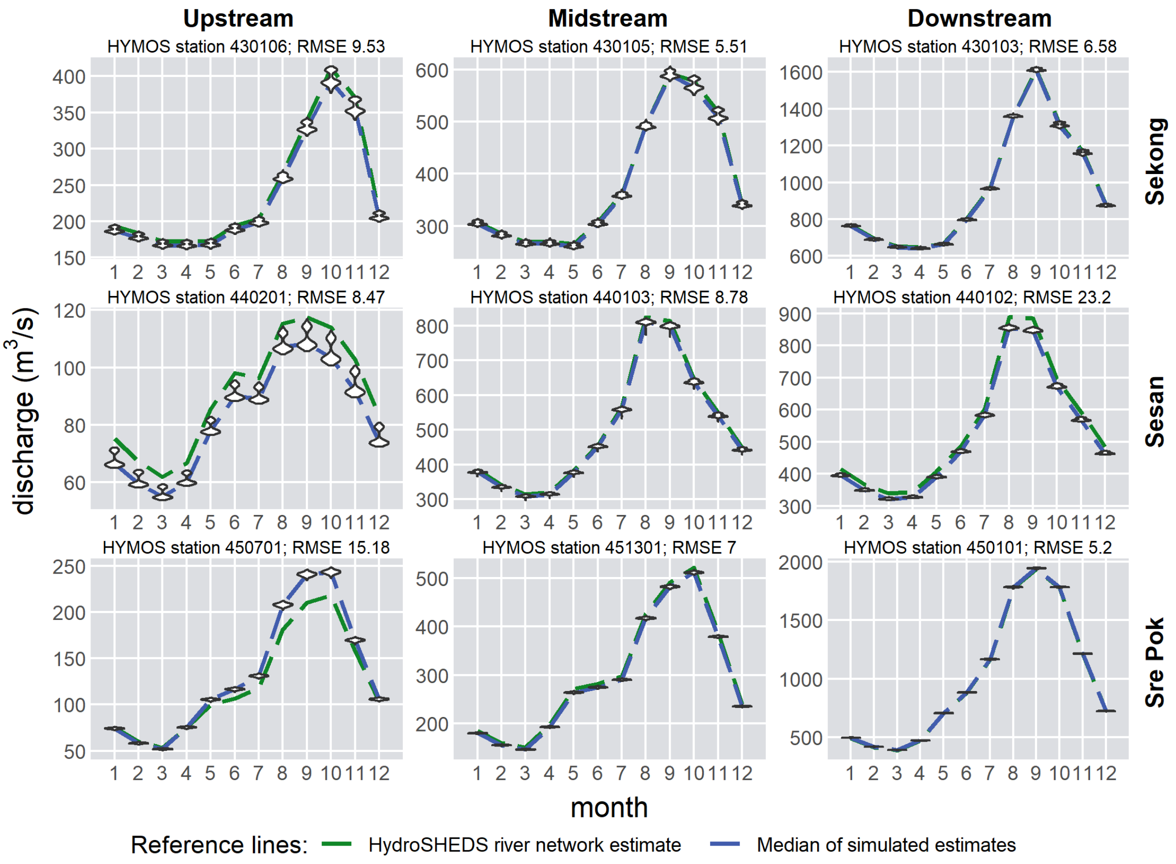


Figure 8. Simulated estimates and reference flow, length weighting

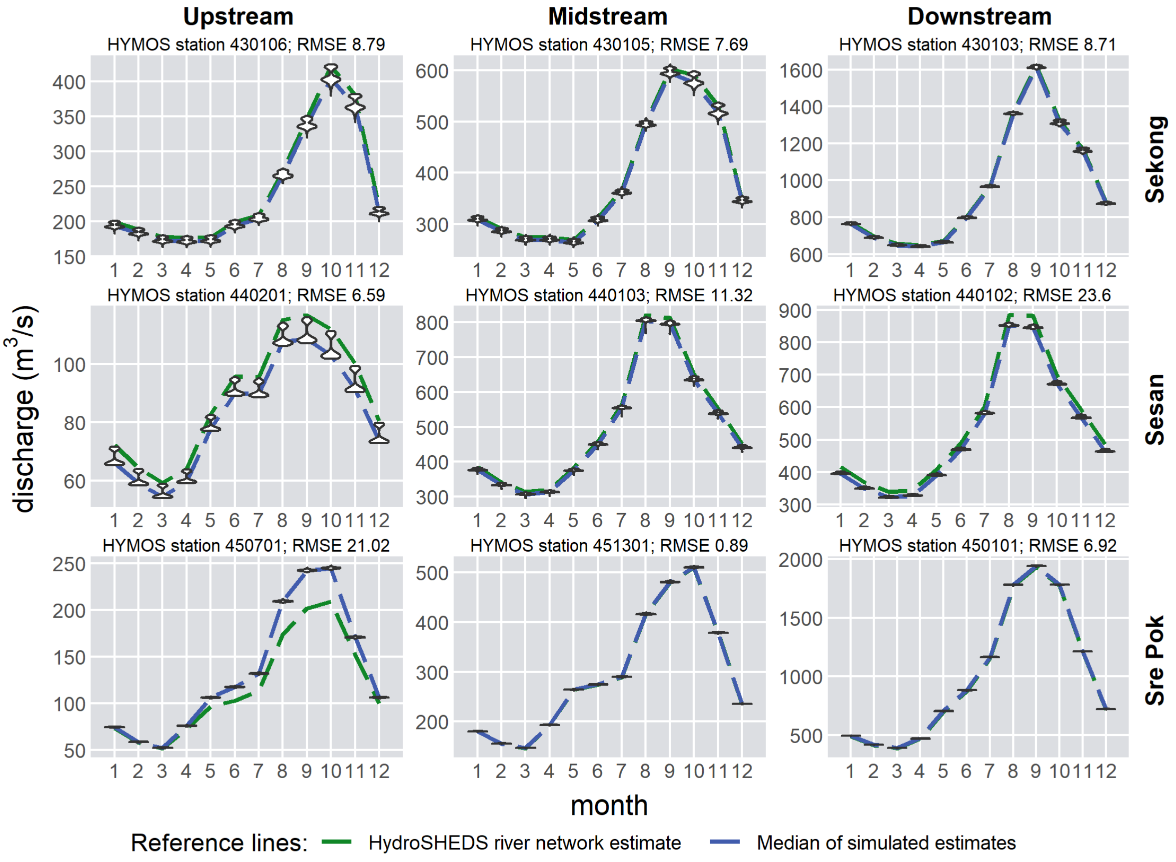


Figure 9. Simulated estimates and reference flow, Voronoi weighting

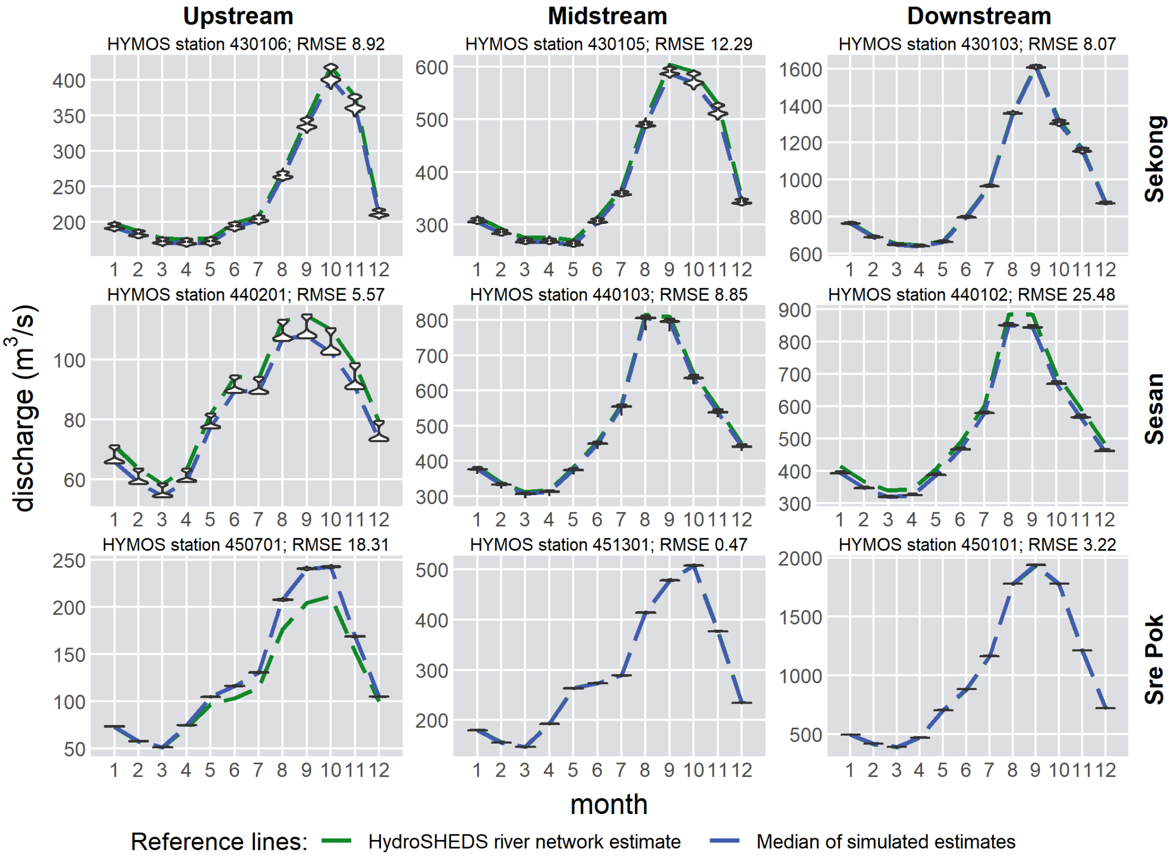


Figure 10. Simulated estimates and reference flow, basin weighting

Table 2 presents the RMSE values for each HYMOS station and weighting method, lowest RMSE value in each station being highlighted. It is observed that the RMSE values are not consistently smaller or larger for any of the weighting methods when compared to each other. This suggests that the uncertainty effects are not fully predictable in punctual observations. Therefore, discharge uncertainty at a punctual location should be addressed case-by-case instead of resorting to total discharge as a defining factor.

Table 2. Comparison of RMSE values for each HYMOS station and weighting method

River	HYMOS station	Length weighting	Voronoi weighting	Basin weighting
Sekong	430106	9,53	8,79	8,92
Sekong	430105	5,51	7,69	12,29
Sekong	430103	6,58	8,71	8,07
Sesan	440201	8,47	6,59	5,57
Sesan	440103	8,78	11,32	8,85
Sesan	440102	23,20	23,60	25,48
Sre Pok	450701	15,18	21,02	18,31
Sre Pok	451301	7,00	0,89	0,47
Sre Pok	450101	5,20	6,92	3,22

Finally, an attempt to generalize the uncertainty according to methodology presented in the end of Chapter 3.2.3 is presented in Figure 11. Recalling that the relative uncertainty was defined as the ratio of the standard deviation of the discharge estimates and the mean of the discharge estimates (distribution width normalized by discharge magnitude), two key observations emerge: 1) relative uncertainty decreases near-exponentially with respect to Strahler order of the river segments and 2) basin weighting provides the least uncertain estimates when applying *hydrostreamer* followed by Voronoi weighting and length weighting nearly tied. Again, the relative uncertainty falls to minimal values when the Strahler order increases. This result clearly suggests that the uncertainty behavior works as expected: headwaters that are low in river hierarchy are the most affected while the uncertainty decreases to negligible values in the overall largest streams, also not depending on the weighting method. This happens despite that the headwater rivers are the most certain in location: as mentioned, changes in small total upstream basin cause relatively larger variation in discharge estimates than similar-sized changes in large upstream basins.

