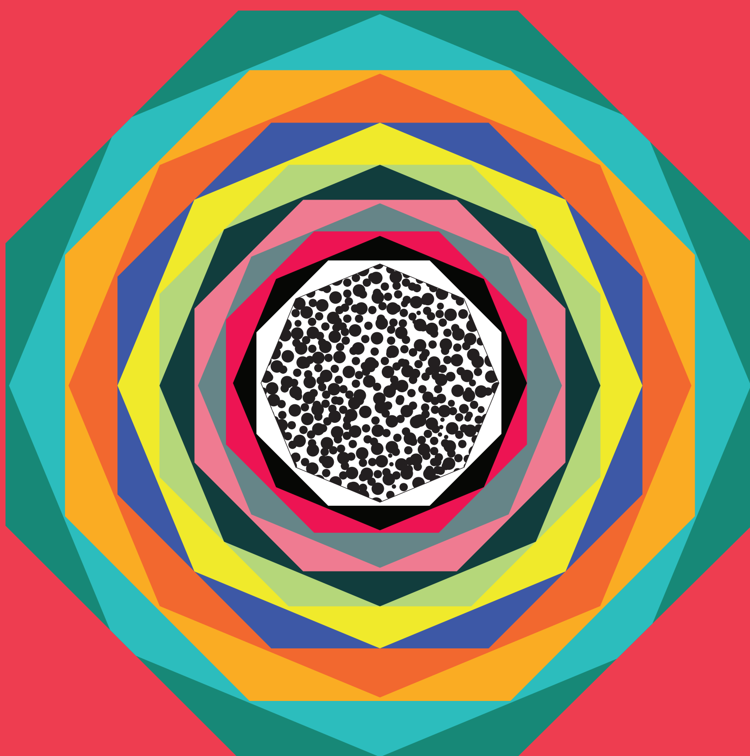
Department of Civil and Environmental Engineering

Spatial Vulnerability Assessments for Water Resources Management

Cases from major Asian river basins with a focus on spatial unit of analysis and the use of big and open data

---

Aura Salmivaara



Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 93/2015

Spatial Vulnerability Assessments for Water Resources Management

Cases from major Asian river basins with a focus on spatial unit of analysis and the use of big and open data

Aura Salmivaara

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Engineering, at a public examination held at the lecture hall M1 of the school on 4 September 2015 at 12.

Aalto University
School of Engineering
Department of Civil and Environmental Engineering
Water and Environmental Engineering Research Group

Supervising professor
Professor Olli Varis

Thesis advisors
Assistant Professor Matti Kummu
Dr. Marko Keskinen

Preliminary examiners
Professor Patrick M. Reed, Cornell University, USA
Professor Bin Chen, Beijing Normal University, China

Opponent
Dr. Robyn Margaret Johnston,
International Water Management Institute, Sri Lanka

Aalto University publication series
DOCTORAL DISSERTATIONS 93/2015

ISBN 978-952-60-6282-2 (printed)
ISBN 978-952-60-6283-9 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-6283-9

Unigrafia Oy
Helsinki 2015

Finland

NORDIC ECOLABEL
441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Aura Salmivaara

Name of the doctoral dissertation
Spatial vulnerability assessments for water resources management – Cases from major Asian river basins with a focus on spatial unit of analysis and the use of big and open data

Publisher School of Engineering

Unit Department of Civil and Environmental Engineering

Series Aalto University publication series DOCTORAL DISSERTATIONS 93/2015

Field of research Water and Enviromental Engineering

Manuscript submitted 10 April 2015 **Date of the defence** 4 September 2015

Permission to publish granted (date) 15 June 2015 **Language** English

☐ **Monograph** ☒ **Article dissertation (summary + original articles)**

Abstract

This dissertation looks at the use of spatial analysis with big and open data for water-related vulnerability assessments in major river basins of Monsoon Asia. Special focus is on the spatial unit of analysis by exploring various ways to define it and by examining systematically the related Modifiable Areal Unit Problem (MAUP).

The extent and availability of spatial data have grown rapidly. This big and open spatial data, when combined, mapped and analysed, increases our understanding of interlinked issues and provides support for decision-making. However, the seemingly transparent way of map overlay and zonal analysis require closer examination. This is particularly important, when GIS and spatial analysis are applied for water resources management, which involves actors, values, and demands from various sectors and drivers of change on multiple scales.

In Monsoon Asia (covering the area from China to eastern Afghanistan, with a population of 3.52 billion) the drivers of change include: climate change, population growth, urbanisation and various development pressures. The region has major and transboundary river basins making management of water resources particularly challenging.

This dissertation includes four case studies that draw findings from three scales: regional, basin and subbasin. Both data-driven and a priori methods were utilised in defining the spatial unit of analysis and new approaches to finding appropriate spatial units of analysis were developed.

Based on the case studies, this dissertation demonstrates that the big and open spatial data is extremely useful for water resources management. Yet, the findings indicate that the scale influences profoundly the applicability and performance of the spatial datasets. Moreover, the spatial unit of analysis through the MAUP has significant influence in the analysis results. A multizonal and multiscale approach was found to minimise the negative effects of MAUP. Through such approach it is possible to find appropriate spatial unit of analysis.

The findings reinforce the importance of reporting explicitly the choices and assumptions behind the spatial units of analysis. Classifying spatial data to avoid accumulation of uncertainty and identification of data gaps is strongly recommended. Finally, simplicity should be emphasised when conducting vulnerability assessments to ensure comparability. However, also more complex methods were found to have potential to support the process of analysis. The findings help to develop spatial approaches to vulnerability assessments, and thus, enhance the applicability of big and open data for water resources management.

Keywords vulnerability assessment, water resources management, spatial analysis, Monsoon Asia, spatial unit of analysis, MAUP, big data, open data

ISBN (printed) 978-952-60-6282-2 **ISBN (pdf)** 978-952-60-6283-9

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2015

Pages 150 **urn** http://urn.fi/URN:ISBN:978-952-60-6283-9

Aalto-yliopisto

Tiivistelmä

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

Tekijä
Aura Salmivaara

Väitöskirjan nimi
Paikkatietoa hyödyntävät haavoittuvuusarvioinnit vesivarojen hallinnassa: Tapaustutkimuksia Aasian merkittävistä jokivesistöistä keskittyen analyysiyksikköön sekä laajan ja avoimen datan käyttöön

Julkaisija Insinööritieteiden korkeakoulu

Yksikkö Yhdyskunta- ja ympäristötekniikan laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 93/2015

Tutkimusala Vesi- ja ympäristötekniikka

Käsikirjoituksen pvm 10.04.2015 **Väitöspäivä** 04.09.2015

Julkaisuluvan myöntämispäivä 15.06.2015 **Kieli** Englanti

☐ **Monografia** ☒ **Yhdistelmäväitöskirja (yhteenveto-osa + erillisartikkelit)**

Tiivistelmä

Väitöskirja käsittelee paikkatietoanalyysin käyttöä, keskittyen laajan ja avoimen datan (big and open data) hyödyntämiseen vesivaroihin liittyvissä haavoittuvuusarvioinneissa sekä analyysiyksikön valintaan ja sen vaikutukseen tuloksissa. Väitöskirja perustuu neljään vertaisarvioituun tieteelliseen julkaisuun sekä näiden pohjalta laadittuun synteesiin. Tutkimusalueena ovat suuret joet ja niiden valuma-alueet erityisesti monsuuni-Aasiassa eli monsuuni-ilmaston alueella Kiinasta itäiseen Afganistaniin. Analyysiyksikön määrittämisessä käytettiin sekä ennalta määritettyjä että datan perusteella tilastollisin menetelmin määräytyneitä kriteerejä. Väitöskirjassa kehitettiin uusia menetelmiä sopivan analyysiyksikön määrittämiseksi.

Paikkatietoaineistojen kattavuus ja saatavuus ovat lisääntyneet viime vuosina huomattavasti. Laajaa ja avointa dataa voidaan hyödyntää yhdistämällä ja kartoittamalla ja täten käyttää tätä dataa monimutkaisten ilmiöiden tarkasteluun ja päätöksenteon tukemiseen. Paikkatietoanalyysin näennäinen yksinkertaisuus vaatii kuitenkin huomiota. Erityisen tärkeää tämä on vesivarojen hallinnassa, johon vaikuttavat monenlaiset toimijat ja tarpeet eri sektoreilla. Vesivarojen hallinta on haastavaa nopeasti muuttuvassa monsuuni-Aasiassa, jonka väkiluku on tällä hetkellä noin 3,5 miljardia. Muutoksia aiheuttavat mm. ilmastonmuutos, väestönkasvu, kaupungistuminen sekä erilaiset kehityspaineet kuten luonnonvarojen hyödyntäminen ja infrastruktuurin rakentaminen. Lisäksi monet alueen suurista jokivesistöistä myös jakaantuvat kahden tai useamman valtion alueelle tuoden lisähaasteita vesivarojen hallintaan. Tätä taustaa vasten paikkaan sidotut haavoittuvuusarvioinnit ja niiden kehittäminen ovat tärkeitä asioita.

Väitös osoittaa, että laajat ja avoimet paikkatietoaineistot ovat erittäin hyödyllisiä veteen liittyvissä haavoittuvuusanalyyseissä. Lisäksi analyysiyksikön valinnalla on suuri merkitys tuloksiin, ja tähän ei ole kiinnitetty vesivarojen hallinnassa riittävästi huomiota. Väitöskirja suositteleekin usean analyysiyksikön ja mittakaavan soveltamista veteen liittyvissä paikkatietoanalyyseissä. Väitöskirja painottaa analyysiprosessiin liittyvien valintojen selkeän ja läpinäkyvän raportoinnin merkitystä. Myös paikkatiedon luokittelu on tarpeen, jotta voidaan huomioida paremmin aineiston puutteet sekä välttää epävarmuuksien kertaantuminen. Vertailtavuus on paikkatietoon perustuvien haavoittuvuusarviointien suurin etu, ja siksi on suositeltavaa käyttää mahdollisimman yksinkertaisia menetelmiä esimerkiksi indikaattorien yhdistelyssä.

Väitöskirjan tulokset edesauttavat paikkatietoon perustuvien analyysimenetelmien kehittämistä veteen liittyvissä haavoittuvuusarvioinneissa ja hyödyttävät laajan ja avoimen datan monipuolisempaa käyttöä vesivarojen kestävän ja kokonaisvaltaisen hallinnan tukemiseksi.

Avainsanat haavoittuvuusarviointi, vesivarojen hallinta, Aasia, paikkatietoanalyysi, spatiaalinen analyysiyksikkö, avoin data

ISBN (painettu) 978-952-60-6282-2 **ISBN (pdf)** 978-952-60-6283-9

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Helsinki **Painopaikka** Helsinki **Vuosi** 2015

Sivumäärä 150 **urn** http://urn.fi/URN:ISBN:978-952-60-6283-9

CONTENTS

ACKNOWLEDGMENTS

I´d say the Water & Development Research Group is one of the best things in Aalto University and it provided great surroundings for my master's thesis and for this doctoral dissertation. It also served as a good base camp gearing me up for conference and study trips to various locations in Southeast Asia and Europe.

This work would not have been started if it hadn't been for Professor Olli Varis and his WDRG group, especially Assistant Professor Matti Kummu and university lecturer Dr. Marko Keskinen. I am grateful for getting the opportunity to join the group and I am very thankful for the positive and inspiring atmosphere as well as the supportive and understanding instructors and colleagues it provided. And naturally the dissertation would have not been finalised without the continuous support from Marko Keskinen, Olli Varis and Matti Kummu.

The articles have each provided a true learning process. Particularly pleasant was the planning and writing of the fourth article, in which doctoral candidate M.Sc. Miina Porkka and Dr. Joseph H. A. Guillaume also contributed as co-authors. I want to express my gratitude for their enthusiasm and engagement.

Other projects, outside this dissertation, have provided very important contribution to the whole PhD process. I want to thank doctoral candidate M.Sc. Mirja Kattelus and Lic.Soc.Sc. Ilkka Mellin for cooperation and enormous support.

I would also like to thank Associate Professor Muhammad Mizanur Rahaman for study ideas concerning Myanmar and Irrawaddy & Salween. I am also thankful for Dr. Ilan Kelman and M.Sc. Armando Lamadrid for the book project on Myanmar, which provided helpful contribution to the dissertation.

Regarding the preliminary examination of this synthesis, I would like to express gratitude for Professor Patrick M. Reed and Professor Bin Chen for a thorough and efficient examination process that helped me to refine the synthesis.

I would also like to thank Dr. Robyn Johnston for accepting the invitation for acting as the opponent for the dissertation.

In addition, I would like to thank people in Water and Environmental Engineering Research Group and in Aalto University for creating such a nice working environment: Miina Porkka, Mirja Kattelus, Suvi Sojamo, Dr. Timo Räsänen, Juho Haapala, Mika Jalava, Dr. Shokhrukh-Mirzo Jalilov, Prof. Riku Vahala, Prof. Harri Koivusalo, Anne-Maj Seppälä, Irmeli Gustavsson, Ari Järvinen, Antti Louhio, Aino Peltola, Pirjo Rantanen, Dr. Hanne Laine-Kaulio, Dr. Nora Sillanpää, Kaisa Västilä, Johanna Jalonen, Matti Keto, Dr. Teemu Kokkonen, Emeritus Professor Pertti Vakkilainen, Dr. Ulla Heinonen, Dr. Ari Jolma + others coming and going and keeping up the liveliness in the hallways of water lab.

During various trips to Southeast Asia I have encountered wonderful people, and I'd like to use this opportunity to thank them: Jorma Koponen, Dr. Juha Sarkkula, Dr. Pech Sokhem, Dr. John Dore, Dr. Lilao Bouapao, Dr. Dipak Gyawali, Kate Lazarus, Dr. Guillaume Lacombe, Dr. Tira Foran, Tarek Ketelsen, Dr. Paradis Someth, Dr. John Ward, Dr. Alex Smajgl, Tony Hell, Yim Sambo, Bernard Cartella, Dr. Chayanis Krittasudthacheewa, Dr. Kim Geheb, Mira Käkönen, and the M-Power network.

Most part of the dissertation was conducted with the funding from the VALUE doctoral program, which brought me to meet fellow doctoral candidates in its seminars and courses, thanks to Dr. Riitta Kamula and Prof. Björn Klöve. For especially great peer support, thank you Anniina Takala, Dr. Mia O'Neill, Dr. Pekka Rossi, and Iisa Orell.

I have encountered many people who have each in some way contributed to this dissertation, sometimes by just being part of my life. For inspiration, fun, humour, sailing, musical amusements, good conversations and support, I would like to express gratitude to my family and friends: particularly Juha Knuuttila, Anna-Liisa & Timo Salmivaara, Assi, Tomaso, Helmi & Carlo Battilana, Inari, Guara & Aruañan Salmivaara da Conceiçao, Liisa Satukangas, Paula Nuorteva, Helga Lahtinen, Piia Kähkölä, Erika Palomäki, Antti Mäntymaa, Tuula Renbi, IPPJR 2012-2014, FC MiksipäEi, and also the scene and provider of all good things and people: folkboat Avanti.

Finally the doctoral studies would have not been possible without the funding support that I received during the years: TKK tutkijankoulutusstipendi, NIAS SUPRA Nordic scholarship, VALUE doctoral school, Maa- ja vesitekniikan tuki ry., "Exploring Tonle Sap Futures" research project, RYM-TO travel grants, Aalto University funds for PhD, Academy of Finland projects 133748 and 269901, and Sven Hallin Research Foundation.

60°11'32.6"N 24°56'15.8"E
25th June 2015

Aura Salmivaara

LIST OF APPENDED ARTICLES

The dissertation consists of this synthesis and the following peer-reviewed scientific articles:

Article I — Varis, O., Kummu, M. & **Salmivaara, A.**, (2012): Ten major river basins in monsoon Asia-Pacific: An assessment of vulnerability. Applied Geography 32(2): 441-454, doi: 10.1016/j.apgeog.2011.05.003.

Article II — **Salmivaara, A.**, Kummu, M., Keskinen, M. & Varis, O., (2013). Using Global Datasets to Create Environmental Profiles for Data-Poor Regions: A Case from the Irrawaddy and Salween River Basins. Environmental Management (2013) 51:897-911, doi: 10.1007/s00267-013-0016-x.

Article III — **Salmivaara, A.**, Kummu, M., Varis, O., Keskinen, M. (2015) Socio-Economic Changes in Cambodia's Tonle Sap Lake Area: A Spatial Approach. Applied Spatial Analysis and Policy (2015), doi: 10.1007/s12061-015-9157-z.

Article IV — **Salmivaara, A.**, Porkka, M., Kummu, M., Keskinen, M., Guillaume, J. H. A., Varis, O. (2015). Exploring the Modifiable Areal Unit Problem in Spatial Water Assessments: A Case of Water Shortage in Monsoon Asia. Water 7(3): 898-917, doi: 10.3390/w7030898.

AUTHOR’S CONTRIBUTION

The contribution of the author to each of the appended articles is presented in detail below:

Article I	The author contributed in collecting data and designing the study. Professor Varis and Assistant Professor Kummu were mainly responsible for the study design. Assistant Professor Kummu conducted the data preparations and GIS analyses and Professor Varis was responsible for the writing the article. The author contributed in writing parts for description and discussion on vulnerability assessments.
Article II	The author contributed in designing the study, for which the preliminary idea came from Professor Varis, Assistant Professor Kummu and Associate Professor Muhammad Mizanur Rahaman. The author is fully responsible for collecting data and conducting analyses. The author is mainly responsible for preparing figures and writing the article. Assistant Professor Kummu, Professor Varis and Dr. Keskinen participated in formulating the focus of the article and assisted in writing the discussion and conclusions.
Article III	The author had the main responsibility of preparing the data and designing and carrying out the analyses, discussing the results, preparing figures and writing the article. Dr. Keskinen had the main responsibility of designing the study related to a wider research project. Dr. Keskinen contributed in formulating the focus for the article and writing parts of the article. Assistant Professor Kummu and Professor Varis participated in formulating the discussion and conclusions parts of the article.
Article IV	The author was fully responsible for designing the study, collecting the data, carrying out the analyses and discussing the results. The author prepared the figures and wrote the manuscript. M.Sc. Porkka and Assistant Professor Kummu participated in designing the study and assisted with the analyses. Author together with M.Sc. Porkka and Assistant Professor Kummu discussed the results jointly and Dr. Keskinen, Dr. Guillaume and Professor Varis contributed in editing and refining the manuscript.

1. INTRODUCTION: Research setting and objectives

Water, the crucial element of all forms of life, varies spatially and temporally. Managing water resources requires spatially-explicit understandings of this complex system, which can be vulnerable to various factors.

This dissertation is generated within the applied science of water resources management under civil and environmental engineering. Generally the research aims to contribute to the sustainable use of natural resources, building on the principles of integrated water resources management (IWRM)[1] and cross- and inter-disciplinary approaches. IWRM is here understood as a baseline framework: it emphasises the importance of looking at water issues from a wider perspective rather than just one sector's point of view, and in such a way that equitable use of water is possible without endangering ecosystems or the possibilities for meeting the future needs for water (GWP-TAC, 2000).

The disciplinary foundations lie mainly in water resources management, and are supported by geoinformatics and vulnerability assessment, with a focus on the use of spatial analysis with big and open spatial data (Figure 1).

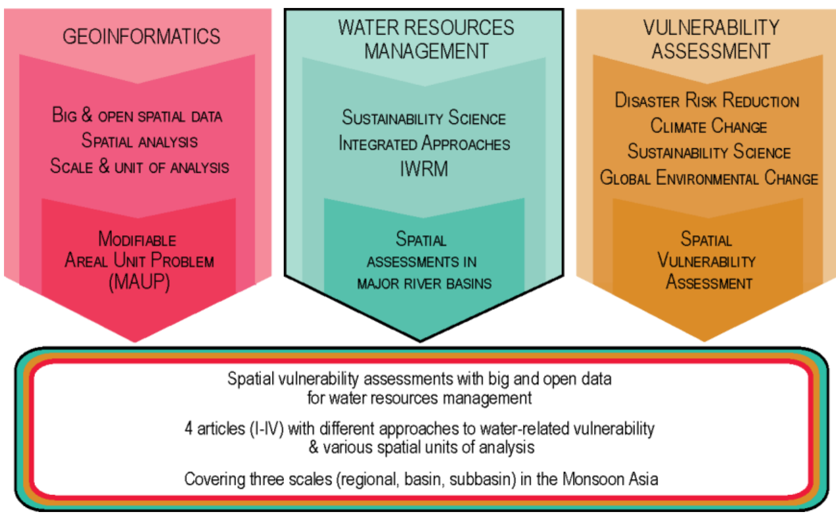


Figure 1. Scientific domains and conceptual framework of the dissertation. Details for each domain are presented in Chapter 2.