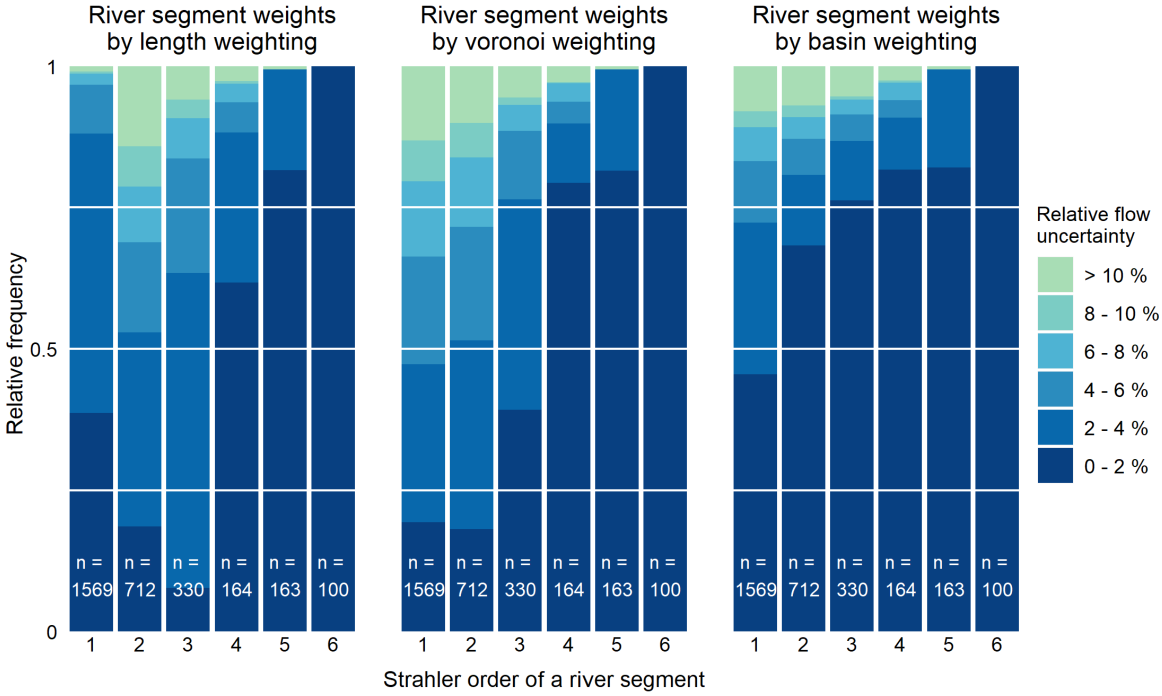


Figure 11. Relative uncertainty as a function of Strahler order

An anomaly related to Strahler order one and length weighting is caused by the method explained before: when only rivers with equal start and end points are compared with each other, the length of the rivers is nearly same for all rivers in the comparison. Again, length weighting takes only this into account, and provides rather equal weights corresponding to the rather similar lengths. Some difference is present though, due to small changes in the river lengths despite that the rivers start and end from the same points and the sum of river lengths inside one runoff cell. Despite this, Figure 11 provides a sought-after generalized outlook on uncertainty: while it may not be applied directly to an ambiguously picked river segment inside a river network, quantified estimates of the DEM-induced uncertainty in the river network can be estimated to an extent when the river network hierarchy is known by Strahler order.

4.2 Survey and interviews

After the survey and the interviews were conducted with materials shown in Appendix 2, Appendix 3, and Appendix 4, the results were summarized by grouping answers to the survey questions and extended interview notes into one and further dividing topics into broad categories partially based on the interview guide of Appendix 4. There was significant overlap between the answers to the survey and the observations made from the interview notes: this was expected because the interviews were to be conducted about the same topics while the survey provided the interviewers with only some preliminary information about the interviewees.

The profiles of the interviewees of this study are presented in Table 3. Two of the interviewees were Cambodian, one Vietnamese, and one Australian citizen, who however has been residing and working in Vietnam and the 3S basin area for a decade. From Table 3, it is observed that the interviewees consisted of technically oriented people; despite one of the aims of this study is to investigate open source solutions in policy-making, the interviewees are not directly in charge of policy-making. However, policy-makers themselves may not apply modelling tools as decision support but rather resort to expert knowledge, which is provided by people like those interviewed in this study.

Table 3. Interviewee profiles

Background	**Main tools/data currently or previously known or used**
Environmental hydrologist, hydrological modeler	GIS data, groundwater models, rainfall-runoff models, SWAT
Researcher, modeler, lecturer	HEC-HMS, neural networks, SWAT, MIKE
Private consultant, hydrological modeler	rainfall-runoff models, water quality models, surface water - ground water interaction models, SWAT
Environmental engineer, private consultant	HEC-HMS, SWAT, VMod

A common denominator for all interviewees is that each of them has used the Soil & Water Assessment Tool (SWAT), which is a comprehensive hydrological model commonly used in watershed management (TAMU 2018). Another common model mentioned by the interviewees, HEC-HMS, is a multiple-use hydrological model that can perform tasks from water quality to sediment transport (HEC 2018). As seen from Table 4 presenting the key observations from the survey and the interviews, most issues related to tools currently in use by the interviewees are related to difficulties in gaining data to be used with the models. The two tools mentioned above require observed data for calibration of parameters. On top of difficulties in the right calibration data even existing or being distributed adequately, the need for expert knowledge in using and calibrating the tools was emphasized.

Table 4. Observations from survey and interviews: the most common observations highlighted

Issues in current tools	**Difficulties in data collection** **Lack of data** Calibration of models Difficult user interface Needs for regionalized data Needs for training and expertise
Hydrostreamer benefits	**Not requiring much in terms of data** **Usage in data sparse areas** Enhancing other models' results Fast and easy estimates Preliminary modelling to support planning Small stream modelling without calibration Usage in ungauged catchments
Hydrostreamer issues	**Only discharge estimates as outputs** **Alone not enough for practical applications** Human impacts not included Coarse temporal resolution for some applications Lack of calibration poses needs for validation Lack of GUI
Uncertainty	**Application-related uncertainties are the most useful** **Estimation helps building confidence in methods** Should be included as a part of modelling work Acceptance limits varying Demonstration with respect to application Method to be selected and explained carefully Simplicity of demonstration for understanding
Open data impacts	**Generally valued and desirable** **No license costs or binding contracts** Major donors moving towards it No needs for difficult field data collection Enables easy learning Maintenance easier

Regarding *hydrostreamer,* the answers from the survey and the interviews seemed to agree on its main benefits: *hydrostreamer* was seen to possess potential in fast computing of discharge estimates to be used in supporting early-stage modelling, especially in areas where data is limited and will make the use of more sophisticated models hard. Lack of calibration requirements was seen to enable *hydrostreamer* to be used in exploratory analysis where gathering data for complicated modelling is not desired, or if the calibration of complex models doesn’t scale down to smaller streams. Main drawbacks of *hydrostreamer* were related to the simplicity of the model: outputting only discharge estimates may be alone not enough for practical applications, and the coarse temporal resolution and lack of human activities in the model eliminate some potential applications, such as dam control operations. One interviewee raised the usability question: creating a graphical user interface (GUI) for *hydrostreamer* would attract more users than the current command-line implementation whereas one interviewee emphasized strongly the demonstration of credibility to attract more users.

It was clear during the interviews that the uncertainty estimation presented in this study was a difficult topic to grasp even for the technically oriented interviewees. Therefore, no deep discussion about the quantitative values of uncertainty related to *hydrostreamer* was conducted but general views on uncertainty in modelling were expressed. The interviewees agreed that uncertainty estimation is important in modelling since it builds confidence on the tool used. However, the method of estimating uncertainty and the limits of acceptable uncertainty should be defined with respect to the application. Lack of calibration, which was also raised as an issue of *hydrostreamer* by one of the interviewees, may be supplemented by conducting a credible uncertainty estimation. Two of the interviewees noted that especially when communicating the uncertainty to non-technical people, the ways of explaining and expressing the methodology and the results should be done very carefully. Complex representation of uncertainty was also raised an issue by one of the survey respondents that wasn’t interviewed.

The fact that *hydrostreamer* is an open-source approach based on open and global data was received positively amongst the interviewees. The two main factors highlighted in Table 4 are quite predictable and expressed by all the interviewees, but especially the fact that two interviewees mentioned the large donor organizations, such as World Bank and Asian Development Bank, are moving towards open-source solutions is a concrete step that shows potential in adopting open-source solutions in larger scale. However, long-term binding and expensive contracts on commercial tools and the “official science” of state being the only kind of science accepted by government may hinder this adoption according to one of the interviewees. Additional strengths of open-source emerged in avoiding tedious data collection and ease of learning and maintaining the tools used. As shown in Appendix 4, the interviewees were asked whether there are any drawbacks caused by *hydrostreamer* being completely open. The interviewees did see that misuse of open source tools is possible but not a very big concern; in larger projects the misuse would be noticed, and the overall impact of open source is strongly on the positive side. An interesting detail mentioned by two of the interviewees was that commercial software are indeed used in the area, but to circumvent license costs, the versions used are illegal copies of the software.

In other things that came up during the interviews, some considerations on the study area characteristics were described. The current development state of the 3S basin countries was reflected especially in data gathering difficulties: either data required for calibration is completely unavailable, or then it exists but is out of reach due to deficient ways of collecting and distributing it. In contrary, extensive hydrological modelling and data sharing infrastructure building may not be a high priority in countries with more urgent development needs. One topic of consideration is also related to by whom is the modelling conducted: some of the interviewees noted that modelling capacity lies within universities and governments, which may be operating with funding from multilateral donor organizations. One of the interviewees raised a concern on the local decision makers’ ability to grasp complex models and rather resorting to simple solutions for decision support, which poses a challenge of structuring information to their needs.

5 Discussion

The discussion chapter of this work answers to the research questions set in Chapter 1, note some of the key limitations of the study, expand the discussion based on the results presented in the previous section, and finally address future study prospects based on this work. First, the research questions are answered in brief, after which the discussion is continued with additional topics of interest.

5.1 Research questions

Research question 1: *How could hydrostreamer be used in supporting water resources management or water governance?*

Answering the first research question can be started from recalling the difference between water resources management and water governance as stated in Chapter 2: water resources management deals with implementing policies whereas water governance is more related to creating the policies to be implemented. Practical applications of *hydrostreamer* presented in Chapter 4.2 are mainly related to small-scale and preliminary planning tasks in data-sparse areas. Simplicity of the tool and issues found through the interviews restrict the use of *hydrostreamer* to this limited number of practical tasks and disable using *hydrostreamer* alone in large-scale planning, which often takes advantage of more comprehensive modelling tools, such as SWAT. Therefore, *hydrostreamer* is not an instrument which directly affects high-level guidelines of water governance, but there exists potential to apply it in water resources management, especially in exploring areas that would not be available for exploration otherwise due to lack of data. In contrary, if the water governance policies guiding water resources management are strictly binding in terms of software to be used or highly specific tasks to be implemented, this exploration may not be possible.

Research question 2: *How could open-source and transparent solutions benefit policy-making that is based on environmental modelling?*

As stated during the initial setting of the research questions, the second research question is an attempt to generalize research question 1 by expanding the scope of environmental modelling tools from *hydrostreamer* to open-source tools in general. Again, it is reminded that the interviewees in this study were not policy-makers: instead, they were individuals who possess expert knowledge, which is inarguably an important part of policy-making and the complete science-policy interaction process. Transparency and open source are both advocated for – in the literature review and in the qualitative results. However, literature sets noble goals of increasing equality and enabling anyone to participate in decision-making while the primary benefits of adopting open-source tools, which were observed during the research are related to the negligible costs of the tools and general appreciation.

Clearly, there is a gap between those two statements of this study. According to the literature review, policies made in high levels of governance through science-policy interaction could benefit from open source and transparency by the inclusion of more stakeholders and more knowledge, but those who act in the science side of the science-policy border in the study area are more pragmatic, and see the benefits of open source primarily through monetary terms and ease of use regarding data collection and maintenance. A prominent observation from the interviews can contribute to closing the gap, though: if major donor organizations, which provide rich economic support for the countries in the study area and other developing regions of the world, are advocating open source, transparency interventions may better succeed. Then, modelers such as those interviewed in this study, can apply open-source tools and create transparent results, which are later used as parts of policy-making in natural resources governance. Whether this science-policy interaction is available to be reviewed and more importantly understood from an ordinary citizens' point of view is questionable due to modelling complexity, but this assessment is already out of the scope of this study.

Research question 3: *How does spatial uncertainty affect hydrostreamer outputs and use?*

Finally, the third research question involves the most technical parts of this study but may also be the simplest to answer. Spatial uncertainty does not play a very significant role in *hydrostreamer* outputs when comparing the median of simulated estimates with the reference estimate. The spatial uncertainty effects are decreasing to minimal amounts when propagating from small upstream rivers to large downstream rivers. Also, the width of the discharge estimates' distribution narrows down when propagating from smaller to larger streams, indicating that the probability of outputting highly deviated discharge estimates decreases despite that the underlying assumption about spatial uncertainty remains constant. Relationship between Strahler order and relative uncertainty in modelled discharge was found, providing a generalized outlook on the uncertainty behavior.

However, not only the quantified values but also the meaning of the uncertainty estimation should be addressed. The uncertainty estimation, which made up a very large fraction of time devoted to this work, was a very complex topic even amongst the technically oriented interviewees. Based on the results, the primary benefit of uncertainty estimation is in building credibility and confidence on the method with respect to specific applications. Therefore, the large time input has not been in vein despite that the end users who only use the information created by applying *hydrostreamer* instead of modelling themselves might not repeat or completely understand the works behind the uncertainty estimation. According to Chapter 2.4, developing environmental modelling tools should include uncertainty estimation as a routine rather than as an addition, but the communication of uncertainty should be considered carefully. The results agree with this, encouraging the uncertainty work to be done.

5.2 Further discussion

Next, some of the key limitations of this study are addressed before further discussion. Regarding research methods, all stakeholder contacts and most of the interviews were conducted over the internet, and no visits to the study area were made during the work. Despite that this may demonstrate the applicability of the work elsewhere since the work is place-independent, the lack of personal contacts and face-to-face interviews is a limiting factor related to collecting qualitative data. In addition, aiming the work for addressing individual issues in water governance would require a deeper understanding of the study area conditions. Therefore, the work has not been directed to solve any specific water governance issues. As stated, the uncertainty estimation was limited to spatial uncertainty, meaning that the uncertain part was only the river network used in *hydrostreamer*. Another main input data set, runoff data, was left out of this uncertainty estimation as were all other kinds of uncertainty beyond statistical uncertainty, too. Furthermore, multiple assumptions about the river network were made to achieve the presented results.

Discussing the study further may be started from addressing the nature of environmental modelling and the abstractions required in producing modelling results from raw data. As stated in Chapter 3.2.1, *hydrostreamer* is a model which gets its inputs from other models (runoff models and river networks produced from DEMs). Furthermore, geographical reality can never be modelled to its fullest extent, and abstractions from a field-based reality to an object-based reality are necessary, also in this study. Each abstraction step and each adoption of a model as an input of another model adds elements of uncertainty to the result. This justifies evaluating the uncertainty of discharge estimates by simulating the whole process instead of only inspecting individual input uncertainties and trying to synthesize the sum of effects by each input. It is recalled that only one component was altered during the simulation: even simulating the statistical uncertainty of this one component required a rather complex process. This highlights the fact that the complete statistical uncertainty related to *hydrostreamer* is a very complicated process that no analytical means can address: instead, stochastic methods are required. It should also be reminded that statistical uncertainty is only a part of total uncertainty related to an environmental modelling process as addressed in Chapter 2.4. Therefore, case-by-case uncertainty evaluation should be the preferred solution in a modelling process that involves as much uncertainty as presented in this study. This emphasizes the fact that the input statistical uncertainty is only a small part of a modelling process, which is again another small part of a policy-making process – however, even the smallest parts should be given attention to in a thoroughly transparent modelling process.

As seen, environmental modelling inevitably deals with abstractions and uncertainties that may be perceived complex even by experts. It is no wonder why difficulties in science-policy interaction have sometimes been contributed to the lack of common terminology and language as stated in Chapter 2.2. Lack of research effort is not a problem in resolving this issue: literature on science-policy interaction and participatory environmental modelling is broad. However, the literature is mostly providing generic frameworks to be tailored with respect to unique application contexts. Applying concepts and frameworks, such as boundary work, knowledge brokering, and co-production of knowledge, should be planned carefully before using in any application rather than using them as an out-of-the box buzzword solution. An environmental model may be used as a common platform for any of these concepts; then, those who understand or develop the models have an important role in enabling science-policy interaction to make all stakeholders feel included and gain benefits from the interaction. In addition to the results being used as decision support, improved system understanding, and potentially avoided conflict are beneficial outcomes of an environmental modelling process. However, technical solutions are not the only solutions for problems with characteristics like those in water governance, and comprehensively good governance requires various other ways of working beyond technical modelling only.

While it is recognized that *hydrostreamer* is a very simple model and cannot provide solutions for complex problems alone, it serves well as an example of the environmental modelling process and the science side of science-policy interaction. With only few participants in this study and other limitations, a thorough knowledge co-production process has not been conducted. However, some aspects of *hydrostreamer* that are currently lacking and need development have been found, which is one of the goals related to participatory modelling. From the interviews, ensuring credibility of *hydrostreamer* emerged strongly; this is also one of the factors that should be addressed when evaluating information with frameworks such as those presented in Chapter 2.2. Finally, common between the interview results and the reviewed theory was that transparency and information sharing are promoted, which can be achieved by using open source and open data solutions, such as *hydrostreamer.*

In open data, not everything is beyond criticism, though. In fact, the most commonly used hydrological modelling software mentioned in the interviews, HEC-HMS and SWAT, are both freely available for download, and even SWAT source code is freely available for download while HEC-HMS disables access to source code. As mentioned in the results, both tools are highly comprehensive and therefore complex in use and technical implementation. Hence, it can be argued whether using complicated models, which require expertise in use and extensive data collection for calibration, is truly committed to transparent and open-source way of modelling and decision support. This argument comes down to a trade-off between openness and applicability: on one hand, *hydrostreamer* is a very simple and very open model that works with global data without hydrological expertise, but on the other hand, it is not as applicable as more sophisticated models and tools that are currently used in supporting policy decisions.

Despite that the practical adoption of open-source tools and open data is a major trade-off situation; underlying implications of openness and transparency are positive. Recalling that Chapter 2.3 concludes participatory and transparent modelling having potential in equalizing power relationships between stakeholders, it certainly is a desirable approach for modelling. Some of the characteristics of water governance issues, such as riparian position, cannot be changed physically, but increasing knowledge and technical capabilities in an accessible and an equal way through open-source modelling is a way to affect power relationships. Hence, giving access to and distributing information about *hydrostreamer* is a naïve solution that does not solve complex problems by itself, but through the literature review and the empirical parts of this study it has become clear that potential for open-source solutions exist as minor parts of major ensembles of governance and management. Future emphasis should be put on ensuring the credibility of the tools and methods used: uncertainty estimation like the one rigorously conducted in this work is one of the means to do this.

Regarding prospective research directions based on this study, two separate but not exclusive approaches could be undertaken: 1) continue the technical parts of the study by repeating the uncertainty estimation in other study areas and by using data sets with varying spatial and temporal resolutions and 2) dive deeper into water governance and water resources management to investigate what are the concrete steps in which transparency and open source would yield the most benefits. First, a more general outlook on the uncertainty could be derived in broader scale than the 3S basin while keeping in mind that the total uncertainty of a modelling process does not equal statistical uncertainty. Auxiliary data sets, such as remote sensing data, could be applied for either ground proofing of the river network or for adding more explaining variables to *hydrostreamer.* Second, more emphasis could be put on investigating the specific governance processes and the role of open-source and transparent solutions, either in the 3S basin area or in other relatively under-developed countries where harnessing natural resources has not yet been completely saturated and room for improvement in governance exists.

6 Conclusion

This work has revolved around *hydrostreamer,* a tool that estimates river streamflow by computing discharge values: it divides low-resolution runoff data into a high-resolution river network and accumulates flow from upstream to downstream. One of the main contributors of statistical uncertainty, namely the spatial uncertainty related to the river network induced by uncertainties in digital elevation models, was thoroughly evaluated and generalized with respect to river network hierarchy. A case study was conducted in the 3S river basin (Sekong, Sesan, and Sre Pok rivers in Southeast Asia) to demonstrate the effects of uncertainty and to provide with an example application of *hydrostreamer*. Furthermore, the potential of open-source solutions in supporting water governance or water resources management was assessed through qualitative research involving local stakeholders and discussion on their views based on a survey and interviews.

The literature review of this work concluded that frameworks to be applied in science-policy interaction are plenty but none of them propose universal solutions, instead, all of them should be tailored with respect to the application context. This reinforced the initial setting of research questions in both quantitative and qualitative topics; no technical solution is perfect in itself to provide aid for highly complex issues, such as those in water governance, and the opinions of those who use the technical approaches should be considered. Finally, investigating both statistical and other kinds of uncertainties related to a modelling process was deemed beneficial and value-adding for the modelling process.

The methodology of the empirical parts consisted of both quantitative and qualitative methods. The main qualitative method, Monte Carlo simulation, was applied to quantify the effects of spatial uncertainty in *hydrostreamer* outputs. The analysis was conducted by using R programming language through a simulation inducing the uncertainty in a digital elevation model and propagating it first through the creation of river networks and further through estimating discharge with *hydrostreamer*. The applied qualitative methods were a survey and semi-structured interviews, in which stakeholders with a relationship to the study area were identified and contacted. Their insights were collected to evaluate the current state of *hydrostreamer* and provide basis for discussion on adopting open-source tools.

Two main results emerged from the uncertainty estimation: 1) relative uncertainty of discharge estimates is inversely dependent on Strahler order of the river and behaves near-exponentially, and 2) the magnitude of effects caused by spatial uncertainty on streamflow is not very significant in *hydrostreamer*, diminishing to negligible values in the overall largest streams. Main observations related to the state of *hydrostreamer* were that it is best used in data-sparse areas, but it is not applicable as the only tool in more complex practical applications. Regarding the potential of open-source solutions, open source is generally appreciated but mostly favored in practical applications due to low costs of adoption.

Finally the research questions set in the beginning were answered and the discussion was expanded to reflect on the reviewed theory. In short, *hydrostreamer* is applicable in supporting water resources management but only in small-scale applications, uncertainty estimation increases the value of a tool, and transparent open-source based solutions are generally valuable in supporting good governance. However, emphasis of future works should be put on ensuring the credibility and truly effective open access on open-source tools to make their adoption successful while respecting application context at the same time.

References

Asian Development Bank (ADB), 2010. Sesan, Sre Pok, and Sekong River Basins Development Study in Kingdom of Cambodia, Lao People's Democratic Republic, and Socialist Republic of Viet Nam (Technical Assistance Consultant's Report).

Armitage, D., de Loë, R.C., Morris, M., Edwards, T.W.D., Gerlak, A.K., Hall, R.I., Huitema, D., Ison, R., Livingstone, D., MacDonald, G., Mirumachi, N., Plummer, R., Wolfe, B.B., 2015. Science-policy processes for transboundary water governance. Ambio 44, 353–366. https://doi.org/10.1007/s13280-015-0644-x

Babel, M., S., Wahid, S., M., 2009. Freshwater under threat South East Asia: Vulnerability Assessment of Freshwater Resources to Environmental Change. United Nations Environment Programme.

Baird, I.G., 2016. Non-government Organizations, Villagers, Political Culture and the Lower Sesan 2 Dam in Northeastern Cambodia. Critical Asian Studies 48, 257–277. https://doi.org/10.1080/14672715.2016.1157958

Beven, K., 2008. Environmental Modelling: An Uncertain Future? Routledge, Oxon, United Kingdom.

Beven, K.J., 2000. Uniqueness of place and process representations in hydrological modelling. Hydrology and Earth System Sciences 4, 203–213. https://doi.org/10.5194/hess-4-203-2000

Bourgine, B., Baghdadi, N., 2005. Assessment of C-band SRTM DEM in a dense equatorial forest zone. Comptes Rendus Geoscience 337, 1225–1234. https://doi.org/10.1016/j.crte.2005.06.006

Bryman, A., 2008. Social Research Methods, 3rd ed. Oxford University Press, Oxford, United Kingdom.

Campo, P.C., Bousquet, F., Villanueva, T.R., 2010. Modelling with stakeholders within a development project. Environmental Modelling & Software, Thematic Issue - Modelling with Stakeholders 25, 1302–1321. https://doi.org/10.1016/j.envsoft.2010.01.005

Cash, D., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Jäger, J., 2002. Salience, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making (SSRN Scholarly Paper No. ID 372280). Social Science Research Network, Rochester, NY.