[1] Described in Chapter 2.1.

The integrative principles of IWRM, concepts from vulnerability science and the increasing availability of big and open spatial data have influenced strongly the case studies (Articles I-IV) that form the four pillars of this dissertation (presented in the appended articles). While the pillars differ and have their limitations, they nevertheless provide a platform for increasing our understanding about the key issues related to the use of spatial vulnerability assessment for water resources management. The appended articles are complemented by this synthesis that brings together the main findings from the case studies and puts them into a wider theoretical context.

The concept of big and open data is here defined as large public datasets that are available free of charge or at minimal cost[2]. The focus is particularly in spatial data. The data used in this dissertation follows the classification for open data and big data by Joel Gurin (2014). Gurin classifies e.g. large datasets from scientific research and large public government datasets, such as census, weather, GPS, as both big data and open data. Classifications are not uniform. Kitchin (2014), for example, defines big data as generated continuously, with fine-grained scope and flexible and scalable production, thus defining census data actually as small data due to its limitations in scope, temporal coverage and size.

This dissertation focuses on the use of structured secondary and tertiary data (Kitchin, 2014), i.e. the data is not generated by the author. The sort of big data described by Kitchin (2014) and Hurwitz, Halper & Kaufman (2013) is yet to emerge in the field of water resources management, and thus, the dissertation treats the data used in the appended articles as big and open data. Furthermore, despite the focus in openness and geoinformatics, the dissertation does not go beyond the data into discussing open source geographical information system (GIS) software or cloud processing applications.

The overall goal for this dissertation is to increase the understanding about the use of spatial analysis with big and open quantitative data for water-related spatial vulnerability assessment. This goal is approached through more detailed research questions (Figure 2):

> **Research Question 1:**
> What is the applicability and performance of big and open spatial data in spatial platforms?

The dissertation aims to increase understanding about the spatial units of analysis used in water-related spatial vulnerability assessments in particular by exploring ways to define appropriate spatial unit of analysis.

---

[2] It should be noted, however, that part of the data used requires registration or is freely available only for educational activities and research.

Research Question 2:
What are appropriate spatial units of analysis for water-related spatial vulnerability assessments?

Related to the previous, the third objective is to explore the role of the spatial unit of analysis for assessment results, particularly in relation to the phenomenon of Modifiable Areal Unit Problem (MAUP).

Research Question 3:
What is the role of the spatial unit of analysis and how the Modifiable Areal Unit Problem (MAUP) influences the assessment results?

Finally, and based on the above three research questions, the fourth objective is to discuss and give practical recommendations for the way forward in water-related spatial vulnerability assessments.

Research Question 4:
What kind of practical recommendations can be given to overcome the challenges in water-related spatial vulnerability assessments?

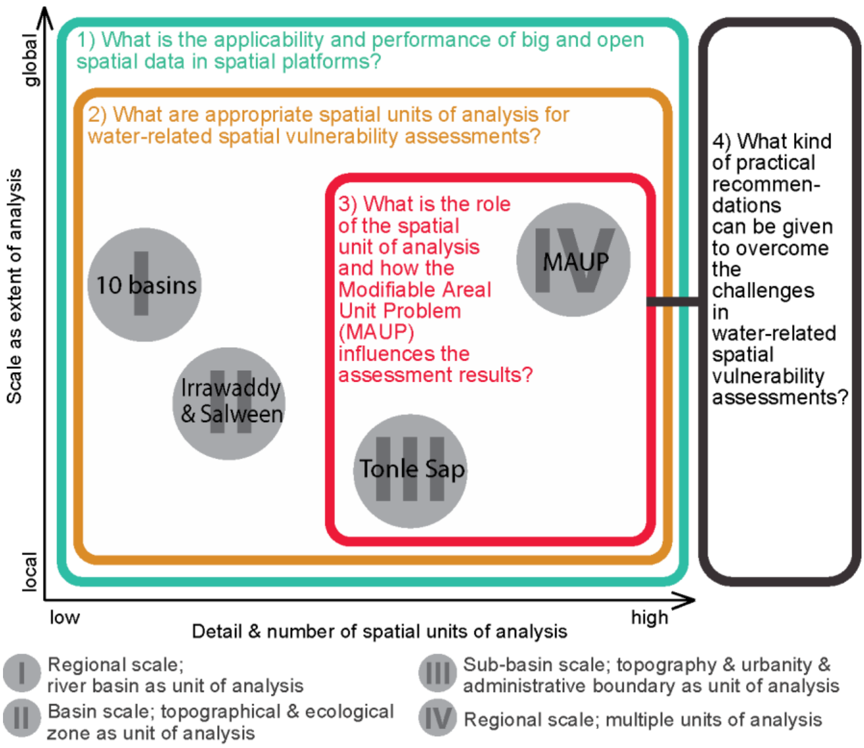


Figure 2. Research questions with the contribution of appended articles (indicated with roman numbers I-IV). The scale as extent of the analysis and the detail and number of the spatial units of analysis are presented on the axes.

The emphasis of this dissertation is on the large-scale quantitative assessments concerning changes and development pressures and their implications for water resources in the major river basins of Monsoon Asia (defined as the area experiencing monsoon climate extending from China to eastern Afghanistan). The results of the dissertation contribute to increasing knowledge on main water-related vulnerabilities and pressures in the studied areas. According to the author's knowledge, these research questions have not been studied to this extent before. Consequently, the appended articles provide an array of new findings from their respective study area. This synthesis focuses on the aspects related to spatial analysis and spatial data, while the actual context-specific findings can be found in the appended articles.

This synthesis has altogether seven sections, including this introductory chapter. The next section introduces the scientific domains forming the disciplinary foundation and summarises the current knowledge related to each domain in terms of this dissertation. Description of focus area of Monsoon Asia and the case studies are presented then in Chapter 3. These are followed by the overview on the methods and data in Chapter 4. Key findings (Chapter 5), discussion with recommendations (Chapter 6) and, finally, conclusions (Chapter 7) form the latter part of this synthesis.

2. BACKGROUND: Introducing the three scientific domains

2.1. Water resources management

Water resources consist of rivers, wetlands, lakes, reservoirs, groundwater aquifers, glaciers, snow covered mountains, marine waters and water in the atmosphere, which are all connected through the hydrological cycle. This cycle is driven by sun energy that causes water to evaporate, transpirate, move, condensate and precipitate, and through interception, infiltration, and percolation water ends up to storage or continues moving through runoff in spatially and temporally varying ways (Costanza et al., 1997; Flügel, 2007). Hydrological cycle is connected to – and strongly influenced by – human activities, land use and climate. Consequently, changes in these factors influence the hydrological cycle (Flügel, 2007; Vörösmarty, Green, Salisbury, & Lammers, 2000; Vörösmarty et al., 2010). Water is the primary medium through which climate change influences the ecosystems and people's livelihoods and wellbeing (Gain, Giupponi, & Renaud, 2012).

Water resources can further be seen as a complex system of individuals, organisations, society and environment (Gain et al., 2012; Simonovic, 2009). Lakes, rivers and wetlands are typical freshwater ecosystems that sustain many social and economic processes. Management of water resources must therefore consider multiple decision criteria and competing demands. These are often accompanied with high uncertainty and complex interactions at various spatial and temporal scales.

Management includes activities such as planning and policy-making, which involve a continuous task of identifying multiple impacts and trade-offs of present-day and future (Loucks, 2000). Political processes of policies, strategies, programs, plans and projects of various scales (WWAP, 2012) and legislation then address these impacts and trade-offs, aiming to solve situations of conflicting interests (Loucks, 2000).

The implementation of management varies depending on the scale, timeframe and scope of management, which Keskinen (2010) classifies in three dimensions, based on the work by Sutherland (1983) and Varis (1996):

operational, tactical and strategic levels. This dissertation concerns mainly the strategic level, leaving day-to-day management decisions and management routines out, and focuses on long-term planning and assessment of water resources.

During last decades, the drive towards integrated management of rivers, lakes, groundwater with the consideration of social, economic and environmental aspects has been strong (GWP-TAC, 2000; Keskinen, 2010; Nilsson, 2003). Integrated Water Resources Management (IWRM) is widely accepted as the way forward for efficient and equitable management of water (Gain et al., 2012; GWP-TAC, 2000; Nilsson, 2003; UNEP, 2012; WSSD, 2002). The Global Water Partnership (GWP) defines IWRM as a *"process, which promotes the coordinated development and management of water, and related resources, in order to maximize the resultant economic and social welfare in an equitable manner without compromising the sustainability of vital ecosystem"* (GWP-TAC, 2000). Thus, management of water consists of more than just concerns over physical water quantity and quality (Nilsson, 2003).

Sustainable management of water resources (i.e. acknowledging and balancing social, economic and environmental aspects) requires information that considers the spatially and temporally varying nature of water resources and their linkages to society and environment (Liu, Zehnder, & Yang, 2009; Perveen & James, 2010; Raskin et al., 1997). Due to the multiple scales, information related to water comes in many forms and from many sources. At the same time, both data collection and analysis capacity are increasing, resulting in growing information (re)sources.

Integrative drive has emphasised the importance of considering and managing water based on the hydrological boundaries (Molle, 2009). The argument for using river basin scale is based on the notion that freshwater is generated, transported and stored within river basins, thus making basin as an appropriate unit for effective research, assessment and management (Alcamo & Henrichs, 2002; Lundqvist, Lohm, & Falkenmark, 1985; Millennium Ecosystem Assessment, 2005; Montgomery, Grant, & Sullivan, 1995; WWAP, 2009). It should be noted, however, that water can also be transported between basins. Besides physical diversion, water can be stored in various products that involve water use in certain basin, and this water along with the products can be transported to another basin. Moreover, Mollinga et al. (2007) argue that it is difficult to build a managing organisation for a river basin, particularly if it is assumed that it should follow the physical boundaries of the basin. Instead, the problems reach beyond the boundaries of watersheds creating a problemshed (Mollinga et al., 2007), or its optimistic counterpart solutionshed (Keskinen, 2010).

Major part of the continental water system consists of rivers. The activities and phenomena occurring in the river basin area strongly influence the water quantity and quality in the river as well as the social, economic and environmental processes linked to the river (Revenga, Murray, Abramovitz, & Hammond, 1998). Thus, Integrated River Basin Management, which is closely related to IWRM, is also recognized as an important strategy for managing waters (Gain et al., 2012; Keskinen, 2010; Nilsson, 2003; Perveen & James, 2010; Revenga et al., 1998; Wolf, Natharius, Danielson, Ward, & Pender, 1999).

The basins consist of socially, economically and environmentally different areas that have different relationships to the river. Such areas include upstream and downstream areas, mainstream, tributaries and headwaters, and rural and urban areas. Population growth, climate change or development projects will influence the areas within a river basin differently. This variation may pose challenges for managing the water resources, especially in transboundary basins that are shared by two or more countries. In these basins, the boundaries of hydrology and administration can differ and overlap. However, large river basins even within one country may include various interests from sub-national administrative units similarly as in international river basins. As a result, such large basins are areas where water-related spatial vulnerability assessments are particularly important.

2.2. Geoinformatics

Spatial assessments have emerged as a response to the need to analyse interlinked drivers of change and challenges regarding complex systems by utilising increasing data resources and spatially-explicit characteristics. The generation, use and computing capacity of spatial information have significantly increased in recent years (Dark & Bram, 2007; del Campo, 2012). The geographical information system (GIS) and related computer-based decision-support systems enhance planning and assessment processes by providing tools for recording, storing, processing, and integrating of various data sources (Burstein & Holsapple, 2008; Chapman & Thornes, 2003; Dark & Bram, 2007; del Campo, 2012; Díez & McIntosh, 2009; Volk, Lautenbach, van Delden, Newham, & Seppelt, 2010).

Spatial data and GIS provide rapid information generation and prediction of potential (cumulative) environmental effects in a graphical form that enables visual and comparative examination (del Campo, 2012; Volk et al., 2010). GIS can be used to draw information on the hydrological, ecological and economic consequences of different management strategies (Volk et al., 2010).

GIS and spatial analysis often have a dual role, meaning both creating and analysing data. For example, these can be used for creating continuous data layers from point data or they can be used for applications which combine many spatial datasets to analyse them statistically (Chapman & Thornes, 2003). GIS is often separated from remote-sensing, even though these are closely related. Remote sensing is about the acquisition of large-scale comprehensive datasets, whereas GIS provides means to display and analyse the data. Remote sensing is the basis for many datasets that are used in GIS, e.g. digital elevation models, precipitation data, and land use.

Within GIS, there are two methods of conceptualising phenomena (Couclelis, 1992; Goodchild, 2011). Discrete object conceptualisation treats objects as discrete countable things that maintain their integrity (e.g. buildings, biological organisms), and these can be represented as points, lines, areas or volumes. Other phenomena are conceptualised as continuous fields (mountains, lakes, temperature). Both conceptualisations can be represented as raster (often regular square grid) or vector (shapefile) data. This spatial data can further be visualised in thematic maps (Slocum, McMaster, Kessler, & Howard, 2013):

- a choropleth map (colour tone showing e.g. population density in river basins)
- a map with proportional symbols (size of symbol scaled according to population in a basin)
- an isopleth or isarithmic map (annual precipitation or elevation model with continuous colour tone)
- a dot density map (dots indicating village spots or damages of natural hazards), or
- a dasymetric map (a compromise between choropleth and isopleth map, e.g. using ancillary data on water bodies to show population distribution more realistically).

The degree of detail is explicit in raster data based on the size of the raster cells, while it is more difficult to define for vector data (Goodchild, 2011). This is a strong argument for the scientific research community to use and provide data preferably in raster format (Goodchild, 2011). All data must be resolution-specific, and thus all transformations and analyses must be resolution-specific as well (Goodchild, 2011). The level of resolution available is influenced e.g. by temporal and financial resources or even by the mechanisms of satellite used to produce the data (Dark & Bram, 2007).

Del Campo (2012) has discussed a number of critical considerations that are crucial for successful application of spatial data and GIS. These include pre-requisites for data: availability, accessibility, uniformity of reference systems, accuracy, scale, consistency, completeness and timeliness. Provision of detailed metadata plays an important role in ensuring these prerequisites (del Campo, 2012; del Campo, Gilmer, Foley, Sweeney, & Fry, 2011).

Scale and spatial unit of analysis

One fundamental decision when using spatial analysis involves defining of the unit of analysis and aggregation scheme, particularly when multiple sources of data with various resolutions are used. In the transformation of data to information, the level of aggregation usually increases (Volk et al., 2010). The results are therefore dependent on the choices concerning the unit of analysis and the aggregation scheme.

Decisions on the unit of analysis relate to the definitions of scale. At least three meanings can be given to scale. First meaning stems from cartography and refers to the representative fraction of a map, i.e. the parameter that defines the scaling of the Earth's surface to a sheet of paper and gives the ratio of distance on the map and the corresponding distance on the ground (Goodchild, 2011). With the use of GIS and spatial assessment, however, two other meanings are more relevant and refer to i) the extent of a study area (primarily in space but also to some degree in time or other dimensions) and to ii) the resolution of digital data (primarily in spatial terms) (Goodchild, 2011; Preston, Yuen, & Westaway, 2011). Scale can be understood also as a measurement dimension including physical, social, temporal and additional (e.g. administrative) criteria (Gain et al., 2012). Scale is often mixed with the term 'level', which refers to a region along a measurement dimension (Evans, Ostrom, & Gibson, 2003). For example, macro level refers in general to large regions (continents or global) on spatial scales and to large-scale phenomena. Also in social sciences level is often used in reference to the unit of analysis (e.g. study using data on household level).

Drawing the boundaries of the unit of analysis is not straightforward. Looking at the earth from space, mainly the boundaries between land and water are clear. With a closer look, the boundaries caused by human impacts can also be visible. In the beginning boundaries are often very arbitrary and artificial, but they can be set to enclose certain characteristics and separate different cultures and different management practices. A nation is inherently a process, but through jurisdiction it gets very physical and crisp boundaries (Coombes & Openshaw, 2001; Raper, 2001). However, boundaries will always be prone to uncertainty, as the ecological boundaries, for example, can never be accurately mapped (del Campo, 2012).

Questions related to scale and the unit of analysis are particularly relevant in water-related spatial assessments. When looking at water issues, scale extends from microbiological to urban, local, subbasin, basin, regional and global (Gain et al., 2012). In global scale, water resources are mapped mostly at political or continental scales (Kulshreshtha, 1998; Meybeck, Kummu, & Dürr, 2013; Rockström et al., 2009; Wada et al., 2011; Vörösmarty et al., 2000). One of the most common approaches is using grid-based maps and discussing the results along national, regional or basin boundaries (Arnell, 2004; Falkenmark, 1997; Falkenmark, Rockström, & Karlberg, 2009; Islam et al., 2007; Kulshreshtha, 1998; Kummu, Ward, de Moel, & Varis, 2010; Perveen & James, 2010; Sullivan et al., 2003; WWAP, 2009; Vörösmarty et al., 2000; Vörösmarty et al., 2010). Some studies utilise multiple scales (Balica, Douben, & Wright, 2009; Perveen & James, 2010; Sullivan & Meigh, 2007). The scale of the assessment influences significantly the recommendations for management decisions. For instance, a study finding large inefficiency in irrigation water use at local scale might recommend different decisions compared with a larger-scale study finding the inefficiency smaller due to return flows (Lebel, Garden, & Imamura, 2005).

It seems clear that there is no single scale or data structure that would be suitable for all purposes, and therefore the ability to disaggregate and aggregate data to various scales and to use various data structures is desirable (Perveen & James, 2010). GIS and spatial analysis can be used to demonstrate a variety of possible units of analysis (zonings) that produce a variety of possible results. Some of these results could be falsifying the results of other zonings (Raper, 2001).

Closely related to this falsifiability is the phenomenon of Modifiable Areal Unit Problem (MAUP). The MAUP has two dimensions, the first one related to the size of the unit of analysis and to aggregation of data into a larger scale (scaling effect), and the second one related to the criteria based on which the unit of analysis is chosen (zoning effect) (Dark & Bram, 2007; O'Sullivan & Unwin, 2010b; Wong, 2009). In more detail, the zoning effect stems from the possibility of dividing the study area in various ways even when the scale or the size of the unit of analysis is staying the same. For example, provinces and hydrological subbasins can be very similar in size but different in location and shape. Article IV provides more details on the MAUP.

2.3. Vulnerability assessment

Vulnerability assessments have the potential to provide important information effectively. Vulnerability has therefore been assessed in numerous studies, but there is no universally accepted definition for the concept of vulnerability (Adger, 2006; Gregory Bankoff, 2001; Greg Bankoff, Frerks, & Hilhorst, 2004; Clark et al., 2000; Cutter, 1996; Eakin & Luers, 2006; Füssel & Klein, 2006; Gain et al., 2012; Ionescu, Klein, Hinkel, Kavi Kumar, & Klein, 2009; Luers, 2005; Pelling, 2003; Preston et al., 2011; Thywissen, 2006; Turner et al., 2003). This is partly due to the fact that vulnerability depends on the aspect and focus, i.e. whether vulnerability is considered in terms of nature, people, ecosystems, biodiversity, individuals, or species. Further, the temporal scale influences, as it matters greatly whether we speak about a time span of one year or of hundred years.

In the 1970s, vulnerability was examined in the context of disasters (Adger, 2006; O'Keefe, Westgate, & Wisner, 1976) and through the entitlement approach (Article I; Sen, 1976). Since then vulnerability has emerged in many fields of research, for example concerning resource management, social change, urbanisation and climate change (Adger, 2006).

Contemporary vulnerability assessments utilise various approaches, methods and schools of thought. Debates have thus risen, for example, between objectivist and subjectivist schools of thought, between structuralist and individualistic approaches, between physical scientists and social scientists, and between engineering and human dimension focused approaches (Kasperson et al., 1988). During recent decades, interactive and integrative approaches have been promoted, moving away from the polarisation by debates described above.

At least four schools of thought have influenced the definitions of vulnerability: disaster risk reduction community, climate change adaptation community, the sustainability science community, and global environmental change community (Gain et al., 2012) (Figure 1). The legacy of human geography and human ecology can be seen in the latter two (Adger, 2006). Also the political economy of resource use influences greatly to the views in the environmental change community (Adger, 2006).