Cash, D.W., Borck, J.C., Patt, A.G., 2006. Countering the Loading-Dock Approach to Linking Science and Decision Making: Comparative Analysis of El Niño/Southern Oscillation (ENSO) Forecasting Systems. Science, Technology, & Human Values 31, 465–494. https://doi.org/10.1177/0162243906287547

Clark, W.C., Tomich, T.P., Noordwijk, M. van, Guston, D., Catacutan, D., Dickson, N.M., McNie, E., 2016. Boundary work for sustainable development: Natural resource management at the Consultative Group on International Agricultural Research (CGIAR). PNAS 113, 4615–4622. https://doi.org/10.1073/pnas.0900231108

Constable, D., 2015. Atlas of the 3S basins. IUCN, Bangkok, Thailand.

Cook, B.R., Spray, C.J., 2012. Ecosystem services and integrated water resource management: Different paths to the same end? Journal of Environmental Management 109, 93–100. https://doi.org/10.1016/j.jenvman.2012.05.016

Das, S., Patel, P.P., Sengupta, S., 2016. Evaluation of different digital elevation models for analyzing drainage morphometric parameters in a mountainous terrain: a case study of the Supin–Upper Tons Basin, Indian Himalayas. SpringerPlus 5, 1544. https://doi.org/10.1186/s40064-016-3207-0

Du, X., Guo, H., Fan, X., Zhu, J., Yan, Z., Zhan, Q., 2016. Vertical accuracy assessment of freely available digital elevation models over low-lying coastal plains. International Journal of Digital Earth 9, 252–271. https://doi.org/10.1080/17538947.2015.1026853

Eden, S., Megdal, S.B., Shamir, E., Chief, K., Mott Lacroix, K., 2016. Opening the Black Box: Using a Hydrological Model to Link Stakeholder Engagement with Groundwater Management. Water 8, 216. https://doi.org/10.3390/w8050216

Falorni, G., Teles, V., Vivoni, E.R., Bras, R.L., Amaratunga, K.S., 2005. Analysis and characterization of the vertical accuracy of digital elevation models from the Shuttle Radar Topography Mission. Journal of Geophysical Research: Earth Surface 110. https://doi.org/10.1029/2003JF000113

Felipe-Lucia, M.R., Martín-López, B., Lavorel, S., Berraquero-Díaz, L., Escalera-Reyes, J., Comín, F.A., 2015. Ecosystem Services Flows: Why Stakeholders' Power Relationships Matter. PLOS ONE 10, e0132232. https://doi.org/10.1371/journal.pone.0132232

Franks, T., Cleaver, F., 2007. Water governance and poverty: a framework for analysis. Progress in Development Studies 7, 291–306. https://doi.org/10.1177/146499340700700402

Fung, A., 2006. Varieties of Participation in Complex Governance. Public Administration Review 66, 66–75. https://doi.org/10.1111/j.1540-6210.2006.00667.x

Geheb, K., West, N., Matthews, N., 2015. The invisible dam: hydropower and its narration in the Lao People's Democratic Republic, in: Hydropower Development in the Mekong Region. Routledge, Oxon, United Kingdom, pp. 101–126.

Geophysical Fluid Dynamics Laboratory (GFDL), 2018. Earth System Model [WWW Document]. URL https://www.gfdl.noaa.gov/earth-system-model/ (accessed 12.11.2018).

Grimsditch, M., 2012. 3S rivers under threat. 3S Rivers Protection Network & International Rivers, Ratanakiri, Cambodia.

Gupta, J., Pahl-Wostl, C., Zondervan, R., 2013. 'Glocal' water governance: a multi-level challenge in the anthropocene. Current Opinion in Environmental Sustainability, Aquatic and marine systems 5, 573–580. https://doi.org/10.1016/j.cosust.2013.09.003

Gurstein, M.B., 2011. Open data: Empowering the empowered or effective data use for everyone? First Monday 16. https://doi.org/10.5210/fm.v16i2.3316

Guston, D.H., 2001. Boundary Organizations in Environmental Policy and Science: An Introduction. Science, Technology, & Human Values 26, 399–408. https://doi.org/10.1177/016224390102600401

Global Water Partnership (GWP), 2000. Integrated Water Resources Management (No. 4), TAC Background Papers. Global Water Partnership Technical Advisory Committee (TAC), GWP Secretariat, Stockholm.

Ham, K., Hay, S., Sok, T., 2015. The politics of the Lower Sesan 2 Dam in Cambodia, in: Hydropower Development in the Mekong Region. Routledge, Oxon, United Kingdom, pp. 153–172.

Harris, M., 2016. Diverted Justice: The Lower Sesan 2 Dam and the Role of Law in Cambodian Hydropower Development, in: Water Governance Dynamics in the Mekong Region, M-POWER Book Series on Water Governance in the Mekong Region. Strategic Information and Research Development Centre No. 2, Petaling Jaya, Selangor, Malaysia, pp. 91–122.

Hydrological Engineering Center (HEC), 2018. HEC-HMS [WWW Document]. URL http://www.hec.usace.army.mil/software/hec-hms/ (accessed 15.11.2018).

Hoang, L.P., van Vliet, M.T.H., Kummu, M., Lauri, H., Koponen, J., Supit, I., Leemans, R., Kabat, P., Ludwig, F., 2019. The Mekong's future flows under multiple drivers: How climate change, hydropower developments and irrigation expansions drive hydrological changes. Science of The Total Environment 649, 601–609. https://doi.org/10.1016/j.scitotenv.2018.08.160

Horn, B.K.P., 1981. Hill shading and the reflectance map. Proceedings of the IEEE 69, 14–47. https://doi.org/10.1109/PROC.1981.11918

Jahn, T., Bergmann, M., Keil, F., 2012. Transdisciplinarity: Between mainstreaming and marginalization. Ecological Economics 79, 1–10. https://doi.org/10.1016/j.ecolecon.2012.04.017

Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modelling & Software 21, 602–614. https://doi.org/10.1016/j.envsoft.2006.01.004

Jarihani, A.A., Callow, J.N., McVicar, T.R., Van Niel, T.G., Larsen, J.R., 2015. Satellite-derived Digital Elevation Model (DEM) selection, preparation and correction for hydrodynamic modelling in large, low-gradient and data-sparse catchments. Journal of Hydrology 524, 489–506. https://doi.org/10.1016/j.jhydrol.2015.02.049

Kallio, M., 2018. R package for downscaling off-the-shelf runoff products to explicit river network.: mkkallio/hydrostreamer [WWW Document]. URL https://github.com/mkkallio/hydrostreamer (accessed 19.11.2018).

Kallio, M., Guillaume, J.H.A., Kummu, M., Desalegn, F., Virrantaus, K., 2018. (PDF) Spatial allocation of low resolution runoff model outputs to high resolution stream network. http://dx.doi.org/10.13140/RG.2.2.19122.09921

Karlsson, J.M., Arnberg, W., 2011. Quality analysis of SRTM and HYDRO1K: a case study of flood inundation in Mozambique. International Journal of Remote Sensing 32, 267–285. https://doi.org/10.1080/01431160903464112

Kiamehr, R., Sjöberg, L.E., 2005. Effect of the SRTM global DEM on the determination of a high-resolution geoid model: a case study in Iran. J Geodesy 79, 540–551. https://doi.org/10.1007/s00190-005-0006-8

Kolecka, N., Kozak, J., 2014. Assessment of the Accuracy of SRTM C- and X-Band High Mountain Elevation Data: a Case Study of the Polish Tatra Mountains. Pure Appl. Geophys. 171, 897–912. https://doi.org/10.1007/s00024-013-0695-5

Kosack, S., Fung, A., 2014. Does Transparency Improve Governance? Annu. Rev. Polit. Sci. 17, 65–87. https://doi.org/10.1146/annurev-polisci-032210-144356

Lautze, J., Silva, S. de, Giordano, M., Sanford, L., 2011. Putting the cart before the horse: Water governance and IWRM. Natural Resources Forum 35, 1–8. https://doi.org/10.1111/j.1477-8947.2010.01339.x

Lebel, L., Garden, P., Imamura, M., 2005. The Politics of Scale, Position, and Place in the Governance of Water Resources in the Mekong Region. Ecology and Society 10. https://doi.org/10.5751/ES-01543-100218

Lehner, B., 2013. HydroSHEDS Technical Documentation. Conservation Science Program, World Wildlife Fund US, Washington, DC.

Lemos, M.C., Kirchhoff, C.J., Ramprasad, V., 2012. Narrowing the climate information usability gap. Nature Climate Change 2, 789–794. https://doi.org/10.1038/nclimate1614

Lienert, J., Schnetzer, F., Ingold, K., 2013. Stakeholder analysis combined with social network analysis provides fine-grained insights into water infrastructure planning processes. Journal of Environmental Management 125, 134–148. https://doi.org/10.1016/j.jenvman.2013.03.052

Liu, Y., Gupta, H., Springer, E., Wagener, T., 2008. Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. Environmental Modelling & Software 23, 846–858. https://doi.org/10.1016/j.envsoft.2007.10.007

Lund, J.R., 2015. Integrating social and physical sciences in water management. Water Resources Research 51, 5905–5918. https://doi.org/10.1002/2015WR017125

Meadow, A.M., Ferguson, D.B., Guido, Z., Horangic, A., Owen, G., Wall, T., 2015. Moving toward the Deliberate Coproduction of Climate Science Knowledge. Wea. Climate Soc. 7, 179–191. https://doi.org/10.1175/WCAS-D-14-00050.1

Meyer, M., 2010. The Rise of the Knowledge Broker. Science Communication 32, 118–127. https://doi.org/10.1177/1075547009359797

Michaels, S., 2009. Matching knowledge brokering strategies to environmental policy problems and settings. Environmental Science & Policy 12, 994–1011. https://doi.org/10.1016/j.envsci.2009.05.002

Mouratidis, A., Briole, P., Katsambalos, K., 2010. SRTM 3″ DEM (versions 1, 2, 3, 4) validation by means of extensive kinematic GPS measurements: a case study from North Greece. International Journal of Remote Sensing 31, 6205–6222. https://doi.org/10.1080/01431160903401403

Mukherjee, Sandip, Joshi, P.K., Mukherjee, Samadrita, Ghosh, A., Garg, R.D., Mukhopadhyay, A., 2013. Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). International Journal of Applied Earth Observation and Geoinformation 21, 205–217. https://doi.org/10.1016/j.jag.2012.09.004

Mukul, Manas, Srivastava, V., Jade, S., Mukul, Malay, 2017. Uncertainties in the Shuttle Radar Topography Mission (SRTM) Heights: Insights from the Indian Himalaya and Peninsula. Sci Rep 7. https://doi.org/10.1038/srep41672

Nam, L.D., Hai, B.T., Tu, D.T., Pech, S., 2013. Institutional Arrangements: Policies and Administrative Mechanisms for Water Governance in Vietnam, Mekong Project 4 on Water Governance. Mekong Program of Water, Environment and Resilience (M-POWER).

National Aerospace and Space Administration (NASA), 2015. The Shuttle Radar Topography Mission (SRTM) Collection User Guide [WWW Document]. https://lpdaac.usgs.gov/sites/default/files/public/measures/docs/NASA_SRTM_V3.pdf.

NASA Jet Propulsion Laboratory (JPL), 2013. NASA Shuttle Radar Topography Mission Global 3 arc second sub-sampled. https://doi.org/10.5067/MEaSUREs/SRTM/SRTMGL3S.003

National Geospatial-Intelligence Agency (NGA), 2018. Length of degree calculator [WWW Document]. URL https://msi.nga.mil/msisitecontent/staticfiles/calculators/degree.html (accessed 12.11.2018).

Organisation for Economic Co-operation and Development (OECD), 2015. OECD Principles on Water Governance. Organisation for Economic Co-operation and Development, 7th World Water Forum, Daegu, Korea.

Oksanen, J., Sarjakoski, T., 2005a. Error propagation of DEM-based surface derivatives. Computers & Geosciences 31, 1015–1027. https://doi.org/10.1016/j.cageo.2005.02.014

Oksanen, J., Sarjakoski, T., 2005b. Error propagation analysis of DEM- based drainage basin delineation. International Journal of Remote Sensing 26, 3085–3102. https://doi.org/10.1080/01431160500057947

Olaya, V., 2004. A gentle introduction to SAGA-GIS [WWW Document]. URL https://vorboss.dl.sourceforge.net/project/saga-gis/SAGA%20-%20Documentation/SAGA%20Documents/SagaManual.pdf (accessed 19.11.2018).