Disaster risk reduction community defines vulnerability as *"the conditions determined by physical, social, economic and environmental factors or processes, which increase the susceptibility of a community to the impacts of hazards"* (International Strategy for Disaster Reduction, 2004). Intergovernmental Panel on Climate Change (IPCC) on the other hand, defines vulnerability as *"the degree to which a system is susceptible to, or unable to cope with, adverse effects of climate change, including climate variability and extremes. Vulnerability is a function of the character, magnitude, and rate of climate variation to which a system is exposed, its sensitivity and its adaptive capacity"* (IPCC, 2001). This is very similar to the definition by White (1974) within the sustainability science: *"vulnerability is the degree to which a system, subsystem, or system component is likely to experience harm due to exposure to a hazard, either a perturbation of stress/stressor"* (White, 1974, cited in Preston et al., 2011). The global environmental change and sustainability science communities have introduced the notion of coupled socio-ecological system or human-environment system when conceptualising vulnerability (Gain et al., 2012; Turner et al., 2003).

Despite differences in approaches and definitions, vulnerability is commonly seen to be constituted by components representing the system's exposure and sensitivity to perturbations or external stresses as well as system's capacity to adapt or in other words its resilience (Adger, 2006; IPCC, 2001). Exposure is the nature and degree to which a system experiences the stress, sensitivity is the degree to which a system is modified or affected. Adaptive capacity, or sufficient resilience, means having the capacity to face stressors without significant losses, and to change, learn and evolve in order to accommodate hazards or policy changes and to increase the coping capacity of the system (Folke et al., 2002). The vulnerability components are also defined by the magnitude, frequency, duration and areal extent of stressors (Adger, 2006; Burton, Kates, White, 1993). Vulnerability assessments are often based on the use of proxy indicators representing these components.

Vulnerability factors can be presented in two broad classes of biophysical and socio-economic determinants (Preston et al., 2011). Further, various models (listed in Article I) are used for vulnerability assessments, as vulnerability itself is something that cannot be measured (Adger, 2006; O'Brien, Eriksen, Schjolden, & Nygaard, 2004). Models define how the quantitative indicators representing the determinants are thought to interact and influence the level of vulnerability (Article I; Füssel, 2007; Preston et al., 2011; Turner et al., 2003). In addition to the listed models, simple aggregation and weighed aggregation (based on statistical methods or expert opinions) of indicators to a composite index are used in vulnerability assessments (Article I; Hinkel, 2011; Kappes, Papathoma-Kohle, & Keiler, 2012; Lebel, Nikitina, Pahl-Wostl, & Knieper, 2013; Li, Wang, & Liang, 2006; Wang, Liu, & Yang, 2008).

Practically all vulnerability assessments aim to identify the most vulnerable part of the system. However, the higher-level objective of assessments can differ. For example, assessments can be diagnostic and problem-oriented or decision-support oriented. The first one indicates the need for targeting research or resources and the latter aims for policy intervention (Abson, Dougill, & Stringer, 2012; Adger, 2006; Metzger & Schröter, 2006; Preston et al., 2011; Stelzenmüller, Ellis, & Rogers, 2010). In their current form, vulnerability indicators are appropriate for identification of vulnerable people, communities and regions at local scales (Hinkel, 2011). Also Turner (2003) emphasised the questions of who are vulnerable to multiple environmental changes and where are they located, leading the direction to spatial approaches.

2.4. Putting the three together: Water-related spatial vulnerability assessments

In relation to water resources, the term vulnerability was used in general policy debate for the first time in the 1992 international conference of Dublin (The Dublin Statement, 1992). Water-related vulnerability concerns generally the quantity and/or the quality of water, which is influenced by the temporal and spatial scale of analysis. The description or model for vulnerability varies depending on what is in the focus, e.g. people or biodiversity (eco-environment).

Martin Beniston (2002) describes the vulnerability of water resources:

> *"Vulnerability assessments of water resources integrate multiple interacting stresses and feedbacks from climate change, climate variability, and direct effects of human activities, such as changes in land and water use, changes in distribution and age-profiles of populations, changes in economic activities and settlement patterns and changes in political environments. There is a strong human dimensions component in vulnerability assessments. Exposure to environmental stresses and reaction to such exposure are major factors, but also very important is the potential for adaptation, which is essentially region-specific. … A goal of integrated water resource management is to reduce vulnerabilities."*

This description is found to represent well the approaches in the appended articles of this dissertation. While the indirect effects of human activities are important for the vulnerability of water resources, the appended articles have put more emphasis on direct human activities. In some cases, the inclusion of all direct human activities is challenging, as data is not readily available.

In terms of quantity, vulnerability may stem from the imbalanced supply and use of water or changes in water availability due to changing climate. Further, when the focus is on people, vulnerability can also be thought to concern sea-water intrusion, levee failures, floods or droughts, groundwater contamination, or water supply system failures. More specifically, a certain group of people might be vulnerable due to their social status being poor, elder, children or disabled or strongly dependent on water-dependent livelihoods. In terms of water quality, vulnerability is caused by many different factors including land-use change, population growth, changes in the wetland status and protection, changes in agricultural fertilizer use, and changes in the sediment load. The conflicts often stem from a combination of challenges described above (Yoffe et al., 2004).

Approaching water-related vulnerability through components of exposure, sensitivity and adaptive capacity is strongly case-specific. The exposure can be measured e.g. as extent of changes in hydrology or changes in land-use. The degree of the sensitivity of a user group depends on the significance water has for the group. Further, the adaptive capacity of a group may be enhanced by a range from alternative livelihoods sources and education to compensation mechanisms and access to irrigation.

The role of spatial heterogeneity in physical, socio-economic, and cultural determination of vulnerability is important (Turner et al., 2003). Climate change, for instance, is believed to change the spatial distribution of hazards, creating a need for spatial and localised information (Preston et al., 2011).

Indeed, spatial specificity is crucial in water resources management as the spatial (and temporal) distribution of water varies. Location is important also from the upstream-downstream point of view. In addition, the mitigation, adaptive capacity and exacerbation vary spatially (Cutter, 2003), and planning and management involve reconciling competing demands that depend on the location.

The mapping approach is increasingly used in vulnerability studies, firstly due to its ability to account for spatial heterogeneity and secondly due to the possibilities for an effective communication of the status of complex system and of interaction of determinants with the system components (Preston et al., 2011). Maps are believed to act as powerful visual tools (Abson et al., 2012).

3. CASES FROM THE MAJOR ASIAN RIVER BASINS

3.1. Study context: Monsoon Asia and its major river basins

The oriental biogeographic region of tropical Asia, Monsoon Asia, is in this dissertation defined to consist of the area reaching from China to eastern Afghanistan. It includes ten major river basins and the nations located within those basins in mainland Asia (Figure 3).

In Monsoon Asia there are five large river basins of size over 500×10^3 km^2: Indus, Ganges-Brahmaputra-Meghna, Mekong, Yangtze, Huang He (Yellow) River, and five intermediate river basins of size larger than 100×10^3 km^2: Irrawaddy, Salween, Chao Phraya, Hong (Red) River and Xun Jiang (Pearl) (Kummu, 2008). Seven of the river basins are transboundary, located within two or more nations. Two of the case studies cover the whole study area of Monsoon Asia (Article I & IV) while the other two cases focus either on certain little-studied but major river basins (Irrawaddy & Salween in Article II) or on an important subbasin part of a major river (Tonle Sap from the Mekong in Article III).

Monsoon Asia is an important area as significant quantities of the world’s population and food production is located there (Dudgeon, 2005; Sanyal & Lu, 2004). Many parts of the area have rich biodiversity, which is endangered by various pressures from development and environmental change (Dudgeon, 2005). The complexity of water resources management stems from the need to consider particularly the following drivers of change: population growth, demographic changes, economic development and related need for energy (including hydropower development), changes in food demand and production, and changes in political regime (Biswas, 2005; Kummu, Lu, Wang, & Varis, 2010; Kummu, Ward, et al., 2010; WWAP, 2009, 2012; Vörösmarty et al., 2000; Vörösmarty et al., 2010). The climate (Hurd, Leary, Jones, & Smith, 1999) causes the area to experience high variation and extremes, which are seen in many of the river basins as strong seasonality in water availability (Sanyal & Lu, 2004). Monsoon Asia is further plagued by poverty and the socio-economic challenges brought by it (The World Bank, 2008).

The above-mentioned drivers cause changes in water and land use. Related benefits and losses are not equally distributed in space or time. Particularly, the hydropower development in transboundary river basins is a key challenge having a wide impact on the environment (Dudgeon, 2005) and societies (Keskinen, 2006; Keskinen et al., 2013; Lamberts, 2008).

This context has been the starting point for each of the case studies presented below. However, the issues under closer examination vary depending on the case study. Table 1 provides an overview and summary of the key characteristics of the appended articles.

Table 1. Summary of the appended articles in terms of the general objective of the study, the scale and spatial unit of analysis as well as the approach to vulnerability analysis.

	Article I:	**Article II:**	**Article III:**	**Article IV:**
General approach	Explorative and integrative mapping of socioeconomic-environmental vulnerability status	Explorative mapping of multiple environmental indicators on land use and human activity	Examination of change in main economic sectors, particularly water-related, and changes in demography	Illustration and descriptive statistics of MAUP in water shortage
Scale	Regional	River basins	Subbasin	Regional
Spatial unit of analysis	River basin	Topography & climate-ecological zone	Topography, urbanity & administrative boundaries	Multiple zonings: administrative, hydrological, groundwater basin, climate, & agro-ecological zoning
Vulnera-bility of what/ to what?	Vulnerability of basins to water stress, malnutrition, poverty, hazards, fragile governance, poor environmental system, human activities	Vulnerability of river health and ecological state of wetlands, water quality to pressures from human activities	Vulnerability of water-related livelihoods to changes in hydrology and ecology caused by hydropower development and climate change	Vulnerability of people to water shortage depending on reduced water availability and/or growing population
Assessment approach	Aggregation of averaged and weighed (DPSIR-approach) spatial indicators and population-weighed national indicators for river basins & Comparing basins in terms of vulnerability index and individual indicators	Principal Component Analysis (PCA) applied for indicators, individual PCs examined and combined as composite index & Comparing basins and the sub-areas within the basins with composite index scores	Calculating indicators for two time steps of the Census data, both quantities and proportions. Clustering villages based on several indicators (partly combined with PCA) & Comparing changes and status of sub-areas. Testing aggregation scheme's relevance with *k*-means clustering and Levene's test	Calculating water shortage, and population under high water shortage with multiple zonings, descriptive statistics (average values and coefficient of variation) & Comparing differences in water shortage caused by the zoning scheme visually and with descriptive statistics

3.2. The socioeconomic-environmental vulnerability profiles of the ten major river basins (Article I)

Article I analysed ten major river basins of Monsoon Asia (Figure 3) by considering and identifying the major vulnerabilities in the context of the multiple and interlinked change processes.

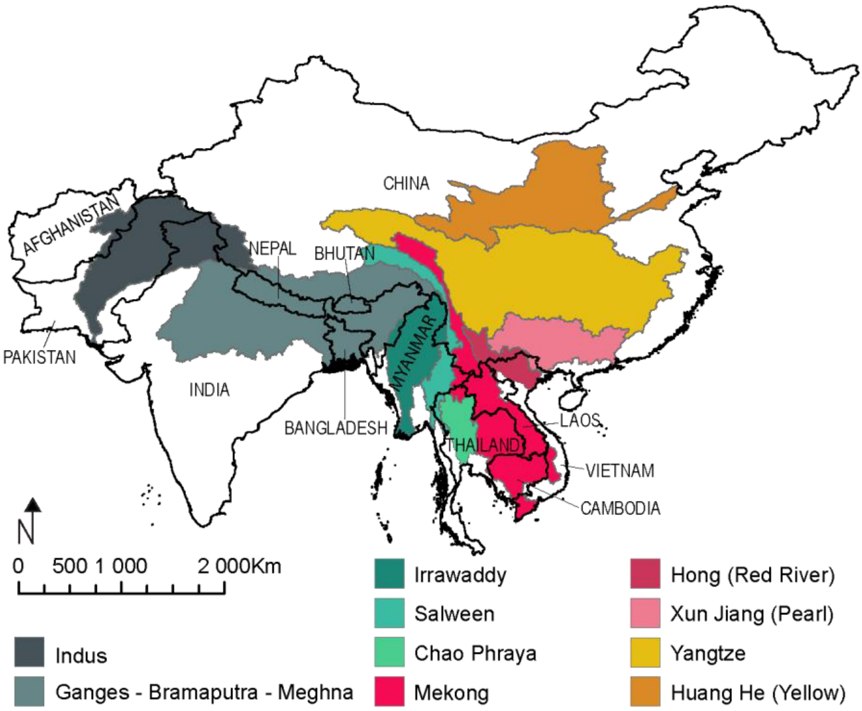


Figure 3. Map of the study area in Article I consisting of 12 countries and 10 major river basins.

National statistical data and globally available spatial datasets on various indicators were spatially analysed with ArcGIS 9.3 software (ESRI, Redlands, CA, USA). Spatial data was used for producing a socioeconomic-environmental vulnerability profile, which included spatial data on water stress, human footprint, multihazards, malnutrition and infant mortality. Data from national and global statistics on Gross National Income per capita, environmental systems, slum population, literacy rate, corruption, political instability, state fragility and multidimensional poverty index were also used. National level data from statistics were weighed with the proportion of national population within the basin. In addition, population density, precipitation, runoff and future population scenarios were examined.

Two approaches were used to build a composite vulnerability index: a simple aggregation of indicators and weighing indicators with a DPSIR method (Drivers – Pressures – States – Impacts – Responses model: OECD, 1993; Rapport & Friend, 1979; United Nations, 2001). The distribution of individual indicators was examined visually in addition to exploring the composite index scores of the basins.

3.3. Utilising global datasets for data-poor areas in Irrawaddy & Salween River basins (Article II)

Representing areas over which less scientific studies are available, the Irrawaddy and Salween River basins are prone to many changes and challenges identified in the region. These include, for example, land use change, hydropower development, population growth, and changes in the political regime.

Article II utilised globally available spatial datasets to create environmental profiles that identify key issues concerning population and land use distribution for Irrawaddy and Salween River Basins (Figure 4).

Gridded data on land cover was examined, particularly the extent and protection status of wetlands, distribution of night-time lights as urban area indicator and extent of croplands. Indicators were built also based on population distribution, estimates for nitrogen load from industrial fertilizer use, potential mobilizable nitrogen load and extent of potential rainfed cropland. Table 1 in Article II describes these indicators in detail.

Indicators were analysed through specially defined spatial units of analysis that was based on a combination of slope and ecological zone. The map algebra of ArcGIS 9.3 software was used to calculate and classify the slope from Digital Elevation Model and the re-classified layer was overlaid with ecological and river basin datasets to create the spatial units of analysis (sub-areas). Then zonal analysis was used to extract indicator values for each sub-area.

Principal Component Analysis (PCA) (Jolliffe, 2002) was conducted to reduce the multicollinearity of indicators in PASW Statistic 18 software (former SPSS Inc, Chigaco, IL, USA, current IBM, Armonk, NY, USA). Individual principal components were analysed and interpreted separately. These were combined to a composite index and mapped to indicate the sub-areas under most pressure.

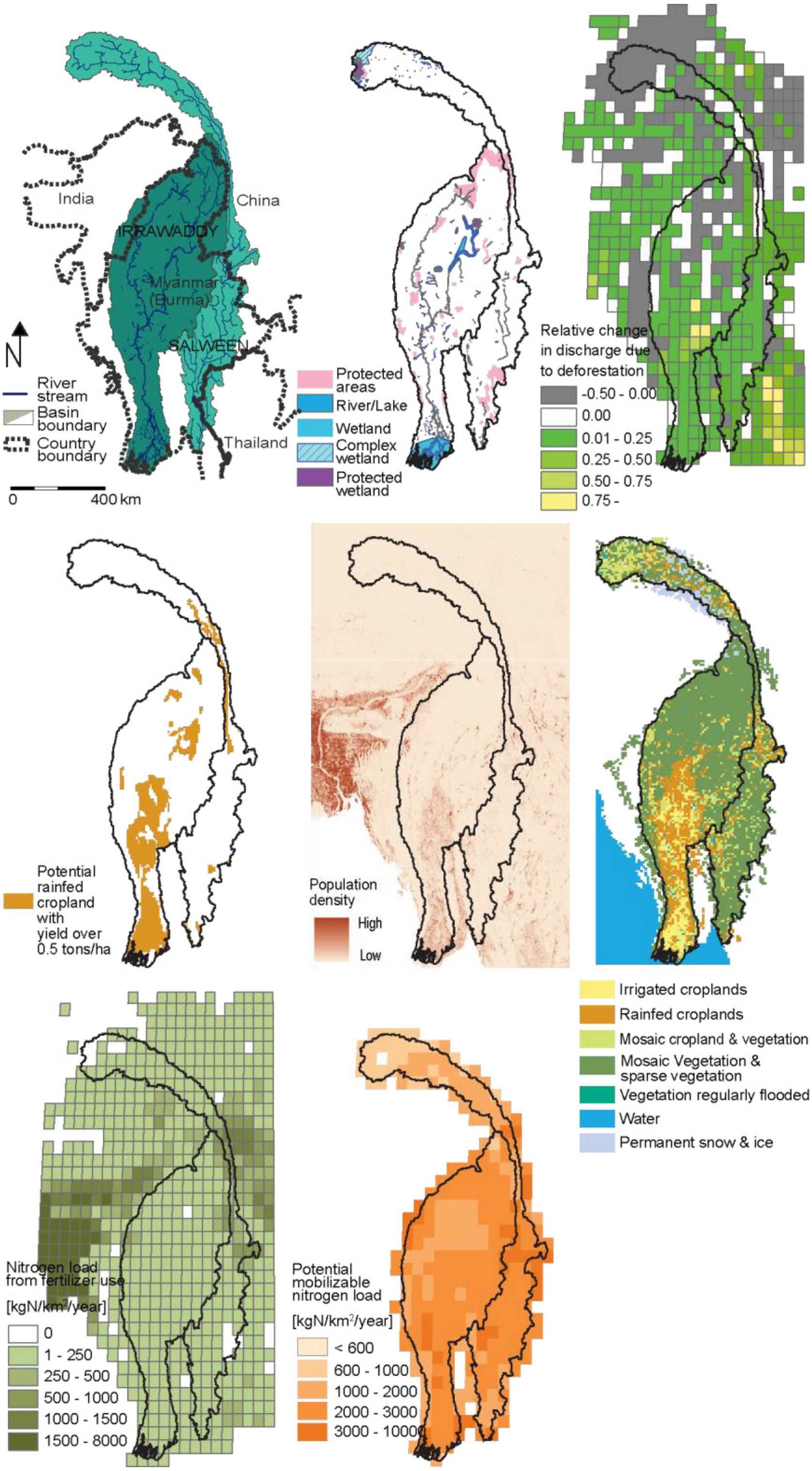


Figure 4. The Irrawaddy & Salween River basins and the global spatial datasets utilised in Article II (modified from Article II).

3.4. Socio-economic changes in Tonle Sap Lake area with a spatial approach (Article III)

The Tonle Sap Lake is a major part of the Mekong River system (Figure 5). The Tonle Sap Lake area is likely to be influenced by the development and changes occurring within the Mekong basin, as Tonle Sap Lake is receiving great quantities of the monsoon season flooding water and sediments of the Mekong River.

Due to the pulsating system, the socioeconomic-environmental system of the lake area is adapted to five-fold change in the lake surface area and to water level changes from one to ten meters. The most common livelihoods, i.e. agriculture and fishing, are strongly dependent on water.

The development pressures together with climate change pose challenges for maintaining the natural state of this river system, which is one of the world's richest areas in terms of biodiversity and fisheries. The area is also strongly influenced by many other changes occurring in the Cambodian society.

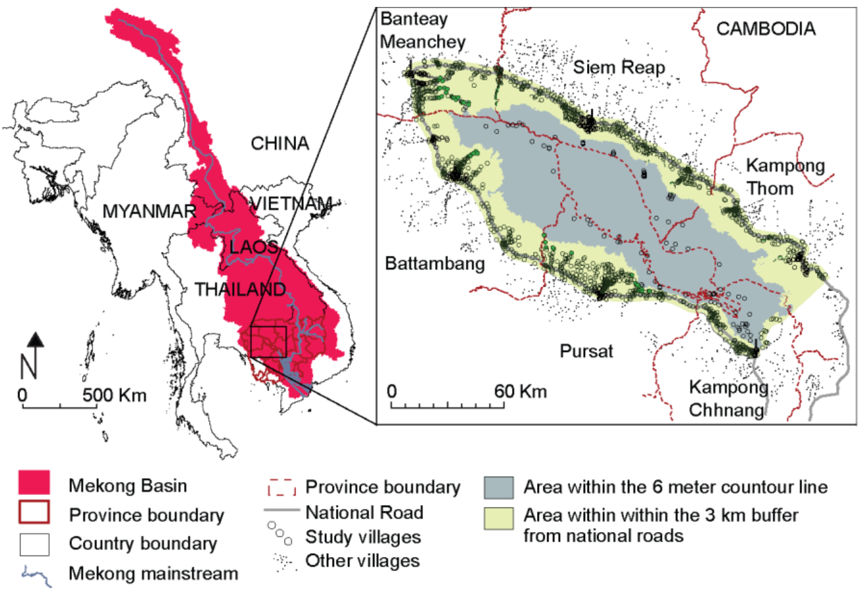


Figure 5. Tonle Sap Lake area (modified from Article III).