OpenStreetMap (OSM), 2018. Planet Dump retrieved from https://planet.osm.org/ [WWW Document]. URL https://www.openstreetmap.org/

Ostrom, E., Burger, J., Field, C.B., Norgaard, R.B., Policansky, D., 1999. Revisiting the Commons: Local Lessons, Global Challenges. Science 284, 278–282. https://doi.org/10.1126/science.284.5412.278

Pahl-Wostl, C., Craps, M., Dewulf, A., Mostert, E., Tabara, D., Taillieu, T., 2007. Social Learning and Water Resources Management. Ecology and Society 12. https://doi.org/10.5751/ES-02037-120205

Pech, S., Ranamukhaarachchi, H., 2013. Water Governance Systems in Cambodia, Lao PDR and Vietnam, Mekong Project 4 on Water Governance. Mekong Program of Water, Environment and Resilience (M-POWER).

Pedregal, B., Cabello, V., Hernández-Mora, N., Limones, N., Moral, L.D., 2015. Information and knowledge for water governance in the networked society. Water Alternatives; Montpellier 8.

Räsänen, T.A., Joffre, O.M., Someth, P., Thanh, C.T., Keskinen, M., Kummu, M., 2015. Model-Based Assessment of Water, Food, and Energy Trade-Offs in a Cascade of Multipurpose Reservoirs: Case Study of the Sesan Tributary of the Mekong River. Journal of Water Resources Planning and Management 141, 05014007. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000459

Reed, M.S., Graves, A., Dandy, N., Posthumus, H., Hubacek, K., Morris, J., Prell, C., Quinn, C.H., Stringer, L.C., 2009. Who's in and why? A typology of stakeholder analysis methods for natural resource management. Journal of Environmental Management 90, 1933–1949. https://doi.org/10.1016/j.jenvman.2009.01.001

Reed, M.S., Stringer, L.C., Fazey, I., Evely, A.C., Kruijsen, J.H.J., 2014. Five principles for the practice of knowledge exchange in environmental management. Journal of Environmental Management 146, 337–345. https://doi.org/10.1016/j.jenvman.2014.07.021

Reed, P., M., Kasprzyk, J., 2009. Water Resources Management: The Myth, the Wicked, and the Future. Journal of Water Resources Planning and Management 135, 411–413. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000047

Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.L., Vanrolleghem, P.A., 2007. Uncertainty in the environmental modelling process – A framework and guidance. Environmental Modelling & Software 22, 1543–1556. https://doi.org/10.1016/j.envsoft.2007.02.004

Rexer, M., Hirt, C., 2014. Comparison of free high resolution digital elevation data sets (ASTER GDEM2, SRTM v2.1/v4.1) and validation against accurate heights from the Australian National Gravity Database. Australian Journal of Earth Sciences 61, 213–226. https://doi.org/10.1080/08120099.2014.884983

Rodríguez, E., Morris, C.S., Belz, J.E., 2006. A Global Assessment of the SRTM Performance. Photogrammetric Engineering & Remote Sensing 72, 249–260. https://doi.org/10.14358/PERS.72.3.249

Rogers, P., Hall, A., W., 2003. Effective Water Governance (No. 7), TEC Background Papers. Global Water Partnership Technical Committee (TEC), GWP Secretariat, Stockholm.

Roux, D., Rogers, K., Biggs, H., Ashton, P., Sergeant, A., 2006. Bridging the Science–Management Divide: Moving from Unidirectional Knowledge Transfer to Knowledge Interfacing and Sharing. Ecology and Society 11. https://doi.org/10.5751/ES-01643-110104

Scott, A., 2000. The dissemination of the results of environmental research - A scoping report for the European Environment Agency (Publication No. 15), Environmental issues series. Science and Technology Policy Research, University of Sussex, Falmer, Brighton.

Seng, S., Keartha, C., Phyrom, S., Pech, S., 2013. Institutional Arrangements: Policies and Administrative Mechanisms for Water Governance in the Kingdom of Cambodia, Mekong Project 4 on Water Governance. Mekong Program of Water, Environment and Resilience (M-POWER).

Sharma, A., Tiwari, K.N., 2014. A comparative appraisal of hydrological behavior of SRTM DEM at catchment level. Journal of Hydrology 519, 1394–1404. https://doi.org/10.1016/j.jhydrol.2014.08.062

Shortridge, A., Messina, J., 2011. Spatial structure and landscape associations of SRTM error. Remote Sensing of Environment 115, 1576–1587. https://doi.org/10.1016/j.rse.2011.02.017

Sisouvanh, A., Bouapao, L., Sayalath, C., Phosalath, S., Phuonvixay, V., Ngonvorarath, V., Lazarus, K., Pech, S., 2013. Institutional Arrangements: Policies and Administrative Mechanisms for Water Governance in the Lao People's Democratic Republic, Mekong Project 4 on Water Governance. Mekong Program of Water, Environment and Resilience (M-POWER).

Stein, C., Ernstson, H., Barron, J., 2011. A social network approach to analyzing water governance: The case of the Mkindo catchment, Tanzania. Physics and Chemistry of the Earth, Parts A/B/C, 11th WaterNet/WARFSA/GWP-SA Symposium: IWRM for National and Regional Integration through Science, Policy and Practice 36, 1085–1092. https://doi.org/10.1016/j.pce.2011.07.083

Strahler, A.N., 1957. Quantitative analysis of watershed geomorphology. Eos, Transactions American Geophysical Union 38, 913–920. https://doi.org/10.1029/TR038i006p00913

Szabó, G., Singh, S.K., Szabó, S., 2015. Slope angle and aspect as influencing factors on the accuracy of the SRTM and the ASTER GDEM databases. Physics and Chemistry of the Earth, Parts A/B/C, Emerging science and applications with microwave remote sensing data 83–84, 137–145. https://doi.org/10.1016/j.pce.2015.06.003

Texas A&M University (TAMU), 2018. SWAT | Soil & Water Assessment Tool [WWW Document]. URL https://swat.tamu.edu/ (accessed 15.11.2018).

Tortajada, C., 2010. Water Governance: Some Critical Issues. International Journal of Water Resources Development 26, 297–307. https://doi.org/10.1080/07900621003683298

United Nations (UN), 2018. Sustainable Development Goal 6: Synthesis Report on Water and Sanitation. United Nations, New York, NY, USA.

Uusitalo, L., Lehikoinen, A., Helle, I., Myrberg, K., 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. Environmental Modelling & Software 63, 24–31. https://doi.org/10.1016/j.envsoft.2014.09.017

van Beek, L.P.., Bierkens, M.F.., 2008. The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification (Report). Department of Physical Geography, Utrecht University, Utrecht, The Netherlands.

Virkki, V., 2018. R Package for downscaling off-the-shelf runoff products to explicit river network.: vvirkki/hydrostreamer [WWW Document]. URL https://github.com/vvirkki/hydrostremer (accessed 19.11.2018)

Voinov, A., Bousquet, F., 2010. Modelling with stakeholders. Environmental Modelling & Software, Thematic Issue - Modelling with Stakeholders 25, 1268–1281. https://doi.org/10.1016/j.envsoft.2010.03.007

Voinov, A., Gaddis, E.J.B., 2008. Lessons for successful participatory watershed modeling: A perspective from modeling practitioners. Ecological Modelling, Special Issue dedicated to the memory of Yuri Svirezhev 216, 197–207. https://doi.org/10.1016/j.ecolmodel.2008.03.010

Wang, L., Liu, H., 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. International Journal of Geographical Information Science 20, 193–213. https://doi.org/10.1080/13658810500433453

Wang, W., Yang, X., Yao, T., 2012. Evaluation of ASTER GDEM and SRTM and their suitability in hydraulic modelling of a glacial lake outburst flood in southeast Tibet. Hydrological Processes 26, 213–225. https://doi.org/10.1002/hyp.8127

Watt, B., 2015. Strategic Priorities for Trans-boundary Water Cooperation in the Sekong, Sesan and Sre Pok (3S) Basins. IUCN (International Union for Conservation of Nature), Bangkok, Thailand.

Wechsler, S.P., 2007. Uncertainties associated with digital elevation models for hydrologic applications: a review. Hydrology and Earth System Sciences 11, 1481–1500. https://doi.org/10.5194/hess-11-1481-2007

Wechsler, S.P., Kroll, C., 2006. Quantifying DEM Uncertainty and its Effect on Topographic Parameters.

Wehn, U., Collins, K., Anema, K., Basco-Carrera, L., Lerebours, A., 2018. Stakeholder engagement in water governance as social learning: lessons from practice. Water International 43, 34–59. https://doi.org/10.1080/02508060.2018.1403083

World Bank, 2018. World Development Indicators | DataBank [WWW Document]. URL http://databank.worldbank.org/data/reports.aspx?source=2&series=EG.ELC.ACCS.ZS&country=KHM,VNM,LAO (accessed 12.12.2018).

World Wildlife Fund (WWF), 2018. HydroSHEDS [WWW Document]. URL http://www.hydrosheds.org/ (accessed 12.11.2018).

Zeitoun, M., Cascão, A.E., Warner, J., Mirumachi, N., Matthews, N., Menga, F., Farnum, R., 2017. Transboundary water interaction III: contest and compliance. Int Environ Agreements 17, 271–294. https://doi.org/10.1007/s10784-016-9325-x

Zeitoun, M., Warner, J., 2006. Hydro-hegemony – a framework for analysis of trans-boundary water conflicts. Water Policy 8, 435–460. https://doi.org/10.2166/wp.2006.054

Zhang, J., Goodchild, M.F., 2002. Uncertainty in Geographical Information. Taylor & Francis, London, United Kingdom.

Ziv, G., Baran, E., Nam, S., Rodríguez-Iturbe, I., Levin, S.A., 2012. Trading-off fish biodiversity, food security, and hydropower in the Mekong River Basin. PNAS 109, 5609–5614. https://doi.org/10.1073/pnas.1201423109

Appendix 1: SRTM case study review

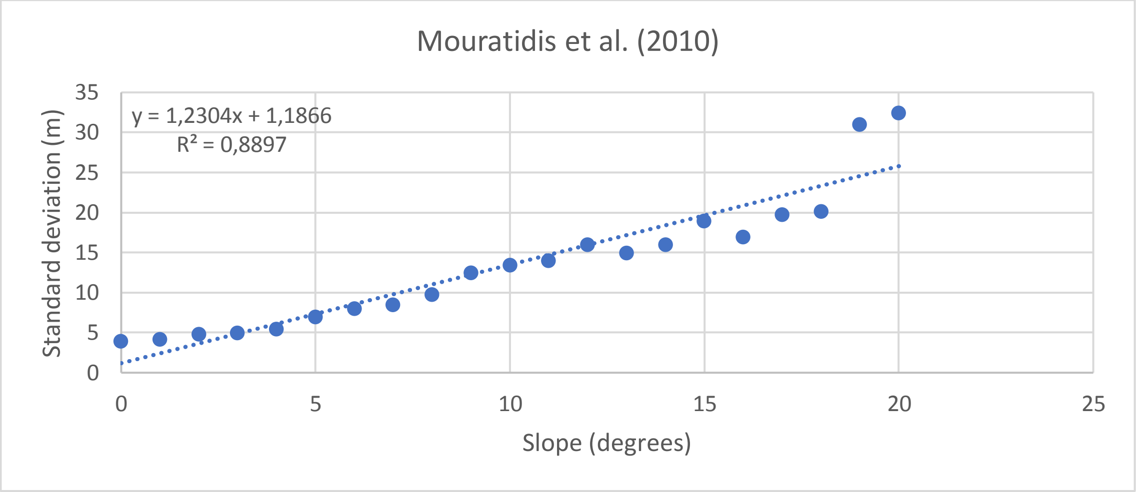
This appendix presents the case studies that were reviewed to build the DEM uncertainty model used in deriving the river networks as explained in Chapter 3.2.3. First, four studies that contain estimating SRTM errors with respect to slope are presented by graphs and Table 1. Table 1 contains the data gathered from the studies from which the following graphs are computed.