Article III looked at the two census datasets and explored how the livelihoods and demographics have changed between 1998 and 2008 (the year of the census data) in the 1555 villages of the study area.

Data from the national statistics was prepared to be comparable between these two years. Data was then spatially aggregated to the specially defined spatial units of analysis that considered topography (in this case meaning the proximity to the lake and flooding), census information on the urbanity of villages and the provincial boundaries. The method was developed further based on earlier studies (described e.g. in Keskinen, 2006; Keskinen et al., 2013). In addition, the robustness of the developed spatial approach was tested with data-driven *k*-means clustering method (e.g. Wang, Huo, Huang, Xu, Sun, Li, 2010) in Matlab R2014a software (MathWorks, Natick, MA, USA).

Spatial analysis was conducted in ArcGIS 10.2 software. For *k*-means clustering, part of indicators were combined with the PCA (Jolliffe, 2002) in IBM SPSS Statistics v20 software to represent an 'urban' indicator.

The status and change of demography and of economic sectors were mapped in a visually effective way to identify areas of stagnation or rapid change. The sub-area division enabled creating a link between the socioeconomics and hydrology, which further enabled discussing the potential impacts and role of future changes in hydrology.

3.5. Exploring water shortage and the Modifiable Areal Unit Problem (MAUP) in Monsoon Asia (Article IV)

Article IV presents a regional scale study with focus on one indicator, water shortage (defined as available water divided by population). Water shortage is one of the key vulnerabilities in the region due to potentially changing water availability caused by growing population and climate change.

The study calculated water shortage based on available global spatial datasets using altogether 21 different criteria for defining the spatial unit of analysis (zonings) (Figure 6). These included delineating the boundaries of the unit of analysis according to nations, river basins, agro-ecological zones, climate zones, groundwater basins, provinces, subbasins and various combinations of these.

Data on average annual surface and sub-surface runoff, surface areas of the grid cells and population were used for calculating water shortage (Falkenmark & Widstrand, 1992) in Matlab R2014a software. Water shortage indicator values were classified to three classes of "no" (indicator value >1700 m^3/cap/year), "moderate" (1000 – 1700 m^3/cap/year) and "high" (<1000 m^3/cap/year) water shortage.

Population under high water shortage was calculated and the differences depending on the zonings were tabulated. The average water shortage, number of zonings under which a grid cell fell under high water shortage and the coefficient of variation for water shortage per grid cell were calculated and mapped to indicate areas that were particularly sensitive for the MAUP.

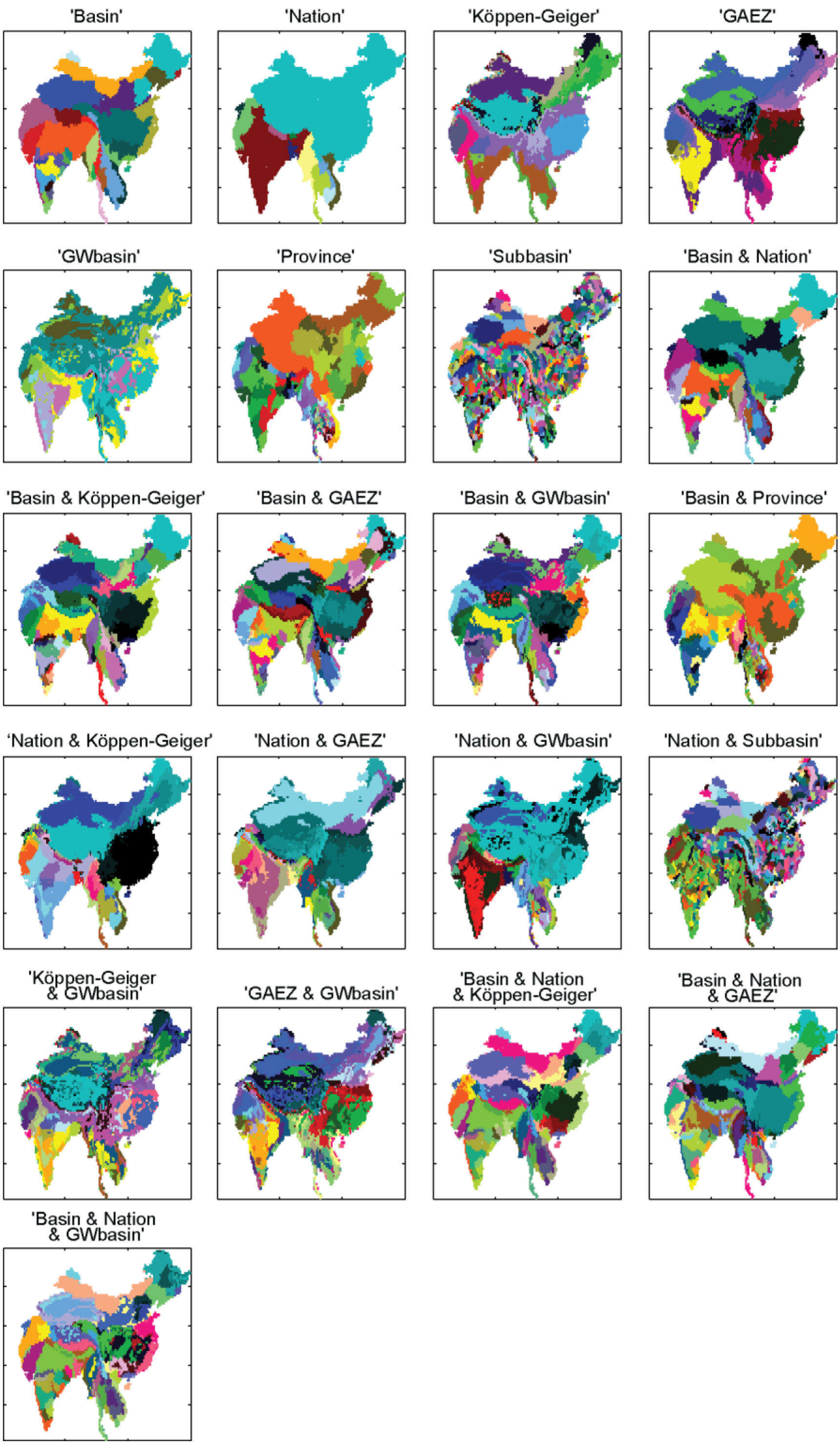


Figure 6. The 21 zonings for calculating the water shortage in Monsoon Asia (modified from Article IV).

4. METHODS AND DATA

4.1. Methods

The overarching methodology used in this dissertation is spatial analysis, which is a broad term identified at least in four different areas of literature (O'Sullivan & Unwin, 2010a). These concern i) spatial data manipulation, ii) spatial data analysis, iii) spatial statistical analysis, and iv) spatial modelling. Most research involves many, if not all, of these. Methodologically, this dissertation is situated between spatial data analysis and spatial statistical analysis, with emphasis on carrying a descriptive and exploratory approach. As such, it reports important first steps for spatial analysis with big and open datasets (O'Sullivan & Unwin, 2010a). Spatial modelling is also considered, as some of the data examined are outputs of modelling approaches.

This dissertation generally applies overlaying, map algebra and zonal analysis of pre-processed spatial datasets, and further conducts spatial analysis of other spatially-referenced data (e.g. attribute data for spatially-referenced objects). Statistical methods, such as Principal Component Analysis (PCA) (e.g. Jolliffe, 2002) and *k*-means clustering (e.g. Wang et al., 2010; Larose, 2005), are also utilised.

The methods are here presented in the context of the first three research questions, which then provide input for the fourth research question. The contribution of articles and related methods used for each of the three research questions are presented in Figure 7.

It must be noted that the appended articles include other methods as well. Article III is part of a larger research project, for which spatial analysis provides only one assessment method. In addition, two articles form a series of studies with other publications, Article I with Kattelus, Kummu, Keskinen, Salmivaara & Varis (2015), Varis & Kummu (2012) and Varis, Kummu, Lehr, & Shen (2014) and Article II with Salmivaara (2009; 2012). Each article was briefly described in Chapter 3 and summarised in Table 1. More details of studies and methods used are available in the appended articles.

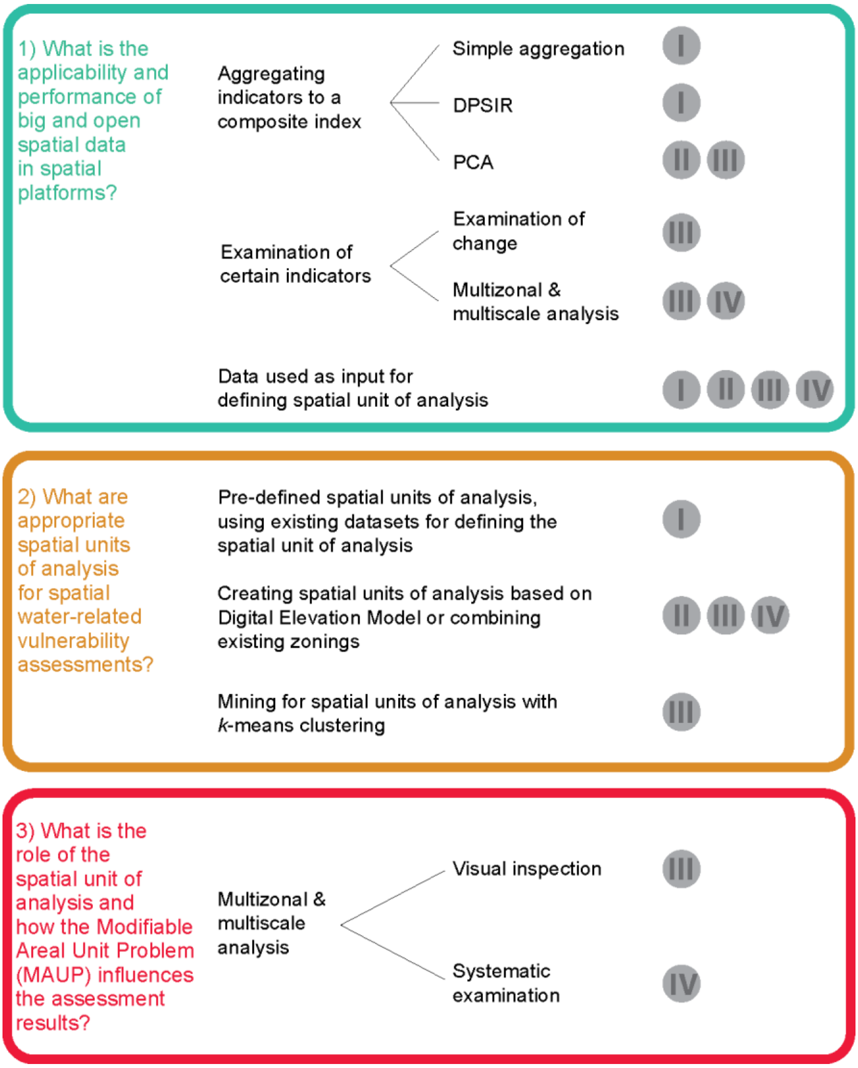


Figure 7. Summary of methods used in Article I-IV to study the three research questions.

To seek an answer to the **first research question**, various sources of big and open data are used (described in Chapter 4.2.). This data is utilised as indicators to build a composite index with simple aggregation (Article I) and with weighed aggregation (Article I and II). Article I applies the DPSIR method (Driver – Pressures – States – Impact – Responses model) (Table 6 in Article I; OECD, 1993; Rapport & Friend, 1979; United Nations, 2001) for weighing, while Article II utilises PCA method, which reduces the collinearity of the data (Jolliffe, 2002). Risks related to collinearity are discussed in Article II. In Article III, a set of indicators are combined with PCA to one "urban" indicator.

Besides aggregation to a composite index, other approaches are also used. In Article III, the national census data is prepared to be comparable over the two available points in time, 1998 and 2008. The change in the demographic and socioeconomic indicators of the census, both in quantities and proportions, is explored using spatial units of analysis allowing exploration on two scales. Article IV explores only one indicator but with multiple spatial units of analysis of various scales.

Third aspect in assessing the applicability and performance of big and open spatial datasets is using those as input to creating various criteria for defining the spatial unit of analysis (i.e. zoning schemes). These further contribute to finding answers to the second and third research question on the spatial unit of analysis.

Related to the **second research question**, the criteria for finding appropriate unit of analysis can be based in theory, i.e. defined *a priori*, inductively, or the unit of analysis can be defined by the system conditions at hand, in which case they are defined by the data, deductively. For example, based on the theories of hydrological cycle (Shaw, 1994), climate has a major role in the generation and distribution of the water resources. Characteristics of climate can, therefore, be used as an argument for looking at variance of water-related indicators across climate zones. On the other hand, the data can be given a 'chance to speak' when statistical data mining methods are applied.

Purely *a priori* defined criteria is used in Article I, in which data is aggregated according to river basins boundaries. More special and case-specific, but still theory-based, criteria are applied in other articles. Article II, for example, combined slope calculated from Digital Elevation Model and ecological zone for distinguishing especially the delta areas. Article IV uses various readily available zonings including nations, river basins, agro-ecological zones, climate zones, groundwater basins, provinces, and subbasins, while also creating combinations of these to find appropriate spatial units of analysis. Thus, the spatial units of analysis are either available as such (basin boundaries in Article I) or they are created by applying calculations (e.g. on slope) and combining available datasets (Article II-IV).

A data-driven method of *k*-means clustering (considering notions related to the method by Wang et al., 2010; Rousseeuw, 1987) is applied in Article III to organise the census data into groups with maximum variance between groups and minimum variance within a group. This data-driven result is then compared with the spatial approach (using topography and census classification on urbanity) to provide information on the robustness of the spatially-defined unit of analysis. However, it also provides information on the applicability of data-driven methods for finding appropriate spatial units of analysis for water-related vulnerability assessments.

Levene's test (Levene, 1960 in Gastwirth, Gel, & Miao, 2009) is applied to define cases where the differences between the *k*-means clusters (data-defined) and the spatially defined zones are statistically significant. The test considers the similarity of variances of datasets in comparison and not only the means or centres. Comparing the variances of two datasets is more informative than comparing their means, as two datasets can have similar mean with very different variances (i.e. values are all close to the mean or values are spread on a wide scale). However, when variances are similar, the datasets have somewhat similar structure. The test is also applicable when data does not follow normal distribution or the sample sizes are unequal (Brown & Forsythe, 1974 in Gastwirth et al., 2009).

The role of the spatial unit of analysis and the related phenomenon of the Modifiable Areal Unit Problem (MAUP), i.e. the subject of the **third research question**, is systematically explored through using multizonal and multiscale approach in Article IV. The variance in the indicator values (caused by the MAUP) is examined by calculating the coefficient of variation, the average water shortage value and occurrence of high water shortage for each grid cell over the 21 different zonings.

Article III contributes also to this research question, as the analysis is to some extent conducted as a multiscale assessment. This enables examining the role of the spatial unit of analysis for results through a visual comparison of maps produced with spatial units of analysis varying in scale.

Furthermore, the appended articles as a combination can be considered as a multiscale and multizonal approach, which provides a method for examining the role of the spatial unit of analysis.

4.2. Data

The availability of data is increasing all the time. Many institutions provide spatial datasets in their digital libraries or data portals, where the user can find data in terms of the area or the theme of interest. For example, various hydrological characteristics are provided by data libraries (e.g. GEO Data Portal, 2014; GWSP Digital Water Atlas, 2014). The increasing availability of globally consistent datasets that are relatively cheap to produce, enables application of such data in research and management of natural resources. Thus, reporting the application of multisource big and open spatial data contributes to the future development of water-related spatial assessments.

The datasets used in the appended articles represent only a fraction of what is currently available. The general types of spatial data are listed in Table 2 with reference to datasets used in the appended articles and general examples of datasets. General data types include digitized maps, data produced with remote sensing technology, data layers produced with spatial interpolation of point data, modelled data based on physical or mathematical models and spatially referenced statistics i.e. census data. Additionally, 'mixed data types' are combinations of the general types of data or involve qualitative measures such as expert estimations or classifications.

It should be noted that remote-sensing data products are increasingly available covering various different components of the hydrological cycle. However, the majority of data products available are mixed data products. While remote-sensed data are increasing, there will be continuous value to mixed data products for assessments aiming at combining multiple and not only measurement-based aspects.

Another thing to note is that while there are potential to provide temporally very detailed data, particularly with remote-sensed data, this is not yet the case with the data available for end-users (Table 2). Data is aggregated and also the calculation of more specific indices requires time and reduces the possibilities for data producers to provide this data. Furthermore, until recently, the demand has not perhaps been there. The spatial scope and scale of the data increasingly cover the whole globe i.e. providing big data. However, many other characteristics usually associated with the term big data, are not fully met with the contemporary data available.

There are issues to consider, particularly concerning the mixed data products. For example, the reason for listing 'river basin modelled data' under the mixed data section is that often the datasets available are aggregations of daily or monthly outputs of models. These can be provided in shapefile format instead of the original model output as raster data, which is relevant to consider when using these datasets for further analyses.

In the case of population models, it should be noted that there are different ways to produce the data: for example LANDSCAN (2008) uses dasymetric modelling with the imagery analysis technologies of a satellite picture to disaggregate census counts, while datasets by GPWfe (2005), GRUMP (2004) and GWSP Digital Water Atlas (2008a, 2008b, 2008c) distribute country level urban and rural population statistics according to night-time lights and points of populated places. These methods result in different distributions, particularly in areas with limited access to electricity.

Finally, the datasets that are produced as combinations of layers and involve special classifications (e.g. climate zones or potential rain-fed cropland in Article II) require careful consideration when they are used in vulnerability assessments. The assumptions behind the classifications need to be in line with the assumptions behind the vulnerability assessment. When these are in line, such datasets are useful as they reduce the workload involved in an assessment process.

Table 2. The main spatial data types. Datasets with bold font were used in the appended articles (I-IV). The table includes also examples of datasets that were not utilised by this dissertation. Source information and acronyms are available in the appended articles. The big data characteristics in terms of spatial scope and scale as well as temporal scale are indicated for the data used [spatial: **G**=global, **R**=regional, **L**=local; temporal: **De**=decadal, **Y**=yearly, **Da**=daily].

Data type:	Description and examples	
Digitized maps	Compilation of maps, aerial photographs, decisions concerning areas	**Lakes and Wetlands (II) [G;De], Protected areas (II) [G;De], Administrative areas (I,II,III,IV) [G;Y]**
Remote-sensing	Satellite sensor of various wavelengths: visual light; infra-red; microwave	**Land cover data (II) [G;Y/Da], Night-time lights (II) [G;Y/Da],** groundwater (GRACE), precipitation (TRMM), evapotranspiration (Landsat, ASTER, MODIS), Soil moisture (TRMM, METOP), water level (TOPEX POSEIDON, ENVISAT)
Interpolation of observations	Interpolating point data into continuous layer	**Digital Elevation Model (DEM) (II,III) [G;Y],** precipitation, temperature
Modelled data	Mathematical model based on physical equations	**Potential mobilizable nitrogen load (II) [G;De],** river discharge
Census data	Usually household level detail, village level aggregate, based on questionnaires covering in theory all households	**Cambodian Census (III) [L;De]**
Mixed data products	National level data gridded	**Nitrogen load from industrial fertilizer use (II) [G;De],** GDP per grid
	River basin (or other system based) modelled and complemented with other ancillary or modelled data then gridded or provided as shapefile	**Water stress (I) [G;De], Average annual surface runoff (I,IV) [G;De], Relative change in discharge due to deforestation (II) [G;De], River basin boundaries (I,II,IV) [G;De],** Annual river discharge
	Population models	**Gridded population (I,II, IV) [G;De/Y]**
	Expert estimates and calculations + compilations of various sources	**Political stability (I) [G;Y], Corruption Perception Index (I) [G;Y], Environmental System component from Environmental Sustainability Index (I) [G;Y], Gross National Income per capita (I) [G;Y], Slum population (I) [G;Y], Literacy rate (I) [G;Y], State fragility (I) [G;Y], Multidimensional Poverty Index (I) [G;Y]**
	Datasets defined and classified by multiple layers and criteria	**Potential rainfed cropland (II) [G;De], Köppen-Geiger Climate Zones (IV) [G;De], Global ecological zones (II) [G;De], Global Agro-Ecological Zones (IV) [G;De], Groundwater resources with annual recharge rates (IV) [G;De],** Water footprint data, Water scarcity

5. FINDINGS

5.1. Applicability and performance of big and open spatial data

The key finding based on the appended articles was that big and open spatial datasets are crucial, perform well and enable large-scale comparative studies concerning major river basins. Most of the assessments could not have been conducted in same extent and efficiency if big and open spatial datasets were not available. Another key finding was the preference for using simple aggregation of the indicators when using multiple spatial datasets, providing more transparent approach compared to using more complex weighing methods (such as the DPSIR method in Article I). Transparency is important, as the available data did not cover all the data that could have been relevant. The applicability and performance of big and open spatial data are elaborated according to each article in detail below.

Geographical information system (GIS) enables consideration of multiple data layers on the same platform. Through the availability of various data, it is possible to go beyond sectoral assessments, as spatial analysis enables a transparent method that puts emphasis on visual inspection of the components used. Thus, taking a multisectoral view, 'looking outside the water box', is possible through the use of spatial analysis. Such a view was taken in Article I, resulting in visually effective profiles for the ten major river basins of Monsoon Asia.

According to Article I, the Ganges-Brahmaputra-Meghna basin and the Indus basin scored the highest in the overall vulnerability assessment. Other basins' vulnerability scores were on the same (lower) level but their profiles differed. For example, the Irrawaddy and Salween basins were vulnerable in terms of hazards and economic factors. The Mekong basin also had the economic factors as cause for vulnerability. The Indus and Yellow River basins were the only ones experiencing water scarcity. More detailed results are presented in Article I.

In Article II, big and open spatial datasets enabled the study of the Irrawaddy and Salween River Basins, which form one of the least studied regions in Monsoon Asia. Article II included a large number of indicators, and Principal Component Analysis (PCA) was applied to avoid multicollinearity and aggregating indicators that are highly correlated, which may occur when using simple aggregation of indicators. PCA reduced the number of indicators to three in the case of Irrawaddy (representing 74% of variation in the original dataset) and to four in the case of Salween (representing 84% of the original variation). Individual PCs were interpreted, and for example, one PC was found to represent the potential for land use change. Irrawaddy delta and Salween river mouth areas scored the highest with most of the PCs and also with the composite index combined from the PCs.

Article II provides support for preferring the simplicity, as the results from PCA on the most vulnerable areas are very similar to results of a simple aggregation of indicators (based on partly the same indicators, presented in Salmivaara, 2012). The PCA generally performs well, when there are multiple indicators that can be interpreted as representing the same phenomenon (e.g. see below for urbanity in Article III). However, the big and open datasets used in water-related spatial vulnerability assessments can be too heterogeneous, as in Article II. In such cases, the complexity and difficulty of interpretation increases to that extent that simple aggregation is more informative and transparent.

In Article III, the big and open data of Cambodian Census was examined according to spatial distribution of the villages. Article III found that there is great potential when more emphasis is put on the spatial dimensions of this socioeconomic data. Using topography together with census data, the livelihood profiles were possible to be grouped in an effective way. This enabled the discussion with more linkages to the hydrology of the Tonle Sap Lake, and addressing the spatial heterogeneity in such a way that would have not been possible if analysed according to provincial or national level statistics. The examination of change between the two census datasets, both in quantities and proportions, provided important information on the areas of change and stagnation.

Article III further found that the main livelihood sectors are fishing and agriculture, which are clearly located in rather distinctive spatial zones, and are thus strongly dependent on the Tonle Sap's unique hydrology. These could be influenced by hydrological changes by hydropower development in upstream of the Mekong or climate change induced changes in the precipitation and hydrological cycle. The special role of the Siem Reap as driving the changes in the area was explicit in the results, which were made visible by using the spatial approach on the census data.

The role of big and open global data sets was crucial in Article IV. To calculate water shortage indicator, an assumption on the spatial unit of analysis is required. This assumption defines the available water resources for the population located in that area, and entails uncertainty, because it is not straight-forward to define accurately who are using which water resources. For example, using the most detailed level (grid-cell level) would require an assumption that the population strictly within that parallelogram[3] would use the water available in that same area.

With lack of detailed information on how much, when and who are using water, deciding over one spatial unit of analysis is not simple. Thus, alternative spatial units of analysis are required. The availability of multiple big, consistent and open spatial datasets for creating alternative spatial units of analysis is thus crucial to overcome the uncertainty in the assumption of water availability and demand within a certain spatial unit of analysis.

5.2. Exploring the ways to find appropriate spatial unit of analysis

One of the most brilliant possibilities of spatial data and spatial analysis is the ability to define the spatial unit of analysis freely. Yet, responsibility follows this freedom due to the Modifiable Areal Unit Problem (discussed in section 5.3). This section takes, however, a more optimistic view and considers the phenomenon as *modifiable areal unit possibility*.

In Article I, the possibility to organise the information and explore the status of various indicators according to the major river basins was informative. Article I also paid attention to the problem of excluding important areas that were very close but not within the physical boundaries of the basin. Including those areas, e.g. Shanghai in case of Yangtze, could have changed the quantitative vulnerability profiles to some extent. With GIS and specially defined spatial units of analysis, that do not necessarily follow the physical basin boundaries, this could be addressed in future spatial assessments. However, considering the scope of Article I, river basin provided an appropriate unit of analysis.

Article II and Article III were both particularly concerned with finding an appropriate spatial unit of analysis. In Article II, the spatial tools were used for calculating the slope from the Digital Elevation Model, which was then used together with data on ecological zones to create sub-areas that share similar biogeographical characteristics. The developed approach enabled distinguishing the important delta areas, particularly in the Irrawaddy basin (Figure 8). This is an important finding as there are no readily available datasets of delta and river mouth areas, even though these areas are generally stated to be very vulnerable (e.g. Ericson, Vörösmarty, Dingman, Ward, & Meybeck, 2006). River delta areas are naturally changing areas, and thus this possibility to utilise spatial tools and

[3] A square grid cell in WGS84 projection dataset would correspond to a parallelogram having upper end shorter compared to lower end, and also grid cell size would decrease when moving north or south from the equator.

data for the specially defined spatial unit of analysis will maintain as a crucial feature for anybody conducting spatial assessments in major river basin scale.

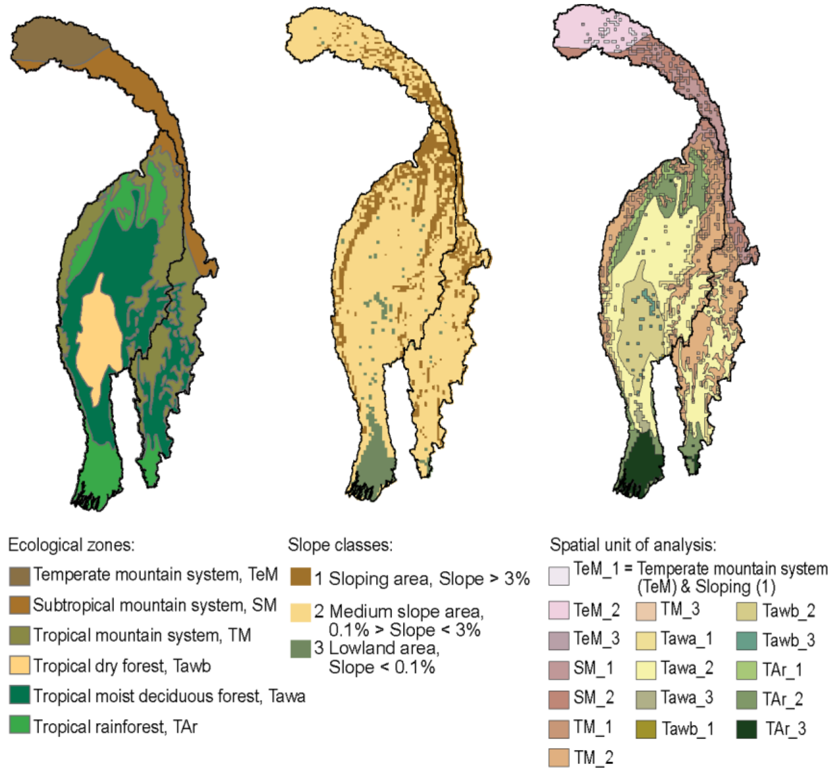


Figure 8. Distinguishing the delta areas (Modified from Article II).

In Article III, appropriate spatial units of analysis were found based on topography (following the contour line of six meters), census information on the urbanity of villages and administrative province boundaries. The advantages of the spatial units were explicit, when results were compared with using mere topographical zoning or mere provincial boundaries. As can be seen in Figure 9, important pieces of information would not have been visible if the spatial units of analysis developed in the study were not used. The spatial dimension improved the informativeness of the Census data. The topographic zoning enabled creating a connection to water level and hydrology, and the inclusion of administrative boundaries connected the results with actual planning processes and more detailed level survey design. The results enabled putting future changes in hydrology into socioeconomic context.

In both Article II and III, the term 'appropriate spatial unit of analysis' entails certain assumptions, which could be different if the general approach of the assessment would be changed. However, when sub-areas are created in a transparent way using big and open spatial data, it is possible to find other appropriate spatial units of analysis as well.

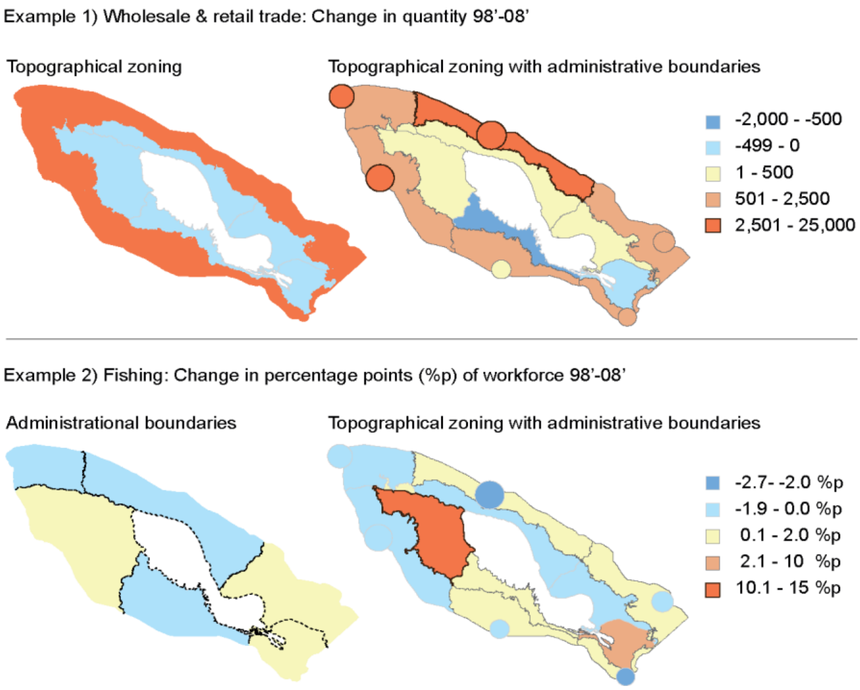
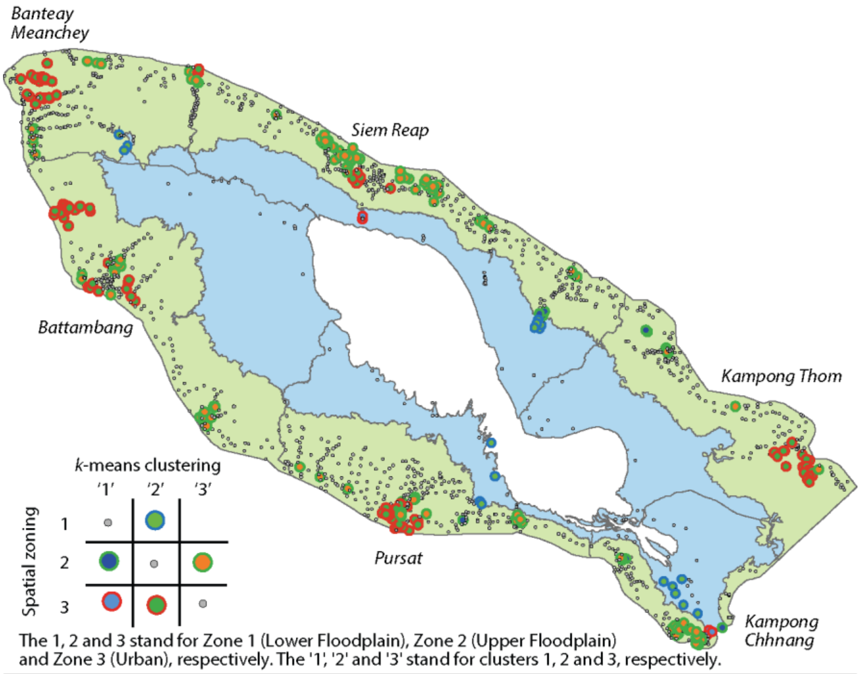


Figure 9. Two examples on the advantages of using a spatial analysis unit that is based on both the topography and the administrative units of analysis, instead of using only one of those. The spatial variation and interesting spatial differences become more visible with the unit of analysis combining topography, urbanity and administrative boundaries. (Modified from Article III).

Transparency and robustness in defining appropriate spatial unit of analysis can be increased when statistical methods are incorporated in the analysis. In Article III, the robustness of using the contour line of six meters and the census classification of urban villages for defining the unit of analysis was compared with the results from the *k*-means clustering. Altogether 14 indicators were used for clustering the villages. One of those was an 'urban' indicator that was a combination of ten indicators (combined with PCA). Various numbers of clusters were tested and the clusterings were performed multiple times to ensure the robustness of the *k*-means clustering (as described in the Article III).

Both methods ended up grouping the villages similarly to three main groups (Table S5 in the supplement for Article III). There was a clear pattern in the Census data, which divided the villages into roughly three groups, particularly distinguishing the villages mainly dependent on fishing and the urban areas as their own groups. The bulk of agricultural villages were clustered as one group. The majority of villages were grouped similarly with both of these methods. Figure 10 shows the differences between the results of the *k*-means clustering and spatial approach. The urban areas were the most difficult to be clustered together, which indicates their high level of heterogeneity that is characteristic of urban areas. The fastest changes are occurring in the Siem Reap area, and the decadal Census cannot keep up with defining the urbanity of the area close to Siem Reap. This explains much of the differences between *k*-means clustering and spatial definition. It is important to note that without spatial dimension, it would have been more difficult to interpret the results from the *k*-means clustering and putting them into context.



Differences in number of villages between spatial zoning (SP) and *k*-means clustering (KM):

Similar in both	Zone 1 (SP) -> 'zone 2' (KM)	Zone 2 (SP) -> 'zone 1' (KM)	Zone 2 (SP) -> 'zone 3' (KM)	Zone 3 (SP) -> 'zone 2' (KM)	Zone 3 (SP) -> 'zone 1' (KM)
Banteay Meanchey	3		12	22	
Battambang			20	26	
Pursat	4	1	20	16	
Kg Chhnang	6	2	22		3
Kg Thom		2	6	19	
Siem Reap	5	3	52	8	5
TOTAL IN ZONE	18	8	132	91	8
Overall differences per zone	Zone 1 21 %	Zone 2 12 %		Zone 3 32 %	

Figure 10. The differences in village grouping between spatial zoning and *k*-means clustering mapped and tabulated. Modified from Article III.

While Articles I, II and III used basically only one criterion for the spatial unit of analysis, Article IV used altogether 21. As a result, Article IV brought one of the key findings that especially in case of water issues, it is often not possible to decide one optimal spatial unit of analysis. In such case using multiple and various spatial units of analysis is required. Further, it is possible to explore various assumptions e.g. regarding climate's influence on water shortage within a nation by calculating water shortage according to national boundaries and climate zones. By conducting an assessment with specially designed spatial unit of analysis, it is possible to include non-quantitative information into the assessment. Examining water shortage according to a climate zone classification provides one example, as it compares water shortages between areas that have different climate characteristics and thus probably also different approaches to management and organization of society (e.g. arid areas versus monsoon areas).

5.3. The role of the spatial unit of analysis and the influence of MAUP for assessment results

The role of the spatial unit of analysis is significant in shaping the assessment results. This statement is supported by the previous section stating the possibility to find certain criteria for spatial unit of analysis that remarkably increases the potential and informativeness of an assessment (especially Article III and II). Due to the influence of the spatial unit of analysis, the related Modifiable Areal Unit Problem requires a systemic analysis. This was conducted in Article IV, which studied the effects of the MAUP.

The key finding of Article IV was that the results change remarkably depending on the spatial unit of analysis (zoning) used (Figure 11). The results reveal how the areas for *high water shortage* (water availability less than 1000 m^3/capita/year) vary throughout the 21 zonings examined. At the same time, it was possible to find areas that fall under this category with most of the zonings. In terms of the MAUP, the most interesting areas were those that have higher average water availability but still fall under the *high water shortage* threshold with some of the zonings (Figure 4b in Article IV, light blue and green areas).

Population under *high water shortage* with different zonings ranged between 782 million (22% of total population) with the Köppen-Geiger climate zoning and 2.11 billion people (60% of total population) with subbasins as the unit of analysis (Table 2 in Article IV). The difference between these two extremes was almost three-fold. Twelve out of 21 zonings resulted in population under *high water shortage* exceeding half of the total population, i.e. between 1.76 billion and 2.11 billion. The scaling effect was visible when nations and provinces and basins and subbasins were compared. In both pairs the former is a direct aggregation of the latter. The multizonal and multiscale approach enabled a more profound and explicit exploration of the effects of the MAUP and brought robustness to the results.

Article III gave a good visual example of how the spatial unit of analysis influences the results (Figure 9). Furthermore, the study examined the status and changes in socioeconomy and demographics on different scales. For instance, the overwhelming role of fisheries sector of the Tonle Sap area for the whole nation was clearly revealed in the study. The trends in demographics and in tertiarisation process were explored according to the main three zones and the more detailed development of main economic sectors was examined according to the 18 sub-areas that revealed the spatial heterogeneity and the special role of the Siem Reap. Thus, Article III supports the finding on the usefulness of the multizonal and multiscale approach, as otherwise important issues would have not been as clear as with the approach used.

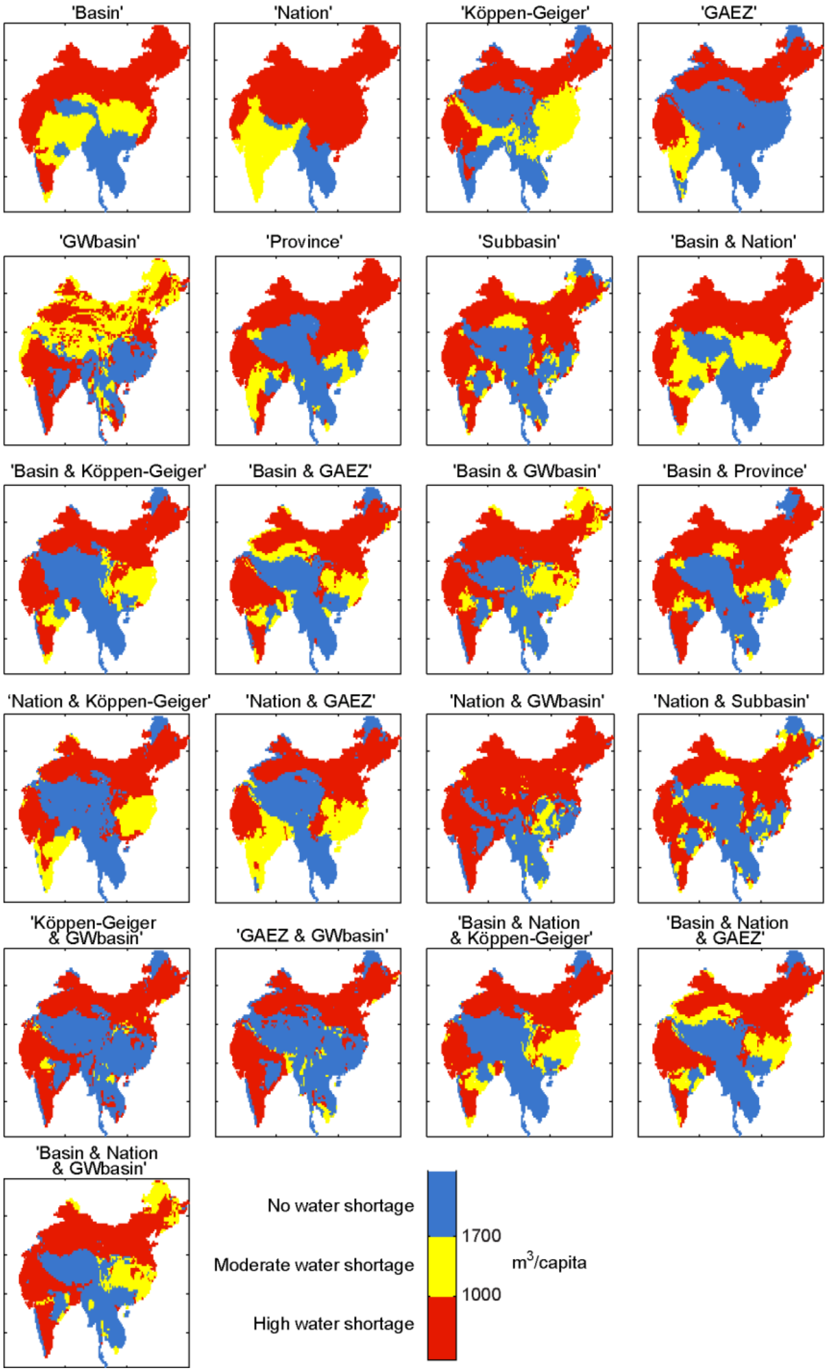


Figure 11. The variation in water shortage over the 21 zonings used. (Modified from Article IV).