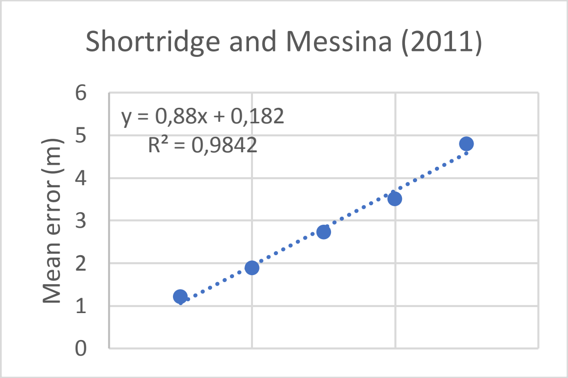
Second, additional studies are presented in Table 2 to demonstrate that SRTM errors tend to be positive, therefore overestimating the surface (Bourgine and Baghdadi 2005; Das et al. 2016; Du et al. 2016; Falorni et al. 2005; Jarihani et al. 2015; Karlsson and Arnberg 2011; Kiamehr and Sjöberg 2005; Kolecka and Kozak 2014; Mouratidis et al. 2010; Mukherjee et al. 2013; Mukul et al. 2017; Rexer and Hirt 2014; Rodríguez et al. 2006; Sharma and Tiwari 2014; Shortridge and Messina 2011; Szabó et al. 2015; Wang et al. 2012).

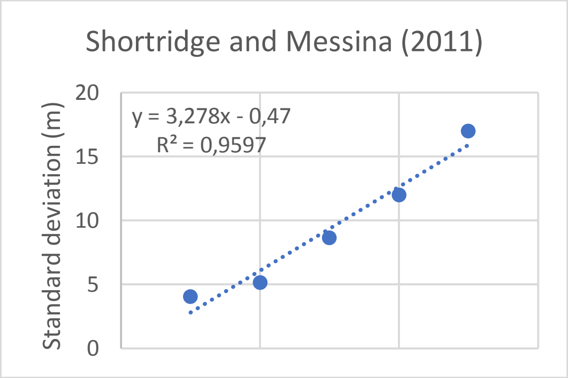
Related to Table 1, Mukherjee et al. (2013) and Shortridge and Messina (2011) classified slopes not to bins of 10-15 and 15-20 degrees but to 10-20 degrees. In Mukherjee et al. (2013), the standard deviation in observed locations having a slope from 10 to 20 degrees is 17,25 m while in Shortridge and Messina (2011) the standard deviation is 14,27 m and mean error is 4,05 m. From these values, the highlighted rows in Table 1 were estimated by noting that the counts of observation points in higher slopes will decrease like reported by Rexer and Hirt (2014). This was done to provide additional data points for the graphs.

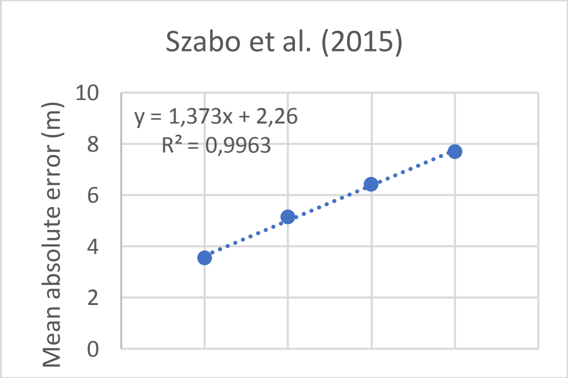
From the graphs, it is seen that the SRTM errors behave near linearly with respect to slope, which justifies the choice of error model made in Chapter 3.2.3. Mouratidis et al. (2010) present the only approach, which considers slope in one-degree bins while in other studies the binning is not as consistent. Therefore, the error model chosen for this study takes its linearity from all of these studies but a value for the slope coefficient from Mouratidis et al. (2010) in the absence of previous studies or ground truth information in the study area.

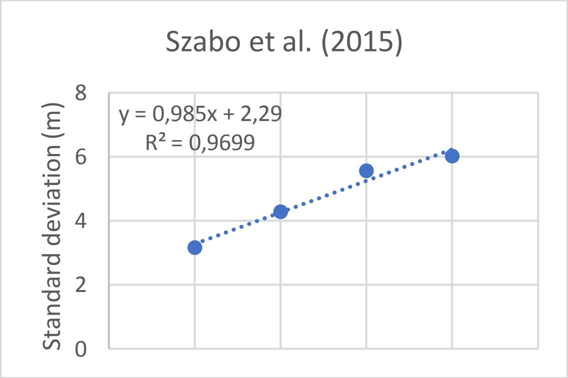
Another error model choice was made related to mean error: the mean slope in degrees for the study area is approximately 6,1 degrees and multiplying this by chosen value 0,2 yields 1,22, which is now the mean error in the study area according to mean error equation defined. This is lower than mean and median errors in Table 2 and is closest to Shortridge and Messina's (2011) mean error for slopes lower than one degree and near to intercept value in the results of Mouratidis et al. (2010). Hence, the mean error may be considered as a conservative estimate, and the resulting errors should not be higher than in reality.

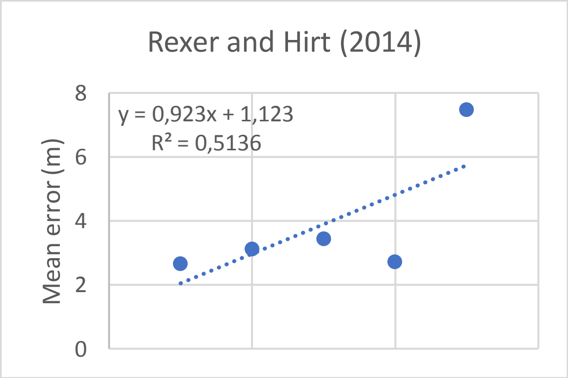

Mouratidis et al. (2010)
y = 1,2304x + 1,1866
R² = 0,8897
Standard deviation (m)
35
30
25
20
15
10
5
0
0
5
10
15
20
25
Slope (degrees)


Shortridge and Messina (2011)
y = 0,88x + 0,182
R² = 0,9842
Mean error (m)
6
5
4
3
2
1
0


Shortridge and Messina (2011)
y = 3,278x - 0,47
R² = 0,9597
Standard deviation (m)
20
15
10
5
0


Szabo et al. (2015)
y = 1,373x + 2,26
R² = 0,9963
Mean absolute error (m)
10
8
6
4
2
0


Szabo et al. (2015)
y = 0,985x + 2,29
R² = 0,9699
Standard deviation (m)
8
6
4
2
0


Rexer and Hirt (2014)
y = 0,923x + 1,123
R² = 0,5136
Mean error (m)
8
6
4
2
0

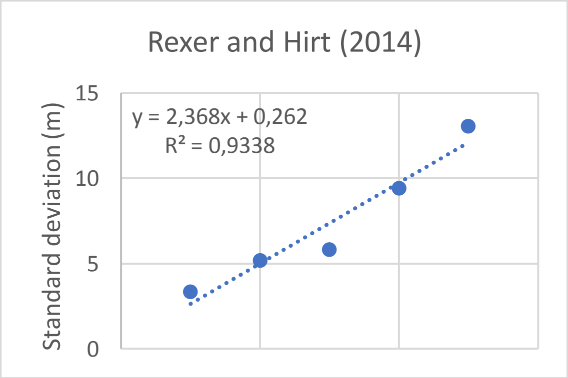

Rexer and Hirt (2014)
y = 2,368x + 0,262
R² = 0,9338
Standard deviation (m)
15
10
5
0

Table 1. SRTM case studies with evaluation of terrain slope as an error source

study	slope (degrees)	standard deviation (m)	mean error (m)	mean absolute error (m)	root mean square error (m)
Mouratidis et al. (2010)					
	0	4			
	1	4,2			
	2	4,8			
	3	5			
	4	5,5			
	5	7			
	6	8			
	7	8,5			
	8	9,8			
	9	12,5			
	10	13,5			
	11	14			
	12	16			
	13	15			
	14	16			
	15	19			
	16	17			
	17	19,8			
	18	20,2			
	19	31			
	20	32,5			
Szabo et al. (2015)					
	< 5	3,16		3,53	
	5-10	4,28		5,14	
	10-15	5,55		6,42	
	15-20	6,02		7,68	
	> 20	8,02		8,97	
Rexer and Hirt (2014)					
	very smooth	3,34	2,67		
	smooth	5,19	3,13		
	rough	5,82	3,45		
	mountainous	9,41	2,72		
	very mountainous	13,07	7,49		
Mukherjee et al. (2013)					
	< 2				9,22
	2-5				9,99
	5-10				12,52
	10-15				15
	15-20				20
	> 20				22,25
Shortridge and Messina (2011)					
	< 1	4,04	1,21		
	1-5	5,14	1,88		
	5-10	8,64	2,72		
	10-15	12	3,5		
	15-20	17	4,8		
	> 20	24,5	5,48		

Table 2. SRTM case studies

study	area characteristics	reference method	reference points	mean error (m)	mean absolute error (m)	RMSE (m)	standard deviation (m)
Bourgine and Baghdadi (2005)	diverse, 1-315 m	1 m accurate DEM	7152	26,50			8,20
Bourgine and Baghdadi (2005)	diverse, 0-450 m	topographic maps	392	13,30			10,20
Das et al. (2016)	mountaineous, 1200-6300 m	GPS	48	19,68			27,15
Du et al. (2016)	lowlands, river delta, 0-1300 m	GPS	76	-1,65			1,68
Falorni et al. (2005)	lowlands	geodetic benchmarks	36	-3,17			7,18
Falorni et al. (2005)		GPS	78	0,36			8,94
Falorni et al. (2005)	hilly, diverse, 10-1800 m	8,6 m accurate DEM	a lot	14,00			
Jarihani et al. (2015)	floodplains, -65-852 m	geodetic benchmarks + ICESat	373 200	2,68		3,25	1,84
Karlsson and Arnberg (2011)	floodplains, < 50 m	topographic maps	200	-1,95			4,21
Kiamehr and Sjöberg (2005)	diverse, -21-2551 m	geodetic benchmarks	260	2,02		6,53	
Kolecka and Kozak (2014)	mountainous, 900-2450 m	photogrammetric DEM	a lot	4,31		14,74	14,09
Mouratidis et al. (2010)	hilly, 0-1200 m	GPS	10 792		0,30		6,40
Mukherjee et al. (2013)	mountainous, 450-800 m	GPS	11	-2,94		9,02	
Mukherjee et al. (2013)		topographic maps	30	2,10		18,94	
Mukherjee et al. (2013)		photogrammetric DEM	a lot	2,37		13,18	
Mukul et al. (2017)	lowlands, < 300 m	GPS	21	3,27	7,01		9,64
Mukul et al. (2017)	foothills, 300-2000 m	GPS	40	11,91	12,26		10,74
Mukul et al. (2017)	mountainous, > 2000 m	GPS	30	14,75	16,32		11,58
Mukul et al. (2017)	diverse, -67-2245 m	GPS	100	10,01	10,44		8,29
Rexer and Hirt (2014)	diverse	geodetic benchmarks	772 696	3,02		6,29	5,52
Rodriquez et al. (2006)	diverse, continental	GPS	445 000	-0,70			3,70
Sharma (2014)	diverse, 130-1220 m	geodetic benchmarks	1 011	6,77	14,48	27,58	
Shortridge and Messina (2011)	diverse	National Elevation Dataset	245 579	2,04			8,31
Szabo et al. (2015)	mountainous	topographic maps	a lot	2,60			4,10
Wang et al. (2012)	mountainous, 4396-4777 m	GPS	5081			13,80	
average				**5,71**	**10,14**	**12,59**	**8,43**
median				**2,68**	**11,35**	**13,18**	**8,25**

Appendix 2: Research booklet

This appendix presents the booklet used to convey information about the study to the potential stakeholders. The stakeholders were prompted to read through the booklet before answering to the survey and the interview. In the study, the booklet consisted of three pages, however for printing purposes of this text, the booklet was expanded into four pages. The content has not been altered during the transformation.

Aalto University Water & Development Research Group Aalto University Geoinformatics Research Group

The value of open source-based river streamflow estimation in Southeast Asia

Water & Development Research Group
Geoinformatics Research Group
Department of Built Environment
Aalto University School of Engineering, Finland

Introduction

Streamflow in rivers relates closely to various water resource management issues, such as hydropower and irrigation systems development. Kallio et al.[1] have developed a tool called *hydrostreamer* which downscales low resolution runoff data by assigning the runoff into river segments and accumulating flow from upstream to downstream. The package can be used with free and openly available data, such as global runoff models, and all source code is freely available and documented. **Input data requirements for *hydrostreamer* have been kept minimal and using it does not require expertise in the field of hydrology.** With these factors in mind, the developers have aimed the tool for fast, understandable computing of streamflow estimates. This study conducts a case study of using *hydrostreamer* and evaluating its value in the water management context in Southeast Asia. In addition, one of the main uncertainties of *hydrostreamer* is quantified to increase its credibility and transparency.

Box 1. The focus of the study

What?
Evaluating the value of open source streamflow estimation in water resource management context. Case study in the 3S river basin.

Why?
To investigate the potential of open source solutions in supporting water resource management.

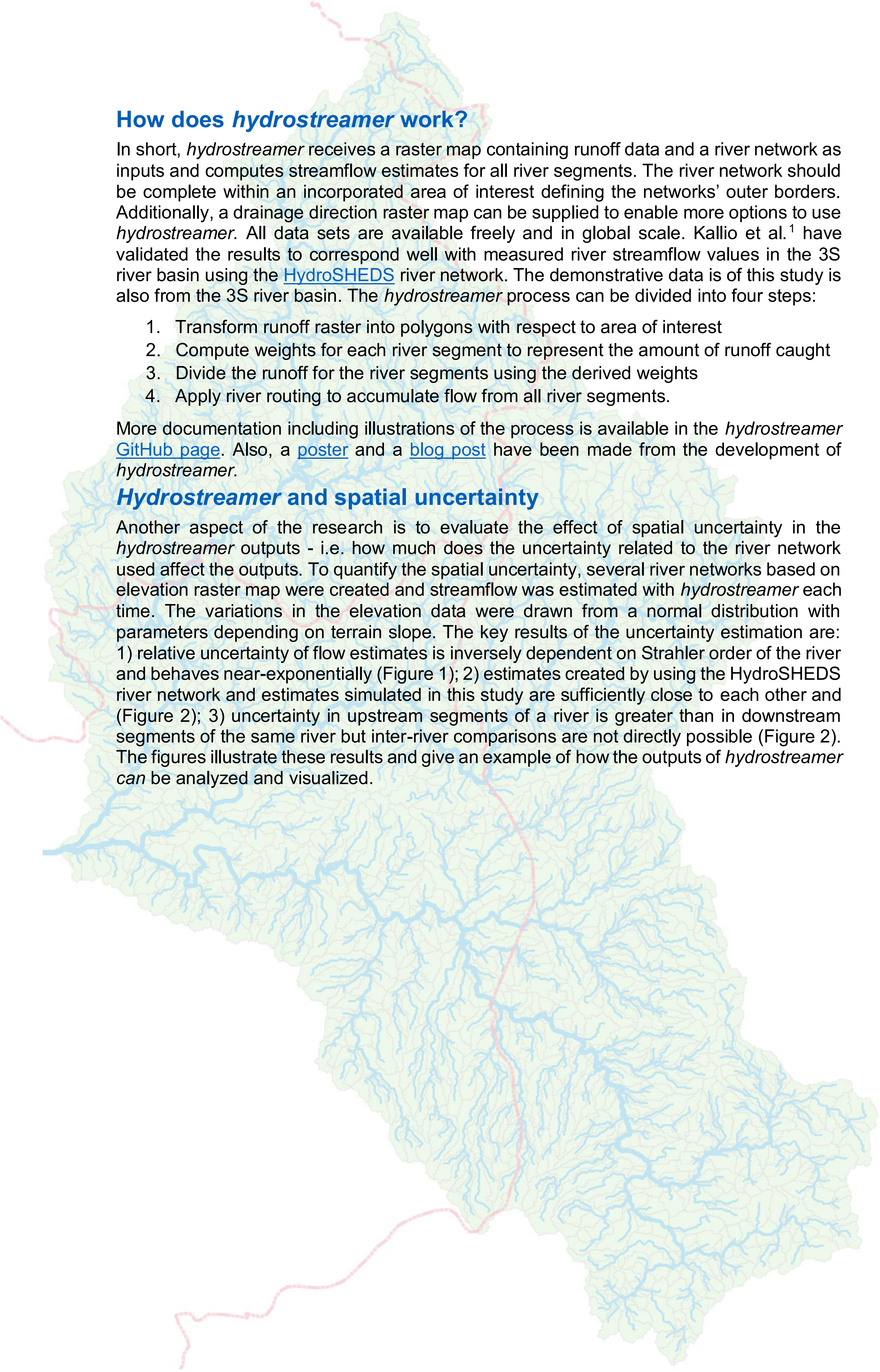
How?
Identifying and interviewing key stakeholders to gather feedback and development proposals based on example data and results provided by the researchers.

Stakeholder participation

The key aims of this study are presented in Box 1. We are searching for potential users, such as researchers and experts, who either apply or develop river streamflow modelling tools, or are a part of an organization that would benefit from the tools of our study. **The value of a tool or a workflow can't be estimated only by researchers looking at the tools and results produced by themselves: therefore, we are conducting a qualitative study consisting of potential end-users in their respective context.** The participants are asked to complete a survey, which is partially based on the information given in this booklet. After the survey, it is possible to arrange an interview between the us and you to gain deeper insights for both parties. The benefits of participating include achieving in-depth knowledge about *hydrostreamer* and the chance to guide further development of *hydrostreamer* by addressing issues.

---

[1] Kallio, M., Guillaume, J., Kummu, M., Desalegn, F., Virrantaus, K., 2018. Spatial allocation of low resolution runoff model outputs to high resolution stream network. https://doi.org/10.13140/RG.2.2.19122.09921



How does *hydrostreamer* work?

In short, *hydrostreamer* receives a raster map containing runoff data and a river network as inputs and computes streamflow estimates for all river segments. The river network should be complete within an incorporated area of interest defining the networks' outer borders. Additionally, a drainage direction raster map can be supplied to enable more options to use *hydrostreamer.* All data sets are available freely and in global scale. Kallio et al.[1] have validated the results to correspond well with measured river streamflow values in the 3S river basin using the HydroSHEDS river network. The demonstrative data is of this study is also from the 3S river basin. The *hydrostreamer* process can be divided into four steps:

1. Transform runoff raster into polygons with respect to area of interest
2. Compute weights for each river segment to represent the amount of runoff caught
3. Divide the runoff for the river segments using the derived weights
4. Apply river routing to accumulate flow from all river segments.

More documentation including illustrations of the process is available in the *hydrostreamer* GitHub page. Also, a poster and a blog post have been made from the development of *hydrostreamer.*

Hydrostreamer and spatial uncertainty

Another aspect of the research is to evaluate the effect of spatial uncertainty in the *hydrostreamer* outputs - i.e. how much does the uncertainty related to the river network used affect the outputs. To quantify the spatial uncertainty, several river networks based on elevation raster map were created and streamflow was estimated with *hydrostreamer* each time. The variations in the elevation data were drawn from a normal distribution with parameters depending on terrain slope. The key results of the uncertainty estimation are: 1) relative uncertainty of flow estimates is inversely dependent on Strahler order of the river and behaves near-exponentially (Figure 1); 2) estimates created by using the HydroSHEDS river network and estimates simulated in this study are sufficiently close to each other and (Figure 2); 3) uncertainty in upstream segments of a river is greater than in downstream segments of the same river but inter-river comparisons are not directly possible (Figure 2). The figures illustrate these results and give an example of how the outputs of *hydrostreamer can* be analyzed and visualized.

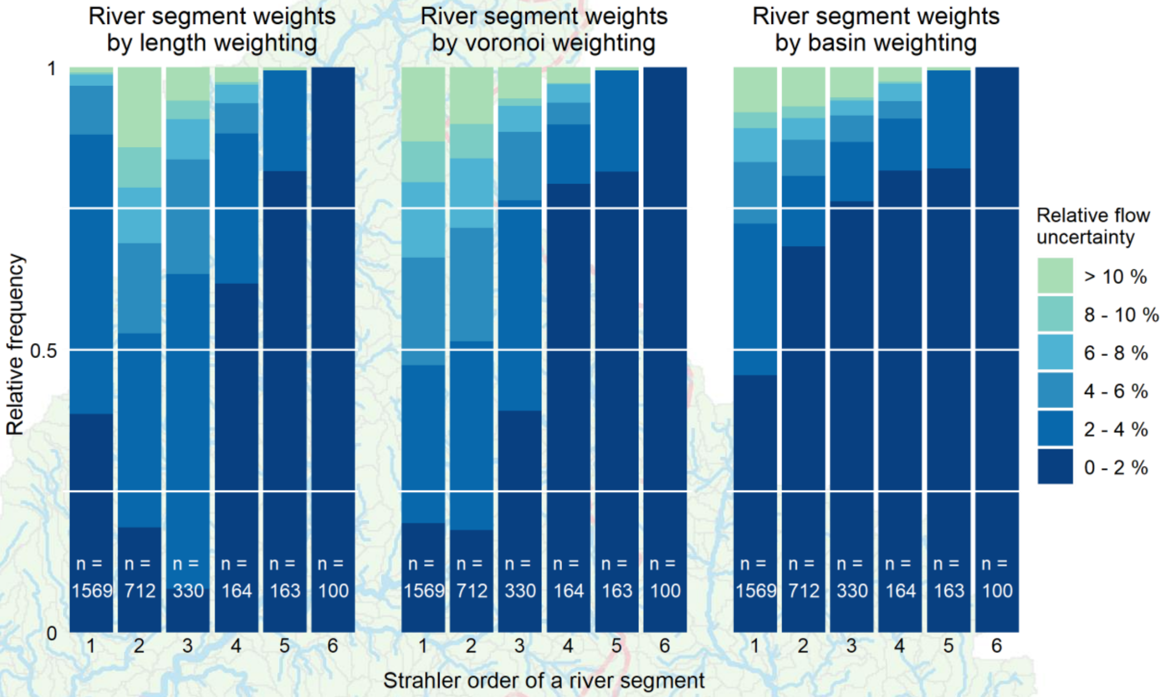


Figure 1. Relative streamflow uncertainty as a function of Strahler order (river size) in the 3S river basin.