6. DISCUSSION AND RECOMMENDATIONS

The key findings of the appended articles were summarised by the three research themes presented in the previous chapter. The research questions addressed issues that have not been studied to this extent before. While big and open data has been widely used in water-related vulnerability assessment (e.g. Vörösmarty et al., 2000; Vörösmarty et al., 2010), the applicability of such data has been less discussed (for exceptions, see e.g. Yamamoto, Fukuda, Nakaegawa, & Nishijima, 2007). Related to the second research question, the criteria for defining the unit of analysis have been less discussed in the water resource management field except for global scale[4]. However, alternative criteria are increasingly used e.g. in poverty mapping (Erenstein, Hellin, & Chandna, 2010) and in the studies of complex agro-ecosystems (Nelson, 2001). In terms of the third research question, the MAUP has been only little studied in the water resources management except for Perveen & James (2010, 2011, 2012).

The presented findings are thus novel and contribute to increasing understanding about the use of spatial analysis with big and open data for water-related spatial vulnerability assessments at multiple scales. Together with this discussion chapter, these findings lead to recommendations for overcoming some of the challenges encountered, thus contributing to the 4th research question (Figure 2).

6.1. Potential and limitations of spatial data

Big and open spatial datasets and census data were used extensively in the case studies providing a cross-cut of available data products and various methodological approaches for utilising those. Due to the extensive geographical and thematic scope of cases, it was not possible to compare the results with data from 'local' or primary sources i.e. to perform some sort of ground truthing. Instead, the dissertation aims at reporting the applicability and performance, for

[4] In global scale, for example, hydrobelts and hydroregions by Meybeck et al. (2013) and food production units (Kummu, Ward, et al., 2010) have been used. Latitudinal bands (Baumgartner, Reichel, 1975; Kummu & Varis, 2011), climate zones (Kottek, Grieser, Beck, Rudolf, & Rubel, 2006), hydro-ecoregions with consideration to fish populations of river basins (Abell et al., 2008) and ecosystem bands (Holdridge, 1967) have also been used (Meybeck et al., 2013).

which the appended articles provide a sufficient set of cases to prove the potential of big and open spatial datasets in water resources management field.

The dissertation extends the discussions related to the use of big and open spatial data in the field of water resources management by Perveen & James (2010, 2011, 2012) and Chapman & Thornes (2003) covering various other datasets and a wider spatial extent in Monsoon Asia. The findings here on the potential and limitations of big and open spatial data for water-related vulnerability assessments are similar to those of the report prepared for UN-Water (FAO, 2006). However, the dissertation provides a more detailed picture with examples applying the data for Monsoon Asia.

Besides potential, there are naturally plenty of limitations. Despite increasing availability of big and open spatial data and transparency in combining and analysing multisource spatial data, one set of limitations stems from the accumulation of uncertainty occurring in the process. Article II discussed this issue in more detail.

Generally adjusting the resolution of a spatial assessment according to the coarsest dataset keeps uncertainty in minimum when using spatial data from various sources (e.g. ArcGIS 9.3 Desktop Help, 2009; O'Sullivan & Unwin, 2010c). This ensures that the ability of spatial analysis of operating at more detailed level is not used in a way that would create a falsified sense of detail. Article II did not quite follow this guideline, and showed that it is possible to conduct seemingly detailed assessment by using a spatial unit of analysis that was in some cases one order of magnitude smaller than the input data. However, the end-user cannot judge the veracity in such case, unless explicitly informed on the resolution mis-match.

This same falsified sense of detail is evident in datasets that are made spatial through distributing national level information according to population distribution. The detail of these datasets is not as high as it seems, and this should be considered when gridded data is used for comparing areas that are within the same national boundary. In such case, areas actually differ only based on the amount of population and not based on the actual indicator.

The benefits of such gridded national statistics do come up when examining transboundary river basins, e.g. the Salween basin (Article II). It differs from the Irrawaddy basin as it is more clearly distributed within several countries and the population distribution varies greatly within the basin. These sorts of datasets, however, do not consider spatial and special characteristics, and thus, lower scores do not always mean less vulnerability, especially in areas with ethnic minorities. In Salween, the ethnic minorities might be more vulnerable due to their restricted possibility to influence policies and participate in decision-making while in quantities they are smaller and could thus get lower quantitative vulnerability scores. Therefore, fitting the national main population averages based solely on the distribution of population might not give reliable information. Even when sub-national data would be available, it is still not clear whether the data truly reflects the internal variability (Kattelus et al., 2015; Varis et al., 2014).

Another set of limitations stems from the representativeness of the data, including the lack of temporal coverage both in terms of seasonal changes and trends as time series. Also the lack of sufficient level of detail (meaning a coarse resolution) limits the applicability of the analysis results. This has been discussed in Salween by Yamamoto et al. (2007) in terms of the applicability of rather coarse GRACE satellite-based dataset. Truly big data with high velocity and detail is thus yet to emerge to cover data needs for spatial vulnerability assessments for water resources management.

The practical usefulness and limitations of proxy indicators is related to representativeness of data. Chape et al. (2005), for example, discuss the effectiveness of using data on protected areas as an indicator for meeting biodiversity goals (also discussed in Article II). In terms of water shortage, Article IV and Varis, Kummu, Lehr & Shen (2014) discuss the effectiveness of indicators to inform us about the actual water shortage. Thus, even though availability of data is increasing, its relevance is another question.

Furthermore, attention must still be given to the unavailability of data. In the appended articles, for example, information on hydropower dams (Article II, III), and fisheries (Article III) or on sub-national governance and socioeconomic indicators are missing from the analyses (Article I, II). Generally, data on water quality would be important to extend the discussion on water availability beyond physical and economical aspects. While the availability of data is of course increasing, it is not likely that full coverage of all relevant data would even in theory be possible. It is important to note that the (un)availability of data influences the assessments in great deals, and the assessments can thus be guided more by the availability of rather than the actual need for information (Article II).

Despite the limitations, the ability to extend the view, 'looking outside the water box' (Article I; Biswas, 2005; WWAP, 2009, 2012), and to integrate socio-economic issues with environmental information possess great potential. As Sullivan (2011) states, assessments combining data from multiple sources make it possible to explicitly recognise that our understanding is not perfect. In order to better manage the accumulation of uncertainty and ensure the representativeness of data, it is recommended to classify the input data in the spirit of Table 2 to keep track on the structure and assumptions behind each dataset and to consider how they go together. Also the gaps in data should be explicitly considered.

6.2. Importance of simplicity and transparency for comparison of assessment results

The most important feature of assessments in macro-scale is the comparativeness. As Article I states, an assessment conducted with simple methods enables comparison to be made more easily, also between other assessments. Complex methods are more sensitive to problem setting and scale (Svarstad, Petersen, Rothman, Siepel, & Wätzold, 2008). However, it should also been noted that simple aggregation of indicators, while easier to communicate and thus widely used, assumes independency of indicators and equal influence of indicators, which should be considered when used for decision-making. However, as discussed later, the application of methods reducing this multicollinearity do not directly solve the issue for enhanced or less-risky decision-making either.

Preference for simpler methods and results (principles of parsimony) are also confirmed by other studies in the water resources management field (Koutsoyiannis, 2009; Pavelic, Xie, Sreedevi, Ahmed, & Bernet, 2014). However, it should be noted that following this principle can also mean that what starts as vulnerability assessment ends up being only a profiling of some sort, and thus not addressing all the adaptive capacity, sensitivity and exposure components of the vulnerability. Moreover the dynamic nature of uncertainty involved in the very idea of vulnerability is not often very much represented in the outcome of vulnerability assessments. On the other hand, the lack of dynamism is very generic problem of static mapping analyses that the case studies in this dissertation represent.

Simplicity involves particularly the definition of the spatial unit of analysis. The various sizes of the spatial units of analysis can make comparison problematic. For example, in Article II altogether 14 sub-areas for both of the basins were created, but only 5-6 areas accounted for the majority of the surface areas of the basins (Figure 8, Figure 3D in Article II). When aggregating indicators and applying the PCA to these sub-areas, comparing areas of 200 km^2 with 20 000 km^2 with similar vulnerability score is challenging. These sorts of difficulties are similar in planning and policy-making, e.g. comparing the status of capital city and remote rural area that might occupy significantly different surface areas.

In the case of Article II, the results of the PCA could have been simpler to interpret and compare, if it had been conducted for original gridded data. In such case the spatial zoning would have not changed the variation in the data to be quite different compared to the variation in the original data. Thus, the combined use of case-specific spatial unit of analysis and complex statistical method, such as PCA, for weighing the indicators makes comparison difficult and new analysis is required, if dataset is updated due to changes in the correlations and other statistical relationships.

Another issue to consider is that when incorporating statistical analysis into spatial studies, the data should be analysed in a form that fulfils the assumptions of the statistical analysis at hand (such as normality, independence) (e.g. Larose, 2005, p.34). This often means that standardisation of the data is required, which can lead to exclusion of the outliers in the data that are actually important parts for a vulnerability assessment aiming to identify extreme areas. Even the simple aggregation of indicators requires scaling of the indicators to the same measurement scale and often also standardisation. This might reduce variation (information) carried by the indicators. In some cases, however, cutting outliers is necessary as otherwise high/low values outside the study area would distort the results (e.g. in case of Article I, high score of an indicator occurring e.g. in South America or in Greenland vs. in Monsoon Asia).

Besides comparability, the limitations mentioned hinder the effective communication of the results. Thus, when using certain pre-defined or specially defined spatial unit of analysis, the methods concerning the indicators should be very simple. Attention should be given on how the indicator's strength in conveying information is altered in the assessment process. The methods can be more complex when the data is allowed to organise itself and the variation and resolution of the original input data has not been changed by applying a spatial zoning beforehand the statistical method (as in Article III).

6.3. Acknowledging MAUP and plausibility of the spatial unit of analysis

The dissertation provided a systematic examination of the role of the spatial unit of analysis and the Modifiable Areal Unit Problem in water shortage assessment (Article IV), contributing to the discussion by Perveen & James (2010, 2011, 2012) on scale effects in water scarcity. Also Article III provided illustrative results showing the influence of the modifiability of the spatial unit of analysis.

In water resources management, the baseline hydrological modelling is already subject to MAUP, as the slope, flow accumulation, flow direction and watershed boundaries are derived from Digital Elevation Models and depend on the resolution of the elevation model used (Dark & Bram, 2007; Goodchild, 2011; Sanyal & Lu, 2004).

There is no way to avoid changing the descriptive statistics calculated for data when using other than the original resolution of the data, as the spatial unit of analysis chosen for the assessment involves neglecting the variation below the level of the chosen unit of analysis. Thus, also the message conveyed differs. For example, comparison of Article I and an article by Kattelus et al. (2015) shows that the upstream-downstream asymmetry is more explicit in the latter due to the more detailed unit of analysis. This was partly also caused by the availability of higher resolution data for some indicators.

One possibility to address the MAUP is using multizonal and multiscale approach (Article IV; Perveen & James, 2010). As such, also the appended articles form a multiscale approach, and convey different messages primarily due to their slightly different focus but also due to differences in scale and detail of the assessments. In Article I, the profiles of the Irrawaddy and Salween River basin are quite different compared to the profiles created in Article II. Multiscale approach is crucial and efficient similarly when exploring the results for the Mekong in Article I and the results for the Tonle Sap in Article III or the water scarcity result of Article I compared with the variation of water shortage in Article IV. While the influence of scale is intuitive, the multiscale approach makes it more tangible.

The possibility to define the spatial units of analysis can generate more complexity, or as George Gore puts it *“every problem that is solved produces other problems to deal with”* (Antiseri, 2006). Openshaw & Alvanides (2001) suggest that the User Modifiable Areal Unit Problem has more degrees of freedom than the classical MAUP and thus even greater possibility to produce a wider range of results. They also note that there is no real knowledge of what the ‘true’ result is, only a glowing awareness of how easy it is to lie with maps of aggregated data. Multizonal and multiscale approaches should therefore be designed carefully not to include unnecessary complexity and uncertainty (e.g. by including zonings that are weakly representative of issues at hand as discussed in Article IV). The spatial units of analysis need to be plausible.

In future, methods to define the unit of analysis in such a way that the influence of MAUP can be minimised might be more easily accessible, e.g. considering the internal homogeneity of a spatial unit of analysis. However, this will not reduce the challenges caused by the indicators (related to the ambiguous assumptions on, for example, water availability, demand and use), which in their complexity leave room for multiple possible spatial units of analysis. Further, defining the spatial unit of analysis should be added in the methodology guidelines for water-related vulnerability assessments, as these are being currently absent (e.g. guidelines by UNEP in Huang & Cai, 2009).

6.4. Potential of spatial vulnerability assessments for planning and policy-making

Articles I, II, and III all confirm the ability of spatial analysis to create a platform to integrate data in an effective and visually inspiring way. The spatial platform thus provides a basis for further discussions and can also be useful for planning and policy-making (del Campo, 2012; Huby, Owen, & Cinderby, 2007; Martinuzzi, Gould, & Ramos González, 2007; Radeloff, Hagen, Voss, Field, & Mladenoff, 2000; Walker & Young, 1997).

Water-related spatial vulnerability assessments can inform water resources management about the heterogeneity of water uses, demands, and the states of environment and the status and needs for quantitative data. Assessments are effective in shaping the bigger picture and trends in terms of various aspects related to water resources management. These assessments enable comparison and provide baseline for further discussions, but are often not directly applicable to local-scale management. Generally, even though aiming more to the strategic level, approaches could benefit from considering more the manageable scale. This means the scale at which actions can be directly implemented (Sullivan, 2011).

Furthermore, assessments are more likely to be influential in shaping the policy if stakeholders simultaneously perceive it as legitimate (how fair information production is in the perspective of different actors), credible (how falsifiable and transparent the methods are), and salient (how relevant information is for decision-making) (Cash & Moser, 2000; Lebel et al., 2005; Social Learning Group, 2001). As shown by the dissertation, the spatial unit of analysis and the scale together with the integrative and comparable data sources play important roles in meeting these conditions.

Policy making most often occurs according to administrative boundaries, which explains partly why issues concerning the scale for management other than nation state have been less addressed (Norman & Bakker, 2009). The basin approach has emerged as a response to this (Gleick, 1993 in Norman & Bakker, 2009; Lundqvist et al., 1985 in Norman & Bakker, 2009). But the question of using e.g. river basin as the spatial unit of analysis touches also upon depolitising issues that are inherently very political (Keskinen, 2010; Käkönen & Hirsch, 2009; Molle, 2009; Mollinga et al., 2007). Water resources cannot be only seen as material and having fixed boundaries due to being a flow resource (Norman & Bakker, 2009; Norman, Bakker, & Cook, 2012). Instead, water resources constitute a dynamic and complex interlinked system of networks, and thus, an issue to note is that the chosen spatial unit of analysis always leaves out something relevant.

Article I touches upon this theme by noting that basin boundaries leave out important areas that can depend on the river and influence on the use of water resources of that basin but do not belong within the physical boundaries of the basin area. The influence of transportation routes for food and goods, electricity grids, and location of remote industrial or high-populated trade areas are important factors and should be considered in delineating the analysis units. This is discussed also by Lebel et al. (2005) who use the term '*politics of place*', by Mollinga et al. (2007), who refer to 'problemsheds', and by Keskinen (2010), who discussed 'solutionsheds' in this regard. Thus, the MAUP related to spatial unit of analysis can stem from possible variation in the basin boundaries due to the politics of space (Cash & Moser, 2000; Lebel et al., 2005; Swyngedouw, 1997).

Generally, the scale could be seen as a socially constructed concept (Brown & Purcell, 2005; Cash et al., 2006; Delaney & Leitner, 1997; Lebel et al., 2005; Norman & Bakker, 2009; Norman et al., 2012; Swyngedouw, 1997, 2010). This is not often the case from the point of view of a practitioner, as scale is seen as fixed and given. Ultimately, the whole practice in science of classifying and defining is conducted by people i.e. socially. Therefore, the processes of basic technical mapping should be considered more carefully as they can too easily produce scales that achieve this stable status. This same goes for spatial unit of analysis, which should not be seen fixed or stable. Thus as Norman et al. (2012) state, there is much room in geography and the social sciences (as well as in the environmental engineering) *"to continue to refine and redefine our understanding of hydrosocial processes and the politics of scale within water governance"*.

This dissertation has contributed to this call by using and promoting the multizonal and multiscale approach that indeed enables to "*disentangle scale as a fixed unit and open up the between-ness of spaces and interrelationships"* and other assumptions related to the spatial units of analysis" (Norman et al., 2012). While this recommendation is seemingly concrete and concerns mapping workflow, it actually also highlights the link from the basic technical assessment processes to the higher decision-making and management level. The multizonal and multiscale approach can reduce the gap between the scale and level at which the problem is experienced, analysed, and discussed and the scale of the decision-making (Lebel et al., 2005; Towers, 2000). Integrating both spatial and discursive concepts in applying the spatial analysis should be developed in the future.

6.5. Way forward and recommendations

Scale-dependent knowledge on the interactions between system heterogeneity, process dynamics and system response over various scales is still incomplete

(Flügel, 2007). In future, however, large-scale and long-term comparative studies and research efforts can be systematically applied for dealing with global change, based on a shared data and knowledge base across the science-policy interface (Pahl-Wostl, Nilsson, Gupta, & Tockner, 2011). Thus, future studies will be heavily dependent on data intensive spatial analysis (Chapman & Thornes, 2003). Also other methods to support spatial analysis, such as Geographically Weighed Regression, Self-Organizing Maps, neural networks analysis, fuzzy logic, Bayesian nets and agent-based modelling, are becoming more convenient to use in GIS environment.

Ever-increasing amount of data and related methods are not, however, fully able to solve the inherent challenges related to vulnerability. The exposure, sensitivity and adaptive capacity components of vulnerability are generally something that cannot be measured. Therefore, getting rid of one-sided profiling of vulnerability can be problematic and even impossible with the means of quantitative spatial analysis. Ultimately, we are dealing with an epistemological question on vulnerability (de Souza Porto, 2012), meaning that even if we had all data and a known stress event, could we think about vulnerability as a deterministic factor that would realise to certain status after a certain stress event?

Without going further into discussing the true meaning of vulnerability and the ability of quantitative spatial assessments to provide deep insight about the state of vulnerability, this dissertation emphasises that numbers can give us ideas for improving our understanding and knowledge. It is becoming increasingly possible to mine and visualise previously unseen patterns in datasets (e.g. Coombes & Openshaw, 2001). While there is potential in more complex methods, it is important to ensure that assessment results are relevant and plausible.

Summary of findings and recommendations

Research Question 1:
There is rather extensive and fast-growing coverage of big and open spatial datasets (on hydrology and also on indicators 'outside the water box') that are applicable and perform well. However, crucial data e.g. on socio-economic aspects of water are still missing. Accumulation of uncertainty and representativeness of data are major hindrances for spatial water-related vulnerability assessments that integrate data from many sources.

Recommendation 1 (Research Question 4)**:**
It is recommended to classify data by resolution, formulation and structure and explicitly present the gaps in data. This can keep accumulation of uncertainty better under control and the assessment is not driven by the availability of the data but also contributes in identifying needs for improving data coverage.

Recommendation 2 (Research Question 4)**:**
The use of simple methods is recommended when conducting explorative quantitative vulnerability assessment that combines data from multiple sources.

Research Question 2:
Spatial analysis methods provide efficient ways to find appropriate spatial units of analysis. For example, approaches utilising topography and various combinations of zoning criteria were developed in the case studies and these proved to be very useful. Methods are, however, strongly location- and case-specific and assumptions behind influence a great deal.