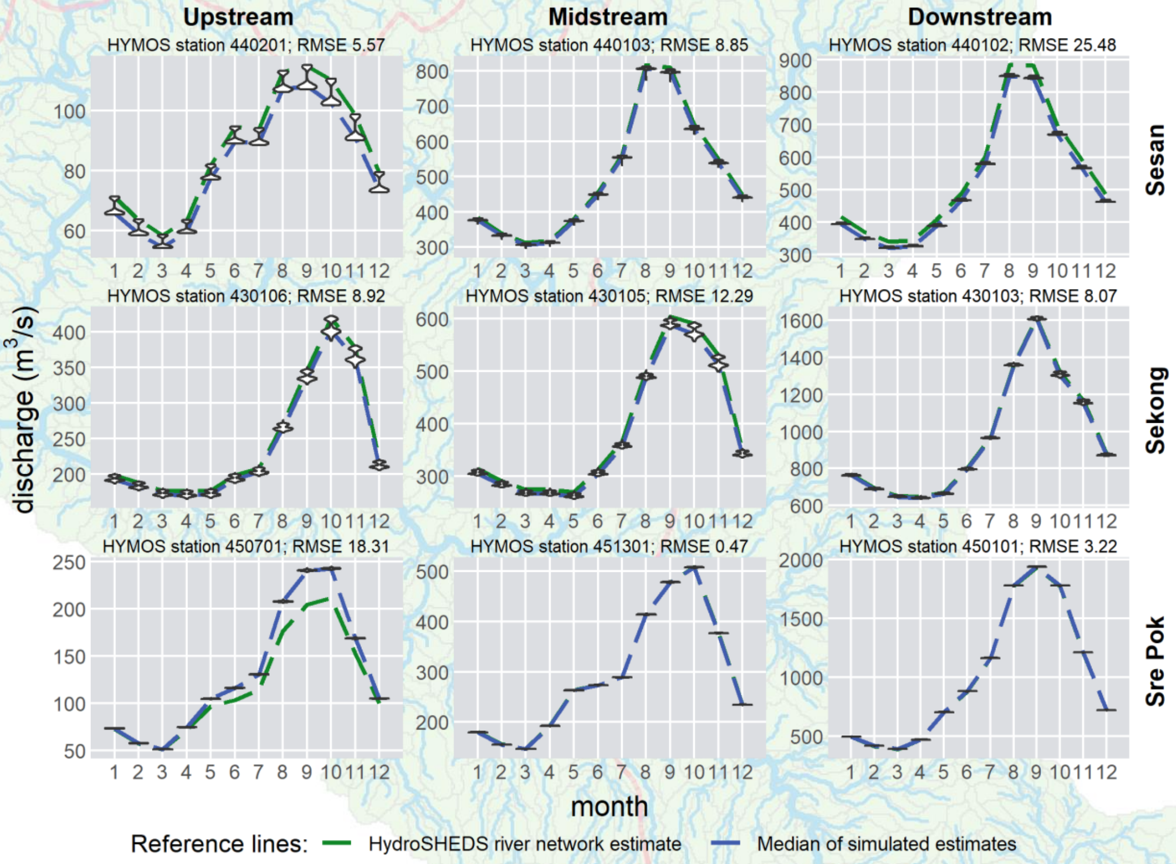
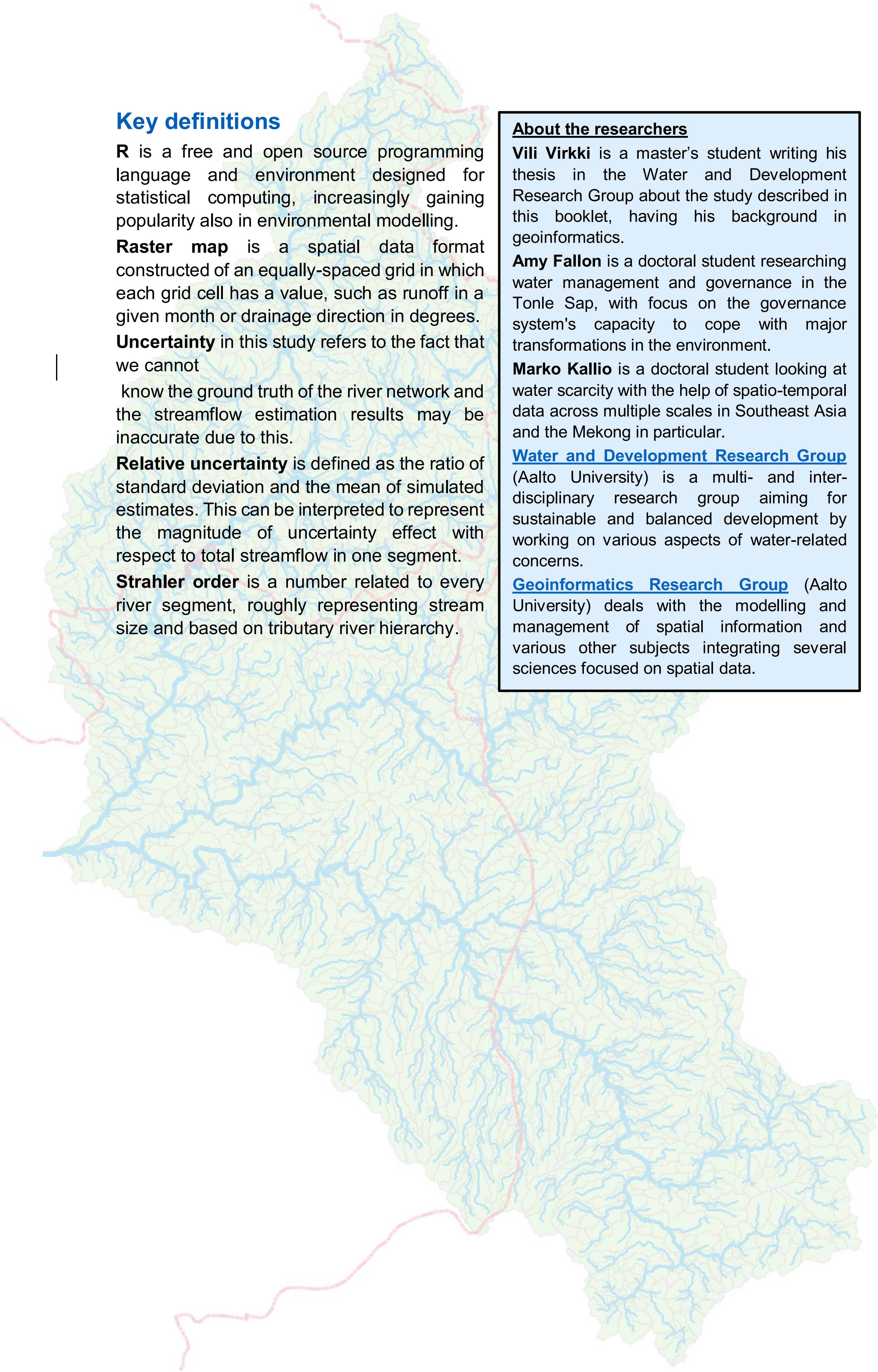


Figure 2. Streamflow in upstream, midstream, and downstream parts of Sekong, Sesan, and Sre Pok rivers. River segment weights computed by basin weighting and root mean square error between reference lines.



Key definitions

R is a free and open source programming language and environment designed for statistical computing, increasingly gaining popularity also in environmental modelling.

Raster map is a spatial data format constructed of an equally-spaced grid in which each grid cell has a value, such as runoff in a given month or drainage direction in degrees.

Uncertainty in this study refers to the fact that we cannot know the ground truth of the river network and the streamflow estimation results may be inaccurate due to this.

Relative uncertainty is defined as the ratio of standard deviation and the mean of simulated estimates. This can be interpreted to represent the magnitude of uncertainty effect with respect to total streamflow in one segment.

Strahler order is a number related to every river segment, roughly representing stream size and based on tributary river hierarchy.

About the researchers

Vili Virkki is a master's student writing his thesis in the Water and Development Research Group about the study described in this booklet, having his background in geoinformatics.

Amy Fallon is a doctoral student researching water management and governance in the Tonle Sap, with focus on the governance system's capacity to cope with major transformations in the environment.

Marko Kallio is a doctoral student looking at water scarcity with the help of spatio-temporal data across multiple scales in Southeast Asia and the Mekong in particular.

Water and Development Research Group (Aalto University) is a multi- and inter-disciplinary research group aiming for sustainable and balanced development by working on various aspects of water-related concerns.

Geoinformatics Research Group (Aalto University) deals with the modelling and management of spatial information and various other subjects integrating several sciences focused on spatial data.

Appendix 3: Survey outline

In this appendix, the survey used in this study is presented. While collecting answers to the survey, the survey website was displayed in four parts: first, the introductory text and information about the survey and then three distinct parts consisting of questions. Each page of this appendix presents a part of the survey; the wordings in this appendix are exactly the same than in the version provided to the respondents. Questions marked with an asterisk (*) were compulsory. If the respondent chose to answer a multiple-choice answer with a "please specify" option, filling of the specification was compulsory. If no choices were provided, the question was an open question.

Introduction

The value of open source-based river streamflow estimation in Southeast Asia

Water & Development Research Group
Geoinformatics Research Group
Department of Built Environment, Aalto University School of Engineering, Finland

This survey aims to gather experts' and practitioners' opinions on the current state and prospects of river streamflow estimation in Southeast Asia. The survey is based on the study described in the attached booklet. Answering the survey may be complemented by arranging an interview with the researchers that will go deeper into the subject partially based on the answers given.

The study aims to evaluate the value of hydrostreamer, a newly developed open source tool to estimate river streamflow. The tool is based on open data, which is available globally, and can provide quick and understandable means of streamflow estimation. This survey asks the participants to browse through the provided booklet and answer questions which are grouped into three parts. The results of the study will be used to collect information about current streamflow modelling tools and the potential use and development proposals of hydrostreamer.

The survey will take 15-20 minutes to complete. Regarding open-ended questions, please keep your answers concise: use of bullet points or lists is encouraged. Once a part of the survey is complete, it is not possible to return and change the answers of that part. All answers will be treated confidentially. It is not compulsory to provide any contact information in the study if you wish to remain anonymous.

The study is an essential part of Vili Virkki's Master's Thesis and builds on Aalto University Water & Development Research Group's previous studies in Southeast Asia.

Thank you for participating in our survey!

Vili Virkki (vili.virkki@aalto.fi)
Amy Fallon
Marko Kallio

Below is a map of our study area with HydroSHEDS river network and the drainage basins for each river segment. The HYMOS streamflow observation stations used for validation are marked and numbered on the map. After the map, the survey starts.

Figure 1. Study area map presented in the survey

Part 1 of 3: Background

The first part of the survey consists of background questions related to streamflow modelling and the present methods and tools used in your organization. Even if your organization doesn't apply streamflow modelling, please still answer to the questions and continue the survey.

1. What kind of an organisation are you representing? *

Academic
Non-governmental organization
Ministry or other governmental organization
Commercial
Other (please specify)

2. Which kind of tools or software do you use to estimate streamflow at the moment? *

My organization does not apply streamflow modelling
The streamflow modelling has been outsourced
My organization applies streamflow modelling with tools specified:

3. On scale from 1 to 5, how satisfied are you with the performance of the current streamflow modelling tools (1 = not satisfied at all; 5 = very satisfied)? *

1
2
3
4
5
Not applicable (no streamflow modelling)

4. Are there any issues in your current tools that lower your satisfaction?

Part 2 of 3: Hydrostreamer

The second part consists of questions related to the usefulness of streamflow estimation and the tool being evaluated in our study. The key definitions and figures represented in the booklet are recapped before the questions to help with this part:

R is a free and open source programming language and environment designed for statistical computing, increasingly gaining popularity also in environmental modelling.

Hydrostreamer is an R package that downscales low resolution runoff data by assigning the runoff into river segments and accumulating flow from upstream to downstream.

Uncertainty in this study refers to the fact that we cannot know the ground truth of the river network and the streamflow estimation results may be inaccurate due to this.

Relative uncertainty is defined as the ratio of standard deviation and the mean of simulated estimates. This can be interpreted to represent the magnitude of uncertainty effect with respect to total streamflow in one segment.

Strahler order is a number related to every river segment, roughly representing stream size and based on tributary river hierarchy.

Figure 1. *See Figure 1 in Appendix 2*
Figure 2. *See Figure 2 in Appendix 2*

5. Would streamflow estimates with quantified uncertainty be useful for your organization? *

No
Maybe
Yes
Unsure, because:

6. If you answered yes to the previous question, what kind of applications would you use the estimates in?

7. How well do you understand the hydrostreamer workflow based on available documents? *

Not at all
Most parts remain unclear; some parts are understood
Most parts are understood; some parts remain unclear
Very well

8. Do you see any issues in gathering the data to use with hydrostreamer? *

No
Yes (please specify)

9. In your opinion, is the demonstrated uncertainty of hydrostreamer within acceptable limits? *

Yes
No (please explain why)
Unsure how to interpret it

10. Can you think of any additions to hydrostreamer that would benefit your organization?

Part 3: Further participation in the study

Finally, the third part queries your or your organizations interest in further participation in our study. It is not compulsory to leave your contact details if you wish not to participate further. However, participation includes benefits also for your organization.

11. Additional comments and queries about the study

12. Would you be interested in further discussion conducted over internet about the study? In doing so, you could be involved in developing hydrostreamer and gain deeper understanding of its use. *

Yes
No

13. Would your organization be interested in receiving more information or for example a demo session on hydrostreamer, free of charge?

Further information/documentation
Demo data
Training session
Other, please specify:

14. If you answered yes to either of the previous questions, please provide your contact information. Answering to this question is not compulsory if you wish to stay anonymous.

Name

Organization

E-mail address

Thank you for taking your time to answer our survey!
Please submit the answers by checking the box and clicking "Submit" below. If you left your contact information in the survey, we will contact you shortly afterwards.

Appendix 4: Semi-structured interview guide

The following interview guide was used in semi-structured interviews. The questions were not asked in exact wordings, but all topics were considered in each interview.

1. **Introductions of participants**

2. **General streamflow modelling things**
   a. Do you model streamflow or other water-related systems?
   b. What kind of tools do you use for modelling?
   c. Is it hard to use the current tools or are there any difficulties at all?

3. **Hydrostreamer**
   a. What kind of applications would you see hydrostreamer to be used in?
   b. Do you find that hydrostreamer would need some additions?
   c. How easy do you find hydrostreamer to use?

4. **Uncertainty**
   a. Can you understand the uncertainty modelling?
   b. How do you see the simulation results, do they increase the value of hydrostreamer?

5. **Open data**
   a. What is the meaning of the tool being completely open? Good or bad?
   b. Is it possible that practitioners could misuse the tool?

6. **Final remarks**
   a. Any other topics you would like to talk about?
   b. Do you know other people or organizations that could be interested in this?