Recommendation 3 (Research Question 4)**:**
Simple methods, transparency and explicit argumentation in defining the spatial unit of analysis are recommended. This maintains interpretability of the results particularly when using pre-defined spatial units of analysis. More complex methods, e.g. statistical methods, can be used as complementary methods or for mining data for alternative unit of analysis.

Research Question 3:
The spatial unit of analysis and related phenomenon of MAUP influence on the results and results depend on zoning and scale.

Recommendation 4 (Research Question 4)**:**
The use of multizonal and multiscale approach is recommended for explicit presentation of the role of the unit of analysis and exploration of scale-dependency.

Recommendation 5 (Research Question 4)**:**
Explicit consideration of the MAUP is recommended with the use of multiscale and multizonal approach with careful consideration and explicit reporting of assumptions of included zonings.

7. CONCLUSIONS

This doctoral dissertation examined four research questions aiming to increase the understanding of the use of spatial analysis with big and open spatial data for water-related spatial vulnerability assessments. The findings of this dissertation contribute to discussion concerning spatial vulnerability assessments in terms of multiple and interlinked change processes[5] and the increasing availability of big and open spatial data. The research was conducted through four case studies addressing the major river basins in Monsoon Asia at three different scales: regional, basin and subbasin.

The appended articles represent new and valuable scientific findings with concrete results for the globally remarkable river basins in Asia. The compilation and analysis of the underlying datasets provide considerable contribution for further development of spatial vulnerability assessments in the context of water resources management. Particularly the identification of data types with work flow recommendations and notions on the spatial unit of analysis form the key contribution of this dissertation.

The research questions were:

1) What is the applicability and performance of big and open spatial data in spatial platforms?
2) What are appropriate spatial units of analysis for water-related spatial vulnerability assessments?
3) What is the role of the spatial unit of analysis and how the Modifiable Areal Unit Problem (MAUP) influences the assessment results?
4) What kind of practical recommendations can be given to overcome the challenges in water-related spatial vulnerability assessments?

The case studies used various big and open spatial datasets to examine different aspects of water-related vulnerability. Starting from the smallest level, the appended articles include examples where data on village level was analysed at sub-province level with subbasin as the boundary of the study (Article III). In addition, grid level data (resolution ranging from 300 m to 0.5 degree or roughly 60 km in the Equator) was analysed at subbasin level with major basins as the

[5] For example, population growth and other demographic changes, climate change, as well as changes in land use, hydropower development, political regimes, economy, demand and consumption of water.

boundary of the study (Article II). On the coarsest scale, data with 30 arc-second to 0.5 degree resolution and national level data was aggregated to basin level, the ten major river basins giving the extent to the study (Article I). In the Article IV, 0.5 degree grid level data was aggregated to subbasin, province, basin, groundwater basin, agro-ecological, and climate zones with the tributary nations giving the boundaries to the study, hence, exceeding the mere basin boundaries.

The approaches ranged from merely exploring the physical water shortage calculated based on annual average water availability and population (Article IV) to looking at multiple indicators with different weightings (Article I), their co-dependencies with PCA (Article II), as well as patterns and change (Article III). Spatial analysis and statistical methods were utilised to create new ways to find appropriate spatial units of analysis and to study the role of the spatial unit of analysis in particular.

The key finding in terms of **the first research question** is that big and open spatial datasets are crucial and facilitate large scale comparative and explorative studies on water-related vulnerability. The extent and increasing detail of such datasets allow freedom in choosing the spatial unit of analysis and scale. Particularly, the spatial dimension enhances the applicability of big and open data, as was found in the case of Cambodian Census. Yet, the accumulation of uncertainty hampers the possibility to use the results for direct guidance for the management and planning decisions.

Spatial datasets with global coverage provide bridges between global scale studies and regional or local scale studies. The spatial vulnerability assessments end up many times being profiling, but succeed in framing the analysis and results in such way that something not quantitatively measurable can still be discussed with the help of quantitative analysis. The power of vulnerability assessments lies in this process rather than in having a quantitative value for vulnerability for a certain area.

In order to manage the accumulation of uncertainty, close attention should be given to classifying the input data in the spirit of Table 2 to keep track on the ‘anatomy and diet’ of the assessment. Also it is important to identify the gaps in data.

While big and open spatial data and the spatial platform are thus considered highly useful, these need to be accompanied with a thorough understanding of the spatial unit of analysis and its influence on the application of methods, as well as on comparativeness and communication of the results.

The findings of this dissertation support the recommendation of simple methods when using certain pre-defined or specially defined spatial unit of analysis. Simple methods are transparent and allow comparison. The methods can be more complex when the data is allowed to organise itself and tell its story without the distortive influence to the variation by the spatial unit of analysis. However, with a careful design, more complex methods can be used to increase the robustness of spatial approach. In such cases, the more complex methods would be separate from the vulnerability scoring and concern more the refinement and special definition of the spatial unit of analysis.

The possibility to explore various scales and non-traditional spatial units of analysis and to give up the stabile and fixed nature of scale is one of the greatest assets of spatial data. But possibility makes way for problems as well. The Modifiable Areal Unit Problem is introduced whenever there are multiple possibilities for spatial unit of analysis and scale.

This dissertation presented examples to answer **the second research question** on the appropriate spatial units of analysis for spatial water-related vulnerability assessments. The answer is that there is no universal way to find one appropriate unit of analysis. The examples were location-specific and particularly context-specific, meaning that a slight change in the focus of the research questions could set the boundaries of the appropriate spatial unit of analysis differently. The examples do show the potential of utilising spatial approach for finding appropriate units of analysis. The ability to utilise topography and other ancillary data to distinguish important areas or to link socio-economy and hydrology better together provided illustrative examples.

The general recommendation following these examples is that a free choice on the spatial unit of analysis comes with responsibility, and this has not been discussed explicitly in the methodological guidelines and spatial assessments concerning water-related vulnerability. The spatial unit of analysis needs to be plausible, meaning that the assumptions behind the choices on the level of variation neglected and the issues left outside the boundaries are stated, well-argumented and considered.

The complexity of water issues can make finding one optimal spatial unit of analysis unfeasible, and thus, the multizonal and multiscale approach is recommended. Various spatial units of analysis can be used if their plausibility is thoroughly considered. This dissertation showed as an answer to **the third research question** that the MAUP influences greatly to the results of assessments. These findings emphasise the important role of the spatial unit of analysis. Recommendation of the use of multizonal and multiscale approach follows also from this finding enabling the explicit acknowledgment of the MAUP.

The fourth research question on what kind of practical recommendations can be given to overcome the challenges in water-related spatial vulnerability assessments has been answered particularly in the previous chapter and repeated in this concluding chapter. One final recommendation is related to the future development of water-related spatial vulnerability assessments.

The future prospects look promising for vulnerability assessments, as the availability of big and open data is increasing along with the development of data mining methods and modelling. However, as K. R. Popper has stated[6]:

"in agreement with Peirce and Socrates,
I also believe that human beings are fallible,
and that they know so little;
I also believe that computers,
which are produced by men, are fallible.
Are they less fallible than men?
Perhaps they are, but they are also much less responsible."

Thus, while it is important to pursue a holistic, integrative approach and explorative curiousness in water-related spatial vulnerability assessments, it is crucially important to keep in mind that no method will ever be able to automatically turn data into well-argumented information or knowledge. Thus, the participation of intellectuals from various scientific backgrounds in solving problems is important. Ensuring this requires integration and systematic applications, which should be utilised to develop water-related spatial vulnerability assessments to better serve the science-policy interface.

[6] Cited from Antiseri (2006, p.13): K. R. Popper, *Perché siamo liberi? Computer, mente, razionalità*, cit., p.1.

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., . . . Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *BioScience, 58*(5), 403-414.

Abson, D. J., Dougill, A. J., & Stringer, L. C. (2012). Using Principal Component Analysis for information-rich socio-ecological vulnerability mapping in Southern Africa. *Applied Geography, 35*(1-2), 515-524.

Adger, N. (2006). Vulnerability. *Global Environmental Change, 16*(3), 268–281.

Alcamo, J., & Henrichs, T. (2002). Critical regions: A model-based estimation of world water resources sensitive to global changes. *Aquatic Sciences, 64*(4), 352-362.

Antiseri, D. (2006). *Popper's Vienna, Chapter One*. Aurora, CO, USA: Davies Group Publishers

ArcGIS 9.3 Desktop Help. (2009). Cell size of raster data. Redlands, CA, USA: Environmental Systems Research Institute, Inc.

Arnell, N. W. (2004). Climate change and global water resources: SRES emissions and socio-economic scenarios. *Global Environmental Change, 14*(1), 31-52.

Balica, S., Douben, N., & Wright, N. (2009). Flood vulnerability indices at varying spatial scales. *Water Science & Technology, 60*(10).

Bankoff, G. (2001). Rendering the World Unsafe: 'Vulnerability' as Western Discourse. *Disasters, 25*(1), 19-35. doi: 10.1111/1467-7717.00159

Bankoff, G., Frerks, G., & Hilhorst, D. (Eds.) (2004). *Mapping vulnerability: disasters, development, and people*. London, UK: Earthscan

Baumgartner, A., Reichel, E. (1975). *The world water balance: mean annual global, continental and maritime precipitation, evaporation and run-off.* 179 pp. New York, NY, USA: Elsevier.

Beniston, M. (2002). *Climatic Change: Implications for the Hydrological Cycle and for Water Management*. Secaucus, NJ, USA: Kluwer Academic Publishers.

Biswas, A. K. (2005). Integrated water resources management: a reassessment. In O. V. A.K. Biswas, C. Tortajada (Ed.), *Integrated water resources management in South and Southeast Asia* (pp. pp. 319–336). Delhi, India: Oxford University Press.

Brown, C. J., & Purcell, M. (2005). There's nothing inherent about scale: political ecology, the local trap, and the politics of development in the Brazilian Amazon. *Geoforum, 36*(5), 607-624. doi: 10.1016/j.geoforum.2004.09.001

Burstein, F., & Holsapple, C. W. (Eds.) (2008). *Handbook on decision support systems 2: variations*. New York, NY, USA: Springer.

Burton, I., Kates, R. W., White, G. F. (1993). *The Environment as Hazard* (Second ed.). New York, NY, USA: Guilford Press

Cash, D., Adger, N., Berkes, F., Garden, P., Lebel, L., Olsson, P., . . . Young, O. (2006). Scale and cross-scale dynamics: governance and information in a multilevel world. *Ecology and Society, 11*(2), 8.

Cash, D., & Moser, S. C. (2000). Linking global and local scales: designing dynamic assessment and management processes. *Global Environmental Change, 10*(2), 109-120.

Chape, S., Harrison, J., Spalding, M., & Lysenko, I. (2005). Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical Transactions of the Royal Society B, 360*(no. 1454), 443-455 doi: 10.1098/rstb.2004.1592

Chapman, L., & Thornes, J. E. (2003). The use of geographical information systems in climatology and meteorology. *Progress in physical geography, 27*(3), 313-330.

Clark, W. C., Jäger, J., Corell, R., Kasperson, R., McCarthy, J. J., Cash, D., . . . Willbanks, T. J. (2000). *Assessing Vulnerability to Global Environmental Risks*. Paper presented at the Vulnerability to Global Environmental Change: Challenges for Research, Assessment and Decision Making, May 22-25, 2000, Warrenton, Virginia, Cambridge.

Coombes, M., & Openshaw, S. (2001). Contrasting approaches to identifying 'localities'for research and public administration. In A. Frank, J. Raper & J.-P. Cheylan (Eds.), Life and Motion of Socio-Economic Units. pp. 293-306. London, UK: Taylor & Francis.

Costanza, R., D'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., . . . Van Den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature, 387*(6630), 253-260.

Couclelis, H. (1992). People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In Frank, A. U., Campari, I., Formentini, U. (Eds.) *Theories and methods of spatio-temporal reasoning in geographic space*. pp. 65-77. Berlin Heidelberg, Germany: Springer.

Cutter, S. L. (1996). Vulnerability to environmental hazards. *Progress in human geography, 20*, 529-539.

Cutter, S. L. (2003). GI science, disasters, and emergency management. *Transactions in GIS, 7*(4), 439-446.

Dark, S. J., & Bram, D. (2007). The modifiable areal unit problem (MAUP) in physical geography. *31*(5), 471-479. doi: 10.1177/0309133307083294

de Souza Porto, M. F. (2012). Complexity, Vulnerability Processes and Environmental Justice: An Essay in Political Epistemology. *RCCS Annual Review*(4). doi: 10.4000/rccsar.420

del Campo, A. G. (2012). GIS in Environmental Assessment: A Review of Current Issues and Future Needs. *Journal of Environmental Assessment Policy and Management, 14*(1), 125007. doi: 10.1142/S146433321250007X

del Campo, A. G., Gilmer, A., Foley, R., Sweeney, J., & Fry, J. (2011). Applying geographic information systems to support strategic environmental assessment: Opportunities and limitations in the context of Irish land-use plans. *Environmental Impact Assessment Review, 31*(3), 368-381. doi: 10.1016/j.eiar.2010.12.001

Delaney, D., & Leitner, H. (1997). The political construction of scale. *Political geography, 16*(2), 93-97.

Díez, E., & McIntosh, B. S. (2009). A review of the factors which influence the use and usefulness of information systems. *Environmental Modelling & Software, 24*(5), 588-602. doi: 10.1016/j.envsoft.2008.10.009

Dudgeon, D. (2005). River rehabilitation for conservation of fish biodiversity in moonsonal Asia. *Ecology and Society, 10*(2).

Eakin, H., & Luers, A. (2006). Assessing the vulnerability of social-environmental systems. *Annual Review of Environment and Resources, 31*(1), 365.

Erenstein, O., Hellin, J., & Chandna, P. (2010). Poverty mapping based on livelihood assets: A meso-level application in the Indo-Gangetic Plains, India. *Applied Geography, 30*(1), 112-125.

Ericson, J. P., Vörösmarty, C. J., Dingman, S. L., Ward, L. G., & Meybeck, M. (2006). Effective sea-level rise and deltas: Causes of change and human dimension implications. *Global and Planetary Change, 50*, 63–82.

Evans, T., Ostrom, E., & Gibson, C. (2003). Scaling issues in the social sciences. In J. Rotmans & D. S. Rothman (Eds.), *Scaling in Integrated Assessment* (pp. 75-106). The Netherlands: Swets & Zeitlinger.

Falkenmark, M. (1997). Meeting water requirements of an expanding world population. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 352*(1356), 929-936.

Falkenmark, M., Rockström, J., & Karlberg, L. (2009). Present and future water requirements for feeding humanity. *Food security, 1*(1), 59-69.

Falkenmark, M., & Widstrand, C. (1992). Population and water resources: a delicate balance. *Population bulletin, 47*(3), 1-36.

FAO. (2006). Water Monitoring: Mapping Existing Global Systems & Initiatives. Stockholm, Sweden: Food and Agriculture Organization. http://www.fao.org/nr/water/docs/UNW_MONITORING_REPORT.pdf

Flügel, W. A. (2007). The adaptive integrated data information system (AIDIS) for global water research. *Water Resources Management, 21*(1), 199-210. doi: 10.1007/s11269-006-9049-8

Folke, C., Carpenter, S., Elmqvist, T., Gunderson, L., Holling, C. S., & Walker, B. (2002). Resilience and sustainable development: building adaptive capacity in a world of transformations. *AMBIO: A Journal of the Human Environment, 31*(5), 437-440.

Füssel, H.-M. (2007). Vulnerability: a generally applicable conceptual framework for climate change research. *Global Environmental Change, 17*(2), 155-167.

Füssel, H.-M., & Klein, R. T. (2006). Climate Change Vulnerability Assessments: An Evolution of Conceptual Thinking. *Climatic Change, 75*(3), 301-329. doi: 10.1007/s10584-006-0329-3

Gain, A. K., Giupponi, C., & Renaud, F. G. (2012). Climate Change Adaptation and Vulnerability Assessment of Water Resources Systems in Developing Countries: A Generalized Framework and a Feasibility Study in Bangladesh. *Water, 4*(2), 345-366.

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. *Statistical Science, 24*(3), 343-360. doi: 10.2307/25681315

GEO Data Portal. (2014). United Nations Environment Programme UNEP/DEWA/GRID-Europe, available at http://geodata.grid.unep.ch/

Gleick, P. H. (1993). Water and conflict: Fresh water resources and international security. *International security*, 79-112.

Goodchild, M. F. (2011). Scale in GIS: An overview. *Geomorphology, 130*(1–2), 5-9. doi: 10.1016/j.geomorph.2010.10.004

GPWfe. (2005). Gridded Population of the World: Future Estimates (GPWfe). Center for International Earth Science Information Network (CIESIN), Columbia University, United Nations Food and Agriculture Programme (FAO), Centro Internacional de AGricultura Tropical (CIAT), Socioeconomic Data and Application Center (SEDAC), Columbia University. Retrieved 18 Mar 2008 from http://sedac.ciesin.columbia.edu/gpw

GRUMP alpha. (2004). Global Rural-Urban Mapping Project, aplha version (GRUMP alpha), Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); The World Bank; and Centro Internacional de Agricultura Tropical (CIAT). 2004. Global Rural-Urban Mapping Project (GRUMP), Alpha Version. Retrieved 21 Apr 2009, from Socioeconomic Data and Applications Center (SEDAC), Columbia University http://sedac.ciesin.columbia.edu/gpw.

Gurin, J. (2014). Big data and open data: what's what and why does it matter?, *15 Apr 2014, The Guardian, Public Leaders Network, Guardian News & Media Limited.*

GWP-TAC. (2000). Integrated Water Resources Management *Background Paper No. 4*. Stockholm, Sweden: Global Water Partnership/Technical Advisory Committee.

GWSP Digital Water Atlas. (2008a). Map 42: Population (Urban) (Dataset) (V1.0). Retrieved 9 Apr 2009, from GWSP Digital WaterAtlas

GWSP Digital Water Atlas. (2008b). Map 43: Population (Rural) (Dataset) (V1.0). Retrieved 9 Apr 2009, from GWSP Digital WaterAtlas

GWSP Digital Water Atlas. (2008c). Map 44: Population (Total) (Dataset) (V1.0). Retrieved 9 Apr 2009, from GWSP Digital Water Atlas

GWSP Digital Water Atlas. (2014). GWSP Digital Water Atlas. http://atlas.gwsp.org/index.php

Hinkel, J. (2011). "Indicators of vulnerability and adaptive capacity": Towards a clarification of the science-policy interface. *Global Environmental Change, 21*(1), 198-208.

Holdridge, L. R. (1967). Life Zone Ecology. San José, Costa Rica: Tropical Science Center,.

Huang, Y., & Cai, M. (2009). Methodologies guidelines - Vulnerability Assessment of Freshwater Resources to Environmental Change. Nairobi, Kenya: United Nations Environment Programme (UNEP).

Huby, M., Owen, A., & Cinderby, S. (2007). Reconciling socio-economic and environmental data in a GIS context: An example from rural England. *Applied Geography, 27*(1), 1-13.

Hurd, B., Leary, N., Jones, R., & Smith, J. (1999). Relative regional vulnerability of water resources to climate change. *Journal of the American Water Resources Association, 35*(6), 11.

Hurwitz, J., Halper, F., & Kaufman, M. (2013). *Big Data For Dummies*. Somerset, NJ, USA: John Wiley & Sons, Incorporated.

International Strategy for Disaster Reduction. (2004). *Living with risk: A global review of disaster reduction initiatives* (Vol. 1). New York, NY, USA: United Nations Publications.

Ionescu, C., Klein, R. T., Hinkel, J., Kavi Kumar, K. S., & Klein, R. (2009). Towards a Formal Framework of Vulnerability to Climate Change. *Environmental Modeling & Assessment, 14*(1), 1-16. doi: 10.1007/s10666-008-9179-x

IPCC. (2001). Intergovernmental Panel on Climate Change: Climate Change 2001: Impacts, Adaptation and Vulnerability (pp. 967). Cambridge: Cambridge University Press.

Islam, M. S., Oki, T., Kanae, S., Hanasaki, N., Agata, Y., & Yoshimura, K. (2007). A grid-based assessment of global water scarcity including virtual water trading. *Water Resources Management, 21*(1), 19-33. doi: 10.1007/s11269-006-9038-y

Jolliffe, I. T. (2002). *Principal Component Analysis*. New York: Springer-Verlag NY Inc.

Kappes, M. S., Papathoma-Kohle, M., & Keiler, M. (2012). Assessing physical vulnerability for multi-hazards using an indicator-based methodology. *Applied Geography, 32*(2), 577-590. doi: 10.1016/j.apgeog.2011.07.002

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., . . . Ratick, S. (1988). The Social Amplification of Risk: A Conceptual Framework. *Risk Analysis, 8*(2), 177-187. doi: 10.1111/j.1539-6924.1988.tb01168.x

Kattelus, M., Kummu, M., Keskinen, M., Salmivaara, A., & Varis, O. (2015). China's southbound transboundary river basins – A case of asymmetry. *Water International, 40*, 113–138. doi: 10.1080/02508060.2014.980029

Keskinen, M. (2006). The lake with floating villages: A socio-economic analysis of the Tonle Sap Lake. *International Journal of Water Resources Development, 22*(3), 463-480. doi: 10.1080/07900620500482568

Keskinen, M. (2010). *Bringing back the common sense? Integrated approaches in water management: Lessons learnt from the Mekong*. Doctoral Thesis, Water & Development Publications, Aalto University.

Keskinen, M., Kummu, M., Salmivaara, A., Someth, P., Lauri, H., de Moel, H., . . . Pech, S. (2013). Tonle Sap now and in the future? Final Report of the Exploring Tonle Sap Futures study. Espoo, Finalnd: Water & Development Publications WD-11.

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Londin, UK: Sage Publication Ltd.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorol. Z., 15*, 259-263. doi: 10.1127/0941-2948/2006/0130

Koutsoyiannis, D. (2009). *Seeking parsimony in hydrology and water resources technology*. European Geosciences Union General Assembly 2009. Geophysical Research Abstracts, vol. 11, Vienna, 11469, http://www.itia.ntua.gr/en/docinfo/906/

Kulshreshtha, S. N. (1998). A global outlook for water resources to the year 2025. *Water Resources Management, 12*(3), 167-184.

Kummu, M. (2008). *Spatio-temporal scales of hydrological impact assessment in large river basins: the Mekong case.* (Doctor of Science), Helsinki University of Technology, Water & Development Publications,TKK-WD-04, Espoo.

Kummu, M., Lu, X. X., Wang, J. J., & Varis, O. (2010). Basin-wide sediment trapping efficiency of emerging reservoirs along the Mekong. *Geomorphology, 119*(3-4), 181–197.

Kummu, M., Ward, P. J., de Moel, H., & Varis, O. (2010). Is physical water scarcity a new phenomenon? Global assessment of water shortage over the last two millennia. *Environmental Research Letters, 5*(3), 034006.

Kummu, M., & Varis, O. (2011). The world by latitudes: A global analysis of human population, development level and environment across the north-south axis over the past half century. *Applied Geography, 31*(2), 495-507.

Käkönen, M., & Hirsch, P. (2009). The Anti-politics of Mekong Knowledge Production. In F. Molle, T. Foran & M. Käkönen (Eds.), *Contested Waterscapes in the Mekong Region - Hydropower, Livelihoods and Governance*. London: Earthscan.

Lamberts, D. (2008). Little impact, much damage: the consequences of Mekong River flow alterations for the Tonle Sap ecosystem. In M. Kummu, M. Keskinen & O. Varis (Eds.), *Modern Myths of the Mekong - A Critical Review of Water and Development Concepts, Principles and Policies* (pp. 3-18). Espoo, Finland: Water & Development Publications - Helsinki University of Technology.

LANDSCAN. (2008). Landscan (TM) Global Population Database. Retrieved 9 Apr 2009, from Oak Ridge National Laboratory http://www.ornl.gov/landscan/

Larose, D. T. (2005). *Discovering knowledge in data: an introduction to data mining*. New Jersey: John Wiley & Sons, Inc.

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*(2), 18.

Lebel, L., Nikitina, E., Pahl-Wostl, C., & Knieper, C. (2013). Institutional fit and river basin governance: A new approach using multiple composite measures. *Ecology and Society, 18*(1).

Li, A. N., Wang, A. S., & Liang, S. L. (2006). Eco-environmental vulnerability evaluation in the mountainous region using remote sensing and GIS - A case study in the upper reaches of Mingjiang River, China. *Ecological Modelling, 192*, 175–187.

Liu, J., Zehnder, A. J. B., & Yang, H. (2009). Global consumptive water use for crop production: The importance of green water and virtual water. *Water Resources Research, 45*(5), W05428. doi: 10.1029/2007WR006051

Loucks, D. P. (2000). Sustainable Water Resources Management. *Water International, 25*(1), 3-10. doi: 10.1080/02508060008686793

Luers, A. (2005). The surface of vulnerability: an analytical framework for examining environmental change. *Global Environmental Change, 15*(3), 214-223.

Lundqvist, J., Lohm, U., & Falkenmark, M. (1985). River basin strategy for coordinated land and water conservation: synthesis and conclusions In Lundqvist, J., Lohm, U., & Falkenmark, M. (Eds.) *Strategies for River Basin Management*. pp. 5-17. Dordrecht, The Netherlands:D. Reidel Publishing Company.

Martinuzzi, S., Gould, W. A., & Ramos González, O. M. (2007). Land development, land use, and urban sprawl in Puerto Rico integrating remote sensing and population census data. *Landscape and Urban Planning, 79*(3–4), 288-297. doi: 10.1016/j.landurbplan.2006.02.014

Metzger, M. J., & Schröter, D. (2006). Towards a spatially explicit and quantitative vulnerability assessment of environmental change in Europe. *Regional Environmental Change, 6*(4), 201-216.

Meybeck, M., Kummu, M., & Dürr, H. H. (2013). Global hydrobelts and hydroregions: improved reporting scale for water-related issues? *Hydrology and Earth System Sciences, 17*(3), 1093-1111. doi: 10.5194/hess-17-1093-2013

Millennium Ecosystem Assessment. (2005). Ecosystems and human well-being: Wetlands and water.Washington DC, USA: World Resources Institute.

Molle, F. (2009). River-basin planning and management: the social life of a concept. *Geoforum, 40*(3), 484-494.

Mollinga, P. P., Meinzen-Dick, R. S., & Merrey, D. J. (2007). Politics, Plurality and Problemsheds: A Strategic Approach for Reform of Agricultural Water Resources Management. *Development Policy Review, 25*(6), 699-719.

Montgomery, D. R., Grant, G. E., & Sullivan, K. (1995). Watershed analysis as a framework for implementing ecosystem management. *JAWRA Journal of the American Water Resources Association, 31*(3), 369-386. doi: 10.1111/j.1752-1688.1995.tb04026.x

Nelson, A. (2001). Analysing data across geographic scales in Honduras: detecting levels of organisation within systems. *Agriculture, Ecosystems & Environment, 85*(1), 107-131.

Nilsson, S. (2003). *The role and use of information in transboundary water management*. Licenciate Thesis. Royal Institute of Technology, Stockholm, Sweden.

Norman, E., & Bakker, K. (2009). Transgressing scales: water governance across the Canada–US borderland. *Annals of the Association of American Geographers, 99*(1), 99-117.

Norman, E., Bakker, K., & Cook, C. (2012). Introduction to the themed section: Water governance and the politics of scale. *Water Alternatives, 5*(1), 52-61.

O'Keefe, P., Westgate, K., & Wisner, B. (1976). Taking the naturalness out of natural disasters. *Nature, 260*(5552), 566-567.

O'Sullivan, D., & Unwin, D. J. (2010a). Geographic Information Analysis and Spatial Data *Geographic Information Analysis*. pp. 1-31. Hoboken, NJ, USA: John Wiley & Sons, Inc.

O'Sullivan, D., & Unwin, D. J. (2010b). The Pitfalls and Potential of Spatial Data *Geographic Information Analysis*. pp. 33-53. Hoboken, NJ, USA: John Wiley & Sons, Inc.

O'Sullivan, D., & Unwin, D. J. (2010c). Putting Maps Together—Map Overlay *Geographic Information Analysis*. pp. 315-340. Hoboken, NJ, USA: John Wiley & Sons, Inc.

O'Brien, K., Eriksen, S., Schjolden, A., & Nygaard, L. (2004). What's in a word. *Conflicting perceptions of vulnerability in climate change research. CICERO Working Papers. Oslo, CICERO.*

OECD. (1993). OECD core set of indicators for environmental performance reviews. *OECD Environmental directorate monographs, no. 83*. Paris: Organisation for Economic Co-operation and Development (OECD).

Openshaw, S., & Alvanides, S. (2001). Designing zoning system for the representation of socio-economic data. In A. Frank, J. Raper & J.-P. Cheylan (Eds.), Life and Motion of Socio-Economic Units. pp. 273-292. London, UK: Taylor & Francis.

Pahl-Wostl, C., Nilsson, C., Gupta, J., & Tockner, K. (2011). Societal learning needed to face the water challenge. *AMBIO: A Journal of the Human Environment, 40*(5), 549-553.

Pavelic, P., Xie, J., Sreedevi, P. D., Ahmed, S., & Bernet, D. (2014). Application of a simple integrated surface water and groundwater model to assess mesoscale watershed development. In V. R. Reddy & G. Syme (Eds.), *Integrated assessment of scale impacts of watershed intervention: Assessing hydrogeological and bio-physical influences on livelihoods* (pp. 85-99). Amsterdam, The Netherlands: Elsevier.

Pelling, M. (2003). Tracing the Roots of Urban Risk and Vulnerability. In M. Pelling (Ed.), *The Vulnerability of Cities: Natural Disasters and Social Resilience*. UK & USA: Earthscan Publications Ltd.

Perveen, S., & James, L. A. (2010). Multiscale Effects on Spatial Variability Metrics in Global Water Resources Data. *Water Resources Management, 24*(9), 1903-1924. doi: 10.1007/s11269-009-9530-2

Perveen, S., & James, L. A. (2011). Scale invariance of water stress and scarcity indicators: Facilitating cross-scale comparisons of water resources vulnerability. *Applied Geography, 31*(1), 321-328.

Perveen, S., & James, L. A. (2012). Changes in Correlation Coefficients with Spatial Scale and Implications for Water Resources and Vulnerability Data. *The Professional Geographer, 64*(3), 389-400. doi: 10.1080/00330124.2011.609783

Preston, B. L., Yuen, E. J., & Westaway, R. M. (2011). Putting vulnerability to climate change on the map: A review of approaches, benefits, and risks. *Sustainability Science, 6*(2), 177-202.

Radeloff, V. C., Hagen, A. E., Voss, P. R., Field, D. R., & Mladenoff, D. J. (2000). Exploring the Spatial Relationship Between Census and Land-Cover Data. *Society & Natural Resources, 13*(6), 599-609. doi: 10.1080/08941920050114646

Raper, J. (2001). Defining spatial socio-economic units: retrospective and prospective. In A. Frank, J. Raper & J.-P. Cheylan (Eds.), Life and Motion of Socio-Economic Units. pp. 4–20. London, UK: Taylor & Francis.

Rapport, D., & Friend, A. (1979). *Towards a Comprehensive Framework for Environmental Statistics: A Stress-response Approach.* Statistics Canada, Statistique Canada, Occasional Papers, pp. 11–510.

Raskin, P., Gleick, P., Kirshen, P., Pontius, G., Strzepek, K., Kuylenstierna, J., . . . Hedman, B. (1997). *Comprehensive assessment of the freshwater resources of the world. Water futures: assessment of long-range patterns and problems.* Stockholm: Stockholm Environment Institute. Available at: http://www.sei-international.org/mediamanager/documents/Publications/SEI-Report-WaterFutures-AssessmentOfLongRangePatternsAndProblems-1997.pdf

Revenga, C., Murray, S., Abramovitz, J., & Hammond, A. (1998). *Watersheds of the world: Ecological value and vulnerability*: A joint publication by the World Resource Institute and the Worldwatch Institute.

Rockström, J., Falkenmark, M., Karlberg, L., Hoff, H., Rost, S., & Gerten, D. (2009). Future water availability for global food production: the potential of green water for increasing resilience to global change. *Water Resources Research, 45*(7).

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*(0), 53-65. doi: 10.1016/0377-0427(87)90125-7

Salmivaara, A. (2009). *Current Vulnerability and Future Challenges of Water Resources in the Irrawaddy and Salween River Basins - Indicator-Based Geospatial Analysis.* (Master of Science in Technology), Master's Thesis, Helsinki University of Technology, Espoo.

Salmivaara, A. (2012). Chapter 10 Myanmar: Assessing Freshwater Vulnerability in the Irrawaddy and Salween River Basins. In A. Lamadrid & I. Kelman (Eds.), *Climate Change Modeling For Local Adaptation In The Hindu Kush-Himalayan Region* (Vol. 11. Community, Environment and Disaster Risk Management). Bradford, UK: Emerald Group Publishing Limited.

Sanyal, J., & Lu, X. X. (2004). Application of Remote Sensing in Flood Management with Special Reference to Monsoon Asia: A Review. *Natural Hazards, 33*(2), 283-301. doi: 10.1023/B:NHAZ.0000037035.65105.95

Sen, A. (1976). Famines as failures of exchange entitlements. *Economic and Political Weekly*, 1273-1280.

Shaw, E. M. (1994). The Hydrological Cycle, Hydrometeorology, and Climate. In Shaw, E.M. (Ed) *Hydrology in Practice*, 3rd ed., pp. 1–35. London, UK:Taylor & Francis.

Simonovic, S. P. (2009). Water Resources Management: A Systems View. Stockholm Water Front Articles, 2009, SIWI, Available at: http://www.siwi.org/documents/Resources/Water_Front_Articles/2009/WRMASystemsView.pdf

Slocum, T. A., McMaster, R. B., Kessler, F. C., & Howard, H. H. (2013). *Thematic Cartography and Geovisualization: Pearson New International Edition, 3rd Edition*. Essex, UK: Pearson Education.

Social Learning Group. (2001). *Learning to Manage Global Environment Risks: a comparative history of social responses to climate change, ozone depletion and adic rain*. Cambridge, Massachussetts, USA: MIT Press.

Stelzenmüller, V., Ellis, J. R., & Rogers, S. I. (2010). Towards a spatially explicit risk assessment for marine management: Assessing the vulnerability of fish to aggregate extraction. *Biological Conservation, 143*(1), 230-238.

Sullivan, C. A. (2011). Quantifying water vulnerability: a multi-dimensional approach. *Stochastic Environmental Research and Risk Assessment, 25*(4), 627-640.

Sullivan, C. A., & Meigh, J. (2007). Integration of the biophysical and social sciences using an indicator approach: addressing water problems at different scales. *Water Resources Management, 21,* 111-128. doi: 10.1007/s11269-006-9044-0

Sullivan, C. A., Meigh, J. R., Giacomello, A. M., Fediw, T., Lawrence, P., Samad, M., . . . Steyl, I. (2003). The Water Poverty Index: Development and application at the community scale. *Natural Resources Forum, 27*, 189–199.

Sutherland, J. W. (1983). Normative predicates of next-generation management support systems. *IEEE Transaction on Systems, Man, and Cybernetics*, *13*(3), 79-297.

Svarstad, H., Petersen, L. K., Rothman, D., Siepel, H., & Wätzold, F. (2008). Discursive biases of the environmental research framework DPSIR. *Land Use Policy, 25*(1), 116-125.

Swyngedouw, E. (1997). Excluding the other: the production of scale and scaled politics. In R. Lee & J. Wills (Eds.), *Geographies of economies* (pp. 171-180). London, UK: E. Arnold.

Swyngedouw, E. (2010). Place, nature and the question of scale: interrogating the production of nature *Diskussionpapier 5*: Berlin-Brandenburgische Akademie der Wissenschaften. http://edoc.bbaw.de/volltexte/2010/1512/pdf/diskussionspapier_5_swyngedouw_online.pdf

The Dublin Statement. (1992). *The Dublin statement on water and sustainable development*. International Conference on Water and Environment, Dublin, Ireland, 29–31 December 1992. http://www.un-documents.net/h20-dub.htm

The World Bank. (2008). World Development Indicators 2008.

Thywissen, K. (2006). Components of Risk: A Comparative Glossary *Studies of the University Research, Counsel, Education, Publication Series No 2*. Bonn, Germany: United Nations University, Institute for Environment and Human Security (UNU-EHS).

Towers, G. (2000). Applying the political geography of scale: Grassroots strategies and environmental justice. *The Professional Geographer, 52*(1), 23-36.

Turner, B. L., Kasperson, R. E., Matson, P. A., McCarthy, J. J., Corell, R. W., Christensen, L., . . . Schiller, A. (2003). A framework for vulnerability analysis in sustainability science. *Proceedings of the National Academy of Sciences of the United States of America (PNAS), 100*(14), 8074–8079.

UNEP. (2012). UN-Water Status report on The Application of Integrated Approaches to Water Resources Management: United Nations Environmental Programme. http://www.un.org/waterforlifedecade/pdf/un_water_status_report_2012.pdf

United Nations. (2001). Indicators of sustainable development: Framework and methodologies. Background Paper No. 3. New York: Division for sustainable development, Department of Economic and Social Affairs. http://www.un.org/esa/sustdev/csd/csd9_indi_bp3.pdf

Wada, Y., Van Beek, L., Viviroli, D., Dürr, H. H., Weingartner, R., & Bierkens, M. F. (2011). Global monthly water stress: 2. Water demand and severity of water stress. *Water Resources Research, 47*(7), W07518. doi: 10.1029/2010WR009792.

Walker, P. A., & Young, M. D. (1997). Using integrated economic and ecological information to improve government policy. *International Journal of Geographical Information Science, 11*(7), 619-632.

Wang, H. B., Huo, D., Huang, J., Xu, Y., Yan, L., Sun, W., Li, X. (2010). An approach for improving K-means algorithm on market segmentation In Proc. International Conference on System Science and Engineering (ICSSE), Tapei, Taiwan, 1–3 July 2010. IEEE Xplore, 2010. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5551709

Wang, S. Y., Liu, J. S., & Yang, C. J. (2008). Eco-environmental vulnerability evaluation in the Yellow River Basin, China. *Pedosphere, 18*(2), 171–182.

Varis, O. (1996). Water quality models: Typologies for environmental impact assessment. *Wat. Sci. Tech., 34*(12), 109-117.

Varis, O., & Kummu, M. (2012). The Major Central Asian River Basins: An Assessment of Vulnerability. *International Journal of Water Resources Development, 28*, 433-452.

Varis, O., Kummu, M., Lehr, C., & Shen, D. (2014). China's stressed waters: Societal and environmental vulnerability in China's internal and transboundary river systems. *Applied Geography, 53*(0), 105-116. doi: 10.1016/j.apgeog.2014.05.012

White, G. F. (Ed.) (1974). Natural hazards: Local, National, Global. New York: Oxford University Press.

Wolf, A. T., Natharius, J. A., Danielson, J. J., Ward, B. S., & Pender, J. K. (1999). International River Basins of the World. *International Journal of Water Resources Development, 15*(4), 387-427. doi: 10.1080/07900629948682

Volk, M., Lautenbach, S., van Delden, H., Newham, L. T., & Seppelt, R. (2010). How can we make progress with decision support systems in landscape and river basin management? Lessons learned from a comparative analysis of four different decision support systems. *Environmental Management, 46*(6), 834-849.

Wong, D. W. (2009). Modifiable Areal Unit Problem. In K. Rob & T. Nigel (Eds.), *International Encyclopedia of Human Geography* (pp. 169-174). Oxford: Elsevier.

WSSD. (2002). Report of the World Summit on Sustainable Development, A/Conf. 199/20.: World Summit on Sustainable Development, http://www.johannesburgsummit.org.

WWAP. (2009). The United Nations World Water Development Report 3: Water in a Changing World. Paris: The United Nation Educational, Scientific and Cultural Organization (UNESCO).

WWAP. (2012). The United Nations World Water Development Report 4: Managing Water under Uncertainty and Risk: World Water Assessment Programme (WWAP), UNESCO, Paris.

Vörösmarty, C. J., Green, P., Salisbury, J., & Lammers, R. B. (2000). Global water resources: Vulnerability from climate change and population growth. *Science, 289*(5477), 284–288.

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., . . . Davies, P. M. (2010). Global threats to human water security and river biodiversity. *Nature, 467*, 555–561.

Yamamoto, K., Fukuda, Y., Nakaegawa, T., & Nishijima, J. (2007). Landwater variation in four major river basins of the Indochina peninsula as revealed by GRACE. *Earth, Planets and Space, 59*(4), 193-200.

Yoffe, S., Fiske, G., Giordano, M., Giordano, M., Larson, K., Stahl, K., & Wolf, A. T. (2004). Geography of international water conflict and cooperation: Data sets and applications. *Water Resources Research, 40*(5), W05S04. doi: 10.1029/2003WR002530

Aalto-DD 93/2015

ISBN 978-952-60-6282-2 (printed)
ISBN 978-952-60-6283-9 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

Aalto University

Department of Civil and Environmental Engineering
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS