Challenges and opportunities in large-scale river routing: Development and application of a hyper-resolution river routing model to assess anthropogenic impacts on freshwater ecosystems

Günther Grill

Department of Geography

McGill University, Montreal

December 2014

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Abstract

Despite significant recent advancements, global hydrological models and their input databases still show limited capabilities in supporting many spatially detailed research questions and integrated assessments, such as required in freshwater ecology or applied water resources management. In this body of research, I analyze the reasons for the lack of modeling support, identify shortcomings and challenges of current models, and design a next-generation eco-hydrological river routing model, termed HydroROUT. Based on the global hydrographic data repository of HydroSHEDS, HydroROUT is the first global hyper-resolution river routing model and includes a nested, multi-scale model approach; advanced implementation of connectivity; and a novel implementation of object-oriented vector data structures in a graph-theoretical framework. I subsequently explore the applicability of the model for different settings and research applications by designing and conducting four case studies related to assessing human impacts on freshwater system integrity. These case studies include both data-rich and data-limited areas, with spatial scales ranging from small headwater streams to large rivers, involving regional to global extents. I apply HydroROUT in two distinct research domains. First, I assess its capacity to model global spatio-temporal patterns of dam impacts. A set of new indicators, including the River Connectivity Index (RCI) and the River Regulation Index (RRI) were developed as part of two case studies, providing previously overlooked insights into intra-basin variability of dam impacts. Second, I apply HydroROUT to assess its capability for water quality modelling, by estimating in-river concentrations and risk of pollutants in freshwater systems using mass balance approaches at large scales. For this purpose, HydroROUT was adapted to function as a chemical fate model, capable of providing first-order risk assessments from point- and diffuse chemical sources. The results of two case studies show significant risk from pharmaceuticals in river reaches of the Saint Lawrence River Basin, Canada, and in extended areas of continental China. These large-scale outcomes of the HydroROUT modelling approach are at a previously unachieved spatial resolution of 500 m and can thus support local planning and decision-making in many of the world's large river basins.

Résumé

Malgré d'importantes avances récentes, les modèles hydrologiques globaux et leurs bases de données présentent encore des capacités limitées pour soutenir l'étude de nombreuses questions de recherche résolues dans l'espace et l'étude d'évaluations intégrées, comme requis dans le domaine de l'écologie des eaux douces ou dans la gestion des ressources en eau. Dans ce corpus de recherche, mon travail est d'analyser les raisons qui justifient les manquements dans les modélisations, d'identifier les lacunes et les défis des modèles actuels, et de concevoir une nouvelle génération de modèle écohydraulique de routage des rivières, appelé HydroROUT. Basé sur le répertoire de données hydrographiques mondiales de HydroSHEDS, HydroROUT est le premier modèle global spatialement hyperrésolu de routage de rivières et comprend une approche imbriquée et multiéchelle; une implémentation avancée de la connectivité; et une nouvelle mise en place de structures vectorielles orientées objet dans un cadre de la théorie des graphes. Par la suite, j'ai exploré l'applicabilité du modèle sous différents paramètres et domaines de recherche et je conçois et mets en œuvre quatre études de cas liées à l'évaluation de l'impact de l'activité humaine sur l'intégrité du système d'eau douce. Ces études de cas comprennent à la fois des zones riches et limitées en données, des échelles spatiales pouvant varier entre de petits de grands cours d'eau, dans un contexte aussi bien local que régional ou même mondial. HydroROUT a été appliqué dans deux domaines de recherche distincts. Tout d'abord, il a été évalué par rapport à sa capacité de modéliser au niveau mondial les tendances spatio-temporelles des impacts des barrages. De nouveaux indicateurs, soit l'indice de connectivité de la rivière (RCI) et l'indice de régulation de la rivière (RRI), ont été développés dans le cadre de ces études de cas et ont fourni de nouvelles connaissances sur la variabilité intrabassin des impacts des barrages, auparavant négligée. Deuxièmement, j'ai utilisé HydroROUT pour évaluer sa capacité de modélisation de la qualité de l'eau, en estimant les concentrations de polluants et les risques de les retrouver dans les systèmes d'eau douce à l'aide d'approches par bilan massique appliquées à grande échelle. À cet effet, HydroROUT a été adapté pour fonctionner comme un modèle de prédiction du sort des produits chimiques, capable de fournir une évaluation de premier ordre des risques de sources chimiques ponctuelles ou diffuses. Les

résultats montrent un risque important associés aux produits pharmaceutiques dans des tronçons de rivière du bassin du fleuve Saint-Laurent, au Canada, et dans des zones étendues de la Chine continentale. Les résultats à grande échelle du modèle HydroROUT sont à une résolution de 500 m, résolution jamais atteinte jusqu'à ce jour, et peuvent ainsi soutenir la planification locale et la prise de décision dans de nombreux grands bassins fluviaux du monde.

Contents

Table of Figures

Supplementary figures

List of Tables

Supplementary tables

Acknowledgements

First and foremost I would like to thank my supervisor Prof. Bernhard Lehner whose insights, perspectives, and humor have been greatly appreciated and made the Ph.D. experience an enjoyable one. My skills as a researcher have grown substantially under his guidance.

I would also like to express my appreciation to the members of my Ph.D. Advisory Committee, Prof. Elena Bennett and Prof. Michel Lapointe for advice, encouragement and critical questions during the early stages of my Ph.D. research.

Thank you to Prof. Tim Moore, Pauline Nesbitt, Nancy Secondo, Elisa David, Joseph Vacirca, Michael Havas, June Connolly, Franca Mancuso, Ruilan Shi, and Paula Kestelman for the departmental support and to Prof. Sébastien Breau, Graduate Program Director, for "demystifying the doctoral thesis process" to get past the last administrative hurdles.

My Ph.D. experience has been especially enjoyable because of colleagues and friends from the 'lab', formally the Global Land and Water Laboratory here at McGill University. The two groups—one representing 'Land' led by Prof. Navin Ramankutty (now University of British Columbia), and the other representing 'Water' led by Prof. Bernhard Lehner, complemented each other greatly and had many enlightening discussions around interacting issues. For the 'Land' side, special thanks goes out to Navin Ramankutty and his research 'crew' Verena Seufert, Jean-Sébastien Landry, Jordan Graesser, Dany Plouffe (thanks for all the chocolate, and especially for translating my abstract into French), Larissa Jarvis, Lisa Ribaudo, Federico Martellozo, Laura Bryson and many, many more. As for the 'Water' side, I'd like to say thanks to Bano Mehdi, Gregory Bunker, Delphine Deryng, Camille Ouellet Dallaire, Etienne Fluet-Chouinard, Joseph Ariwi, Rachel Dryden, Mathis Messager, and Irena Nedeva. Special thanks to Verena Seuffert, Bano Mehdi, and Jean-Sebastien Landry for research discussions and for friendly reviews.

Numerous people and organizations greatly supported this thesis: my co-authors Graham MacDonald (thank you very much as well for feedback and discussions on other chapters of this thesis), Usman Khan, Jim Nicell, Joseph Ariwi, Camille Ouellet Dallaire, Etienne Fluet-Chouinard, Nikolai Sindorf, Catherine Reidy Lierman, and Christiane Zarfl. It was a pleasure to conduct research with you. I would like to thank Dr. Petra Döll, Goethe-Universität Frankfurt am Main, for sharing essential hydrological data and for friendly reviews. Thank you, Mathis Messager, for providing essential lake volume data.

Even though as a global modeller I haven't had a field site I could visit, I was fortunate to have been invited on a few occasions to learn and share with other researchers: I am indebted to Prof. Francisco Olivera (Texas A&M University), Prof. David Maidment, Dr. Timothy Whiteacker and Dr. Cédric David (all University of Texas at Austin) for their hospitality during a research visit in May 2010 and for sharing their knowledge and research with me. I would like to thank William Darwall, Vineet Katariya, Adrian Hughes, and Robert Holland for the invitation to Cambridge to work at IUCN in 2011 on developing connectivity tools, and the interesting insights into global freshwater conservation. I also very much appreciated the invitation of Nike Sommerwerk to visit and present at the Leibniz Institute for Freshwater Ecology and Inland Fisheries (IGB), Berlin in 2012.

Generous funding for my doctoral studies and research came from several sources. This thesis was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). I am grateful for additional funding from the Department of Geography, particularly for the Warren GIS Award and the GREAT award, and I appreciated the networking grant by the Global Environmental and Climate Change Centre (GEC3), Quebec, Canada.

And last but not least, thank you my dear neighbours, Johanne Rabbat and Jan Scheck, for the mental and culinary support, and for taking care of business while away for research trips.

Thesis structure and contributions

This manuscript-based thesis is prepared in accordance with McGill University guidelines. Formatting follows the IOP Harvard (Author-Date) style. The thesis contains five research articles, as well as brief literature review and synthesis chapters. I am leading author on each of the chapters; the first authorship of chapter 2 is equally shared with Bernhard Lehner. In all cases, I led (or in the case of chapter 2 equally led) the research design and topic development, led or conducted the data compilation, led or conducted all manuscript-related programming and modelling, led or conducted all statistical analyses, and led the preparation of the manuscript text. However, several co-authors made important contributions to each of the chapters, which are detailed below.

Chapter 2 has been published:

Lehner B and Grill G* (2013): Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes* **27**. 2171-2186.

**Equally shared first authorship with Bernhard Lehner*; the author order on the manuscript only reflects Bernhard Lehner's longer involvement in the overall research framework (including the creation of global hydrographic baseline data) and his original invitation by the journal editor to contribute this manuscript as part of a special journal edition. Bernhard Lehner (BL) and Günther Grill (GG) equally designed the various research aspects presented in the manuscript and equally contributed ideas for its structure. GG wrote the original manuscript, contributed sections 2.1, 2.2, 2.4, and 2.5, and contributed to section 2.3. BL supported the revision of the final manuscript, in particular sections 2.3, 2.6, and 2.7.

Chapter 3 has been published:

Grill G, Ouellet Dallaire C, Fluet-Chouinard E, Sindorf N and Lehner B (2014): Development of new indicators to evaluate river fragmentation and flow regulation at large scales: A case study for the Mekong River Basin. *Ecological Indicators* **45**. 148-159.

Camille Ouellet Dallaire contributed river classification data and supported the revision of the manuscript. Etienne Fluet-Chouinard provided floodplain data for the river classification and supported the revision of the manuscript; Nikolai Sindorf provided advice on connectivity methodologies. Bernhard Lehner provided advice and guidance on various aspects of the analysis and helped write the manuscript.

Chapter 4 has been submitted to the journal
Environmental Research Letters:

Grill G, Lehner B, Lumsdon AE, MacDonald GK, Zarfl C, Reidy Liermann C (*in press*): An index-based framework for assessing patterns and trends in river fragmentation and flow regulation by global dams at multiple scales.

Bernhard Lehner provided advice and feedback on multiple aspects of the analysis, advised on methodological aspects, provided hydrographic data, and helped revise the manuscript. Alexander Lumsdon provided data on future dams. Graham MacDonald provided advice and feedback on the methodology and analysis, and helped revise the manuscript. Christiane Zarfl helped revise the manuscript. Catherine Reidy Liermann helped revise the manuscript and provided methodological advice.

Chapter 5 is in preparation for journal consideration in
***Environmental Science and Technology*:**

Grill G, Khan U, Lehner B, Nicell J, Ariwi J (*in preparation*): Geospatial contaminant risk assessment in the Saint Lawrence River basin.

Usman Khan provided advice on methodological design and the modelling approach, provided feedback on statistical modelling, and provided pharmaceutical data for the model. Bernhard Lehner and Jim Nicell provided feedback on the study design and helped revise the manuscript. Joseph Ariwi helped in the statistical analysis and the data preparation. Mathis Messager, McGill University, provided lake volume data.

Chapter 6 is in preparation for journal consideration in *Environment International*:

Grill G, Li J, Khan U, Yan, Z, Lehner B, Nicell J, Ariwi J (*in preparation*): A high-resolution contaminant fate model for continental China.

Jing Li provided geospatial data of wastewater treatment plants, provided parameters for the modelling, prepared the Chinese input data, and helped revise the manuscript. Usman Khan contributed to methodological design, conducted the statistical modelling of hormone sources, guided the risk assessment, and helped revise the manuscript. Zhong Yan attributed wastewater treatment plants with values for population served. Bernhard Lehner advised on methodological design and modelling, contributed data, and helped revise the manuscript. Jim Nicell helped revise the manuscript. Joseph Ariwi helped in the data preparation and geoprocessing.

1 Introduction

1.1 Background

Freshwater ecosystems are vital to human societies through the provision of freshwater from rivers, lakes, and wetlands as well as by delivering goods and services that contribute to the livelihood of local communities via fisheries, irrigation, and floodplain agriculture (Millennium Ecosystem Assessment, 2005). It is essential to understand, protect and improve ecosystem functioning to guarantee that these services continue to be provided. Anthropogenic pressures, however, have diminished the Earth's capacity to maintain many ecosystem services (Bagstad et al., 2013; Helfenstein and Kienast, 2014; Serna-Chavez et al., 2014) as we change our environments increasingly fast and at larger scales.

To minimize extensive impacts, it is imperative to understand the effects of human pressures and alterations on freshwater ecosystems from a large-scale perspective. Freshwater systems around the world experience multiple ongoing pressures (Vörösmarty et al., 2010), and changes to the hydrological cycle due to climate change may cause additional future stress (Kundzewicz et al., 2007). Aquatic systems have been greatly affected by physical alterations, with more than 6500 large dams built (Nilsson et al., 2005; Lehner et al., 2011), and more than 3700 large hydropower dams currently under construction or planned (Zarfl et al., 2014). Substantial amounts of water are diverted within and between basins to supply agriculture (Rost et al., 2008) and cities (McDonald et al., 2014). Nutrient contamination has also become an increasingly recognized global water quality issue in lakes, estuaries, and coastal areas (Bennett et al., 2001; Mayorga et al., 2010). The continuous release of contaminants from agriculture, industry, and household sources has now led to persistent water quality problems in many rivers around the world (Millennium Ecosystem Assessment, 2005; Schwarzenbach et al., 2010; Kümmerer, 2011).

Large-scale modelling is becoming increasingly important as a means to understand these pressures and their impacts at the global scale. Global hydrological models (GHMs) have emerged as a preferred research tool to analyze water resources and related impacts from humans. These GHMs are typically tailored towards the assessment of global freshwater resources and/or water scarcity by simulating basic processes of surface runoff generation and horizontal flow routing at global, regional, or national scales (e.g., Vörösmarty et al., 2000b; Alcamo et al., 2003; Arnell et al., 2004; Oki and Kanae, 2006; Rost et al., 2008; Hanasaki et al., 2010; Siebert and Döll, 2010; Haddeland et al., 2011).

However, hydrological modelling at large scales has traditionally been hampered by a number of issues, including general limitations in our knowledge and understanding of the underlying processes for many regions of the Earth; incomplete, inconsistent, or highly uncertain data collections; a lack of spatial integration between models and datasets; and the difficulty to create models that support multi-scale or coupled approaches (Lehner and Grill, 2013). There are five major challenges in hydrological and river routing models, which are related to *resolution*, *structure*, *scale*, *connectivity,* and *integrated modelling*. Current large-scale models generally operate at scales incompatible with other research domains, e.g. local ecological research. Many ecological or environmental management applications require accurate representations of stream and watershed attributes, but the relatively coarse spatial resolution of existing global models introduces bias and misrepresentation of hydrological processes and river topology. This can result, for example, in underestimation of river length and travel times (Gong et al., 2009; Verzano et al., 2012) and can lead to subsequent inaccuracies of travel-based attributes, e.g. sediment delivery to the ocean (Vieux and Needham, 1993). From a modelling perspective, the incorporation of local or fine-scale information in integrated models, such as the explicit identification of habitat or flow characteristics for individual tributaries, or the linkage of species information to river reaches and small sub-basins is difficult, if not impossible, with coarse scale models. Calls for modelling hydrological processes at a global scale with substantially higher spatial and temporal resolutions than currently available, so-called ‘hyper-resolution’ models, have been made but remain a ‘grand challenge’ in the hydrological modelling community (Wood et al., 2011).

An important component of GHMs is river routing. River routing models are specialized components that focus on the horizontal transport processes in river networks, e.g. they are frequently used within GHMs to translate runoff into discharge. However, the routing algorithms are often simplified in GHMs. More sophisticated types of routing models are decoupled from the model component that generates the runoff, and are instead built to improve the accuracy of flow routing, e.g., by using a higher resolution river network to make better estimates on river velocity. More specialized routing models have been developed to take into account additional important aspects, for example, the effect of floodplains on the hydrograph (Yamazaki et al., 2011), but considerable opportunity exists for improving other aspects of river routing models. The need for more capable routing models is frequently mentioned as a way to improve the results of coarse scale hydrological models (e.g., He et al., 2011). These more sophisticated river routing models are particularly important for expanding the types of applications for which GHMs can be used by making their results more relevant through increased resolution and better representation of critical in-stream processes.

I argue that a new generation of integrated models is required that support the linkage of hydrological processes with other fields, such as ecology, biogeochemistry, and water management. There are existing approaches, such as the use of GIS-based watershed models in an integrated assessment framework (e.g., Aspinall and Pearson, 2000), but applications that assess large-scale effects of human alterations to the water system, such as the impact of global reservoir and dam constructions on downstream river ecosystems (Lehner et al., 2011), are relatively sparse. There is currently no framework for conducting global-scale eco-hydrological modelling at high spatial resolution, i.e., a resolution that supports the accurate integration of local scale objects such as dams, wastewater treatment plants, sampling locations, gauging stations, etc. A number of continental type frameworks have been developed (Pistocchi and Pennington, 2006; Stein et al., 2014), yet these only exist for data-rich countries with good availabilities of high-quality geospatial resources such as Australia, Brazil, United States, and European countries.

Given the above outlined shortcomings and the lack of integrated modelling support at large scales, the focus of my thesis is to (1) *identify the challenges of large scale hydrological and routing models*, and (2) *to develop solutions in response to these challenges that contribute to the next generation of integrated large-scale eco-hydrological models*. A description of specific objectives is outlined in section 1.2.

This focus will contribute to a more integrated data and modelling framework, which facilitates an easier linkage between hydrological and ecological information that is geared towards eco-hydrological research, applications, and management. A key characteristic of my approach is to utilize a harmonized database of hydrographic baseline information (i.e., river network, sub-basin delineations, and linked features such as dams, lakes, and gauging stations) and to develop customized assessment tools that couple hydrological model results and ecological information within this data framework.

In this thesis, I describe the development and application of a new global, seamless river routing model, called HydroROUT (see 1.4 for more details) that is fully implemented in a geographic information system (GIS). HydroROUT is built on the foundation of the HydroSHEDS database (Lehner et al., 2008), combining existing and newly developed global scale hydrographic baseline data with a dedicated river network and routing model. HydroROUT provides vector-based routing capabilities and is integrated in the widely used GIS software package ArcGIS (ESRI, 2011). HydroROUT's river network consists of a spatial graph (Bunn et al., 2000) with more than 17 million river reaches. The main information linked to each river reach is length, upstream watershed area, monthly long-term average discharge, river reach volume, and statistics on upstream and downstream connectivity. HydroROUT can operate at multiple scales, through the implementation of hierarchical, nested hydrological subdivision. Additional high-resolution datasets such as gauging stations, locations of large dams and reservoirs, floodplain extents, and a river classification are also integrated. The main functions of HydroROUT include accumulation functions, using reach-to-reach processing; a number of sub-modules to simulate in-stream processes (e.g., decay in river reaches and lakes) and upstream/downstream tracing algorithms, which can operate with distance, attribute- or location-based barriers.

The research opportunities from such a model are manifold, but the actual usefulness can only be assessed if the model framework is applied to specific eco-hydrological research. I therefore *conducted four representative studies related to assessing human impacts on freshwater system integrity in different settings*. These include both data-rich and data limited areas, spatial scales ranging from small headwater streams to large rivers, and involving regional to global extents. More specifically, I focus on two sets of applications:

First, in thesis chapters 3 and 4, I assess and examine HydroROUT regarding its capacity to *model global spatio-temporal patterns of dam impacts*. Two of the largest consequences of dam construction are river fragmentation and flow regulation (Nilsson et al., 2005; Lehner et al., 2011; Birkel et al., 2014). River fragmentation leads to a decline in connectivity (Pringle, 2003), which affects species migration and dispersal (Fukushima et al., 2007; Ziv et al., 2012); community structure and biodiversity patterns (Altermatt, 2013); and the rivers' function as transport pathway of organic and inorganic matter downstream into floodplains (Vörösmarty et al., 2003; Syvitski et al., 2009). Dam operation, particularly temporary water storage, is the main contributor to flow regulation, and may disrupt ecological functioning (Ward and Stanford, 1995; Pringle et al., 2000; Carlisle et al., 2011). As part of two case studies, one for the Mekong River Basin and another for the globe, I demonstrate the application of novel indicators that advance our understanding of dam impacts: the River Connectivity Index (RCI) and the River Regulation Index (RRI).

Second, in thesis chapters 4 and 5, I apply HydroROUT to assess its capability for *large-scale water quality modelling*, by estimating in-river concentrations and risk of pollutants in freshwater systems using mass-balance approaches. This fills an important research gap because water quality modelling is typically conducted at small to medium sized watersheds, but has rarely been conducted at larger scales, e.g., entire basins or continents (Wang et al., 2000; Pistocchi et al., 2009). This is particularly true for micropollutants—chemicals found in the environment at low concentrations, but high toxicity. I adapted HydroROUT to act as a Contaminant Fate Model (CFM; Feijtel et al., 1998) for such pollutants to assess first-order ecological risk in river networks at large spatial scales. HydroROUT is the first of its kind hyper-resolution CFM to be applied in two case studies for Canada and China.

1.2 Research objectives

The rationale and main research components of this thesis are illustrated in a conceptual model (Figure 1.1). The overarching goal of this thesis was *to contribute to the development of a new generation of integrated, large–scale, hyper–resolution river routing models for eco–hydrological applications.* To achieve this goal, the thesis addresses two key research questions across four subsequent chapters:

1) What are the key challenges in large–scale river routing models that need to be addressed in order to improve their applicability for assessing the state of global freshwater systems (chapter 2)?

I identify key challenges for designing customized modelling solutions within a broad methodological framework focused on river routing approaches. Responding to these challenges (see dark blue components in Figure 1.1), I designed and assembled the first version of a new river routing model, which included the main building blocks (i.e., the river network and its spatial graph), key information associated with the river network, and routing algorithms. I then implemented new methodological techniques related to river network analysis. Both tracing operations as well as cumulative functions of the routing model were developed and tested.

2) What are the appropriate scale, resolution, and data structures to assess multiple anthropogenic impacts on freshwater systems (chapters 3, 4, 5, and 6)?

After having assembled the first version of HydroROUT in chapter 2, I conducted a number of applied studies to further develop and improve the routing model in different geographic contexts. These studies relate to the five grand challenges surrounding hydrological and routing models that I identified earlier, i.e. resolution, data structure, scale, assessment of connectivity, and integrated modelling. Each study addressed specific aspects of these challenges and contributed analytic insights into how adequately these challenges can be addressed. All of the studies made use of an advanced implementation of connectivity, and each case study represents a component of a larger integrated model—where data and processes from different ecologically-related research domains are integrated into the river routing model.

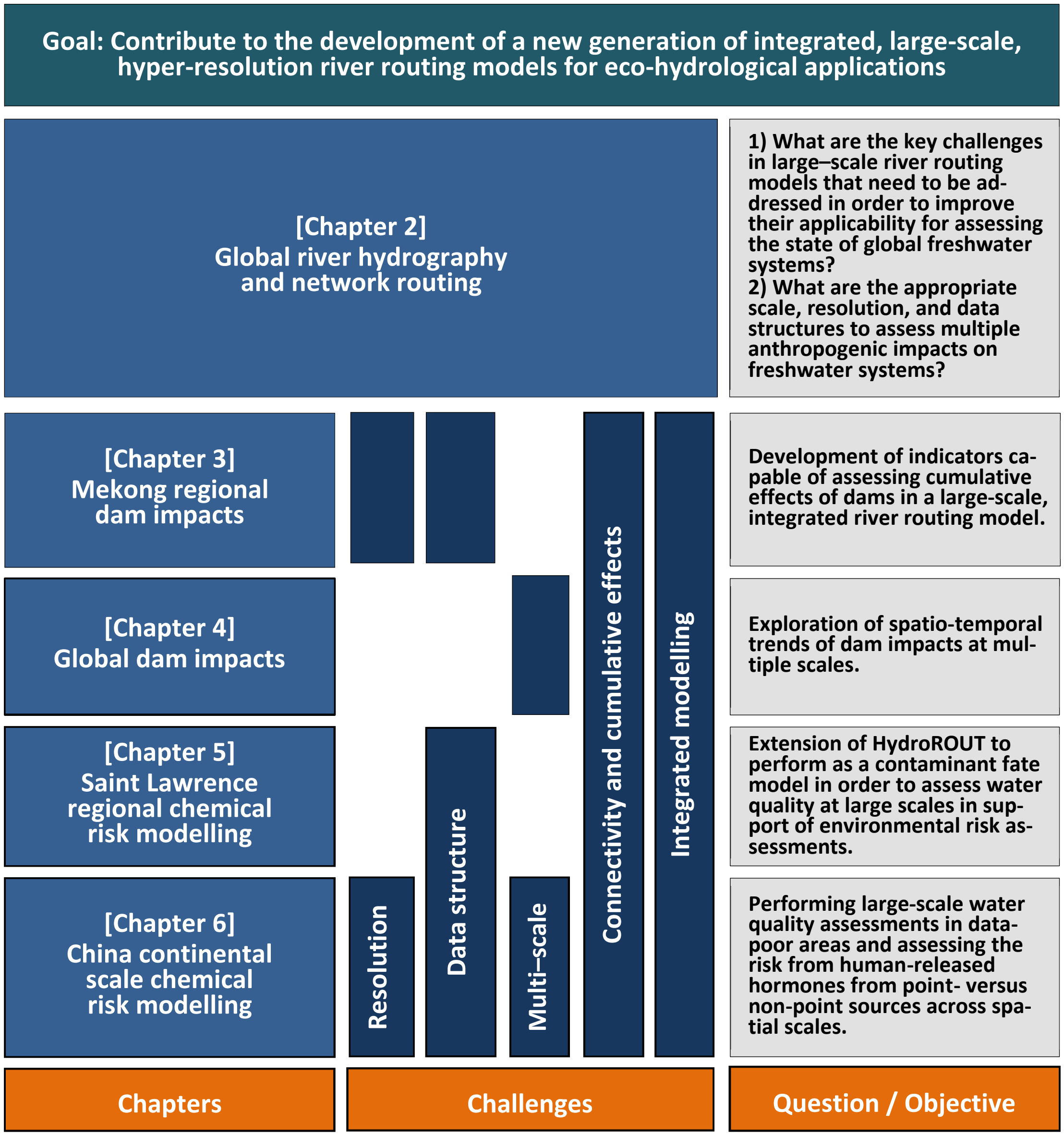


Figure 1.1: Conceptual diagram of relationships between overall goal, main research questions, objectives and chapters of this thesis. The extent of the dark blue components (main challenges identified in chapter 2) indicates in which other chapters these are primarily addressed.

In addition to these broader questions and objectives about large-scale river routing models and their advancements, each of the studies in chapters 3 to 6 address four more specific research objectives representing different components of GHMs and contributing to the five challenges (resolution, data structure, multi-scale assessments, connectivity, and integrated assessment) in different ways:

Development of indicators capable of assessing cumulative effects of dams in a large-scale, integrated river routing model (chapter 3).

There is currently no integrated large-scale modelling framework available to assess cumulative impacts of dams at resolutions appropriate for eco-hydrological applications. In this context, 'integrated' refers to the ability to link adverse types of data from various domains, e.g., ecology, hydrology, socio-economics, etc., into a common modelling framework to conduct assessments. In this chapter, I integrate different types of data (species distributions, floodplains, river classifications) to develop a series of eco-hydrological indicators of dam impacts—in particular, several variations of the River Connectivity Index (RCI) as well as the River Regulation Index (RRI)—and to critically compare and contrast their specific characteristics. This type of modelling and sensitivity analysis iteratively helped to determine the level of complexity necessary and feasible for assessing dam impacts on river systems.

Exploration of spatio-temporal trends of dam impacts at multiple scales (chapter 4).

Large-scale eco-hydrological modelling is typically conducted at the river basin scale, while smaller scales and cross-scale effects are neglected. Hence, of particular interest in this research application was the question of how HydoROUT would perform at scales other than the large river basin scale. I applied hydrological nesting approaches, which allowed identifying intra-basin dam impacts regarding river fragmentation and flow regulation, while traditionally only inter-basin perspectives were provided.

Extension of HydroROUT to perform as a contaminant fate model in order to assess water quality at large scales in support of environmental risk assessments (chapter 5).

Very limited research has been conducted to assess the contamination risk from micropollutants at larger scales in a spatially explicit manner. In this chapter, I explore the feasibility to simulate contaminant risk in extensive river networks from point sources in a major North American river basin. This chapter not only contributed to validating the river discharge of the routing model, but also provided opportunities to develop methods to integrate and model the effect of lakes in the routing model. The case study furthermore resulted in the integration of in-stream processes, such as environmental decay.

Performing large-scale water quality assessments in data-poor areas and assessing the risk from human-released hormones from point- versus non-point sources across spatial scales (chapter 6).

Water quality modelling rarely includes contributions from diffuse sources, such as contaminants released by human populations into small headwater streams without prior sewage treatment. In this chapter, I developed new ways to integrate high-resolution raster data as a source layer required for river routing in basins where non-point sources play an important role (river basins of China). I used hormones as a case study, because their release is strongly correlated with population density, as every human releases certain amounts of hormones, making these contaminants suitable for spatial assessments and comparisons to more localized contaminant sources (wastewater treatment plants). A downscaling method from administrative regions to small river reaches was developed. After integration, I assessed the risk of multiple chemicals in the river network, in particular the relative contribution and sensitivity of the distributed, small-scale sources in respect to other sources within the model.

1.3 Literature review

1.3.1 Hydrological and routing models

Hydrological modelling and existing river routing models

The hydrological cycle continuously recycles water through the processes of precipitation, interception, flow, evaporation, and condensation. These processes connect landscapes with the atmosphere, biosphere, and oceans through the transfer of water, energy, material and organisms. Hydrological models simulate these processes to generate spatio-temporal trends of river flow. A typical hydrological model consists of two components: 1) a Land-Surface component, which estimates the Earth's vertical water balance by simulating the processes occurring between the Earth's surface and the atmosphere, such as precipitation, evaporation, etc., and which produces spatio-temporal trends of runoff; and 2) a horizontal routing component, which accumulates the calculated runoff downstream across the Earth's surface to estimate river flow.

Hydrological models can be classified into three main groups based on how they represent processes: simple models (mostly empirical approach; "black box"), mid-range models (combining empirical and mechanistic approaches; "grey box"), and detailed models (mostly mechanistic approach; "white box") (Abbott and Refsgaard, 1996). Hydrological models are also divided into groups based on the amount of spatial heterogeneity considered. Specifically, hydrological models can be either "lumped" (variables homogeneous throughout the study area), "distributed" (full spatial heterogeneity of variables), or "semi-distributed" (some heterogeneity considered) (Olivera and Maidment, 1999). The spatial scale considered in hydrological models ranges from the patch to global sales. Most models operate at the catchment scale; however, in the last two decades, regional, continental, and global-scale models emerged (Anderson and Bates, 2001). Finally, a distinction can be made upon the type of coupling between the runoff generation scheme and the routing model. Early global flow routing models were developed to evaluate Global Climate Models in a decoupled way, and this type of model continues to be applied (Yamazaki et al., 2013). Another line of research developed dedicated large-scale and global hydrological models, which explicitly couple the land-surface models with a river routing model to derive discharge (Döll et al., 2008; Haddeland et al., 2011).

Spatial resolution and scale

Global hydrological models and their routing schemes have been developed at different spatial resolutions, ranging from aggregated basin scale models ("lumped") to the most commonly applied 0.5 degree (~50 km x 50 km at the equator) resolution, with various lateral routing schemes developed accordingly (Oki and Sud, 1998; Graham et al., 1999; Renssen and Knoop, 2000; Vörösmarty et al., 2000a; Döll and Lehner, 2002).

At the catchment scale, the numerous hydrological and routing models that exist (e.g., SWAT, Mike-SHE, INCA, HEC—described below) are difficult to apply at larger scales for numerous reasons that relate to model complexity, input data requirements, parameterization and calibration. For example, the Soil Water Assessment Tool (SWAT; Arnold, 1998), a semi-distributed, physically based flow and water quality model is commonly used to simulate river and groundwater flow in river catchments at small to medium scales (i.e. catchments of 10^2-10^4 km^2 in basin size). SWAT has modules to route water, sediments, nutrients and organic chemicals, but requires extensive parameterisation of detailed processes. A few studies have been conducted at larger scales (Gosain and Rao, 2006; Schuol et al., 2008), albeit with significant difficulties to provide the numerous required parameters and variables. Similarly, MIKE-SHE (Abbott et al., 1986) is a distributed, integrated surface-groundwater hydrological model that has been extensively used in engineering and water management at scales ranging from individual soil profiles to catchments and larger basins (e.g., Senegal Basin; Andersen et al., 2001). MIKE-SHE utilizes basic Muskingum-type flow routing, which can be extended by the module MIKE 11 (Havnø et al., 1995) for inclusion of a full range of flow routing applications. A link to a generic ecological model (ECO Lab; Butts et al., 2012) has been established to model stream temperature and stressors from pollution in river catchments as part of the Danish project 'Riskpoint' (http://www.risk-point.dk/).

At global scales, a number of mainly physically based, distributed hydrological models, using grid boxes as their modelling units at various resolutions (usually 0.5° lat/lon or coarser) have been developed since approximately the 1990's. These developments were driven by the need for the validation of global climate models (Russell and Miller, 1990), and the lack of models for water resources management at large scales (Döll et al., 2008). The rather coarse modelling units have been applied mostly for reasons of technical feasibility (i.e. computational demand) and due to the fact that important input information, foremost climate data, has traditionally been offered at these resolutions. An example of a model that was designed to improve the hydrology of global climate models is the Variable Infiltration Capacity model (VIC), a soil-vegetation-atmosphere-transfer model (Liang et al., 1994). The model has evolved rapidly and is now capable of evaluating the effect of land use and vegetation changes on hydrology at the regional and global scale (Haddeland et al., 2007).

A more integrated, physically based and distributed global hydrologic model that takes socio-economic factors into account is WaterGAP (Alcamo et al., 2003; Döll et al., 2003). The WaterGAP global hydrological model is described here in slightly more detail than other models as it provides the discharge values that were downscaled and used in HydroROUT (see 1.4 below). WaterGAP calculates the water balance on a 0.5° x 0.5° raster surface, taking into account natural processes and anthropogenic water use. Discharge routing is accomplished using a kinematic routing scheme. In contrast to earlier versions of the model which used a constant velocity, velocity is now derived from a combination of channel slope and channel bed roughness (Fiedler and Döll, 2010). Large lakes and reservoirs are also considered. The model is tuned to match the long term annual discharge at selected gauging stations within an error of 1%; however, locations far from tuning stations perform less well. Many other models exist in this category, e.g. LPJmL (Rost et al., 2008), WBM (Vörösmarty et al., 2000b), TRIP (Oki et al., 2001), or MAC-PDM (Gosling and Arnell, 2011).

In contrast, the “Terrestrial Hydrology Model with Bio-geochemistry” (THMB; formerly HYDRA; Coe, 2000; Coe et al., 2008), differs from many other models by a) separating the routing scheme from the simulation of the hydrological water balance; b) simulating potential lake and wetland areas based on topography (Coe, 1998); and c) operating at a relative high resolution of 5 x 5 arc minutes (~10 x 10 km). Newer versions of the model use high resolution sub grid topography based on SRTM elevation data (Farr et al., 2007) to improve the simulation of flooding in the Amazon. In recent developments, other hydrological routing models are moving also towards higher resolutions, e.g., 15 arc minutes (Yamazaki et al., 2009; Yamazaki et al., 2013) and 5 (or 6) arc minutes (e.g., Wisser et al., 2010; Schneider et al., 2011).

Multi-scale applications

Although some of the models mentioned above could be modified for application at different scales, few models are designed by default to be applied over multiple scales. Beighley et al. (2009) used a delineation process based on the Pfafstetter method (Verdin and Verdin, 1999) to derive irregular hydrological units (catchments) nested in increasingly larger hydrological units at coarser scales. The irregular units were then transformed into regular rectangular units from which routing parameters were developed. This type of framework allows for a flexible modelling at multiple scales. The Hydro1K database provides subunits based on the Pfafstetter concept at the global scale (USGS, 2000). Stein et al. (2014) used a similar approach to divide Australia into homogeneous hydrological units.

River routing models

The term "routing" generally refers to the simulation of transportation processes over space and time and typically refers to the procedure by which the change in magnitude, speed and form of a flood wave at any point in a hydrological network is calculated, resulting in a hydrograph (Maidment, 1993). A flow routing scheme, the mathematical method to simulate the movement of water over the landscape, translates runoff into river discharge by passing the runoff generated in a specific landscape unit (e.g. subwatershed or pixel) to the next downstream unit. The translation process from one unit to the next within hydrologic models, executed by the flow routing scheme, can be 'hydrological' (relatively simple based on a mass conservation approach) or 'hydraulic' (more dynamic, mass and energy is conserved). The routing scheme of hydraulic routing models can be kinematic, diffusive, or fully dynamic (Maidment, 1993). These three schemes increase in level of complexity (respectively) as they more fully represent the physical processes and take more variables into account (e.g., bedslope friction, pressure, convection, and acceleration) (Feldman, 2000). Hydrodynamic processes such as backwater effects and lateral flooding have critical implications on ecosystem functioning, and the inclusion of hydrodynamic principles allows for more advanced routing models. A recent study for the Amazon River Basin employed a diffusive wave routing approach in the CaMa-Flood model (Yamazaki et al., 2011) to demonstrate that floodplains have strong attenuating effects on flood waves and flood peaks in large basins. A fully dynamic routing scheme was developed as part of a large-scale hydrological model, the MGB-IPH, where routing occurs on vector-based hydrological response units (Paiva et al., 2011; Paiva et al., 2013a).

Depending on the spatial data structure (raster/vector) and the complexity of the routing scheme (hydrological/hydraulic), three basic routing approaches are most widely used: a) the cell-to-cell routing, often applied in distributed, raster based models (e.g., WaterGAP, LPJmL, THMB, etc.); b) the source-to-sink approach (rarely used in practice), whereby runoff is calculated at one location (e.g. continental sink) as a result of all contributing sources (Olivera and Famiglietti, 2000); and c) an object-to-object routing approach, where routing occurs between vector hydrological objects, such as river reaches or watersheds. This type of routing is increasingly applied in large-scale models, and is the type of routing implemented in HydroROUT (see section 1.4).

Cell-to-cell routing models are based on raster data structures and have traditionally been the preferred choice of hydrological modellers, as the integration of topography in the form of rasterized Digital Elevation Models (DEMs) is a key requirement for many routing algorithms that trace water along drainage direction maps. Furthermore, many ancillary datasets stem from remotely sensed raster information (e.g., land cover, precipitation, etc.) and these grids are relatively easy to process in any modern GIS.

However, raster based routing models tend to be less efficient than vector based object-to-object routing models. Yamazaki et al. (2013) tested a global vector-based model using unit-catchments of ~400km (equivalent to a grid based model of 15 arc minutes resolution) and showed substantial improvements in computational speed over raster models without reducing modelling accuracy. David et al. (2011) and David et al. (2013) developed a vector-based, Muskingum-Cunge type routing scheme for runoff routing in the U.S. based on the NHDPlus stream network (USEPA and USGS, 2008) and tested its performance in Texas using different runoff raster grids. Paiva et al. (2013b) used the global hydrographic framework HydroSHEDS to create a river network of reaches with an upstream area of ~1000 km^2 in the upper Amazon River, and employed a fully-dynamic flow routing scheme to simulate the complex flow dynamics of the Amazon. Whiteaker et al. (2006) developed a schematic processor for ArcHydro databases (Maidment, 2002)—a conceptualized, vector river network that is able to link output hydrographic databases from ArcHydro to the Hydrologic Engineering Center (HEC) flow routing system (Feldman, 2000).

Basic in-stream processes, such as accumulation and solute transport along the river network, have also been implemented in a number of global hydrological routing models. Solute transport is often modeled simultaneously (e.g. SWAT, INCA), or can be decoupled from the flow routing processes. INCA is an example for a semi-distributed routing model coupled to a nutrient model (Wade et al., 2002b). It combines a global climate model with land use data, and uses simple kinematic routing schemes to calculate nitrogen (Whitehead et al., 1998), phosphorus (Wade et al., 2002a), and sediments (Lazar et al., 2010). The HYDRA routing model (Coe, 2000) was modified to simulate nitrate export variability in the Mississippi River basin (Donner et al., 2002). Dedicated nutrient export models exist to model the export of multiple nutrients at global scales which implement a mass-balance approach to route nitrogen, phosphorus, and carbon along a river network at 0.5 degree raster resolution (Harrison et al., 2010; Mayorga et al., 2010; He et al., 2011).

River tracing and graph theoretical models

Most applications of routing models focus on downstream routing of runoff and can therefore be considered single-direction models. The use of multi-directional approaches in routing models has been much less common. Multi-directional models make full use of connectivity in both upstream and downstream directions. Such models are frequently used in landscape ecology, where the functional and structural connectivity between habitat patches is assessed based on a graph-based framework (Bunn et al., 2000; Urban and Keitt, 2001). An advantage of a graph-based framework is the possibility to include network analysis, allowing the study of objects as part of a connected system. A large number of network analysis functions (search algorithms, analysis of network properties) are available to conduct a wide range of applications in modelling, within ecological, social, transportation, and utility networks.

The application of graph theory in fluvial landscape ecology is less common than in terrestrial landscape ecology. Graph-theoretic models of rivers combine vectorized links and nodes, which may represent river reaches and confluences, respectively, into a spatial graph (Galpern et al., 2011; see also section 1.4.2 for illustrations and further explanations). The spatial structures of river networks modify processes on several levels from individuals to meta-populations and from local fluvio-geomorphology to flow regimes (Brown et al., 2011). It is thus important to understand interactions, functioning and changes in river systems (Campbell Grant et al., 2007). In practice, information about species distribution and richness or other features can be allocated to the links or nodes, and connectivity measures, such as the distance between nodes, can be derived. For example, Erős et al. (2011a) showed examples of the application of “network thinking” to model functional connectivity using patch-based models (see also the review about functional connectivity below) in river networks. Schick and Lindley (2007) examined changing patterns in connectivity and the isolation of salmon populations due to dam construction in California’s Central Valley. Their application has been limited to smaller river systems since the computational requirements increase exponentially with the number of network reaches. (Erős et al., 2011b) used the “habitat availability approach” (Pascual-Hortal and Saura, 2006) to prioritize river conservation areas, and similar techniques were used by Segurado et al. (2013) to determine the best option to increase connectivity by removing individual dams.

1.3.2 Anthropogenic impacts of dam building on river fragmentation and flow regulation

Dam construction is amongst the most extensive human alterations of freshwater ecosystems. Dams have substantial impacts on the ecological integrity of rivers, wetlands and floodplains (Tockner and Stanford, 2002; Arthington et al., 2010; Poff and Zimmerman, 2010; Richter et al., 2010). A hierarchical framework of upstream and downstream impacts of dams on river ecosystems has been developed by Berkamp et al. (2000) identifying first-, second-, and third-order impacts (sensu Petts and Gurnell, 2005). First-order impacts are the immediate abiotic effects on ecosystems as a direct consequence of the barrier (e.g., changes in flow, water quality and sediment load) and lead to second- and third-order impacts. Second-order impacts are the abiotic and biotic changes in ecosystem structure and primary productivity caused by first-order impacts that take place over longer time scales. Third-order impacts are the long-term biotic changes on higher trophic levels that result from the integrated effects of first- and second-order impacts (e.g., effects of changes in invertebrate communities on fish, birds and mammals). River fragmentation and flow regulation are commonly considered as the most substantial first-order impacts (Nilsson et al., 2005; Lehner et al., 2011; Birkel et al., 2014).

River fragmentation and loss of connectivity

A core application of my thesis is to allow better understanding of the cumulative fragmentation effects from dams in river networks. Fragmentation of river networks causes isolation and diminishes river connectivity. River connectivity is an important concept in freshwater ecology, but it lacks a clear definition across disciplines. Most commonly, connectivity is further qualified into structural and functional connectivity, with hydrological connectivity as a special form of structural connectivity (Fullerton et al., 2010).

Structural connectivity has a longitudinal aspect that connects upstream and downstream ecosystems (Vannote et al., 1980; Wiens, 2002), a lateral dimension by linking riverine systems with wetlands and floodplains (Tockner et al., 1999), and a vertical component that connects surface water with groundwater flows

(Renard and Allard, 2013). Longitudinal connectivity is particularly important for river ecology because of its relation to species migration and dispersal (Fukushima et al., 2007; Cote et al., 2009; Ziv et al., 2012), its role in community structure and biodiversity patterns (Altermatt, 2013), and its function as transport pathway of organic and inorganic matter downstream and into flood-plains (Vörösmarty et al., 2003; Syvitski et al., 2009). Structural connectivity has been estimated in the past using increasingly complex methods (Fullerton et al., 2010). Basic indicator methods have been applied in global and large-scale studies. Those include the number of dams within a watershed (dam density), the total length of a river (in km) upstream from each dam, the proportion of the river network inaccessible from the sea, or the total length of swimmable distance from each point of the network (Nilsson et al., 2005; Anderson et al., 2008; Lassalle et al., 2009; Vörösmarty et al., 2010). In an attempt to capture the obstruction of large river systems by dams, Reidy Liermann et al. (2012) measured the length of the longest undammed stretch of the five largest rivers in each 'freshwater ecoregion' (as defined by Abell et al., 2008) to derive the percentage of free-flowing rivers.

Hydrological connectivity is a special form of structural connectivity, by adding a temporal/dynamic component to structural connectivity. The temporal aspect of hydrological connectivity is apparent on two levels: a) the pulse level, where the annual changes in river discharge create distinct, pulsing patterns, which drive ecosystem level processes, and b) regime level, or historical processes over timescales and decades, which influence large-scale fluvio-geomorphological patterns (e.g., incision) that may alter connectivity as well (Amoros and Roux, 1988; Amoros and Bornette, 2002). Hydrological connectivity could be defined simply as water flow between components of the hydrological cycle. As such, studies often treat hydrological connectivity as binary characteristic, which either exists or not. Simple indicators exist to measure hydrological connectivity, e.g., at the river reach scale the number of months a river reach falls dry can indicate 'intermittency', or at the network scale the ratio between actively flowing and potentially flowing stream reaches can be calculated (Phillips et al., 2011). If combined with other hydrologic metrics, intermittency can inform flow regime classifications (Kennard et al., 2010).

Barrier effects by dams are not the only component that influences species dispersal and occurrence in river systems. More broadly—from a landscape ecology perspective—connectivity can be defined as the degree to which a landscape facilitates or impedes movement of organisms among resource patches in a structurally connected river system (Tischendorf and Fahrig, 2000)—often referred to as *functional connectivity*. The ability to move between patches depends on optimal environmental conditions regarding water quality, chemistry, sediment load, fluvio-geomorphological characteristics, and flow conditions and are population specific. Functional connectivity is important in determining the distribution of species (Bunn et al., 2000; Watts and Handley, 2010; Wainwright et al., 2011) and is a fundamental concept in metapopulation biology, which is concerned with the distributions of populations and the gene flow between spatially distinct subpopulations of a larger metapopulation (Moilanen and Hanski, 2001; Muneepeerakul et al., 2007; Mari et al., 2014).

Functional connectivity assessments are species- or population-specific and are frequently used in terrestrial landscape modelling, where patch based models are common to determine connectivity between and within patches (Walker et al., 2007; Watts and Handley, 2010). Other approaches define connectivity on a landscape level, that is, the connectivity of the entire river network is expressed, with no intra-network variability (Moilanen and Hanski, 2001).

In their review of connectivity measures and models, Kindlmann and Burel (2008) provide a strong case for the need to further model species movement in landscape ecology, which would also apply to fish species movement. The ability to move between different patches can be simulated using dynamic species models. This has, for example, been applied to an integrated analysis of the trade-offs between energy production, food supply, and biodiversity in the Mekong River (Ziv et al., 2012). In response to the call for increasing complexity, however, Calabrese and Fagan (2004) argue that more complex connectivity models (species movement and dispersal models) are also very data intensive and are therefore generally restricted to smaller areas and landscapes.

A good compromise between realism and information requirement seems to lie in graph-theoretical methods (Calabrese and Fagan, 2004), which are also implemented in HydroROUT (see section 'River network and spatial graph' in 1.4.2 below). These types of models are based on network theory and use spatial graphs (Bunn et al., 2000). A number of functional connectivity assessments (some were mentioned above) use this approach. In addition, the approach by Padgham and Webb (2010) is worth noting as a more general framework to quantify the network-scale effect of changes in the spatial structure of a river network. They use transition probability matrices to model habitat availability between pairs of river reaches within stream networks.

The Dendritic Connectivity Index (DCI, Cote et al., 2009) is a graph-based index that reduces complexity by determining connectivity from characteristics of network fragments—i.e., sections of the river network separated by dams. The DCI uses network analysis to evaluate the cumulative impact of dams on the life history of potamodromous and diadromous fish at a river network (landscape) level based on the number, permeability, and location of barriers. In this thesis, I use an adaptation of the DCI concept to develop new indicators of river fragmentation (see chapter 3 and 4). This adaptation is similar to the DCI in the way graph-theoretical principles are applied, but my approach is to extend the DCI to allow indicator weighting based on hydrological and ecological variables, and to use 'river volume' as an improved proxy for habitat availability (Grill et al., 2014).

River flow regulation and consequences

The dynamics of river flow are central in maintaining the ecological integrity of aquatic ecosystems (Poff et al., 1997). Flow variability imposes changing physical conditions on the system, with ecological consequences at local to regional scales, and at temporal scales ranging from days to millennia (Naiman et al., 2008). The characteristic "natural flow regime" of a river is described by the magnitude, frequency, duration, timing and the rate of change of its discharge (Poff et al., 1997). Flows and floods of different magnitude reoccurring at certain time intervals shape the composition and structure of aquatic habitats and communities (Ward et al., 2002). Natural flows also provide environmental cues for many species to start spawning, egg hatching, or migration (Poff et al., 1997).

The introduction of a dam into a river system can lead to alterations of the natural flow regime and subsequently to adverse effects on the structural and functional integrity of river systems. If a dam leads to the formation of a reservoir, the natural flow regime is often drastically changed from a free flowing river to a water body that shows both river and lake characteristics. Water retention time (hydraulic residence time) increases significantly with consequences on sediment retention, nutrient cycling, productivity, and water chemistry (Gordon, 2004). Further impacts of reservoir creation are loss of river habitat for mammals and birds (Nilsson and Dynesius, 1994), deoxygenation (Pringle et al., 2000), changes in nutrient budgets (Pinto-Coelho, 1998; Friedl and Wüest, 2002), thermal stratification (Friedl and Wüest, 2002), greenhouse gas emissions (Friedl and Wüest, 2002), water loss from evaporation (Shiklomanov, 2000), and contamination of food webs with substances such as methylmercury (Rosenberg et al., 1997).

Dam operations often lead to stabilizing flow variability as they tend to reduce peak flows and increase minimum flows. The stabilization of river dynamics has been associated with adverse medium to long-term ecological effects on flora and fauna (Johnson, 1992; Poff et al., 1997; Jansson et al., 2000; Dudgeon et al., 2006; Gordon and Meentemeyer, 2006) and are likely to affect biodiversity negatively (Xenopoulos et al., 2005; Xenopoulos and Lodge, 2006). From a fluviogeomorphological perspective, channel forming high flows and channel migration create important aquatic environments such as abandoned channels, backwater reaches, and oxbow lakes, while low flow conditions provide unique environments for species (Shields Jr et al., 2000).

The effect of dams on the regulation of downstream flows can only be fully assessed if the operation rules of the reservoirs are known. Current practical assessments aim to create a hydrological baselines of undisturbed flow, classify flow based on statistical flow variables, determine the level of deviations between the baseline and the disturbed flow, and establish relationships between flow alteration and ecological response (Richter et al., 1996; Biggs et al., 2005; Arthington et al., 2006; Poff et al., 2010). McLaughlin et al. (2014) determined regional flow regimes of rivers at several sites across Canada and analysed deviations caused by flow regulation from hydropower dams. Although this approach could be extended to other parts of the world, the required detailed flow measurements are rarely available in less 'data-rich' regions. Furthermore, the downstream propagation of flow regulation, and the complex cumulative (or interacting) effects of flow regulation from multiple dams compared to single dams (McManamay, 2014) are not assessed by this approach. As a provisional solution in the absence of dam operation rules, the Degree of Regulation (DOR), i.e., the proportion of a river's annual flow that can be withheld by a reservoir or a cluster of reservoirs has been suggested as a first-level approximation of the potential impact of dams on downstream flows. Despite its limitations, the DOR has been a key component of seminal studies on flow regulation (e.g., Nilsson et al., 2005; Lehner et al., 2011) or has been analyzed in terms of the hydrologically equivalent 'change in residence time' or 'water aging' (e.g., Vörösmarty et al., 1997).

The DOR concept, however, is only applicable at the reach scale, and provides no basin-wide (landscape-scale) measure of flow regulation. Such a 'cumulative' DOR has—to my knowledge—not been developed yet. In chapter 3 and 4, I apply the DOR concept for determining reach-scale impacts, and also develop a new indicator based on DOR which assesses flow regulation from dams for entire basins.

1.3.3 Contaminant fate models for large scale water quality and environmental risk assessment

In order to effectively manage our freshwater resources and ensure the health and safety of humans and vulnerable aquatic ecosystems, we require a deep understanding of the sources and distribution of potentially harmful substances in our streams, rivers, and lakes. Furthermore, we need to understand where these substances originate, the magnitude of the risk they pose, and how they can best be managed to maintain the integrity of freshwater systems. Environmental risk assessment is of paramount importance to prioritize and identify chemicals that yield undesirable effects on aquatic life and human health (Ambrose and Clement, 2006; EMEA, 2006). Aquatic systems are especially vulnerable, serving as an endpoint for industrial and municipal effluents and a source of drinking water for downstream communities (Fent et al., 2006; Hernando et al., 2006). However, current risk assessment methodologies, based mainly on *in-situ* measurements, have major limitations, in particular for emerging contaminants that are analytically challenging to measure (e.g., nanoparticles), whose environmental quality standard is below current analytical capabilities (e.g., Ethinylestradiol, EE2), and those known to act as mixtures (e.g., estrogens) (Carlsson et al., 2006; Khan, 2014).

Contaminant fate models (CFMs), which combine spatially explicit hydrological models with environmental risk assessment methodologies, can address these shortcomings. CFMs identify contaminant pathways, take into account accumulation from upstream contaminations, can be deployed for multiple chemicals and at large scales, and can predict environmental concentrations with reasonable accuracy under various hydrological scenarios (Johnson et al., 2008; Cowan-Ellsberry et al., 2009). They are frequently used to conduct risk assessments for substances from wastewater treatment plants or other sources (Feijtel et al., 1998; Wang et al., 2000; Anderson et al., 2004).

These models share common assumptions and similar key mechanisms. Reasonable per capita emission estimates of contaminant mass entering individual wastewater treatment plants (WWTPs) can be gained using the average per capita consumption of a compound of interest and adjusting for human metabolism (for more detail see chapter 5 and 6). The contaminant mass released by individual WWTPs into specific river reaches can be estimated using the average per capita emissions, knowledge of the local population served by a WWTP, and adjusting for removal during treatment, where relevant (Keller et al., 2006). As for chemical routing in the hydrological system, advection is assumed to be the dominant dilution mechanism, which can be modeled effectively using stream length, velocity, discharge, and a decay function (Pistocchi et al., 2010). Predicted environmental concentrations (PECs) are subsequently based on accumulated load and discharge at the river reach scale.

CFMs are used to predict single substance concentrations in surface waters; examples include GREAT-ER (Feijtel et al., 1998), ISTREEM (Wang et al., 2000), LF2000-WQX (Williams et al., 2012), P*h*ATE (Anderson et al., 2004), MAPPE

(Pistocchi et al., 2012) and GWAVA (Johnson et al., 2013). Such models are particularly well-designed for household chemicals and have been applied to chemical exposure assessments in various studies (Atkinson et al., 2009; Cunningham et al., 2009; Hannah et al., 2009; Ort et al., 2009; Cunningham et al., 2012; Hosseini et al., 2012; Johnson et al., 2013).

CFMs have a strong spatial component in that they take into account population size and distribution, river network complexity and its accumulating effects, and local river flow dynamics that can differ substantially from the emission point and across the stream network. Typically, CFMs operate at the scale of river catchments, but a few have been applied at larger scales (Wang et al., 2000; Pistocchi et al., 2009).

As an example for the capability of current CFMs, the routing of substances such as nutrients and pollutants is used in a European context in the MAPPE model (Pistocchi et al., 2011b). Within a GIS, map algebra is used to model steady state concentrations of persistent organic pollutants, pesticides, volatile organic compounds, pharmaceuticals and other household chemicals (Pistocchi et al., 2009; Pistocchi et al., 2010; Pistocchi et al., 2011a). The spatial extent of MAPPE is the European continent with a grid resolution of 1 km^2. MAPPE can back-calculate emissions in catchments upstream given measured contaminant concentrations in rivers (Pistocchi et al., 2012). An example for the use of MAPPE in an Integrated Environmental Assessment is given in Marinov et al. (2014), where the model was used to estimate river loads of chemicals under different policy implementation scenarios.

To my current knowledge, advanced contaminant fate models such as GREAT-ER or P*h*ATE were applied for specific small-to-medium sized watersheds, but have not yet been applied over larger areas. Furthermore, none of the examined models considered contamination from untreated wastewater, for example, direct discharge of untreated wastewater by rural populations (for more details see chapter 6).

1.4 Development of a hyper-resolution river routing model

1.4.1 New approaches in river routing

More often than not, the term river routing has been used interchangeably with the term "flow routing", i.e. the accumulation of runoff across the landscape. In this thesis, I suggest to define the term "river routing" more broadly—in addition to the flow of water, river routing may also include the movement of other elements and objects (e.g., nutrients, plants, and species) *sensu* Pringle (2003), where hydrological connectivity is the flow of material, energy, and organisms within and between components of the hydrological cycle.

I further propose that for eco-hydrological applications, routing (and the associated models) should (i) include the possibility to conduct routing of water and substances along the hydrological flow paths; (ii) represent lakes, dams, floodplains, and wetlands in addition to river reach routing; and (iii) be based on a powerful tracing algorithm for upstream and downstream routing to represent both active and passive dispersal of matter or species. Furthermore, a routing model should be flexible for use across multiple scales.

As such, and to widen the scope of routing in hydrological models, I define the term 'river network routing' following the hydrological connectivity concept of Pringle (2003) as 'the simulation of movement of energy, material and species within a hydrological object space, based on flow routing, tracing and in-stream processing.'

HydroROUT, as a new type of river network and routing model, was designed around this definition with the goal to support eco-hydrological applications at regional to global scales. In these applications, the primary focus shifts from simulating dynamic river flow to representing a larger set of interacting processes (including, e.g., ecological, fluvio-geomorphological, and hydrological processes). This poses challenges for current hydrology-focused models, as the added complexity can render the model too inefficient to be useful as a simulation tool. Although the dynamic modelling of river flow is imperative for many eco-hydrological applications (e.g., to estimate the effects of land use on hydrology), for many integrated modelling applications a reduced representation of the hydrological processes (with less temporal detail of flow simulation or pre-computed discharge) may still be sufficient if larger spatial patterns and steady-state conditions are the focus (see applications in chapter 3 to 6).

In the following sections, I first describe the main design characteristics of HydroROUT, outline how HydroROUT resolves some of the challenges identified earlier, and provide a concise technical and methodological description of the model (i.e., going beyond the shorter explanations provided in the subsequent manuscript chapters).

1.4.2 Model structure and functionality

Data structure and key data layers

A key characteristic of my approach is to utilize a harmonized database of hydrographic baseline information, i.e. a river network and its catchments, sub-basin delineations, and linked features such as dams, waterfalls, and lakes (see Figure 1.2 for a graphical representation of main features, and Table 1.1 for a list of integrated layers).

In response to the identified challenges with data structures, I designed HydroROUT around a hybrid raster/vector data structure to benefit from the advantages of both raster and vector concepts. Nevertheless, vector data is the dominant data structure of HydroROUT (Figure 1.2a) and is derived from the underlying high resolution hydrographic raster database (HydroSHEDS; Lehner et al., 2008). As mentioned earlier, vector routing can significantly increase model execution time compared to cell-based models (Yamazaki et al., 2013) because of the inherently simple 'object-to-object' routing scheme.

A second advantage of the hybrid structure is improved data integration. The integration with underlying raster data at 500 m resolution allows maintaining linkages to other models (e.g., land surface or land use change models) and to transfer and integrate additional, finer scale information from external raster sources to the river network using geospatial statistics tools. Data integration across spatial scales (using the set of hierarchical sub-basins) is made possible using a coding system similar to the Pfafstetter method (Verdin and Verdin, 1999).

Spatial resolution

HydroROUT uses an object-oriented approach, whereby river objects—river reaches and their respective catchments (Figure 1.2a; green lines and thin black outlines)—are created from the high resolution (500 m x 500 m) raster database HydroSHEDS (Lehner et al., 2008). The applied algorithm to create these objects groups cells that are connected along the drainage network into distinct objects (a river reach or catchment), which substantially reduces the number of routing objects compared to the cell approach, while conserving the spatial accuracy of the object.

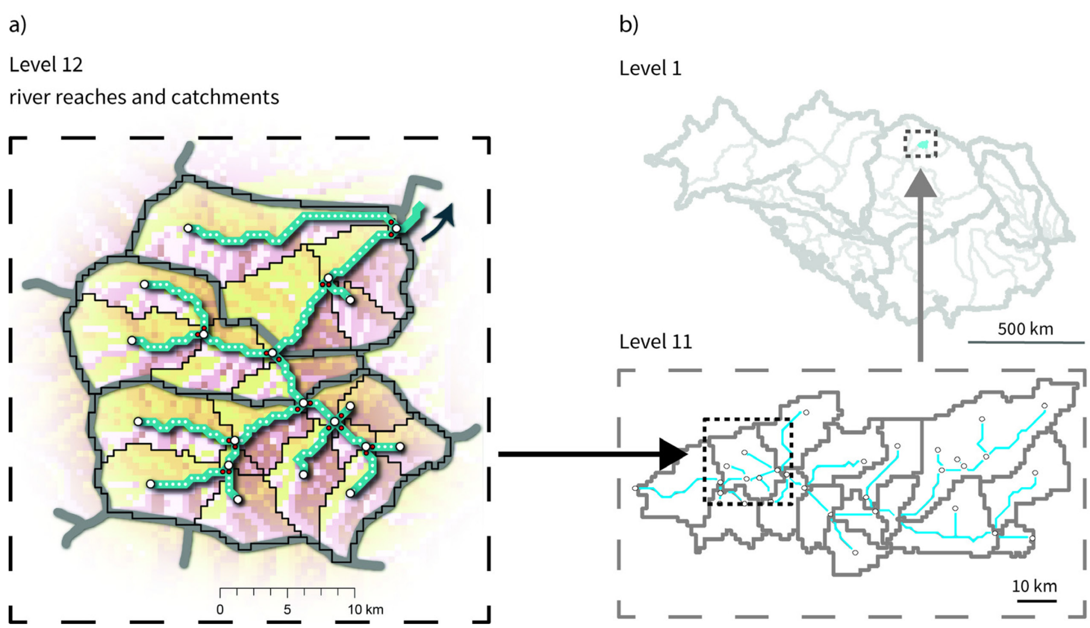


Figure 1.2: a) Flow direction raster (colored gridded surface; raster resolution ~500m), river reaches (green lines), and reach catchments (thin black lines). The reach catchments are used in conjunction with spatial extraction tools to link objects, or information from other models to the HydroROUT river network. The thicker grey lines represent a higher level of spatial aggregation, and each incorporates several smaller reach catchments. b) Sub-basins at different, nested hierarchical levels (HydroBASINS; Lehner and Grill, 2013).

Table 1.1: Key vector datasets integrated in HydroROUT.

Layer	Description
River reaches and associated catchment	Derived from the HydroSHEDS database (Lehner et al., 2008), the river network consists of 17770043 global river reaches (polylines) with an average length of 2.7 km, located in a Geodatabase in the form of a Geometric Network (see Figure 1.2a). Each river reach is linked to a polygon of its hydrological catchment, with an average size of ~12 km^2.
Sub-basin delineation	In addition to the river catchments, HydroROUT is linked to a set of predefined nested subbasin delineations termed HydroBASINS (Lehner and Grill, 2013). A total of 12 levels of basin subdivisions are available for routing at increasingly larger scales.
River discharge	External runoff estimates provided by the global hydrological model WaterGAP at 0.5° raster resolution (Alcamo et al., 2003; Döll et al., 2003), were spatially downscaled to fit the 500m resolution of HydroROUT (Lehner *et al.,* in prep.). Each river reach includes an estimate of discharge for each month (long-term averages), and one yearly long-term average.
River volume	Based on above discharge estimates and simplistic hydraulic geometry laws (sensu Allen et al., 1994), a first-level approximation of the dimensions of channel width and depth were derived for each river reach. These values were then used to calculate river volumes.
River velocity	The estimates of river velocity associated to the river reaches are a simplistic first-level approximation based on hydraulic geometry laws (sensu Allen et al., 1994) and are directly derived from the long-term average discharge value related to each river reach.
Lakes	I integrated a database of more than half a million lake surfaces larger than 1 km^2 (NASA/NGA, 2003) with HydroROUT. For application in chapter 5, lake volumes were estimated using GIS methods (Messager *et al.,* in prep.), based on correlations with lake surface area and surrounding topography (Pistocchi and Pennington, 2006; Hollister and Milstead, 2010).
Dams and reservoirs	Almost 7000 large dams and reservoirs based on the Global Reservoir and Dam database (Lehner et al., 2011), and more than 3700 future hydropower dams (Zarfl et al., 2014) have been registered to the river reaches of HydroROUT. These dams include estimates about the storage volume of each reservoir, which was used for calculating flow regulation indicators.
Waterfalls	A first-time global map of the location of 4588 major waterfalls and cascades (HydroFALLS; Lehner *et al.,* in prep.) has been integrated in HydroROUT.
River classification	Preliminary river classification data based on hydrological, physio-climatic, geomorphological, chemical, biological, and anthropogenic parameters (Ouellet Dallaire *et al.,* in prep.) were used as a proxy for river habitats in combination with the connectivity indicators.
Floodplains	A new global floodplain and inundation map is currently produced at HydroSHEDS' 500 m pixel resolution by combining global topography information with coarser scale inundation maps derived from remote sensing imagery (Fluet-Chouinard *et al.,* in prep.). A draft map of these floodplains was used as part of the river classifications.

Multi-scale modelling

To better support modelling at multiple scales, I created linkages between all objects—from individual raster cells at the highest resolution to corresponding river reaches, catchments, or entire basins (Figure 1.2). A hydrological nesting approach links each river reach and its catchment to predefined hydrological subunits. These nested subunits enable simple aggregation at the next larger scale (Figure 1.2b). This allows conducting object-to-object routing at increasingly larger scales. The nested watershed approach furthermore allows translating spatial information between scales (aggregation/disaggregation), which is an important element in integrated modelling.

River network, routing and connectivity

River network and spatial graph

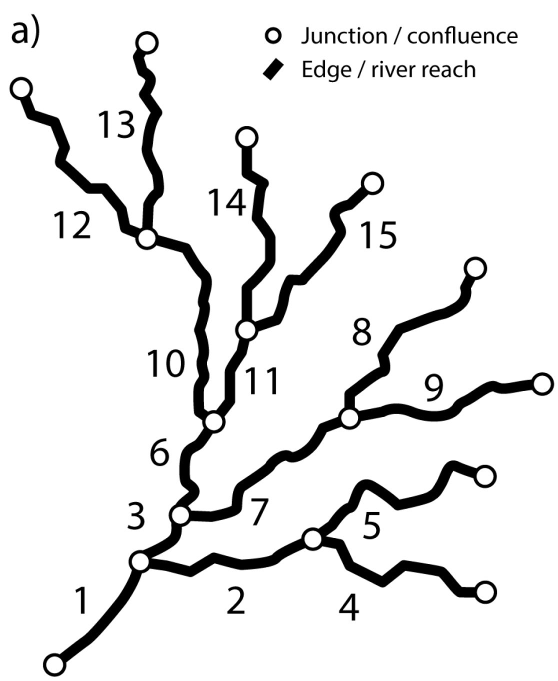
The river network for HydroROUT was generated from HydroSHEDS' flow direction maps using both area (>10 km^2) and discharge thresholds (>0.1 m^3/s). The delineation process resulted in a set of linear vector shapes, which were converted into an ArcGIS Geometric Network (ESRI, 2011), a specific data structure that supports network theory and its application. Geometric networks are typically used to model infrastructures, such as electric utility networks and sewer systems, or transportation networks, but they are also well suited to represent connectivity within dendritic river networks. Geometric networks are collections of line objects (e.g., river reaches/edges) and point objects (e.g., locations of confluences of two river reaches/links), and can furthermore maintain connectivity relationships to multiple layers of information (lakes, dams, and gauging stations, etc.).

HydroROUT can be described as a semi-complex graph-based river routing model (sensu Calabrese and Fagan, 2004), where the structure and spatial relationships of the elements of the river network are expressed as a spatial graph. The basic principles of graph theory consist of the construction of a graph, G=(N,L), where N is a set of nodes (in ecological applications frequently representing distinct habitat patches), connected by links L (which may be attributed with specific characteristics about the type of linkage). In river systems, this terminology can translate to reaches represented as the nodes—linked by river confluences (Segurado et al., 2013), or as links that connect runoff-generating sub-basin nodes (Phillips et al., 2011) depending on the application. In the case of HydroROUT, I follow the approach by Segurado et al. (2013), but define the river reaches as edges (equal to nodes), which are connected through junctions (equal to confluence points or links; see Figure 1.3a).

Lakes were linked to HydroROUT by the lake's flow outlet point defined as the most downstream location in the lake, i.e., the point within the lake having the largest upstream watershed according to HydroSHEDS's flow accumulation map.

Network connectivity

The connectivity information between river reaches and other objects are stored in connectivity tables, the so-called logical network (Figure 1.3b). Unlike traditional approaches that use the "*fromNode—toNode*" concept (Maidment, 2002), HydroROUT stores the object identifier of the downstream reach ("*toNode*") but



b)

Edge	U'stream	D'stream
1	2,3	-
2	4,5	1
3	6,7	1
4	-	2
5	-	2
6	10,11	3
7	8,9	3
8	-	7
9	-	7
10	12,13	6
11	14,15	6
12	-	10
13	-	10
14	-	11
15	-	11

Figure 1.3: Schematic representation of river network as a spatial graph with edges and junctions (a) and logical network table representing upstream and downstream connectivity (b).

additionally stores all upstream object identifiers as a sequence in a separate column. In this way, the connected upstream reaches of one particular reach can be determined with only one database query, and one recursive query will identify all upstream river reaches for a given location most efficiently.

The river network representation is relatively simple compared to more complex networks such as utility networks or social networks. Maximum cardinality, the number of adjacent river reaches connected to another river reach cannot exceed one in the downstream direction—therefore multiple branches of the same river cannot be represented. If the river reach is a sink (inland basin or ocean), cardinality in the downstream direction is zero. In the upstream direction, maximum cardinality can reach eight due to the fact that the underlying flow direction grid allows a maximum of eight connections per cell.

Routing functionality

Geometric networks are computationally optimized for routing and tracing. Connectivity relationships are based on the 'ForwardStar' concept, which is considered the most efficient network representation (Ahuja et al., 1993; Cherkassky et al., 1996). Connectivity based on a 'ForwardStar' also allows flexible routing in both directions, during which the 'cost' or 'resistance' of traversal is being taken into account through network weights. Barriers that stop the routing may also be included. This 'resistance' is important for modelling functional connectivity, where the dispersal and distribution of species is frequently associated with specific river conditions or properties.

In the current iteration of HydroROUT there are six main routing functions implemented, which can be executed with single or multiple objects of origin (Figure 1.4; a-f). During the routing operation, additional processes may be executed if needed, e.g. an in-stream decay function may eliminate contaminant mass of a pollutant as it propagates downstream (described below).

- **Upstream tracing (Figure 1.4a)** searches for all river reaches upstream of one or multiple objects of origin and selects them for further processing. This function was frequently used to extract summary statistics of upstream sections in chapter 3 and 4.

- **Downstream routing (Figure 1.4b)** iteratively identifies the next downstream river reach from one or multiple source locations. This function was used for calculating the RRI index (chapter 3 and 4) and in the routing and accumulation of substances downstream the river network from wastewater treatment plants and other sources (chapters 5 and 6). Downstream routing was also used to calculate the distance from any location to the downstream sink, which was required for statistical and visualization purposes.
- **Upstream and downstream routing (Figure 1.4c)** combines functions a) and b) and selects both upstream and downstream sections. This function is able to determine upstream source areas, and identify potentially affected areas (sinks) downstream in a river network.

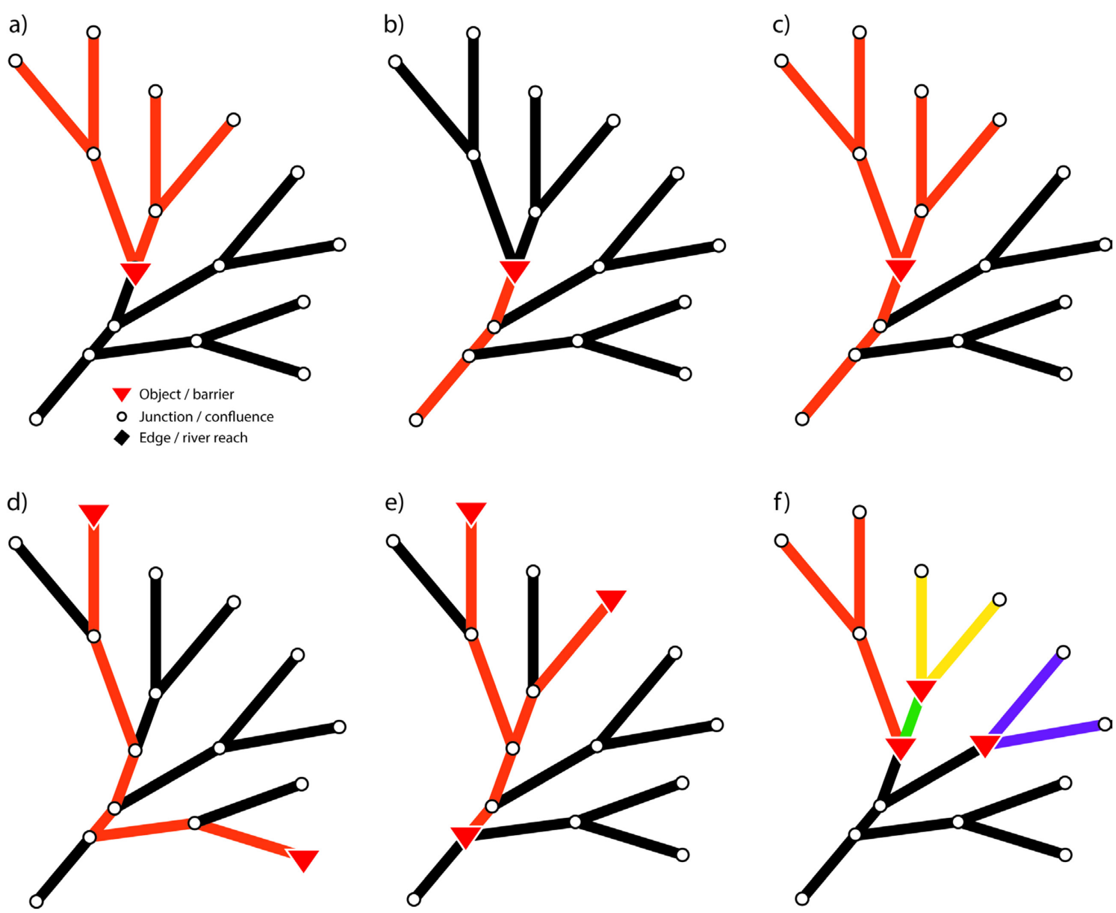


Figure 1.4: Schematic river network illustrating routing functions currently implemented in HydroROUT: a) upstream routing, selecting all reaches upstream of an object/barrier (red triangle); b) downstream routing; c) combined upstream/downstream routing; d) find shortest connected path between two objects; e) shortest path between multiple objects (minimum spanning tree); and f) routing with multiple barriers ('find all connected').

- **Shortest path routing (Figure 1.4d)** finds the path between two network locations, for example to calculate the distance between network locations or to identify possible fish migration routes.
- **Multiple paths routing (Figure 1.4e)** finds the shortest paths between multiple network locations (spanning tree). This function was used to calculate simple species ranges from species occurrence data for migratory fish in the Mekong basin (chapter 3).
- **Routing connected features with barriers (Figure 1.4f)** was used in chapter 3 and 4 to finds all connected features upstream and downstream of dams. This supported the identification of network sections (colored reach groupings), which were statistically analysed as part of the RCI indicator.

Besides the previously described reach-to-reach (or reach-to-lake) routing and tracing, the routing can also occur between larger scale hydrological units, e.g., sub-basins (see levels in Figure 1.2b). As such, researchers can choose the appropriate spatial scale at which they wish to conduct disaggregated spatial analysis in river basins, while maintaining attributes at finer underlying spatial scales.

The advanced implementation of routing in HydroROUT supports a variety of network analysis functions as identified in Peterson et al. (2013), and as such support ecological modelling in river networks from various different perspectives (i.e. geostatistical and process-oriented, etc.).

Network processes

HydroROUT can apply basic network processes during the routing if needed. For example, in chapter 5 and 6, I simulate in-stream first-order decay of pollutants as these are transported downstream, and apply a “reactor-model” to simulate decay in lakes.

In rivers, an exponential decay model diminishes the load passed to the next downstream river reach as $d = e^{-kt}$ where t is the time a plug of water needs to travel through the river reach (based on water residence time), and k is a positive number called the first order rate constant, which determines the environmental decay. Travel time t is derived by dividing river reach length by the average velocity within the river reach. This corresponds to the average retention time in each individual river reach, i.e. the time a plug of fluid needs to travel from the beginning to the end of the river segment. I estimated velocity using the empirically derived formula by Allen et al. (1994):

$$v = 1.07 \times Q^{0.1035} \qquad 1.1$$

where v is the velocity in m/s within the river reach and Q is the bankfull discharge in m^3/s. It must be noted that due to the lack of bankfull discharge estimations in HydroROUT, average discharge is used instead. The next iteration of HydroROUT is anticipated to include attributes of variable velocities both in the HydroSHEDS database and the HydroROUT model, using similar approaches as suggested by Ngo-Duc et al. (2007), Fiedler and Döll (2010), and Verzano et al. (2012). Velocity simulations can further be improved by additional inundation modelling in large floodplains (Yamazaki et al., 2011).

A different process is applied in lakes. During the routing process, lakes and reservoirs are considered objects treated differently than river reaches, primarily due to their different nature representing mixing, flow velocity, depth, and volume. The most simplified way to treat chemical load entering a lake is to let it pass unchanged to the downstream river (worst-case scenario) or completely eliminate the load (best-case). A more advanced and commonly used approach is to treat the lake as 'completely stirred reactor' (Butkus et al., 1988; Whiteaker et al., 2006). To enable this concept in HydroROUT, the concentration of constituents in a lake is determined by the inflow concentration, a decay factor, and the volume of the lake (Mihelcic et al., 2010):

$$PEC(S)_{LAKE,OUT} = PEC(S)_{LAKE,IN} \times \frac{Q_{LAKE}}{Q_{LAKE} + kV_{LAKE}} \qquad 1.2$$

where $PEC(S)_{LAKE,OUT}$ is the concentration of substance s at the lake outlet, $PEC(S)_{LAKE,IN}$ represents the inflowing concentration from all river sources, Q_{LAKE} is the discharge in m^3/s entering the lake, k is the lake's decay constant, and V_{LAKE} is the lake volume, estimated by a geospatial volume estimation model (Messager *et al.,* in prep.).

A third example for a process applied during the routing is the accumulation of distributed sources in the river network in the downstream direction, which consists of multiple downstream routing operations in sequence sorted from upstream to downstream reaches. I use these functions frequently in chapter 3, 4, 5, and 6, where the cumulative discharge volume of wastewater, the cumulative mass of pollutants, or the cumulative storage capacity of dams are necessary elements to support my applications.

Integrated modelling support

The above-described characteristics make HydroROUT well suited as a framework to conduct integrated eco-hydrological modelling, especially geared towards assessing anthropogenic impacts. Throughout the applications in chapters 3 to chapter 6, I demonstrate how diverse objects and processes can be linked and interact within HydroROUT. For example, lakes linked to the river network, interact with anthropogenic contaminants released into the environment, which trigger specialized decay functions during the routing process (chapter 5 and 6). In chapter 3 and 4, I demonstrate how ecological information (river classifications and species range models) as well as hydrological flow data can be linked for application within an integrated assessment of dam impacts.

Connecting statement (Ch. 2)

Chapter 1 provided the background and historical basis for global hydrological and routing models, and their applications, including the current state of the art in river routing. I also summarized the main challenges of global hydrological and routing models related to inadequate spatial resolution, inflexible data structures, missing multi-scale support, limited support for connectivity, and the lack of integrated modelling, and I gave an overview of HydroROUT's development, properties and features.

The next chapter provides a deeper illustration of the five identified challenges for large scale hydrological and river routing models, and discusses improvements to alleviate these challenges. Taking into consideration each of these challenges, I designed and assembled the first version of HydroROUT and conducted pilot applications for dam impact assessments and environmental impact assessments. This step provided an important prerequisite for a more comprehensive application of the HydroROUT framework in chapters 3 to 6.

2 Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems

Hydrological Processes **27** (2013) 2171-86

Bernhard Lehner[1]*, Günther Grill[1]*

* equally shared first authorship, see contribution statement above

[1] Department of Geography, McGill University, 805 Sherbrooke Street West, Montreal, Quebec, H3A 0B9, Canada

Abstract

Despite significant recent advancements, global hydrological models and their input databases still show limited capabilities in supporting many spatially detailed research questions and integrated assessments, such as required in freshwater ecology or applied water resources management. In order to address these challenges, the scientific community needs to create improved large scale datasets and more flexible data structures that enable the integration of information across and within spatial scales; develop new and advanced models that support the assessment of longitudinal and lateral hydrological connectivity; and provide an accessible modelling environment for researchers, decision makers, and practitioners. As a contribution, we here present a new modelling framework that integrates hydrographic baseline data at a global scale (enhanced HydroSHEDS layers and coupled datasets) with new modelling tools, specifically a river network routing model (HydroROUT) that is currently under development. The resulting 'hydro-spatial fabric' is designed to provide an avenue for advanced hydro-ecological applications at large scales in a consistent and highly versatile way. Preliminary results from case studies to assess human impacts on water quality and the effects of dams on river fragmentation and downstream flow regulation illustrate the potential of this combined data-and modelling framework to conduct novel research in the fields of aquatic ecology, biogeochemistry, geo-statistical modelling, or pollution and health risk assessments. The global scale outcomes are at a previously unachieved spatial resolution of 500 m and can thus support local planning and decision making in many of the world's large river basins.

2.1 Introduction

Water is a vital necessity to sustain life on Earth, and many Earth system processes depend on the spatial and temporal distribution of freshwater resources. Not surprisingly, the hydrological cycle and its underlying physical processes have been studied intensely over the past decades, mostly with a focus on small- to medium-sized catchments. In recent years, however, the increasing pressures on freshwater resources from a multitude of complex and interacting factors that span all scales led to a growing recognition of the importance of large river systems. Consequently, many research applications require hydrological or water resources information at regional to global extent. Prominent examples include the estimation of future climate and global environmental change effects related to floods, droughts, water supply, or hydropower generation; the analysis of global biogeochemical cycles, carbon and nutrient budgets; the sustainable management of international freshwater resources; the estimation of possible limits in global food production due to constrained water availability; the systematic planning for freshwater biodiversity conservation; or the assessment of regional health risks due to water-borne diseases or water quality issues.

As in situ measurements of many hydrologic variables at large spatial and temporal scales are difficult and expensive, global hydrological models (GHMs) have emerged as our preferred research tool to analyze current and future world water resources. Haddeland et al. (2011) reviewed eleven global water models that have recently been studied in the Water Model Intercomparison Project. They included five global land surface models, which historically have been designed for improving the vertical water and energy balances and their representation in atmospheric circulation models. In contrast, GHMs are typically tailored towards a more detailed assessment of global freshwater resources by simulating explicit processes of surface runoff generation and horizontal flow routing. Major goals of these models include the estimation of water availability, water use, and/or water scarcity at global or regional scales (e.g., Vörösmarty et al., 2000b; Alcamo et al., 2003; Arnell et al., 2004; Oki and Kanae, 2006; Rost et al., 2008; Hanasaki et al., 2010; Siebert and Döll, 2010).

More specialized GHMs consider the transportation of constituents other than water, e.g. sediment, contaminants, and nutrients, and simulate in-stream processes during the routing process. Such models improve our understanding of biogeochemical cycles and budgets in surface waters, for example the role of rivers in processing and transporting carbon, phosphorus, nitrogen, and silica to the ocean (Alexander et al., 2009; Alvarez-Cobelas et al., 2009; Beusen et al., 2009; Mayorga et al., 2010; Aufdenkampe et al., 2011). Similar models can be used to predict the transport of sediments in the river network to help understand the effects of dams on sediment retention and on coastal and delta geomorphology (Vörösmarty et al., 2003; Syvitski et al., 2009; Kummu et al., 2010). Chemical fate models have been developed to conduct risk assessments for substances from wastewater treatment plants or other sources (Feijtel et al., 1998; Wang et al., 2000; Anderson et al., 2004), yet few have been applied at larger scales (Pistocchi et al., 2009). In a recent integrated approach, Vörösmarty et al. (2010) conducted a study that combines various disciplines, models, and data sources in an attempt to holistically assess anthropogenic threats to global scale freshwater biodiversity and river systems.

Despite the increased attention, hydrological modelling at large scales has traditionally been hampered by a number of issues, including general limitations in our knowledge and understanding of the underlying processes for many regions of the Earth; incomplete, inconsistent, or highly uncertain data collections; a lack of spatial integration between models and datasets; and the difficulty to create models that support multi-scale or coupled approaches. Not surprisingly, the improved monitoring and modelling of global land surface hydrology at high spatial and temporal resolutions has been named as one of the 'grand challenges' for assessing the Earth's freshwater budget, and a call to strengthen engagement in this effort has been made upon the international hydrologic community (Wood et al., 2011). To achieve this goal, new and enhanced global datasets and tools describing the geographical distribution of hydrological features, characteristics, and processes will be required as this information is currently often unavailable or at low quality. The current research status is even more incomplete when looking at large scale effects of human alterations to the water system, such as the impact of global reservoir and dam constructions on downstream river ecosystems (Lehner et al., 2011). While significant progress has recently been made in the development of global hydrographical data (see Data section below), many challenges remain for model development and improvement. In particular, a new generation of integrated models is required that support the linkage of hydrological processes with other fields, such as ecology, biogeochemistry, and water management (Aspinall and Pearson, 2000). In this paper, we argue that there is currently no comprehensive framework that can fully support integrated hydro-ecological modelling at the global scale; yet the building blocks of such a framework already exist. We will illustrate the current status and outline new directions for global scale hydro-ecological modelling to study the world's large river systems. We propose an approach that combines existing and newly developed global scale hydrographic baseline data with a dedicated river routing model in a Geographic Information System (GIS) environment, and we will present preliminary results of our own model development. Finally, we will show that such a framework can enable a broad range of hydro-ecological applications and operate at scales at which local decision making and management becomes feasible.

2.2 Challenges in global scale hydroecological modelling

In this section, we formulate and briefly discuss five main challenges that need to be addressed in order to improve our ability to integrate global hydrological and ecological modelling approaches. The solutions we offer may also be applicable for other, related applications, such as biogeochemical or sediment transport models.

Challenge 1: Spatial resolution

The current spatial resolution of GHMs is not appropriate for many ecological or environmental management applications because of difficulties to accurately represent stream and watershed attributes and to precisely locate individual objects within the river system. GHMs have been developed at different spatial resolutions, ranging from basin scale lumped models to the most commonly applied 0.5 degree pixel resolution, and various lateral routing schemes have been developed accordingly (Oki and Sud, 1998; Graham et al., 1999; Renssen and Knoop, 2000; Vörösmarty et al., 2000a; Döll and Lehner, 2002). These rather coarse scales have been applied mostly for reasons of technical feasibility (i.e. computational demand) and due to the fact that important input information, foremost climate data, has been offered at these resolutions. In the most recent iterations, GHMs are moving towards 5 (or 6) arc-minute spatial resolution (e.g., Wisser et al., 2010; Schneider et al., 2011) but despite this improvement, two major limitations remain: First, the relatively coarse spatial resolution introduces bias and misrepresentation of hydrological processes and river topology. This can result, for example, in underestimation of river length and travel times (Gong et al., 2009; Verzano et al., 2012) and can lead to subsequent inaccuracies of travel-based attributes, e.g. sediment delivery to the ocean (Vieux and Needham, 1993). Second, and more importantly from a modelling perspective, the integration of local or fine-scale information is difficult, if not impossible. While sub-grid parameterizations allow representation of some hydrological processes, ecologists and water managers have long criticized large scale models for being incompatible with their more localized needs, such as the explicit identification of habitat or flow characteristics for individual tributaries, or the linkage of species information to river reaches and small sub-basins. In raster format, the resolution needs to be particularly high to explicitly model river channels, which are the most important structures controlling hydrodynamics.

A possible solution is to move towards hyper-resolution hydrological modelling (Wood et al., 2011). Global digital drainage networks exist at fine scales, such as HYDRO1k at 1-km resolution (USGS, 2000) or HydroSHEDS at up to 90-m resolution (Lehner *et al.,* 2008; Figure 2.1) and are detailed enough to reference objects such as dams precisely to the corresponding rivers. Gong et al. (2011) demonstrated for two medium-sized basins that routing performance increases with the higher resolution hydrographic data of HydroSHEDS. However, computational demand to execute the necessary runoff-response-functions and cell-response functions at a daily time step was high, and their model has yet to be tested at the global scale.

Given these technical difficulties, an alternative solution towards higher spatial resolution may be offered by integrating raster with vector concepts in the modelling framework (see also next challenge). In a raster environment, higher resolution can only be achieved by increasing the number of pixels, while in a vector environment, a subbasin's geometrical resolution can be improved by simply applying a more detailed polygon outline; in this latter approach, the number of modelling objects remains constant while resolution improves (see Figure 2.2).

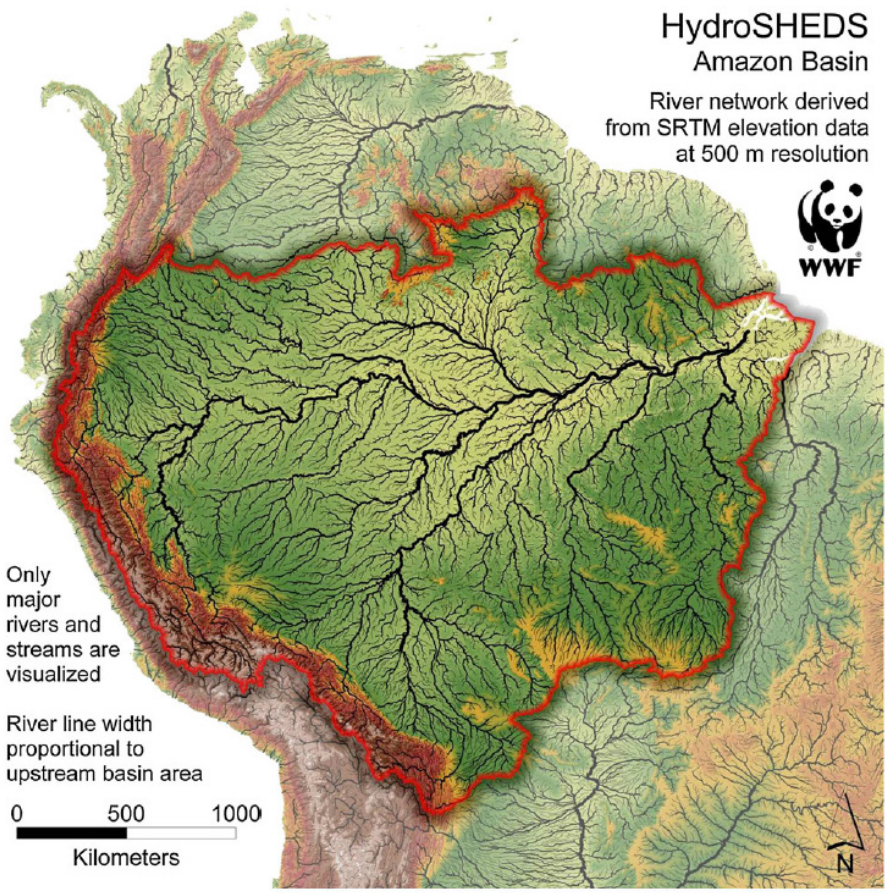


Figure 2.1: HydroSHEDS river network for Amazon Basin at 500m resolution.

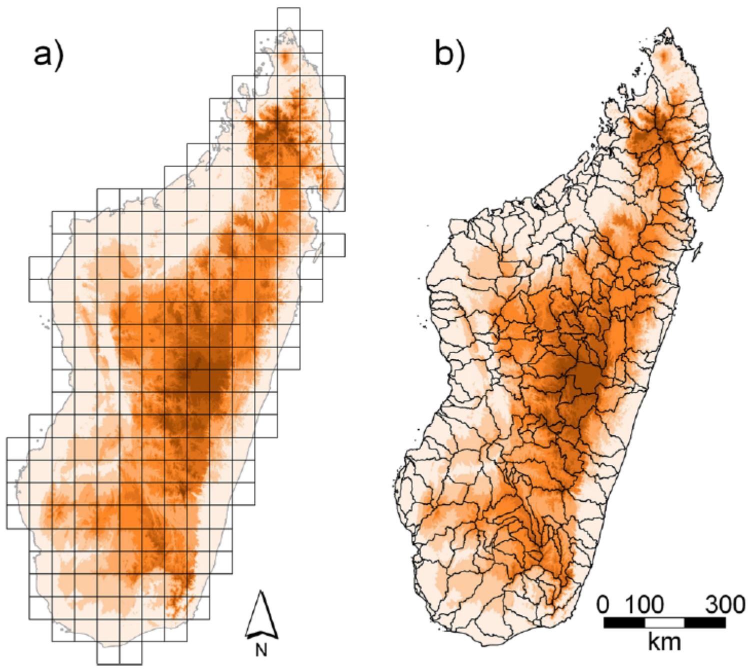


Figure 2.2: Sub-unit representation in half-degree grid format (a) and 500m vector format (b) for Madagascar. There are a total of 250 half-degree pixels (a) versus 282 sub-basin polygons (b). Polygons are derived from the HydroSHEDS database

Challenge 2: Data structure (raster vs vector concepts)

Spatial information in GHMs is typically represented as layers of uniformly sized grid cells, also known as raster datasets. Vector datasets on the other hand consist of objects, represented as points at a certain location, or as a series of point locations chained together to form a line or polygon. The differences between raster and vector concepts in environmental models are frequently debated in the modelling community. Both approaches have advantages that could be harnessed in a hybrid modelling framework.

Raster models have traditionally been the preferred choice of hydrological modellers since the integration of topography in the form of rasterized Digital Elevation Models (DEMs) is a key requirement for many routing algorithms that trace water along drainage direction maps. Furthermore, many ancillary datasets stem from remotely sensed raster information (e.g. land cover, precipitation, etc.), and grids are relatively easy to process in any modern GIS.

A major limitation of the raster concept, however, is the difficulty to integrate objects or information from different sources or formats (such as the discharge time series of a gauging station; a reservoir outline; or a fish migration route) into one common framework, as linkages would have to be established to individual pixels. A vector dataset, on the other hand, can represent multiple objects (points, lines, or polygons) each of which can have multiple attributes associated in a related table. Within a GIS, links between the objects of different vector layers can be created similar to an 'object-oriented' relational database. If such a structure is established, river network routing using the vector structure greatly facilitates the integration of multiple objects from different domains.

Ideally, a modelling framework should support both raster and vector representations of the respective data layers to benefit from the advantages of both. This could be achieved by maintaining fundamental information at high spatial resolution as raster datasets, and to develop tools that allow for an easy transfer or linkage of this information to vectorized river reaches, sub-basins, or point objects. Modellers can use these tools whenever new information needs to be integrated into the model. This conceptual approach is followed, for example, by ArcHydro, a geospatial framework and toolbox within ESRI's ArcGIS software package (Maidment, 2002), which allows the user to create a set of hydrological vector datasets (river network, basins, etc.) from a DEM and to store the information in an object oriented geodatabase.

Challenge 3: Multi-scale approach

A vision to promote hydro-ecological and other integrated models requires means to analyze and simulate processes at multiple spatial scales (Lowe et al., 2006). Yet, difficulties arise from inherent data and modelling characteristics. For example, certain types of geospatial hydrological data, such as runoff coefficients or flow directions, cannot be easily scaled by spatial resampling techniques. Scaling methods for rasterized river networks may provide a solution (e.g., Fekete et al., 2001; Yamazaki et al., 2009), but currently most data layers are developed for a specific resolution, and even small changes in scale require intense data reprocessing or even the development of new datasets from scratch.

Thus tools and concepts need to be developed that support modelling at multiple sales and help transferring information between scales - from individual raster cells at the highest resolution, to river reaches, catchments, or entire basins. This can be achieved by means of hydrological nesting approaches, e.g. by applying hierarchical coding schemes to predefined sub-units, or by allowing accumulation and conglomeration of objects (e.g. sub-basins) based on their topology and connectivity.

Challenge 4: Representation of hydrological connectivity through river network routing

Hydrological connectivity is an important concept in both hydrology and ecology. Flows in traditional GHMs that include lateral transport occur on the premise of a river continuum whereby conditions at every location in the river network are influenced by upstream processes (Vannote et al., 1980). A more ecology oriented concept of spatial integration has been developed by Pringle (2003), who defined hydrological connectivity as the flow of mass, energy, and organisms in surface and groundwater. In this definition, hydrological connectivity occurs in both upstream and downstream direction (longitudinal); between rivers, wetlands, and floodplains (lateral); and between groundwater, surface water, and the atmosphere (vertical).

The advanced implementation of hydrological connectivity in a modelling framework is an attractive concept because of its ability to integrate processes from multiple disciplines (such as terrestrial, climate, ecological and hydrological); because of its easiness to represent spatially explicit topology information of a river network and watershed landscape (up- and downstream relations); and because of its intrinsic ability to establish connections across multiple scales - from point locations and river reaches to nested catchments and entire basins. Hydrologic connectivity can be modeled in various ways. We propose that for hydro-ecological applications, models should at a minimum (i) include the possibility to conduct routing of water and substances along the hydrological flow paths; (ii) represent lakes, dams, floodplains, and wetlands in addition to river reach routing; and (iii) be based on a powerful tracing algorithm for up- and downstream routing to represent both active and passive dispersal of matter or species. We are currently not aware of models that can fully implement such advanced up- and downstream routing in a global framework.

The term ‘routing’ generally refers to the simulation of transport processes over space and time. For example ‘flow routing’ simulates the movement of water over the landscape and translates runoff into river discharge by passing the runoff generated in a specific landscape unit (e.g. sub-watershed or pixel) to the next downstream unit. A distinction is typically made between hydrodynamic routing and hydrological routing (Maidment, 1993). In hydrodynamic routing, flow is described by the Saint-Venant equations of mass and momentum conservation. The first equation governs the translation of mass between sections of the river, while the second relates changes in momentum to the applied forces (in particular pressure gradient, convection, acceleration). In hydrological routing the Saint-Venant equation of mass conservation still applies, but the momentum term is replaced by empirical parameters.

In GHMs such as TRIP (Oki and Sud, 1998; Ngo-Duc et al., 2007), HYDRA (Coe, 2000), WaterGAP (Alcamo et al., 2003; Döll et al., 2003), WBM/WTM (Vörösmarty et al., 1989; Fekete et al., 2006), and LPJmL (Rost et al., 2008), simple hydrological routing is implemented, where flow is passed along storages (e.g. rivers, lakes) in a linear sequence, and residence time is often modeled by assuming constant flow velocities. For variable representation of velocity, the introduction of simple parameters based on slope (or more comprehensive ones based on slope and friction) are used (Coe et al., 2009; David et al., 2011; Verzano et al., 2012). Basic in-stream processes, such as accumulation and solute transport along the river network, have also been implemented in these types of GHMs (e.g., Donner et al., 2002).

Hydrodynamic processes such as backwater effects and lateral flooding have critical implications on ecosystem functioning, and the inclusion of hydrodynamic principles allows for more advanced routing models. A recent study for the Amazon River Basin employed a diffusive wave routing approach in the CaMa-Flood model (Yamazaki et al., 2011) to demonstrate that floodplains have strong attenuating effects on flood waves and flood peaks in large basins. These improvements, however, come at the cost of additional and more complex computational efforts, model design, preparation, and parameterization. Thus, for the primary goal of a first-level integration of baseline hydrology with ecological and other applications at the global scale, we believe that the full implementation of hydrodynamic processes, though desirable, is currently not a stringent priority.

Beyond the type of routing, the direction of movement plays an important role. Multi-directional routing approaches, also characterized by the term 'tracing', are procedures that select a set of network elements based on certain connectivity rules, and then process the selected set in a predefined (directional) sequence. Within such a framework, many hydro-ecological connectivity applications can be supported, for example by simulating the effects of dams on plant dispersal through hydrochory and on the resulting riparian flora (Andersson et al., 2000) the distribution of fish species in river and lake networks in response to environmental factors and anthropogenic pressures (Lassalle et al., 2009) and the fragmentation of river networks and impedance of ecosystem connectivity through dams (see Applications section below). Advanced tracing can also help to better understand the structure of and processes in river networks through geostatistical modelling: with the emergence of detailed hydrologically connected network datasets, a new class of geo-statistical tools can be developed, which uses distances along a curvilinear flow path, instead of the Euclidean distance space (Ganio et al., 2005). Network tracing using weights and barriers, with subsequent routing, can provide a powerful framework for a wide range of applications that involve hydrological connectivity.

In conclusion and to widen the scope of routing, tracing, and connectivity in hydrological models, for this paper, we define the term 'river network routing' following the hydrological connectivity concept of Pringle (2003) as 'the simulation of movement of energy, material and species within a hydrological object space, based on flow routing, tracing and in-stream processing.'

Challenge 5: An integrated data and modelling framework

Different GHMs operate at different resolutions, have been designed for different purposes, use their own specific routing schemes, and are based on different input for the representation of climate or topography. Thus, the coupling of these GHMs with ecological models or databases, such as fish species distributions, requires extensive individual data processing and adjustments in order to align spatial resolutions and data formats. We here propose to design a more integrated data and modelling framework which facilitates an easier linkage between hydrological and ecological information and is geared towards hydro-ecological research, applications, and management. A key characteristic of this approach is to utilize a harmonized database of hydrographic baseline information (i.e. river network, sub-basin delineations, and linked features such as dams, lakes, and gauging stations) and to develop assessment tools that couple hydrological model results and ecological information within this data framework. Using a common data space has the advantage that, for example, species data can be easily mapped to the same units as hydrological information, which facilitates the analysis of impacts, coupled processes, or feedbacks. In the following two chapters, we describe and outline the characteristics of existing global hydrographic baseline data, as well as a new approach to utilize these data within a custom-made river network routing model.

2.3 New Global Hydrographic Data Developments

Large scale hydrological modelling critically depends on the availability of adequate input data. Climate and land surface data and parameterizations are key elements for any water balance model, while many integrative hydro-ecological applications require complementary data sources, such as locations and amounts of human water use, the origin of point- or non-point-source pollution, and biodiversity and species information. Additionally, hydrologic measurements such as provided by gauging stations are important for validation purposes.

Tremendous improvements have been made over the past years in the availability, quality, and resolution of large scale hydrographic datasets, not least due to the increased operational monitoring of the entire Earth surface via satellite remote sensing. There is a long list of recent data developments for many hydrology-relevant themes, including (to name only a few): new land cover data derived from remote sensing (Loveland et al., 2000; Bontemps et al., 2010) historical and current land use data (Ramankutty and Foley, 1999; Ellis and Ramankutty, 2008; Monfreda et al., 2008; Portmann et al., 2010; Goldewijk et al., 2011; MacDonald et al., 2011) lake, reservoir, and wetland inventories (Lehner and Döll, 2004; Lehner et al., 2011) irrigation maps (Siebert et al., 2005) water use estimates (Wisser et al., 2008; Döll et al., 2009) soil parameterizations (FAO et al., 2012), and many more.

There are various data portals and compilations available online to serve and distribute these data, e.g. the Digital Water Atlas of the Global Water System Project (http://atlas.gwsp.org/). Vörösmarty et al. (2010) offer a suite of 23 maps of driver sources at 0.5° pixel resolution globally which were part of their study on threats to global river systems (http://www.riverthreat.net/). We also increasingly utilize indirect measurements to derive hydrological variables, such as the derivation of discharge or lake volume changes based on high-accuracy, real-time altimetry measurements of water surfaces and other remote sensing information (Sahoo et al., 2011; Seyler et al., 2013). Even the measurement of incremental changes in the Earth's gravitational field as provided by the GRACE project enabled us to interpret changes, e.g. in groundwater storage or snow water equivalent at a planetary scale (Llovel et al., 2010)

Among the many input data layers, special attention is often placed on hydrographic baseline information in the form of river networks and watershed boundaries as they form the backbone or 'fabric' of the modelling framework. Examples of highly advanced regional versions include the US National Hydrography Dataset (NHD; developed by US EPA and USGS; http://nhd.usgs.gov/applications.html), a comprehensive set of digital geospatial data about surface water features such as streams, rivers, and lakes; or its successor NHDPlus, which incorporates the US National Elevation Dataset, and the Watershed Boundary Dataset. The Australian Hydrological Geospatial Fabric (Atkinson et al., 2008) has been developed as a knowledgebase of the features within the Australian water system and their interactions, such as catchments, streams, aquifers, floodplains, storages, and wetlands. At the European scale, Vogt et al. (2007) created a consistent database of drainage networks and catchment boundaries for hydrological assessments and reporting within the Water Information System for Europe.

At the global scale, however, there are only limited sources for seamless high-quality hydrographic information. Data for many large international river basins are still patchy, and remote areas are often poorly mapped. Also, the hydrographic information is required in strictly defined digital formats that allow for flow routing along streams and watershed identification. HYDRO1k (USGS, 2000) offered a first version of such information, yet its quality shows strong regional variation. Below, we discuss the HydroSHEDS database (Lehner et al., 2008) which represents the most recent attempt to fill this data gap.

HydroSHEDS is a hydrographic mapping product created by World Wildlife Fund that provides river and watershed information for regional and global-scale applications in a consistent format (Lehner et al., 2008). It offers a suite of geo-referenced datasets at various resolutions ranging from 3 arc-second (approximately 90mat the equator) to 30 arc-second, including river networks, watershed boundaries, and drainage directions. HydroSHEDS is based on high-resolution elevation data obtained during NASA's Shuttle Radar Topography Mission (SRTM) in February 2000. The extent of HydroSHEDS is near-global, currently only excluding regions above 60northern latitude due to the lack of SRTM source data; the global extent is scheduled to be completed by inserting alternative elevation data within 2013. The data is available to the scientific community at http://www.hydrosheds.org.

Besides its core layers, HydroSHEDS is currently undergoing expansion to include a suite of attribute layers and to establish linkages to auxiliary datasets. Some efforts are already completed, some are in progress for release within the next year or two, and some are in proposal stage; Table 2.1 provides an overview of the prime developments. Consistency between the layers is ensured in terms of spatial alignment, and quality indicators are provided where possible. For example, the point locations of gauging stations have been snapped in a best-fit process to HydroSHEDS pixels and corresponding river reaches in a manually controlled and supervised process, and uncertainty has been assessed via discrepancies between reported and modeled watershed areas.

Table 2.1: Recently completed and planned new data layers of the HydroSHEDS database. The attribute data is assigned to the river or sub-basin network at 15 arc-second (500m) resolution.

Layer	**Description**	**Status***
Global Reservoir and Dam (GRanD) database	Coordinated by the Global Water System Project (GWSP) and based on a variety of sources, the locations of nearly 7000 of the world's largest reservoirs were georeferenced, and attribute data were compiled, including storage capacity and main purpose (Lehner et al., 2011). Corresponding dams are linked to the HydroSHEDS stream network via their coordinates. The data is available at http://www.gwsp.org/85.html.	completed
GRDC gauging stations	The location of more than 7000 gauging stations (provided by the Global Runoff Data Center, Koblenz, Germany) were verified and co-registered to the HydroSHEDS river network. The data is available at http://grdc.bafg.de/.	completed
Sub-basin delineation	Hierarchical nesting of sub-watersheds can support multi-scale hydro-ecological analyses (Fürst and Hörhan, 2009). HydroSHEDS has been enhanced by a vectorized watershed layer that subdivides basins into units of approximately 130 km² area. A coding scheme (following the 'Pfafstetter' concept; Verdin and Verdin, 1999) allows for topological up- and downstream queries as well as hierarchical aggregation.	completed
Discharge	Estimates of long-term (1961-90) monthly discharge averages are derived through a downscaling procedure from the 0.5º resolution discharge layer of the global integrated water model WaterGAP (Alcamo et al., 2003; Döll et al., 2003). Individual river reaches can be distinguished down to stream sizes of 1 l/s average discharge. Although no global quality assessment has been completed yet, visual inspections show realistic results and patterns.	Release after validation in 2013
Habitat volume	Based on discharge estimates and simplistic hydraulic geometry laws (sensu Allen et al., 1994), a first-level approximation of the dimensions of channel width and depth will be derived for each river reach. These values are then used to calculate habitat volumes (i.e. in-stream habitat space).	2013
Names	River and sub-basin names will be provided.	2013
Global inundation mapping	River floodplains and inundated areas provide critical aquatic habitat for many species, yet there is currently no consistent, high resolution map of inundation extents available at a global scale. A new global floodplain and inundation map is currently produced at HydroSHED's 500 m pixel resolution by combining global topography information with coarser scale inundation maps derived from remote sensing imagery (Prigent et al., 2007).	2013
Lake surface areas and volumes	The extent of open lake surfaces and the amount of water stored in lakes plays an important role for human water supply and many ecosystem processes. Information on lake volumes, however, only exists for the very largest of lakes while the vast majority of smaller lakes have never been assessed in a spatially explicit way (Pistocchi and Pennington, 2006; Hollister and Milstead, 2010). As part of the SRTM project, NASA provided a mask of more than half a million lake surfaces (NASA/NGA, 2003) which will be utilized in combination with statistical approaches and global elevation data to create a first time	2013-2014

	estimate of water volumes for each individual lake. The approach has been successfully tested for a selection of 166 European lakes. Results will be integrated in the HydroSHEDS database.	
Flow velocity	Using discharge, stream gradients, and global estimates of simplified Manning coefficients, Verzano et al. (2012) derived estimates of flow velocities. A similar methodology will be applied to estimate average velocities for the HydroSHEDS stream network.	2013-2014
Global River Classification (GloRiC)	River classifications provide researchers (and water practitioners) with basic modelling and planning units; deliver groups of similar or distinct river types to be applied in analyses of biodiversity patterns or threats; help prioritizing protection or restoration efforts; and support the development of guidelines and regulations for freshwater management purposes. The goal of the GloRiC project is to develop a first-time, high-resolution, spatially explicit global map of river typologies. Different river classes are distinguished based on a variety of criteria, including hydrological, physio-climatic, geomorphological, chemical, biological, and anthropogenic parameters.	2013-2014
Aquatic biodiversity information	Freshwater biodiversity information (i.e. species lists) are currently coupled to the sub-basins of HydroSHEDS by partners such as the International Union for the Conservation of Nature (IUCN), Cambridge, UK. The efforts are anticipated to result in comprehensive and spatially explicit freshwater biodiversity maps at a global scale.	2013-2014
Global waterfall mapping	Waterfalls and cascades are important indicators of natural discontinuities in the river network. A first-time global map of the location of major waterfalls and cascades is currently created through a combination of GIS procedures and extensive manual investigations. The waterfalls and cascades will be linked with the river reaches of HydroSHEDS to support studies of the longitudinal connectivity or discontinuities along the river network.	2014
Water temperature	Water temperature is a critical environmental factor used to distinguish different types of river habitats or to assess the status and quality of riverine ecosystems. Using established modelling approaches, temperature ranges can be derived based on combining global air temperature regimes with the flow routing scheme and discharge estimates of HydroSHEDS for each reach of the global river network.	Proposal stage
Urban water use patterns	Cities play a special role in terms of water management as they represent locations of highly concentrated water demand within the river network and their wastewater discharge can compromise the downstream freshwater quality. To allow managers to explicitly include cities in their large scale planning, a first-time global assessment of the main water supply and use patterns will be compiled for the world's largest cities. The location of each city will be co-registered to HydroSHEDS to enable studies regarding the impact of cities on the global river network.	Proposal stage

* years indicate expected completion date (end of year)

2.4 Development of a global river routing model (HydroROUT)

In nearly all GHMs, the implemented routing models operate on raster data. Only recently, some large scale models started to shift towards a vector environment (e.g., Paiva et al., 2011) or use advanced sub-grid and/or hybrid structures (e.g., Yamazaki et al., 2011). We here propose to continue this transition towards a vectorized model framework at a fully global extent, and to implement versatile, multi-directional routing based on point, line, and polygon objects, such as gauging stations, river reaches, and sub-basins. One advantage of this approach is that vector routing can significantly reduce the required computational resources as compared to cell-based models because of the inherently simple 'object-to-object' routing scheme. In a cell-based approach, large homogeneous objects, such as lakes, consists of hundreds or even thousands of grid cells, each of which is treated as an individual 'cell object', which increases the number of processing steps - often without added benefit.

We are currently developing a new river network routing model (HydroROUT; Grill *et al.,* in prep.) which provides vector-based routing capabilities and is fully integrated in the widely used GIS software package ArcGIS (ESRI, 2011). Figure 2.3 presents a schematic overview of the implementation of HydroROUT within the HydroSHEDS data framework. Only a small number of vector routing models have been developed so far (e.g., Feijtel et al., 1998; Wang et al., 2000; David et al., 2011; Paiva et al., 2011), and HydroROUT is conceptually similar to that of Whiteaker et al. (2006), who developed a processing engine - the 'schematic processor' - to accomplish basic river routing for vectorized river networks derived from ArcHydro.

The first function of HydroROUT is to establish hydrologic connectivity which is achieved by creating links within and between river reaches and sub-basins following the basic principles of Whiteaker et al.'s routing scheme. However, while their 'schematic processor' uses linear vector networks with connectivity relationships built via traditional attribute tables (FromNode-ToNode adjacency), our network is based on the concept of a 'geometric network', i.e. a directed network graph located in an ArcGIS geodatabase. Geometric networks are normally used to model infrastructures, such as electric utility lines and sewer systems, but they are also well suited to represent connectivity within dendritic river networks. Geometric networks are collections of line objects (e.g. river reaches) and point objects (e.g. locations of confluences of two river reaches) that possess a connectivity relationship based on the coincidence of the start- and endpoints of the river reaches. The connectivity information between river reaches and other objects are stored in connectivity tables, the so-called logical network. River reach geometries can thus be treated as individual elements for use in tracing and flow operations.

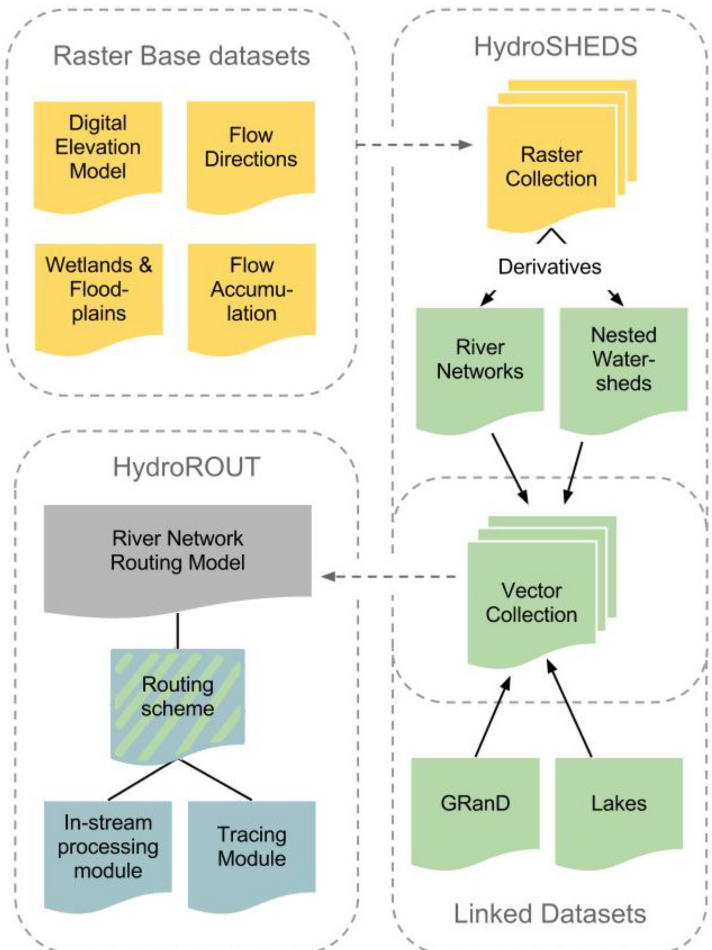


Figure 2.3: Overview of datasets, concepts, and models within the HydroSHEDS framework. HydroROUT is linked to the collection of vector datasets of HydroSHEDS through its routing scheme. The routing scheme currently consists of tracing and in-stream processing modules.

Geometric networks are optimized for routing and tracing. Connectivity relationships are based on the 'ForwardStar' concept, which is considered the most efficient network representation (Ahuja et al., 1993; Cherkassky et al., 1996). It is due to this effectiveness that we are able to conduct river routing at the global scale with several million objects on a single processor. Connectivity based on a 'ForwardStar' also allows flexible routing in both directions, during which the 'cost' or 'resistance' of traversal is being taken into account through network weights, and barriers that stop the trace may be included.

The second function of HydroROUT's processing engine is the routing of substances downstream the river network, which may include the accumulation of mass from different distributed sources in the river network (e.g. wastewater treatment plants) and/or constant or time dependent decay functions, which diminish the substance to be accumulated gradually along its path. For example, for the routing of substances such as nutrients and pollutants, the dominant dilution mechanism is advection, which can be effectively modeled using stream length, velocity, discharge, and a decay function (Pistocchi et al., 2010). A 'plug-flow' model has been chosen similar to Feijtel et al. (1998) and Whiteaker et al. (2006), which is an adequate and frequently used approach in routing at the river reach level (Chapra, 1997; Anderson et al., 2004; Pistocchi et al., 2009). If no biogeochemical processes diminish the substance mass while traveling downstream (e.g. decay, nutrient uptake), then the value received from upstream is equal to the value passed downstream. If, however, dissipation of a biogeochemical substance is assumed, the substance is expected to decrease either by a 0th-order decay function (i.e. a constant decay amount) or at a rate proportional to its value, and can therefore be represented through an exponential decay model.

During the routing process, lakes and reservoirs are considered objects treated differently than river reaches, primarily due to their different nature representing mixing, flow velocity, depth, and volume. Lakes are more appropriately modeled as 'completely stirred reactors' (Butkus et al., 1988; Whiteaker et al., 2006). The concentration of constituents in a lake is determined by the inflow concentration, a decay factor, and the volume of the lake.

As for future developments, we consider lateral routing into floodplains and wetlands an important next step. Beyond their role for regulating high discharges, floodplains and wetlands are ecosystems with increased biogeochemical activity. For example, deposition of nutrients and pollutants attached to sediments during a flood removes those substances from the river temporarily (or permanently), which allows for longer decay or uptake times and hence alters downstream biogeochemical budgets (James et al., 2008). Another future development includes the application of the routing model at different spatial scales, which has rarely been attempted in global models. Beighley et al. (2009) use basin subdivisions calculated by the Pfafstetter method (Verdin and Verdin, 1999) to represent the land surface with varying degrees of resolution to derive 'irregular computational grids'. In HydroROUT, we plan to enable routing at the sub-basin scale and use the connectivity information between the sub-basins as the routing directions. A total of 12 levels of basin subdivisions are available in the HydroSHEDS database. Depending on the complexity of the model, researchers can choose the appropriate spatial scale at which they wish to model, while maintaining attributes at finer spatial scales.

HydroROUT is currently running as a beta version and required extensive data pre-processing and model development. A global river network with routing capabilities was created based on the 15 arc-second DEM of HydroSHEDS. River reaches are the finest scale on which HydroROUT operates, but routing or topology queries can also be performed at the level of sub-basins, e.g. by identifying all sub-basins upstream of a given location. At the 500m resolution, there are more than 17million global river reaches with an average length of 3.6 km, and more than 1 million pre-defined sub-basins with an average area of 30 km^2.

In the current version of HydroROUT, discharge is derived by accumulating land surface runoff along the river network, yet the underlying simulation of runoff generation is not performed within the model itself. Instead, we employ decoupled, external runoff estimates provided by the GHM WaterGAP at 0.5 raster resolution, which we spatially downscale to fit the 500m resolution of HydroSHEDS (see Table 2.1; for advantages and disadvantages of this model design see Discussion section below). In a first validation of the downscaling results, we compared the discharge of 73 randomly distributed stream gauges for the provinces of Ontario and Quebec from the Canadian HYDAT network (Environment Canada, 2012) to the discharge of the closest HydroROUT river reaches. To avoid spatial mismatch, stations were selected based on agreement between reported and simulated catchment areas and only those with a discrepancy of less than 10% were accepted. For long-term (1961-90) annual average discharge (Figure 2.4a), we found a very strong correlation (adjusted R^2 of log-linear model = 0.982; $p<0.0001$). We also conducted a comparison for low-flow conditions as these are commonly used in risk assessments. We aimed to test an index similar to Q90, i.e. the discharge that is exceeded at 90% of time. However, as the climatic input of the underlying WaterGAP model is given in monthly time steps, daily flow estimates, as typically applied for Q90, are not realistic. Instead, we calculated the long-term (1961-90) flow regime and then adopted the lowest month of the year as a first-level proxy for low-flow conditions. The comparison between modeled and observed values (Figure 2.4b) revealed a slightly lower correlation than for average flows (adjusted R^2 = 0.936; $p<0.0001$).

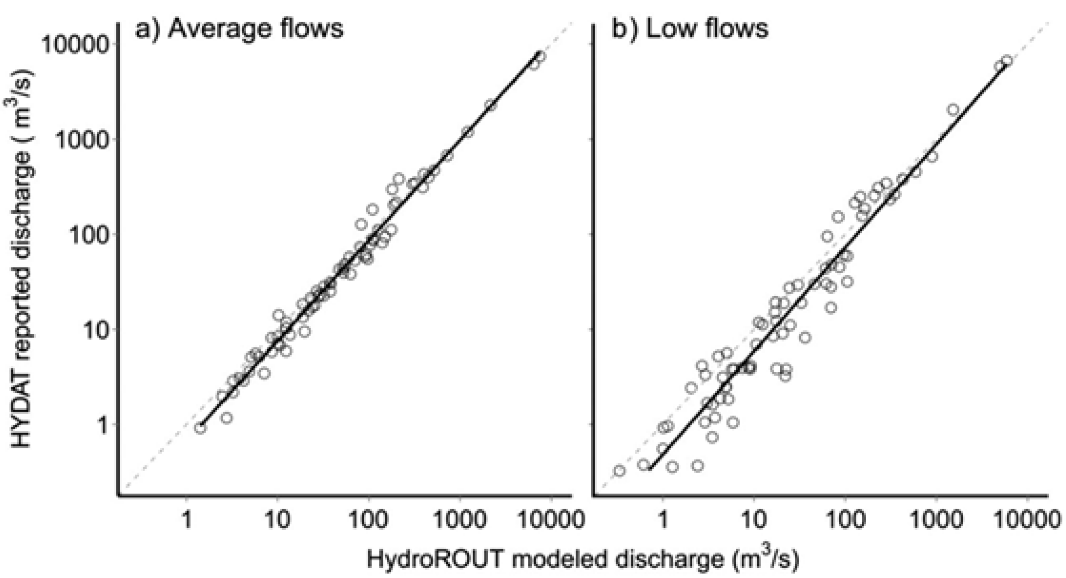


Figure 2.4: Comparison of discharge observations (as reported at 73 HYDAT gauging stations in Ontario and Quebec) and model results (HydroROUT; based on WaterGAP runoff estimates) for long-term (1961-90) average flows (a) and lowest monthly flows b). For more explanations, see text.

Despite these good findings, it should be noted that some results contained significant errors, in particular for smaller streams, and that the model showed a tendency to overestimate low flows. More complex discharge parameters, such as the shape of the annual flow regime, inter-annual variability, monthly time series, or extreme events, are yet to be evaluated. Nevertheless, long-term averages and low flows can provide a first avenue for new applications.

2.5 Applications

In this chapter, we illustrate three case studies performed with HydroSHEDS data and the HydroROUT model that demonstrate the capability of tracing operations; vector routing using river reaches, lakes, and dams; as well as accumulation and decay functions.

2.5.1 Degree of regulation from dams at a global scale

The alteration of the downstream river flow regime is widely recognized as one of the main adverse environmental impacts of dam and reservoir construction (Bunn and Arthington, 2002). In the absence of operational dam release rules, the proportion of a river's annual flow that can be withheld by a reservoir or cluster of reservoirs, i.e. the degree of regulation (DOR), can serve as a first-level approximation of the potential impact on flow regulation (Lehner et al., 2011). A high DOR value indicates an increased probability that substantial discharge volumes can be stored throughout a given year and released in a managed (i.e. non-natural) behaviour. Following this approach, DOR values have been analyzed in a recent global study by Lehner et al. (2011). For this purpose, the new Global Reservoir and Dam database has been coupled with the HydroSHEDS river network (Table 2.1), and DOR values were calculated as 'total upstream storage capacity divided by total annual flow volume' for every reach of the river network (Figure 2.5). Results show that 7.6% of the world's rivers with average flows above 1 m^3/s are affected by a cumulative upstream reservoir capacity that exceeds 2% of their annual flow. The impact is highest for large rivers with average flows above 1000 m^3/s, of which 46.7% are affected. In a related assessment,

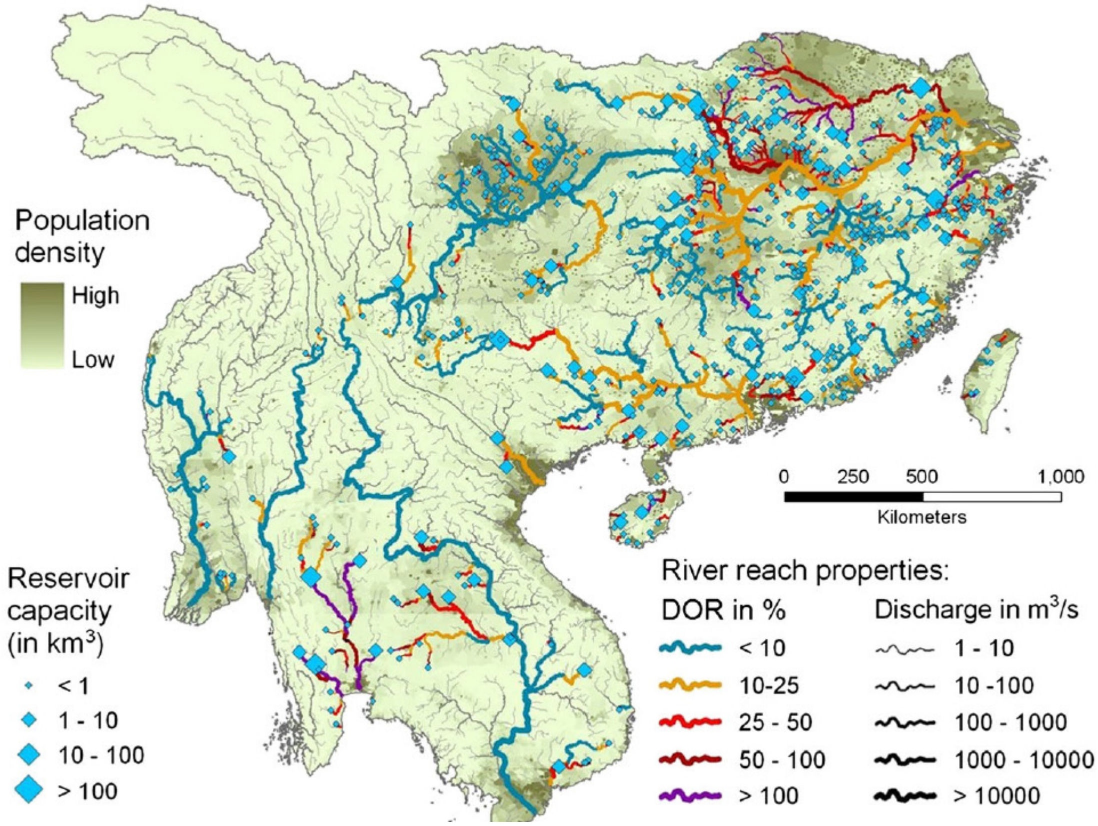


Figure 2.5: Degree of regulation (DOR) for Southeast Asia. The DOR ratio measures the total upstream storage capacity divided by the long-term average discharge at each river reach. Discharge estimates are taken from HydroSHEDS database; reservoir locations and storage capacities are taken from the Global Reservoir and Dam (GRanD) database (Lehner *et al.,* 2011); background population density is from GRUMP database (http://sedac.ciesin. columbia.edu/gpw/)

Richter et al. (2010) have further analyzed these results to estimate the global number of potentially affected people living downstream of dams. These examples demonstrate the capability of river network routing using basic accumulation procedures.

2.5.2 Ecosystem fragmentation from dams in the Mekong River Basin

A new study of ecosystem connectivity at the scale of the entire Mekong River Basin (Grill *et al.,* in prep.) provides an example where river network routing includes extensive tracing operations. Using HydroROUT, we calculated the individual and cumulative impact of 84 proposed dams on ecosystem connectivity in the Mekong Basin (Figure 2.6). The model used tracing operations to distinguish network fragments, calculated statistics for each fragment (such as habitat volumes and number of connected ecosystems), and finally derived several cumulative indices, including the Dendritic Connectivity Index (Cote et al., 2009). The overall result is an index-based ranking for the individual dams, which may provide guidance for decision makers wishing to include basin wide effects into dam planning. The model results illustrate the importance of considering hydrological connectivity, expressed by the location of dams, both individually and in relation to other already existing dams. First results from a subsequent global comparison (Figure 2.7) show that if all 84 additional dams that are currently under consideration were built, the Mekong River Basin would experience strong alterations in the fragmentation index over the next decade, placing the basin among other heavily impounded rivers in the world.

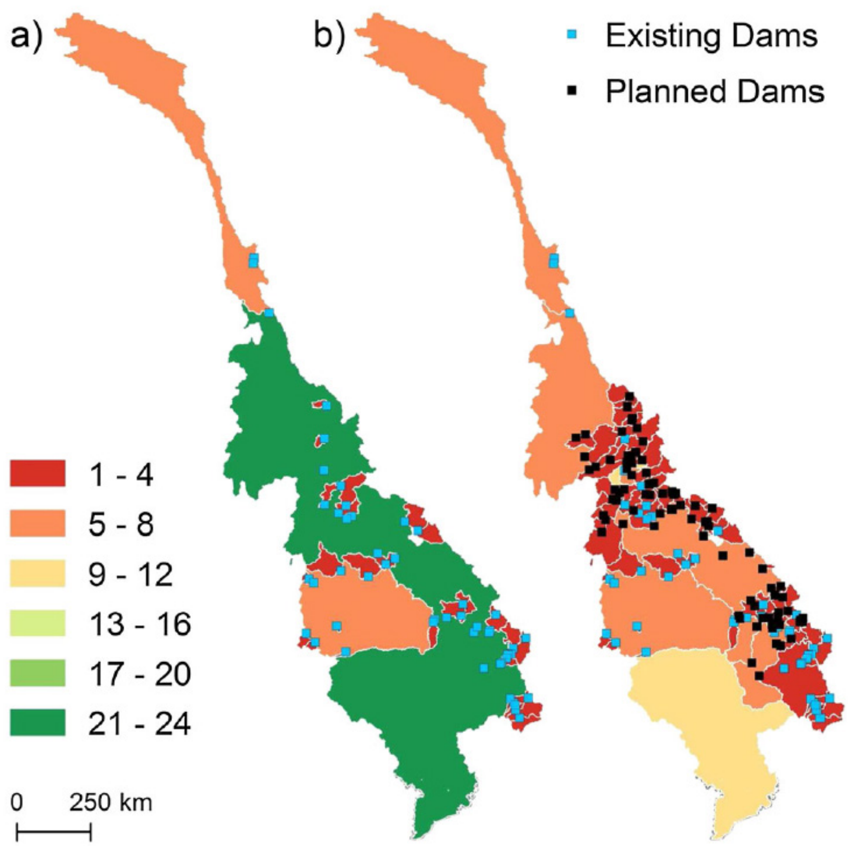


Figure 2.6: Overview of ecosystem connectivity in the Mekong River Basin for today (2011) and the future (2022) if 84 proposed dams were built (MRC, 2009). Colored regions show the number of different habitat classes found in the remaining connected river network sections. The number of connected ecosystems is strongly reduced for the future development scenario.

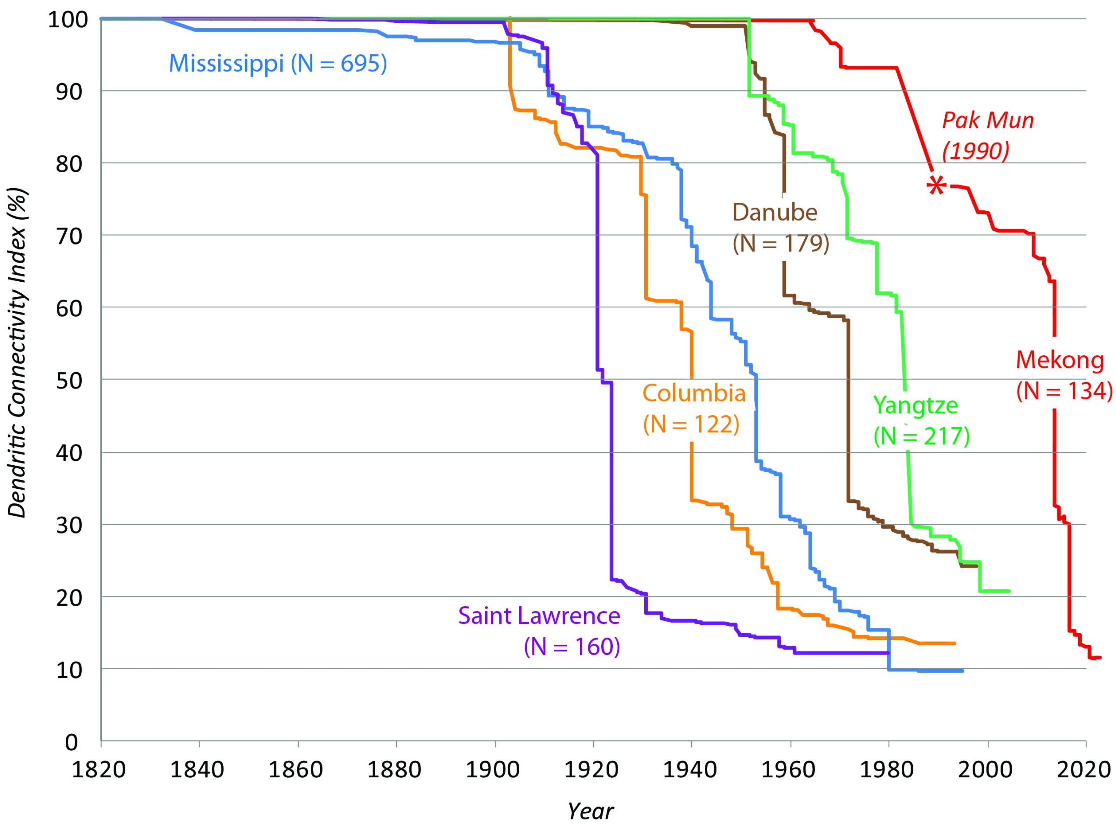


Figure 2.7: Fragmentation history for selected large river basins. The Dendritic Connectivity Index (DCI; Cote et al., 2009) decreases over time as dams are built in the river network. Historic dam constructions prior to 2011 are based on the GRanD database (Lehner *et al.,* 2011), while future connectivity in the Mekong is calculated based on a database of 84 proposed dams with commission dates (MRC, 2009). Connectivity decreases rapidly until 2022 if dam development proceeds as planned. N represents the total number of investigated dams in the basin.

2.5.3 Geospatial fate and transport modelling for pharmaceuticals in the St. Lawrence River Basin

In a third study, HydroROUT was applied to model the fate of contaminants, specifically the pharmaceutical 'Diclophenac', a common anti-inflammatory drug, in the river network of the lower St. Lawrence Basin (Grill *et al.,* in prep.). We calculated the spatial distribution of in-stream concentrations by combining the distribution of river discharge with the substance load at the outlet of sewage treatment plants and accumulated the loadings downstream. A new layer of lakes (as described in Table 2.1) was integrated, and we allowed photo-degradation in both lakes and rivers along the flow path (Buser et al., 1998). In addition to concentrations, a risk index for each river reach was calculated using predicted no-effect concentrations of the contaminant, and risk hot-spots were identified in the basin (Figure 2.8). By mapping concentrations along flow paths, we identified river confluences with unusually high chemical concentrations as well as locations at which major concentration spikes occurred due to inflows from urban conglomerations (Figure 2.9).

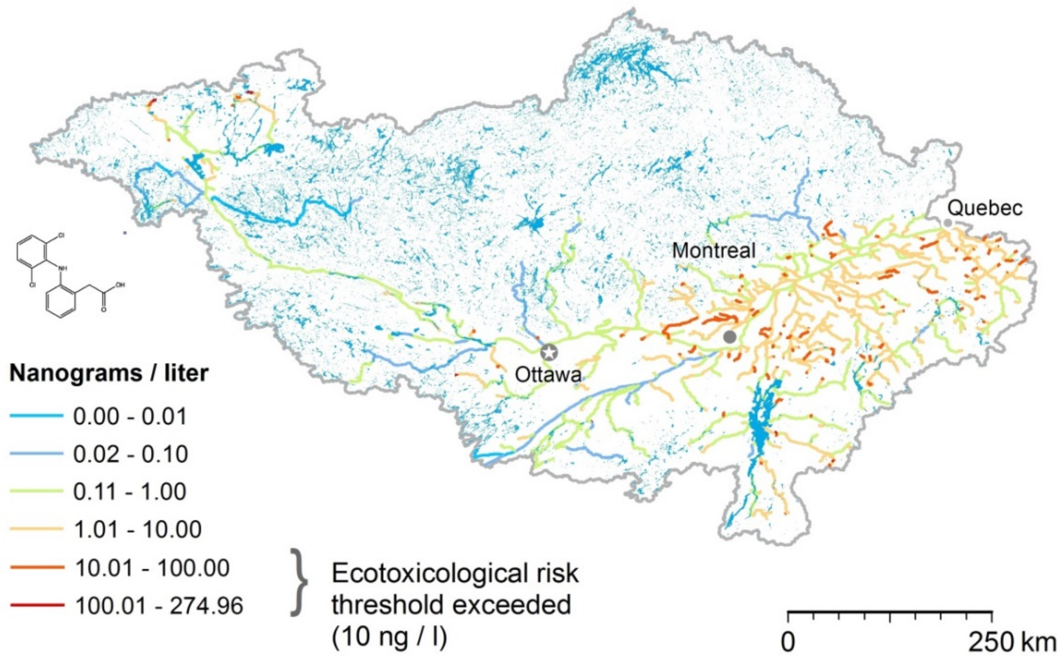


Figure 2.8: Pharmaceutical concentrations in the St. Lawrence River Basin (downstream of Great Lakes). Predicted Diclophenac concentrations in surface waters based on HydroROUT model results.

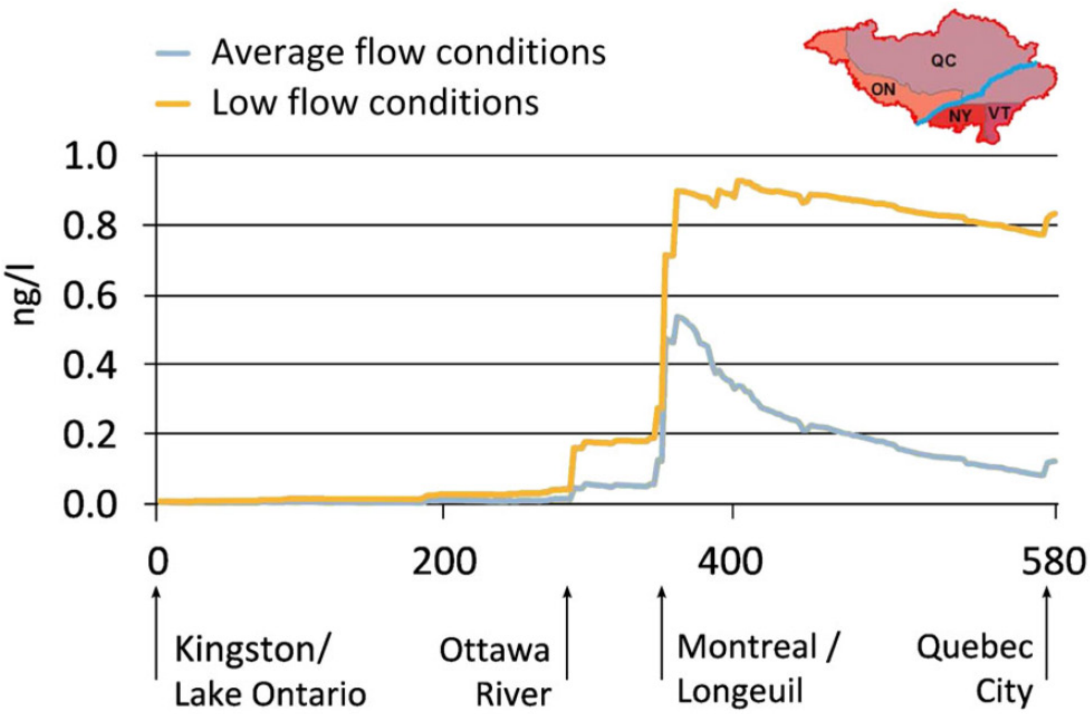


Figure 2.9: Modeled pharmaceutical concentrations in the St. Lawrence main-stem (downstream of Kingston, Ontario) based on HydroROUT results. Scenario 1 predicts concentrations based on long term average discharge values. Scenario 2 calculates concentrations under low flow conditions (monthly Q90 flow).

2.6 Discussion

The data and model descriptions outlined above, as well as our case study examples, are aimed at demonstrating the direction and potential of recent advancements in large scale hydrological and hydro-ecological modelling. Yet this review is far from being complete, and there are many other developments targeting similar goals. In particular, due to the continued increase in computational power, existing models move towards higher resolution both in space and time (Wood et al., 2011) and are enhanced by more complex process representations, such as hydrodynamic routing (Yamazaki et al., 2011) or the inclusion of advanced water management schemes (e.g., Hanasaki et al., 2006; Döll et al., 2009). As for our suggested approach to couple new hydrographic datasets with a tailor-made, vector-based flow routing scheme, a number of limitations must be noted regarding the described data and model performance. These limitations, at the same time, provide avenues for future improvements and research requirements.

Our definition of river network routing transcends beyond discharge routing, as it encompasses the movement of energy, material, and organisms. Yet although the primary goal was to develop novel tools for assessing hydrological connectivity, discharge remains a master variable in the model. In the current version of HydroROUT, discharge is based on external runoff calculations that are spatially downscaled to fit our high-resolution river network. An advantage of this decoupled approach is that different global land surface and hydrological models, or even ensembles, can be employed to provide runoff estimates using independent settings and parameterizations, allowing for comparisons and uncertainty assessments. A major disadvantage, however, is that feedbacks between climate, landscape, and hydrological conditions cannot be modeled directly in HydroROUT, rendering it less dynamic than coupled models which include runoff generation, floodplain inundation and evaporation, and hydrological routing simultaneously. A further source of uncertainty in the discharge estimates is the currently applied temporal resolution of monthly averages and the very preliminary validation of long-term conditions only, while inter-annual variability or extreme events are not tested yet. Hence, time-sensitive model results such as seasonal water quality or low-flow substance concentrations need careful interpretation. Finally, the runoff downscaling approach itself may not be able to correctly reproduce local conditions, adding small-scale uncertainties in the discharge estimates.

Another limitation of our current 'stand-alone' routing model is the rather simplified routing mechanism: the applied 'plug-flow' and 'completely stirred reactor' models simulate the transport of constituents as a linear storage-based routing process. This type of model is only valid where uncertainty from friction and backwater effects can be tolerated. At large scales, these effects may be less dominant, and the error terms may be reasonable for applications that aim at general ecological management rather than highly accurate hydraulic flow predictions. For certain regions, however, such as large floodplains, hydrodynamic processes should be integrated in the model to enable a more realistic simulation.

Associated with the routing scheme is uncertainty from simulating flow velocity. River velocities can greatly vary between river types, e.g. steep mountain streams versus large lowland rivers. In the examples we presented, we set velocity to a constant value of one meter per second, due to lack of reliable river velocities at the global scale - a widespread limitation in large-scale routing models. The next iteration will include attributes of variable velocities both in the HydroSHEDS database and the HydroROUT model, using similar approaches as suggested by Ngo-Duc et al. (2007), Fiedler and Döll (2010), and Verzano et al. (2012). For large floodplain areas, additional inundation modelling should be applied to improve flow velocity simulations (Yamazaki et al., 2011).

There are various limitations related to the underlying baseline hydrographic data of HydroSHEDS. Besides the current lack of high-quality coverage for regions above 60 northern latitude (see Data section above), HydroSHEDS follows a simple 'D8' concept to represent flow direction: each pixel points towards exactly one of its 8 neighboring pixels as the next downstream one. While this concept is mathematically easy to calculate and allows for proper simulation of main-stem river channels, it cannot represent river bifurcations (i.e. splits into multiple flow channels), braided river systems, or the secondary channels in river deltas. Also, even at 90mresolution, highly detailed and complex topographic features such as floodplain channels that regulate local hydrological connectivity are still not adequately represented (Yamazaki et al., 2012).

Finally, there are limitations related to the new lake database that is used in HydroROUT. While surface areas are mapped at very high resolution and quality, lake volumes are only coarse estimates (see Table 2.1) and are associated with difficult-to-assess uncertainty. Initial tests against a selection of 166 European lakes indicated an acceptable overall match in average lake volume, yet the values of individual lakes may be greatly over- or underestimated. Despite the many current limitations related to the quality and implementation of both the HydroSHEDS database and the HydroROUT modelling framework, results have to be judged against large-scale needs of ecologists or water resources managers. For example, even if discharge values are prone to substantial uncertainty and error, ecological changes in the river network are often most pronounced at confluences between small tributaries and large main-stem rivers, where flow magnitudes can differ by one or several orders of magnitude. Many critical characteristics along the river network, such as highly altered conditions, disruptions in connectivity, 'swimmable river length', or contributing catchment areas are well represented in the current model version, despite errors in discharge, due to the very high spatial resolution of the hydrographic baseline data. Major changes between river orders or within geographical regions can easily be mapped and related to species distributions; and detailed objects, such as effluents from a sewer plant, can be included as part of the assessments, even if uncertainties in the exact hydrological values are present.

2.7 Conclusions

This paper has been prompted by the challenge to strengthen the currently limited capabilities of GHMs in conducting more integrated hydro-ecological studies - i.e. studies that are able to support comprehensive water resources management; systematic freshwater biodiversity and conservation planning; health assessments related to waterborne diseases and water quality; or risk analyses of future climate and global change impacts on society (e.g. water supply, floods, or hydropower generation). The main limitations identified are related to spatial resolution, data structure, quality of data, inclusion of hydrological connectivity, and support of multi-scale and integrated modelling approaches. In response, we proposed a versatile global hydrological modelling framework that addresses these limitations, providing support for a broad range of applications. We described the enhancement of a global hydrographic database (HydroSHEDS) and the coupled development of a new river network routing model (HydroROUT) as the backbone of our framework. The main novelty over existing approaches is given by the combination of the following characteristics:

• The extensive development of new global data layers (e.g. dams, a hierarchical sub-basin breakdown, lakes, floodplains, etc.) and their harmonized integration with existing HydroSHEDS layers provides a unique baseline geospatial fabric.

• The hybrid model architecture supports linkage and integration of both raster (e.g. DEM, land cover) and vector layers (e.g. river reaches, nested sub-basins, point features). The vector structure enables routing at a spatial precision that is orders of magnitudes higher than in current global pixel-based models, supporting local-scale decision making. Vector routing is also fast to process and allows for more complex analyses (e.g. repeated execution; tracing using weights or barriers) and more natural, object-oriented modelling.

• The river network routing model can couple river reaches with lakes, reservoirs, dams, and floodplains and provides the potential for simulating various routing processes (transport, diffusion, mixing) at multiple spatial scales.

• Powerful routing and tracing capabilities in both upstream and downstream direction provide support of hydrologic connectivity in a broad ecological sense and for a wide range of objects, such as organisms, nutrients, and pollutants; and for the routing of more abstract concepts such as stressors or human impact indicators.

• The implementation of the modelling framework in a high-resolution GIS-based computing environment increases the suitability for hydro-ecological applications, which typically require river-reach scale resolution.

We argue that our integrated data and modelling framework supports a novel set of integrated studies, specifically to estimate the impact of human activities on hydrological functioning and connectivity, and on ecosystem services at large. We summarized ongoing research in support of such studies, e.g. the impact of dams on natural flows and ecosystem fragmentation, and the effects of anthropogenic pollutants on water quality in river networks. We believe that many more applications can be facilitated by our model framework in a variety of related fields, including aquatic ecology, biogeochemistry, geo-statistical modelling, and health risk assessments.

Our case studies showed the potential of the model and data development to facilitate and conduct hydro-ecological research at the global scale. The outcomes may also be used for general education and mapping purposes. New global information portals and data repositories started to incorporate parts of the HydroSHEDS database, including a planned integration into web services such as Google Earth or World Water Online (by ESRI). We hope that this user-friendly and accessible dissemination of data and information, together with the high spatial resolution of the results will support research, planning, and decision-making efforts for many large river basins in the world.

Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada; World Wildlife Fund, Washington, DC; the EU FP7 BioFresh Project, Berlin, Germany; the Global Environmental and Climate Change Centre (GEC3), Quebec, Canada; and McGill University, Montreal, Canada. The authors wish to thank their many colleagues who contributed to this research and the manuscript.

Connecting statement (Ch. 3)

The previous chapter provided a detailed assessment of the inherent challenges of large-scale hydrological and routing models and provided an overview of new data availability that facilitated the development of HydroROUT. Chapter 2 closed by outlining possible applications of HydroROUT.

In the next chapter, I apply HydroROUT to explore anthropogenic alterations of hydrological connectivity and river flow regulation at large scales resulting from dam construction, and I analyse what indicators can be used to assess these impacts. This chapter focuses on introducing and testing two key indicators: the River Connectivity Index (RCI) and the River Regulation Index (RRI).

I illustrate these applications with an in-depth case study of the Mekong River Basin in Southeast Asia. This research addresses key issues that provide a foundation for the research presented in chapter 4 by identifying new ways to model anthropogenic alterations of surface water systems at large scales. The results of this chapter help to determine the level of complexity necessary and feasible in indicator development and provide theoretical and methodological insights toward applying these indicators at the global scale.

3 Development of new indicators to evaluate river fragmentation and flow regulation at large scales: A case study for the Mekong River Basin

Ecological Indicators **45** (2014) 148-59

Günther Grill[1*], Camille Ouellet Dallaire[1], Etienne Fluet-Chouinard[1], Nikolai Sindorf[2], Bernhard Lehner[1]

[1] Department of Geography, McGill University, 805 Sherbrooke Street West, Montreal, Quebec, H3A 0B9, Canada

[2] Conservation Science Program, World Wildlife Fund, 1250 24th Street NW, Washington, DC 20037, USA

Abstract

Large hydropower schemes have recently gained renewed interest as a provider of efficient and renewable energy, particularly in developing countries. However, some dams may have widespread effects on hydrological and ecosystem integrity, which reach beyond the scales addressed by typical environmental impact assessments. In this paper we address two main ecological impacts—reduced river connectivity and changes in the natural flow regime—at the scale of the entire Mekong River Basin as an important component of dam evaluations. The goal is to improve our understanding of the effect of individual dams as well as clusters of dams at a very large scale. We introduce two new indices, the River Connectivity Index (RCI) as a tool to measure network connectivity, and the River Regulation Index (RRI) as a measure of flow alteration, and calculate the individual and cumulative impact of 81 proposed dams using HydroROUT, a graph-theory based river routing model. Furthermore, we demonstrate how quantitative weighting, e.g. based on river habitat characterizations or species distribution models, may be included in dam impact assessments. A global comparison of large rivers shows that the Mekong would experience strong deterioration in the fragmentation and flow regulation indices if all dams that are currently under consideration in the basin were built, placing it among other heavily impounded rivers in the world. The results illustrate the importance of considering the location of dams, both relative in the network and relative to other already existing dams. Our approach may be used as an index-based ranking system for individual dams, or to compare basin-wide development scenarios, with the goal of providing guidance for decision makers wishing to select locations for future dams with less environmental impacts and to identify and develop potential mitigation strategies.

3.1 Introduction

Dams have played a key role in regional development in recent history in many parts of the world by providing a source of renewable energy, water for irrigation and human consumption, and by protecting human settlements against floods and droughts. In addition to these intended and desirable effects of dams, numerous studies have also documented adverse, often unwanted consequences both locally and regionally (World Comission On Dams, 2000), including effects on hydrological connectivity (Vörösmarty et al., 2010), flow regulation (Lehner et al., 2011), sediment delivery (Syvitski et al., 2009), and biodiversity (Reidy Liermann et al., 2012). Dam construction and reservoirs have caused the displacement of millions of people worldwide and altered the livelihoods of river dependent societies (Richter et al., 2010).

A prominent example of the challenging trade-offs between benefits and risks related to dam construction is apparent in the Mekong River Basin (MRB). On the one hand, the international conservation community pays close attention to developments affecting the MRB as currently the river is widely considered "free-flowing" due to the fact that the entire lower portion of the main stem is unobstructed by dams. As a result, the Mekong's hydrological flow pattern has largely remained unchanged from its natural dynamics causing distinct seasonal patterns of discharge and flooding throughout the region. The natural flow regime is a main trigger of the migration of numerous fish species either longitudinally (up- and downstream) or laterally (river-floodplain) (Baran, 2006), some of which can travel several hundred kilometers up and down the Mekong.

On the other hand, at least 81 different proposals for dam development are currently under consideration for the lower Mekong region, of which 11 are located on the main stem (MRC, 2010; Figure 3.1). Facing poverty, population pressure and rising energy demands, the main stakeholder countries of the MRB, i.e. Laos, Vietnam, Cambodia, Thailand, and China, plan to tap into the large hydropower potential of the Mekong River (MRC, 2010) to accelerate their economic development. However, if dams were built on the main stem, migratory fish populations are expected to decline, resulting in a significant loss in fish catch and consequently leading to reduced income and food supplies upon which the majority of the basin's population rely either directly or indirectly (Hortle, 2009; Baran, 2010; Dugan et al., 2010; Orr et al., 2012).

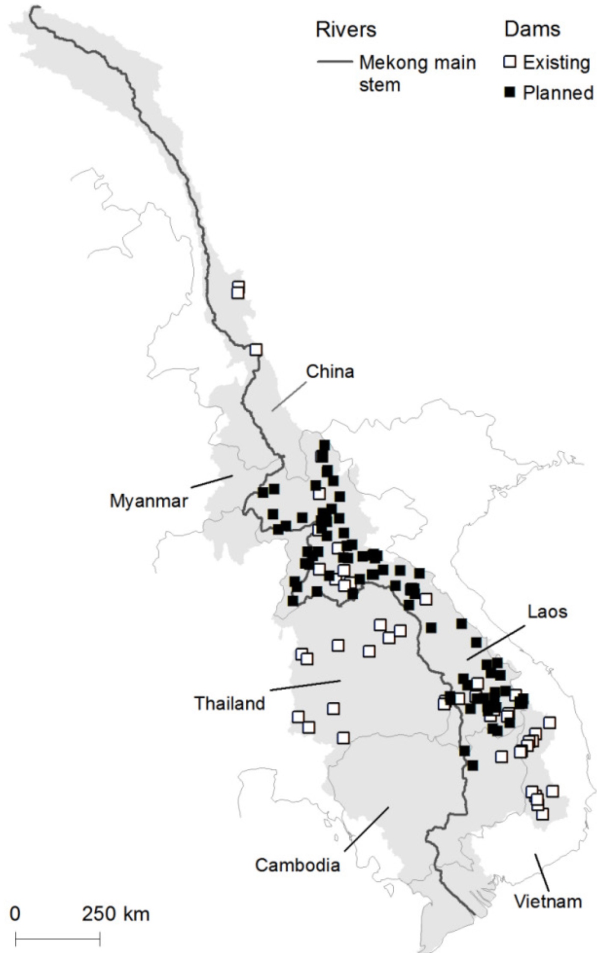


Figure 3.1 : Overview of the Mekong River Basin with existing and planned dams.

As both the plans for economic development and the goal to conserve the ecological integrity of the river system are implemented on large scales, a sustainable hydropower strategy is needed for the MRB in which the risks of dam developments are assessed for the entire basin. While most studies, such as environmental impact assessments for individual dams, focus on isolated, small scale impacts around the site of a dam or in its close downstream vicinity, only very few studies have attempted to estimate combined effects of multiple dam developments on ecosystems at larger scales. For example, the strategic environmental assessment carried out for the Lower Mekong River Basin in 2010 (ICEM, 2010) includes a number of development scenarios for pre-defined sets of dams, but as the approach is not spatially explicit it does not allow decision makers to evaluate individual dams regarding their specific role in the network.

Recently, new approaches have emerged that seek to "optimize" the development, distribution and operation of dams by assessing and prioritizing them based on their geospatial location. A more holistic 'river network mindset' is required for this approach and river basin development and management plans should take advantage of newly available data resources and computer software tools to reveal the cumulative effects of dams on the entire river system, helping to identify important linkages and critical thresholds (Lehner et al., 2011). Such prioritization efforts could also indicate dams where changes in release patterns and operation schemes, or technical interventions such as fish bypass facilities, would be most likely to improve environmental flows and/or ecosystem services.

The goal of this study is to explore some of these "optimization" strategies for the specific setting of the MRB and to report on the feasibility of applying these methods in more detail in the future. We present a model that is based on two major groups of effects: i) the effect of dams on longitudinal connectivity, which includes barrier effects and cut-offs between rivers, floodplains and wetlands. And ii) the effect of dams on the natural flow regime through water storage and retention, which includes changes in the magnitude and timing of flows, as well as associated changes in water quality. We examine both groups of effects in a single framework and can thus improve our understanding of trade-offs between different types of dams in relation to their societal benefit, which in the MRB is typically related to energy production.

The intention of our research is to provide scientific and methodological advances toward the inclusion of quantitative measures and expert knowledge into a common framework to support integrated dam assessments. In particular, we believe that our approach has the potential to improve decision making processes by adding perspectives from a larger scale context. It should be emphasized, however, that by no means our methodology is supposed to substitute or replace local environmental impact assessments, but to expand our understanding by adding the large scale as an explicit new layer of consideration. A second goal of our study is to design indicators that can be derived rapidly as a first-order proxy even in data-poor settings, and to provide a proof-of-concept that the methodology is feasible. While the described framework is, in theory, capable of providing individual dam evaluations, the choices and weights of our data and derived indicators will need critical review and validation by local experts before being reliable for political discussions and decision making. Given this current lack of local validation, it is not within the scope of this study to provide final results for immediate use in policy planning. For this reason we refrain from referring to explicit future dams and their names throughout this report.

3.2 Theoretical background

3.2.1 River fragmentation

Connectivity—or inversely fragmentation—has been estimated in the past using increasingly complex methods (Fullerton et al., 2010). Basic indicator methods have been applied in global and large-scale studies. Those include the number of dams within a watershed (dam density), the total length of a river (in km) upstream from each dam, the proportion of the river network inaccessible from the sea, or the total length of swimmable distance from each point of the network (Nilsson et al., 2005; Anderson et al., 2008; Lassalle et al., 2009; Vörösmarty et al., 2010). In an attempt to capture the obstruction of large river systems by dams, Reidy Liermann et al. (2012) measured the length of the longest undammed stretch of the five largest rivers in each 'freshwater ecoregion' (as defined by Abell et al., 2008) to derive the percentage of free-flowing rivers.

So-called graph-theoretic approaches combine a network of links and nodes, which represent river reaches and confluences, respectively, into a network (Bunn et al., 2000). Information about species populations or other features can be allocated to the links or nodes, and connectivity measures, such as the distance between nodes, can be derived. For example, Schick and Lindley (2007) examined changing patterns in connectivity and the isolation of salmon populations due to dam construction in California's Central Valley. A simple yet elegant example in this category is the Dendritic Connectivity Index (DCI, Cote et al., 2009; see also explanation in Methods section). It is based on the proportion of the length of the disconnected network fragments in relation to the entire network, and it can be applied to river networks of different scales. A disadvantage of this method, however, is that it can lead to the same index if a barrier is placed very high upstream in the network or very low, as long as the disconnected network fragments have the same length. Yet in ecological applications it is commonly argued that barriers placed further downstream pose a higher (or at least different) threat than dams in headwater reaches as the former disconnect large portions of the river network from the critical ecosystems of the main stem or the delta, and the barrier directly affects the generally higher diversity of large rivers (Kanno et al., 2012).

In the Mekong River, the connectivity of migration corridors for fish species, both in lateral and longitudinal direction, is of major concern. Given the absence of spatially explicit and reliable data on migratory species, an alternative is to use simulation models. Population models have been developed for the MRB to estimate dam impacts (e.g., Ziv et al., 2012), yet the underlying assumptions about population size and species migration patterns are difficult to verify. Due to the lack of monitoring data for a wide array of aquatic species, the focus on certain fish guilds, such as longitudinal migratory fish species, may result in solutions that are crafted toward those species, but may neglect the impact on other ecosystem inhabitants. A holistic view should attempt to extend dam assessments to all species that depend on rivers and their dynamics.

A second option to address the absence of species data is to use species distribution modelling (SDM) to determine the species range within the river system (Bahn and McGill, 2007; Elith and Leathwick, 2009). Species data, such as presence/absence surveys, and increasingly also presence-only data can be used as input to the distribution model. The model then predicts the occurrence of the species spatially based on the environmental conditions found at its sample locations. For example, Fukushima et al. (2007) used this method to determine the impact of dam construction on different species in Hokkaido, Japan. Similar statistical measures are used in Hermoso et al. (2012) and were applied to prioritize species as conservation targets. This method, however, has been criticized for not taking into account the underlying spatial structure of the actual distribution pattern, and Bahn and McGill (2007) have shown that simpler models, e.g. those purely based on spatial proximity can predict species distributions equally well. For example, as the habitat range of a migratory fish species is generally related to the distance it travels through the river network, a highly interconnected (or 'auto-correlated') distribution range is likely to exist between locations of observed species presence that are within that distance from each other.

A third way that addresses the absence of species data is to use ecosystems or habitats as a proxy. There is general agreement within the conservation community that protecting representative ecosystem types, or 'coarse-filter' targets, should conserve common species and communities, the ecological processes that support them, and the environments in which they have evolved (e.g., Higgins et al., 2005). Coarse-filter targets can be derived through the development of river or habitat classifications, which can then serve as a proxy for ecosystem types. The abundance and distribution of system classes within the basin can act as a surrogate for the actual species distributions (Sindorf and Wickel, 2011).

3.2.2 Flow regulation

Besides fragmentation, flow regulation is generally considered the second main adverse ecological consequence of dams and reservoirs (Poff et al., 1997; Bunn and Arthington, 2002). The goal of many dam operations is to eliminate peak flows, to stabilize low flows, or to impound or divert river flows partially or entirely. Alterations to natural flow patterns as a result of the operation of dams may disrupt the life cycles of aquatic species and ecological processes (Pringle et al., 2000; Dudgeon et al., 2006; Carlisle et al., 2011), and may cause the loss of endemic species and the invasion of exotics (Bunn and Arthington, 2002). The effect of dams on the regulation of downstream flows can only be fully assessed if the operation rules of the reservoirs are known; yet this information is rarely available at larger scales. As a provisional solution, the Degree of Regulation (DOR), i.e. the proportion of a river's annual flow that can be withheld by a reservoir or a cluster of reservoirs has been suggested as a first-level approximation of the potential impact of dams on downstream flows. Despite its limitations, the DOR has in one form or another been a key component of seminal studies on flow regulation (e.g., Nilsson et al., 2005; Lehner et al., 2011) or has been analyzed in terms of the hydrologically equivalent 'change in residence time' or 'water aging' (e.g., Vörösmarty et al., 1997).

The DOR as calculated by Lehner et al. (2011) is an index that is derived for each individual river reach and therefore does not capture the effect of dams on the entire network in one single value. To address this problem, we here expand the DOR concept and propose a new measure of flow regulation, the River Regulation Index (RRI), which captures flow regulation at the network scale.

3.3 Data, models and methods

3.3.1 Data and models

All calculations and simulations were performed using the HydroROUT river routing model (Lehner and Grill, 2013) within a Geographic Information System (ESRI ArcGIS10.1). HydroROUT is based on the HydroSHEDS database at 500m (15 arc-second) spatial resolution (Lehner et al., 2008). HydroSHEDS is a comprehensive, global scale inventory of spatial hydrographic and hydrological information, and provides core data of river networks, basin and sub-basin delineations and flow regimes. HydroROUT uses a graph-theoretical framework to calculate connectivity measures within a dendritic linear network. We used a total of 36,384 river reaches in the MRB with an average length of approximately 5 km. Each river reach has a simulated long-term average discharge value assigned which has been derived from runoff estimates of the global hydrological model WaterGAP for the time period 1961-90 (Alcamo et al., 2003; Döll et al., 2003); as well as an estimate of 'river volume', i.e. the water volume within the river channel at average discharge based on an approximation of channel width and depth following Allen et al. (1994). Information about dams was compiled by merging the Global Reservoir and Dam (GRanD) database (Lehner et al., 2011) with a database of hydropower projects in the lower MRB (MRC, 2009). A total of 52 existing and 81 planned dams (commission date later than 2011) were included in the analysis.

3.3.2 Calculating river fragmentation

River Connectivity Index (RCI and RCI_{VOL})

Since basic indicator approaches, such as dam densities, provide only a general measure of fragmentation and population modelling approaches are difficult to implement and verify, we used a GIS based framework combined with a state-of-the-art river routing model (Lehner and Grill, 2013) to estimate fragmentation based on structural connectivity of the river network. Following the approach by Cote et al. (2009), we define a River Connectivity Index (RCI) that is calculated based on the premise that each dam creates distinct river fragments (Figure 3.2) made up of multiple connected river reaches. We here assume that a dam completely eliminates species permeability from one fragment to another, although some degree of passability may exist in reality (see Discussion section).

The size of the disconnected river fragment created by a dam in relation to the total size of the original river network is the main descriptor of the index. In addition to fragment size, we introduce a weighting scheme following suggestions by Sindorf and Wickel (2011), in which each of the fragments can be weighted by quantitative measures, such as the number of unique river classes within each fragment (where 'class' refers to rivers with similar hydrological and geomorphological characteristics; see further definition below). As such, our index is flexible in its use and can be modified by stakeholders in subsequent analyses.

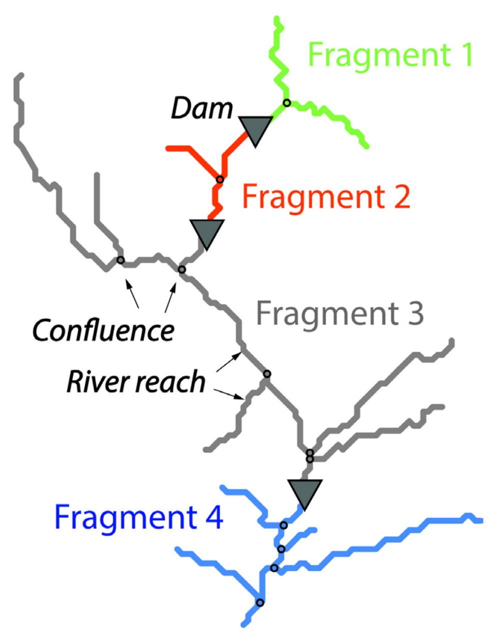


Figure 3.2: Basic concept of the River Connectivity Index (RCI). Dams partition the network into fragments with specific characteristics (e.g., length, volume, class count, etc.).

The RCI is based on the 'Dendritic Connectivity Index' (DCI; Cote et al., 2009). The DCI uses network analysis to evaluate the cumulative impact of the number, permeability, and location of barriers on the life history of potamodromous and diadromous fish. If permeability is set to zero (i.e. impassable, which we consider reasonable for large dams such as those assessed in this study), the formula to calculate DCI can be simplified to:

$$DCI = \sum_{i=1}^{n} \frac{l_i^2}{L^2} \times 100 \qquad 3.1$$

where n is the number of fragments; l_i is the total river length of the contiguous network fragment i that is disconnected by one or more dams (i.e., the fragment can be up- or downstream of a dam or in-between dams); and L is the total length of the entire river network. The DCI of an unfragmented river network is 100%, whereas each subsequent dam reduces the DCI, depending on the size distribution of the fragments. A single dam in a previously undisturbed network leads to the maximum fragmentation if it splits the network into two equally sized fragments (defined by river length), in which case the DCI falls to 50%.

For the most basic version of our new River Connectivity Index, termed RCI_{VOL}, we simply replaced the 'river length' measure of DCI with 'river volume', i.e. the volume of water that is available to a fish in a river or river reach as potential roaming or habitat space. As the river volume typically increases downstream due to increasing discharge and channel dimensions, a dam's impact is accordingly weighted higher in downstream reaches of the river network. This increases the relative impact of dams located on larger or main stem rivers. The RCI_{VOL} follows as:

$$RCI_{VOL} = \sum_{i=1}^{n} \frac{v_i^2}{V^2} \times 100 \qquad 3.2$$

where n is the number of fragments; v_i is the total river volume of fragment i; and V is the total river volume of the entire river network.

The RCI_{VOL} is useful in itself as an unweighted index, but following suggestions of Sindorf and Wickel (2011), we included a weighting scheme which allows inclusion of important characteristics of the river network. To illustrate two examples of weighting, we below introduce RCI_{CLASS} as a measure of river connectivity that incorporates the diversity of river classes, and RCI_{RANGE} as a measure of connectivity for migratory fish species. We then apply these indicators to determine the cumulative impact of dams in the MRB and compare them to the original Dendritic Connectivity Index (Cote et al., 2009).

River Class Connectivity Index (RCI_{CLASS})

The intention of RCI_{CLASS} is to include the number of (dis-)connected aquatic ecosystem types as a weight to compute river connectivity. However, ecosystems are typically defined using both abiotic and biotic information, i.e. physical habitat characteristics and species data. As extensive, spatially explicit species data was not available for our analysis, we here use 'river classes' as a proxy for ecosystems. River classifications can reveal the spatial and hydrological configuration of a river system and thus help to better understand and recognize the various characteristics of habitats and their inter-connections. Two methodologies have been widely applied in previous research: 1) "controlling factor" approaches which distinguish similar ecosystem types by subdividing and grouping environmental factors (e.g., climate, topography, and geology) into characteristic regions (e.g. dry mountains, tropical plains; Snelder and Biggs, 2002) and 2) "ecoregion" approaches, where distinct regions are formed based on unique combinations of environmental factors within each region, often with the help of manual adjustment based on expert knowledge (e.g. Freshwater Ecoregions of the World, FEOW; Abell et al., 2008).

Preliminary river and habitat classifications have been prepared or suggested for the MRB in earlier efforts (see e.g. Sindorf and Wickel, 2011). We slightly expand on these efforts and produced an updated classification map (see supplementary online material for more details). It should be noted that it was not an explicit goal of this project to develop a validated and final river classification for the MRB; rather we attempt to provide a reasonable classification scheme only as a prerequisite for the subsequent connectivity analysis.

We derived a total of 27 individual river classes for the MRB by combining 7 hydrologic river types with 6 ecological regions, incorporating the distribution of major floodplains and carbonate outcrops (Figure 3.3). We then included the number of connected river classes as the weight function in the calculation of RCI to derive an index that puts emphasis on connectivity of different system types:

$$RCI_{CLASS} = \sum_{i=1}^{n} \frac{v_i^2 \times c_i}{V^2 \times C} \times 100 \qquad 3.3$$

where n is the number of fragments; v_i is the total river volume of network fragment i; V is the total river volume of the entire river network; c_i is the total number of distinct river classes in network fragment i; and C is the total number of distinct river classes found in the entire river network. It follows from this equation that the more river classes become disconnected due to barrier effects, the higher the river gets fragmented.

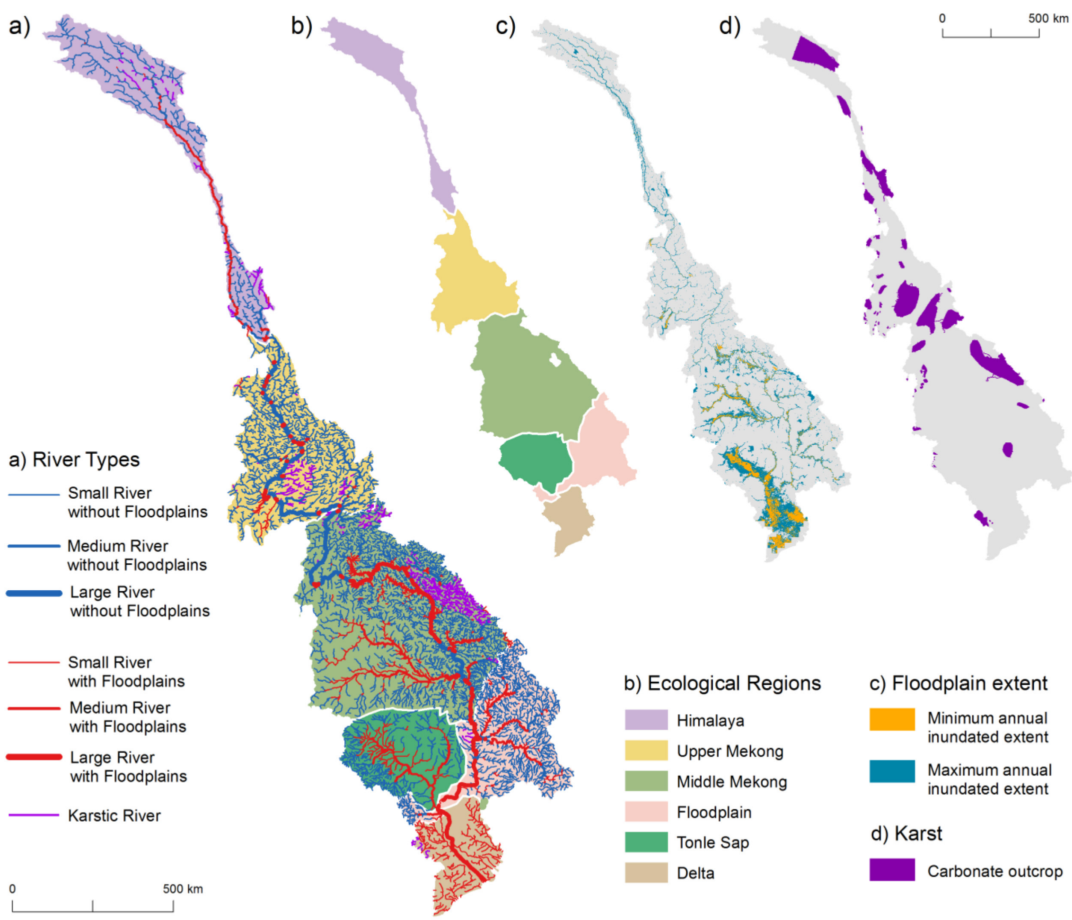


Figure 3.3: River reach classification of the Mekong River Basin. The resulting river classes are defined based on a combination of hydrological river types (a), ecological regions (b), floodplain extent (c) and carbonate outcrop (d); e.g. there are 'Medium rivers without floodplains in the Upper Mekong', etc. Floodplain extent (c) based on Fluet-Chouinard et al. (in prep.). Carbonate outcrop map (d) based on Williams and Ford (2006).

River Migration Connectivity Index (RCI_{RANGE})

In a second example, we weighted the calculation of RCI to represent long-distance migratory fish species. The Mekong Fish Database (Visser et al., 2003) includes information for 924 river species found in the Mekong. We selected migratory species based on a list of important long-distance migratory fish (MRC, 2003; Visser et al., 2003). In order to estimate species ranges from the given point data, we constructed 'minimum spanning trees' based on shortest distances, i.e. we connected the species occurrences in the river network using the shortest distances along the hydrologic flow paths (river network) between each of the observed species locations. A total of 425 sample locations representing 25 selected species were georeferenced to the river network and the number of observed species at each location is shown on Figure 3.4a. Figure 3.4b depicts the resulting range map for the example of *Panagasius gigas* (Mekong giant catfish), and illustrates that while the fish occupies only 12% of the total river network length of 27,845 km, it utilizes almost 72% of the available river volume of 8956 million m^3. By combining the ranges of all 25 species we derived a migratory species 'heatmap' of the number of species per river reach (Figure 3.4c) which confirms the important role of the Mekong main stem as a gateway for migratory fish species.

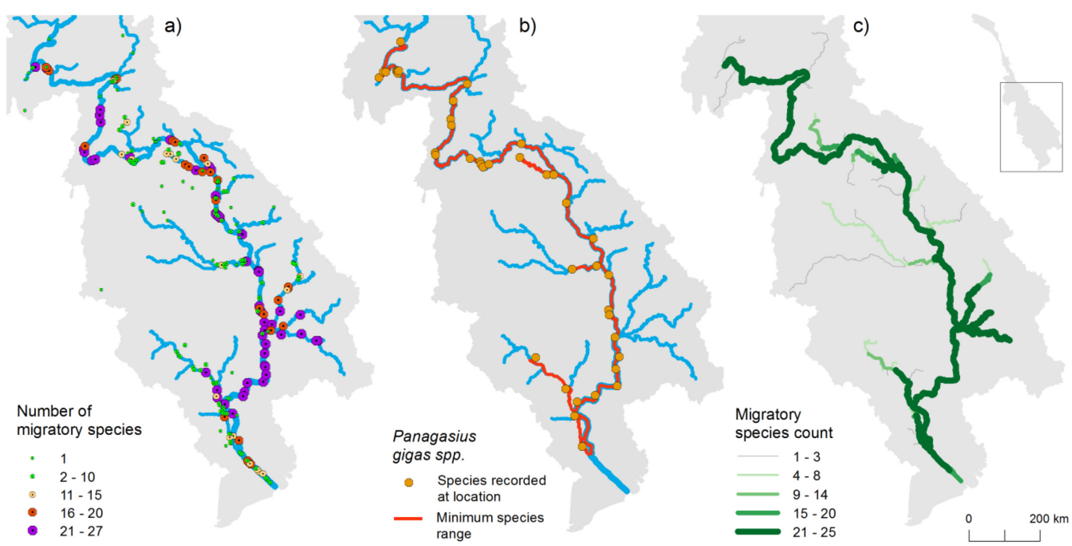


Figure 3.4: Range model of migratory fish species in the MRB: Locations and migratory species count based on Visser et al. (2003)(a); locations of occurrence and minimum spanning tree (representing the species range) for *Panagasius gigas* in the Mekong River Basin (b); combination of the ranges of 25 migratory species in a migratory species "heatmap" (c).

In order to use the migratory species ‘heatmap’ in the calculation of RCI, each species is numerically represented by its individual ‘migration range’, i.e. the river length that is available for the particular species to migrate. The combined migration range of all species is then derived as the sum of individual ranges. Note that we also tested river volume instead of river length to define the migration ranges, yet we believe that river volume (or size) is already represented in the combined migration range due to the number of migratory species (i.e., larger rivers are assumed to have more migratory species and thus larger combined migration ranges), hence we used river length as a less correlated variable to measure the combined migration range. The combined RCI_{RANGE} for all species is defined as:

$$RCI_{RANGE} = \sum_{i=1}^{n} \frac{m_i^2}{M^2} \times 100 \qquad 3.4$$

where n is the number of fragments; m_i is the sum of migration ranges (in terms of river length) of all migratory fish species in network fragment i; and M is the total sum of migration ranges (in terms of river length) of all migratory fish species in the entire river network as calculated by the species range model. It follows from this equation that a dam in the center of the combined migration range (i.e. the ‘heatmap’) will lead to the strongest reduction in connectivity, while a dam at the edge (or outside) of the combined migration range has the smallest effect (or none).

Combined River Connectivity Index ($RCI_{COMBINED}$)

In an attempt to create a combined indicator that includes information on both river classes and migration ranges, we define $RCI_{COMBINED}$ as the arithmetic mean of RCI_{CLASS} and RCI_{RANGE}. We comment on the results of this simple approach in the Discussion section.

3.3.3 Calculating flow regulation

To quantify flow regulation, we calculate the Degree of Regulation (DOR) as described and applied in Lehner et al. (2011) to assess the potential effect of dams on the natural flow regime. After linking the dams and their respective storage capacities to the HydroROUT network, we calculated spatially explicit DOR values for each river reach as:

$$DOR = \frac{\sum_{i=1}^{n} s_i}{D} \times 100 \qquad 3.5$$

where n is the number of dams upstream of the river reach, s is the storage capacity of dam i, and D is the total annual flow volume at the river reach. A high DOR value indicates an increased probability that substantial discharge volumes can be stored throughout a given year and released at later times. For example, 10% is used as a threshold in Lehner et al. (2011) to mark the possibility of substantial changes in the natural flow regime to occur. In particular, multi-year reservoirs (DOR > 100%) have the ability to release water in accordance with an artificial, demand-driven regime, often with the explicit goal to supply water in contrast to natural expectations, such as by increasing dry-season flows or eliminating flood peaks.

The DOR as defined above calculates an individual regulation index for every river reach of the network, accounting for all dams upstream of the reach. In order to quantify the overall impact on the basin in a single index value, we propose a new River Regulation Index (RRI) which is calculated by first weighting the DOR value of each individual reach with its corresponding river volume, and then aggregating the results for the entire basin:

$$RRI = \sum_{i=1}^{n} DOR_i \frac{rv_i}{V} \qquad 3.6$$

where n is the number of reaches in the network; DOR_i is the DOR value of river reach i; rv_i is the river volume of reach i; and V is the total river volume of the entire river network.

3.4 Results

3.4.1 River connectivity

For a general comparison of the Mekong with other large river basins in the world, we calculated changes in connectivity over time by using the original DCI approach (Figure 3.5). Dam development in the Mekong started relatively late with first major dam constructions in 1965. According to the GRanD database, only 13 large dams have been built since then, and the DCI stays relatively high for the times before 2011. The Pak Mun Dam, located 5 km west of the confluence of the Mun and Mekong rivers and operational since 1994, reduced connec-

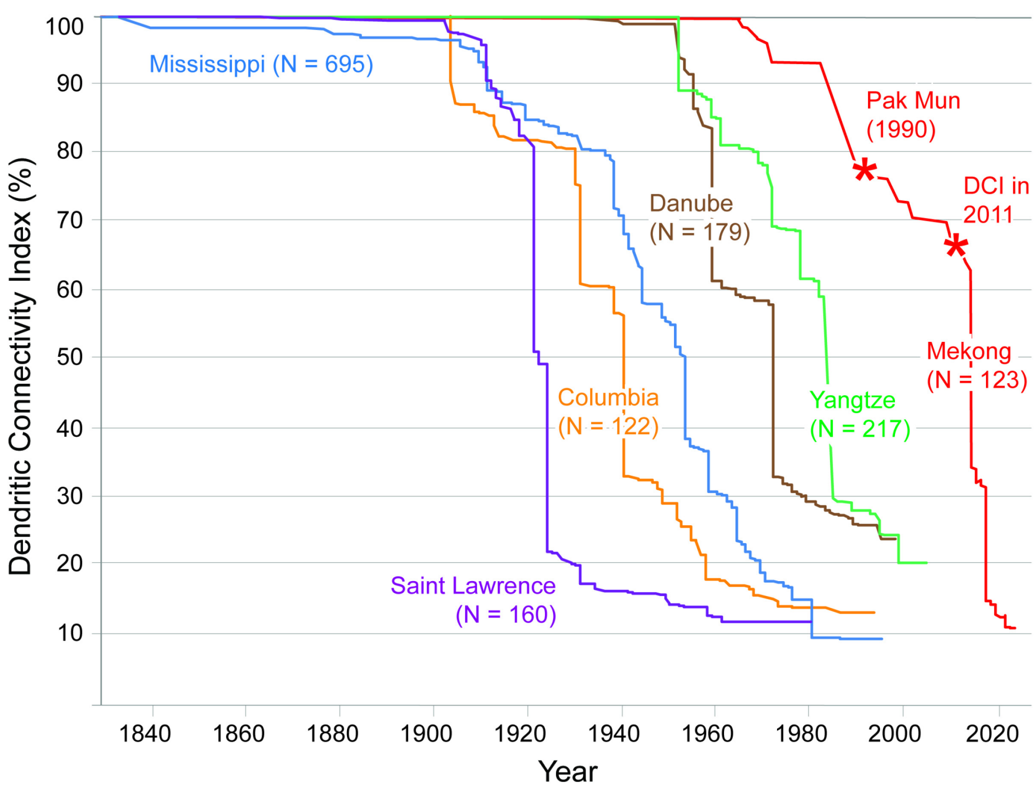


Figure 3.5: Fragmentation history for selected large river basins. As there are currently no consistent river habitat classification or migration range maps available on a global scale (a prerequisite for the calculation of RCI_{CLASS} or RCI_{RANGE}), we calculated only the original DCI (Cote et al., 2009) in this assessment. The DCI decreases over time as dams are built in the river network. For the MRB, historic dam constructions prior to 2011 are based on a combination of the GRanD and MRC databases (MRC, 2009; Lehner et al., 2011), while future connectivity is based on a list of proposed dams with commission dates (MRC, 2009). Connectivity decreases rapidly until 2022 if dam development proceeds as planned. *N* represents the total number of investigated dams in the basin.

tivity from 93% to 77%. Between 1994 and 2011, 29 smaller dams (as provided by MRC, 2009) have also been included in the assessment, which reduced the DCI to 66%. If then all 81 dams that are currently under consideration were constructed, connectivity would further decrease to approximately 11% by 2022. The strong future decrease in connectivity in the Mekong Basin is mainly caused by main stem dams for which their location rather than storage volume is important. Note that the value of 66% for today's conditions differs from the result shown elsewhere in this paper (44.2%); the reason is that for the global assessment a slightly generalized river network was used and DCI is sensitive to the total length of the applied river network (see also discussion section).

RCI_{CLASS} and RCI_{RANGE}

Currently almost 65% of the area of the MRB is connected in the largest contiguous fragment that is still unimpeded by dams, and 24 out of 27 river classes are found within this fragment (Figure 3.6a). This translates to a relatively high RCI_{CLASS} of 70.5%. RCI_{RANGE} is even higher at 93.2% as nearly all migration ranges included in our assessment are currently unobstructed by dams.

In the future scenario, if all 81 considered hydropower dams are included (Figure 3.6b), the entire basin is partitioned into much smaller fragments due to the large number of newly constructed dams. The largest fragment that remains connected accounts for only 20% of the total basin area, mostly due to dam developments on the main stem of the Mekong. The number of connected river classes in this largest fragment is reduced to 12. In this scenario, the RCI_{CLASS} falls to 7.6%, nearly a tenth of its original value. RCI_{RANGE} values are also strongly reduced to 20.8% in the future mostly due to several main stem dams that fragment the migratory species ranges.

In contrast, the DCI, which only considers the lengths of disconnected river fragments and is not sensitive to river volume, the distribution of river classes and migration ranges, starts at a lower value of 44.2% for the current scenario. In the future, it decreases to about a fourth (10.7%), suggesting a smaller relative change than both the RCI indices.

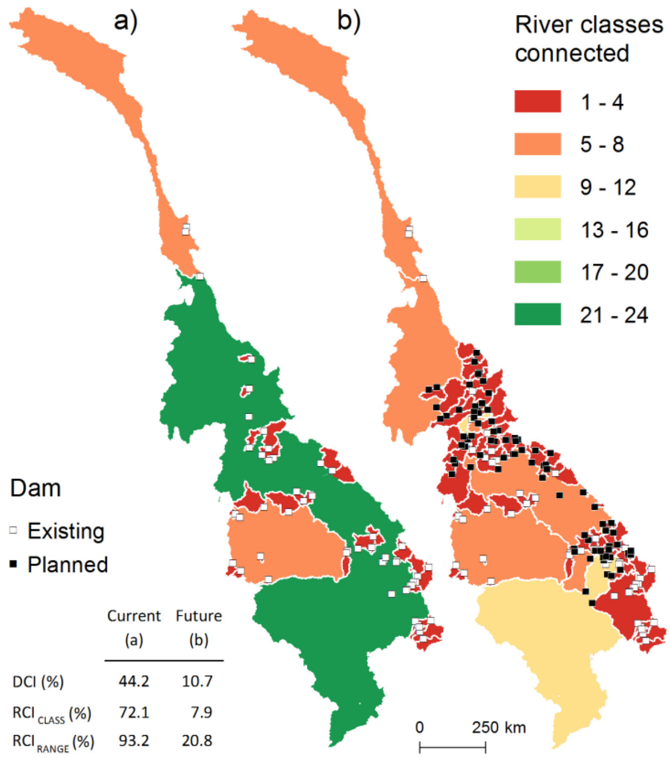


Figure 3.6: Overview of ecosystem connectivity in the Mekong River Basin for today (2011) and the future (2022). Colored regions show the number of different river reach classes found in the contiguous river network fragments. DCI, RCI_{CLASS}, and RCI_{RANGE} values are calculated for reference. All indices as well as the number of connected river classes are strongly reduced for the future development scenario compared to the current situation.

Comparison of indicators

To provide more insight into the behavior of the different connectivity indices, we calculated DCI and three RCI values (RCI_{VOL}, RCI_{CLASS}, RCI_{RANGE}) for 10 of the 81 planned dams individually (Figure 3.7). The results for each selected dam reflect the change in basin-wide connectivity if the dam was built in addition to the 42 existing ones as of 2011.

Figure 3.7b shows that the general trend of all DCI and RCI-type values is similar in that tributary dams remain at relatively high connectivity values whereas main stem dams lead to significant reductions in connectivity. As the DCI does not account for the river volume or topologic position in the network (i.e. up- or downstream), the index is similar for both dam [Id 4] on the upper Mekong main stem and dam [Id 9] on the lower Mekong main stem. Both dams cut off a similarly large proportion of the river network in the headwater and the delta regions, respectively. The RCI accounts for this difference by weighting the effect of dam [Id 9] stronger due to its larger impact on river volume.

Main stem dams can individually reduce DCI by up to a half from the initial value of 44.2%, e.g. dam [Id 7] reduces DCI to 24.4%. When river volume and river classes are taken into account, the calculated RCI_{CLASS} can show an even larger relative decrease in connectivity from the original 72.1% to levels of a third and even below. For example, if dam [Id 9] was built, RCI_{CLASS} would decrease to 18.7% because the dam is located on the large Mekong main stem and divides the system into two fragments with notably fewer connected river classes. In contrast, dams built on smaller headwater streams have less impact on RCI_{CLASS} because a large proportion of the calculated river volume and class connections remain intact.

Finally, the results for RCI_{RANGE} indicate yet again a stronger contrast between tributary and main stem dams. RCI_{RANGE} values for all tributary dams remain high because, according to our species range model, the available contiguous migration ranges of our selected migratory fish species are only little affected. In

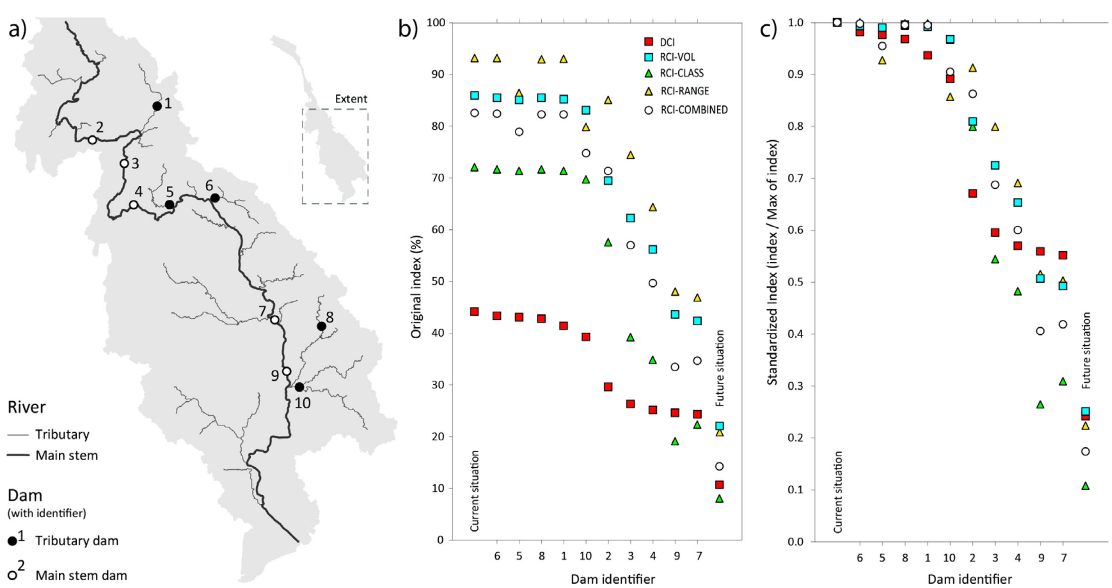


Figure 3.7: The effect of selected dams on river connectivity indices in the Lower Mekong River Basin. Comparison between Dendritic Connectivity Index (DCI; Cote et al., 2009) and four different RCI indices. Dams in diagrams are sorted by decreasing DCI values. See text for explanations.

contrast, the main stem dams obstruct the core migration corridor and consequently reduce connectivity drastically. At tributaries with a significant occurrence of migratory species, for example at the locations of dams [Id 5] and [Id 2], connectivity is more strongly decreased according to RCI_{RANGE} than shown by any of the other indices; this suggests that the index is indeed better suited to reflect the special cases of migratory species occurrences.

In order to further compare the behavior of the indices in relation to each other, we standardized the values of each index to a common starting point (Figure 3.7c). By examining dam [Id 2] it becomes apparent that the DCI shows a larger reduction of connectivity than RCI-based indices due to the fact that at that location, the river volume is still relatively small compared to dams further downstream. In general, RCI-based indices produce a stronger differentiation between main stem dams and tributary dams than the DCI.

3.4.2 Flow regulation

Figure 3.8 shows that the cumulative effects of all existing dams on flow regulation extend far downstream along the river network, with a current cumulative DOR of 6.5% in the Mekong Delta. In the future scenario (right panel), several additional tributaries would be affected if all 81 planned dams were constructed, and DOR values would increase significantly in many already regulated reaches. In particular, the values along the Mekong main stem increase and DOR more than doubles to 17.2% in the Mekong Delta.

Many dams show a large DOR directly downstream of their location; the highest reaching 390% for the Houayho Dam on the Xekong River (see inset panel in Figure 3.8). Further downstream, DOR values can decline if increased inflows from unimpeded tributaries "dilute" the flow regulation effect; while they can increase where the effects of multiple dams coincide.

The total network regulation as measured by the RRI is currently at 5.6%. It should be noted that this value is calculated as an average for the entire river network, including the reaches upstream of dams where no flow regulation occurs. The absolute value is thus dependent on the chosen extent of the river net-

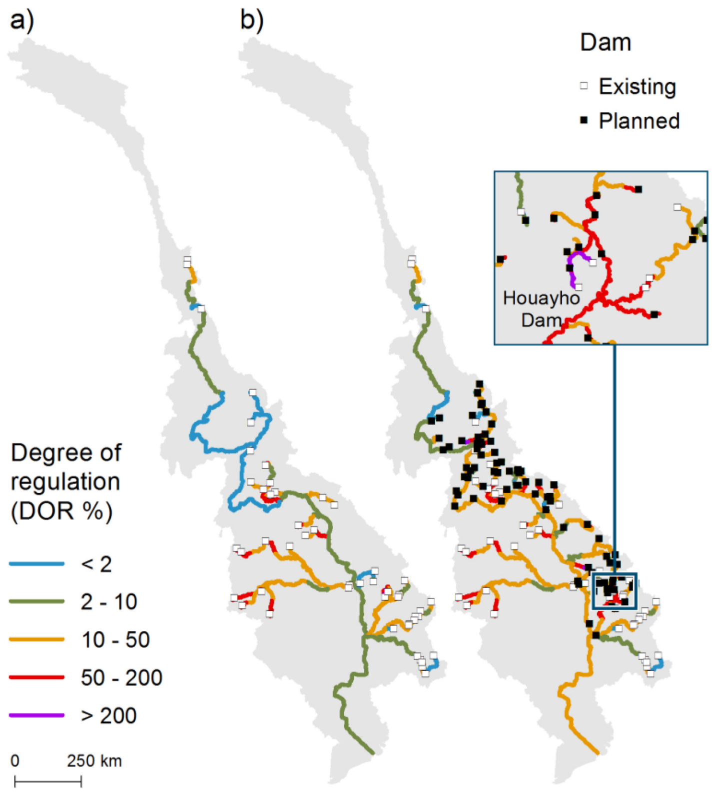


Figure 3.8: Degree of regulation (DOR) downstream of dams for today (2011; panel a) and the future scenario (2022; panel b) if 81 new dams were built.

work and should not be interpreted in the same way as the DOR for single reaches. However, in terms of relative change, the RRI increases to 15.2% if all 81 future dams were built, almost three times higher than the level of current river flow regulation.

3.4.3 Integrated assessment of dam effects

In an attempt to simultaneously assess river fragmentation and flow regulation, Figure 3.9 combines results of RCI_{CLASS} and RRI calculations in a single graph. Many dams remain along the 72.1% line of current connectivity (RCI_{CLASS}), which means that they individually do not significantly alter the overall basin fragmentation (yet note that their cumulative effect is not assessed in this graph). These dams are typically located along smaller headwater streams, meaning that they disconnect only small tributaries from the river network. However, many of these headwater dams strongly contribute to river flow regulation (RRI) mostly due to their large storage capacities, but also as their location further upstream in the river network causes them to affect more downstream river volume than a dam closer to the delta.

Main stem dams typically show more severe impacts on the fragmentation index than on the regulation index. Virtually all dams on the Mekong main stem, which are mostly designed for high energy production, show a very strong reduction in river class connectivity, while indicating relatively small effects on flow regulation due to their limited reservoir storage capacities.

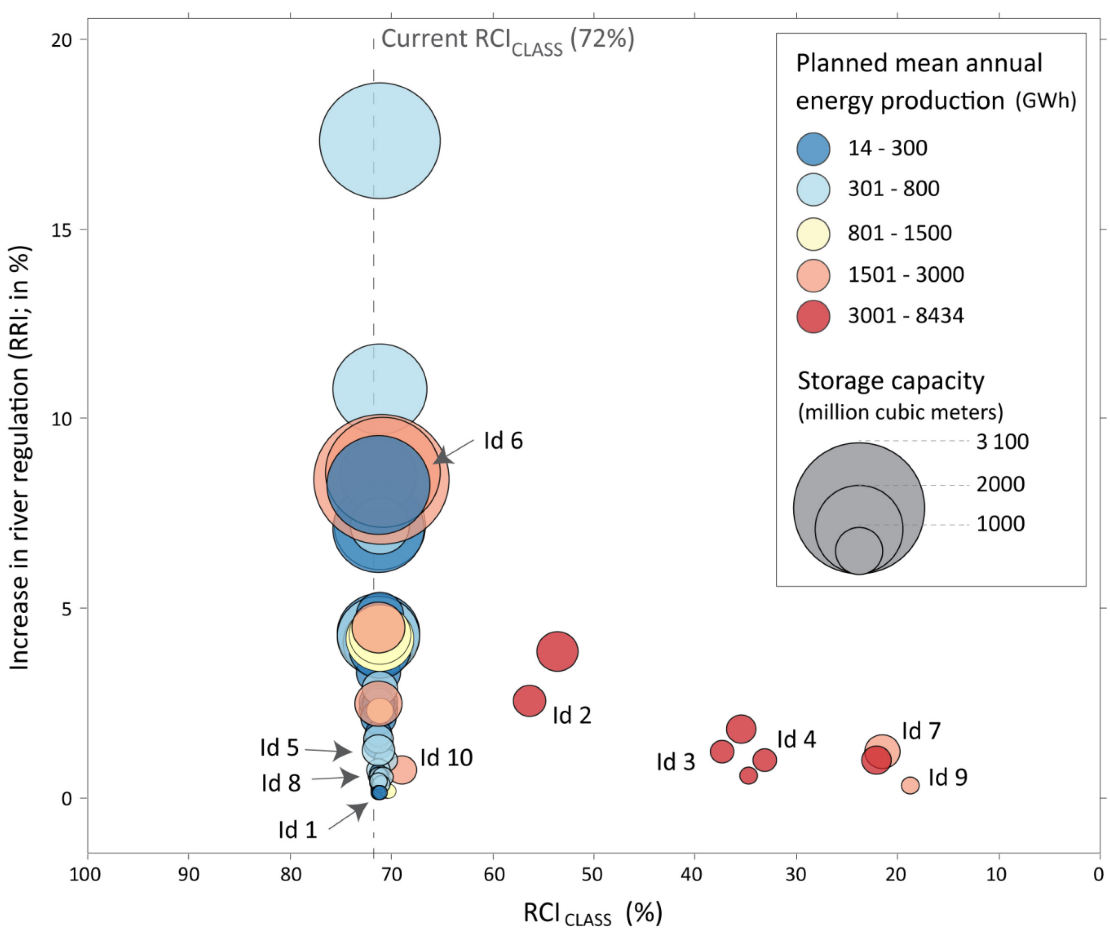


Figure 3.9: Effect of individual dams on the basin-wide River Connectivity Index (RCI$_{CLASS}$) and on the River Regulation Index (RRI). Each of the 81 planned dams is represented by a circle with a size proportional to its storage capacity, and a color shade according to its expected annual hydropower production. The x-axis shows the effect of the dam on river connectivity while the y-axis shows its effect on flow regulation. The ID numbers refer to those used in Figure 3.7.

3.5 Discussion

Using a river routing model, we calculated the individual and cumulative impact of 81 proposed dams on river connectivity and on flow regulation in the Mekong River Basin. We provided an assessment among selected dams in terms of their expected effects following different indices that have been proposed and tested in literature and that we adapted or refined for the purpose of this study. Our findings show that if all dams were built, the current flow regulation and connectivity levels would strongly deteriorate for all tested indices.

By design, the applied connectivity indices inherently condition main stem dams to have a larger impact on fragmentation than dams on tributaries. Although this characteristic has been postulated by many experts and practitioners, others argue that tributary dams block access to important recruitment areas, hence those dams could cause most harm to biodiversity and fisheries (e.g., Ziv et al., 2012). This ambiguity implies that the successful application of any index or model will largely depend on the agreement of experts on the most appropriate parameter settings and weighting factors for a given study area. The weighting component of RCI_{CLASS} can be used in that regard as a flexible "tool" to explore the effects of different weights in dam impact assessments, potentially as part of a collaborative process between stakeholders. For example, we believe that strong weighting of fish recruitment areas in our model (e.g., by introducing specific river classes in RCI_{CLASS}) could lead to similar results and conclusions as found by Ziv et al. (2012).

The importance of integrating ecological information, such as species distributions, abundance, diversity, and dynamics in dam impact assessments has rarely been questioned. Yet using actual species data in suitable models is typically complicated by a general lack of sufficient and reliable observations, and by difficulties to describe the habitat requirements and dynamics of fish species in response to dam effects. We showed how the possible migration ranges of several species can be created through relatively simple species distribution modelling techniques and be included in dam impact assessments (RCI_{RANGE}). Our chosen focus on selected long-distance migratory species, however, might underestimate the biological importance and value of several other important freshwater species impacted by dam alterations, such as short distance and lateral migrants, invertebrates, and insects. We believe that better species mapping and monitoring could provide more useful information for species distribution and population models, and thus improve ecology-based dam assessment models.

As shown in the comparisons of Figure 3.7, all investigated connectivity indices provide somewhat different results; yet their relative changes for specific dams indicate similar trends and patterns. We thus believe that the general DCI- or RCI-type approach of measuring river fragmentation is suitable to distinguish major differences between individual dams based on their specific location within the river network. The original DCI incorporates river length as the only criterion to describe fragmentation; while this index is easy to calculate, it is less sensitive to the topological location within the network (up- or downstream), and it is more susceptible to the spatial extent of the underlying river network (e.g., the addition of smaller rivers significantly alters the total river lengths in the equation). The alternative RCI_{VOL} is more robust in this respect as the addition of smaller streams is less significant for the total river volume. Ideally, RCI_{CLASS} (looking at the entire river network, including volume and ecological weighting), and RCI_{RANGE} (focusing on migration ranges) could be combined, for example by calculating the arithmetic mean of both results; Figure 3.7 depicts this option as $RCI_{COMBINED}$. The resulting index takes into account both river class diversity and migratory species ranges, and is thus more capable of identifying special or unique constellations. However, migratory species information and reliable river classifications are not available for many large-scale river basins. We thus propose that RCI_{VOL}, which overall delivered comparable results to $RCI_{COMBINED}$, could serve as a substitute to measure river fragmentation in the absence of more detailed ecological information.

We identified several dams that, if built, would contribute significantly to increased flow regulation and/or fragmentation within the MRB, depending on their size and location in the network. It is important to note, however, that a new dam located on a tributary with an already pre-existing dam will have a much lower additional impact on the overall connectivity index than the same dam being built on a previously unobstructed river. The analysis therefore supports the hypothesis that building "dam cascades" on carefully selected tributaries, i.e. multiple dams in close succession, may have less impact on basin-wide ecosystem integrity than spreading them over multiple tributaries, or constructing even a single dam on the main stem. In future assessments, our index could potentially include optimization strategies, for example to place a certain number of dams in the network until a predetermined energy goal is satisfied, while minimizing the effect on overall flow regulation and connectivity.

Our calculations clearly show that there is no one index or model that can fully describe the complex effects of dam constructions on downstream flows and river connectivity. Most dams showed strong alterations in *either* the fragmentation index *or* in the regulation index. Both represent different types of impacts, thus incorporating them simultaneously in the assessment (Figure 3.9) allowed for a much refined interpretation. The differentiation between fragmentation and flow regulation may also provide an avenue for identifying and developing potential mitigation strategies. If a dam fares low on the connectivity scale but shows little impact on flow regulation (i.e. low RCI and low RRI), then technical solutions regarding passage, such as fish ladders or diversions to increase permeability, may be particularly important to support migration and species exchange. On the other hand, for a dam with strong impacts on flow regulation (typically those with large storage reservoirs, i.e. high RRI) the main focus could be on optimizing their operation schemes and to align the release patterns in accordance with environmental flow requirements. In this sense, our proposed approach could also be applied to evaluate existing dams for their mitigation potential with a basin-wide perspective.

Despite the success in demonstrating the capability of model simulations for the assessment of dam effects on large-scale river basins, the results need careful interpretation to avoid misleading generalizations, and several shortcomings and limitations need to be taken into account. Similar to the original DCI, our presented RCI indicators are sensitive to boundary conditions, such as the composition of the river network. For example, if the river network contains less headwater tributaries (i.e., a higher discharge cutoff is used), the indicator can give noticeably different results for the same set of dams compared to a more detailed stream network. For this reason, we do not suggest fixed thresholds but rather recommend that the settings and results are discussed in a region-specific context and in collaboration with local experts. Also, the effects of individual dams or sets of dams should be interpreted as relative to each other (i.e. "an RCI of 0.5 is higher than 0.3"), and not in an absolute way (i.e. "negative environmental impacts start at an RCI of 0.5").

In our connectivity model, we assume that a river network without any dams is fully connected, which obviously fails to acknowledge natural barriers and discontinuities. Khone Falls, for example, represents a natural barrier in the upstream direction for all but the strongest fish (Sverdrup-Jensen, 2002). If such natural barriers are added to the fragmentation analysis, this will decrease the fragmentation values of dams built in their vicinity. To include a more differentiated view of natural discontinuities, more data on their location as well as a better understanding of their effect on passage of species up- and downstream is required.

We also assume that any dam, regardless of its design or operation, compromises connectivity, and we assume zero passability for each dam in our model. However, the permeability of barriers for migrating fish and other species depends on the type of dam. For example, a run-of-the-river dam may prevent migration of species in the upstream direction, but can still allow for downstream passage. Smaller dams (albeit not explicitly considered in this study), might be passable by jumping fish in both directions. Technical modifications of dam infrastructure, such as fish ladders, can increase passability for some fish populations. In a recent study, however, some of the planned Mekong dams were examined for their design and permeability for migrating fish and the results showed that some species may not be able to successfully pass the barriers (Halls and Kshatriya, 2009); and there were also indications that the effect may be aggravated for larger fish species because the proposed passage structures are less effective for them. In its original definition, the DCI as developed by Cote et al. (2009) is capable of incorporating relative passability as a percent value (the authors used 50% passability in their example study). A similar concept could be adopted in our RCI calculations if reasonable passability estimates were available for each dam.

Our current study includes only large dams from the global GRanD database and large future hydroelectricity schemes, while many smaller dams are missing. Small dams can have a significant cumulative effect on flow regulation (Lehner et al., 2011) and although researchers have recently collected information on the geographic location of smaller dams worldwide (M. Mulligan, King's College London, personal communication) there is a lack of information about their storage volumes, purposes, and intended operation rules. Nevertheless, future studies should attempt to include smaller dams and their appropriate information.

Finally, our simplified hydrological model only provides long term average flows, hence special systems, such as the Tonle Sap Lake where flows reverse direction in the flood season and connectivity may change, are not properly represented. While we weigh our basin-wide calculations by river volume, the impacts and consequences of flow regulation may additionally vary for different river size classes depending on their role and importance. Obviously, the main stem Mekong is of high regional importance, including far-reaching ecological and socioeconomic aspects, and strong dependencies between agricultural practices, fisheries, and natural river flows exist. Yet smaller streams can provide important local ecosystem services, serve as ecological refuges, or may represent headwater reaches important for municipal water supply. Overall, it will largely depend on the individual reservoir operation scheme and/or additional effects, such as the level of water abstraction, whether the downstream flows are compromised. For some dams, the implementation of environmental flow standards may mitigate the indicated effects. We recognize that ecological effects may vary and that some river habitats may be more threatened than others by a certain level of regulation. Clearly, more research is needed regarding the associated ecological consequences.

3.6 Conclusion

According to our model results, the Mekong is currently still highly connected both structurally and ecologically, despite some dam developments in the past. Most of the river classes identified by our river reach classification are currently connected. Corridors for selected important long-distance migratory fish species as calculated by our range model are currently only little to moderately obstructed by dams. The degree of regulation, i.e. the level of anthropogenic impact on the flow regime is relatively low as well. However, if all currently planned dams were to be constructed, the Mekong would experience similar degradation from flow barriers as other large and heavily impacted river systems of the world experience today. More importantly, all of our indicators agree that even a single dam constructed on the main stem of the river typically results in a significantly higher fragmentation score than dams placed on tributaries.

The case study of the Mekong River Basin represents a complex and unique ecological system which is relatively little impacted today. If ecosystem integrity gets compromised through dam constructions, related ecosystem services are also likely to deteriorate, placing both biodiversity and the food supply of local populations at risk. The strong dependency on fishery resources from the river, and the particular importance of long-distance migrations, makes a large-scale vision of protecting the hydro-ecological functionality of the MRB mandatory. A basin-wide strategy is needed to provide economic development without jeopardizing ecological integrity, and each dam needs to be evaluated carefully for its individual and cumulative impacts and benefits. We believe that the indicators and integrated assessment methods drafted in this paper are one step towards achieving this goal, both in the MRB or elsewhere.

At this stage, our assessment should be interpreted as preliminary and our ranking and evaluation should be seen as proof-of-concept rather than as final answers. The complex task of dam impact assessments requires careful selection, interpretation and weighting of a multitude of biotic and abiotic information by means of models and expert knowledge and should provide a balanced view on the cumulative costs and benefits of single and multiple dams. Before the results of our model can be implemented in a decision making process, researchers, policy makers and water managers need to agree upon a set of important ecological, biological, and hydrological information and need to collect the required data systematically at the basin scale.

Acknowledgements

The authors wish to thank Bart Wickel, Robin Abell and Michele Thieme of WWF US for their previous work and reports related to the Mekong River Basin that provided an important basis for this study. Both the Mekong River Commission, WWF Greater Mekong (Marc Goichot, Trang Dang Thuy), and local experts (Peter-John Meynell) are gratefully acknowledged for their support in sharing data and information to be used in this report. Funding for this study was provided in part by WWF Greater Mekong; the Natural Sciences and Engineering Research Council of Canada (RGPIN 341992-07); the EU FP7 BioFresh Project, Berlin, Germany; the Global Environmental and Climate Change Centre (GEC3), Quebec, Canada; and McGill University, Montreal, Canada.

Supplementary Information

Preliminary river reach classification for the MRB

The goal of river reach classifications is to define groups of river reaches that are more similar to each other than to the reaches of other groups. Using an iterative method based on both supervised and non-hierarchical clustering techniques (Snelder et al., 2005; Olden et al., 2012) we classified the river reaches of the MRB based on abiotic (mostly physical) parameters derived from the main categories of hydrology, geomorphology, and ecology. Information on species distributions or water chemistry was not included due to lack of explicit basin-wide data. The river reach classification is based on the following main criteria:

Discharge. River size is frequently used in aquatic habitat classifications, sometimes substituted by river length or upstream basin area if discharge information is unavailable. We used long-term average discharge based on runoff estimates from the global WaterGAP model (Döll et al., 2003) as a first-order proxy for hydrology and river reach morphology. We categorized three stream orders: small streams (<100 m^3/s), medium tributaries (100-1000 m^3/s), and the Mekong main stem (≥1000 m^3/s).

Floodplains. Floodplain systems represent highly distinct aquatic ecosystem types in the MRB due to their complex hydrological and ecological dynamics (e.g., channel morphology, lateral migration of biota, sedimentation). We derived associated floodplain coverage for each river reach based on the maximum flooding extent as delineated on a global map by Fluet-Chouinard *et al.* (in prep.). The presence of more than 50% of floodplain area within a 5km buffer around each river reach was used to distinguish major floodplain influence.

Karst. Karstic environments have been acknowledged as another indicator of highly unique aquatic ecosystem types in the MRB. These environments are considered hotspots for endemism because of their capacity to easily create cave systems, i.e. environments prone to specialization of species (Williams, 2008). Following the approach by Sindorf and Wickel (2011), we used carbonate outcrops as a proxy for karstic environments based on a global map by Williams and

Ford (2006). We classified reaches as 'karst' types if more than 75% of their respective upstream watershed area is covered by karst.

Ecological regions/Elevation. Elevation is frequently used as a major descriptor of largely different aquatic subregions such as the Himalayan headwaters and the middle and lower parts of the MRB. Here, we decided to use 'ecological regions' as a more holistic indicator that combines topography with available information on the general ecology of the MRB. Following other authors (Poulsen et al., 2002; Baird and Flaherty, 2005; Gupta and Liew, 2007; Campbell, 2009; Kang et al., 2009; Dugan et al., 2010), we defined ecological regions as units where the biota have to adapt to similar constraints and resources, which can result from climate, river regime, and geomorphology among other factors. We used breaks, or thresholds, located along the main stem of the Mekong to distinguish six different ecological regions: Himalayas, Upper Mekong, Middle Mekong., Large Floodplains, Tonle Sap, and the Delta.

Combination of classes. The majority of karstic rivers are small streams, thus we refrained from distinguishing size classes or floodplain extents for this river type. The combination of 7 river types (3 size classes with and without floodplains, plus karstic rivers) and 6 ecological regions results in a total of 42 possible class combinations. Yet as several combinations do not occur in reality, e.g. there is no main stem class in the Himalayan region, a final number of 27 classes was created.

Connecting statement (Ch. 4)

The previous chapter demonstrated an extensive application of HydroROUT for river network routing, including tracing operations, which provided the foundation for the development of two novel indicators of river fragmentation and regulation—the River Connectivity Index (RCI) and the River Regulation Index (RRI). Emphasis was placed on testing and evaluating these and other indicators at large scales using the case study of the Mekong River.

In chapter 4, I extend this effort both spatially and temporally by developing and applying a truly global scale integrated eco-hydrological model. The focus of this chapter is on implementing multi-scale and multi-indicator modelling approaches in river networks. This study combines multiple impacts of dams, namely flow regulation and fragmentation of rivers in one framework. While the two effects were previously assessed either separately or in the form of a lumped indicator, they were here combined to form an integrated "Dam Impact Matrix" which helps to improve our understanding of trade-offs between different types of dams in general and in relation to their societal benefit (i.e. energy production; chapter 3).

This chapter also includes an initial exploration of the effects of waterfalls on river connectivity at the global scale, which, to my knowledge, has not been assessed in previous research. The incorporation of waterfalls offers important new future research avenues to assess the combined anthropogenic and natural determinants of river connectivity.

4 An index-based framework for assessing patterns and trends in river fragmentation and flow regulation by global dams at multiple scales

Environmental Research Letters (in press)

Günther Grill[1*], Bernhard Lehner[1], Alexander E. Lumsdon[2,3], Graham K. MacDonald[4], Christiane Zarfl[2,5], Catherine Reidy Liermann[6]

[1] Department of Geography, McGill University, 805 Sherbrooke Street West, H3A 0B9, Montreal, Canada.

[2] Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB), Müggelseedamm 310, 12587 Berlin, Germany.

[3] Institute of Biology, Freie Universität Berlin, Germany

[4] Institute on the Environment, University of Minnesota, 1954 Buford Avenue, Saint Paul, Minnesota 55108, US.

[5] current address: Center for Applied Geoscience, Eberhard Karls Universität Tübingen, Hölderlinstr. 12, 72074 Tübingen, Germany.

[6] Center for Limnology, University of Wisconsin, Madison, WI, US.

Abstract

The global number of dam constructions has increased dramatically over the past six decades and is forecast to continue to rise, particularly in less industrialized regions. Identifying development pathways that can deliver the benefits of new infrastructure while also maintaining healthy and productive river systems is a great challenge that requires understanding the multifaceted impacts of dams at a range of scales. New approaches and advanced methodologies are needed to improve predictions of how future dam construction will affect biodiversity, ecosystem functioning, and fluvial geomorphology worldwide, helping to frame a global strategy to achieve sustainable dam development. Here, we respond to this need by applying a graph-based river routing model to simultaneously assess flow regulation and fragmentation by dams at multiple scales using data at high spatial resolution. We calculated the cumulative impact of a set of 6374 large existing dams and 3377 planned or proposed dams on river connectivity and river flow at basin and subbasin scales by fusing two novel indicators to create a holistic dam impact matrix for the period 1930 to 2030. Static network descriptors such as basin area or channel length are of limited use in hierarchically nested and dynamic river systems, so we developed the River Fragmentation Index (RFI) and the River Regulation Index (RRI), which are based on river volume. These indicators are less sensitive to the effects of network configuration, offering increased comparability among studies with disparate hydrographies as well as across scales. Our results indicate that, on a global basis, 48% of river volume is moderately to severely impacted by either flow regulation, fragmentation, or both. Assuming completion of all dams planned and under construction in our future scenario, this number would about double to 93%, largely due to major dam construction in the Amazon Basin. We provide evidence for the importance of considering small to medium sized dams and for the need to include waterfalls to establish a baseline of natural fragmentation. Our versatile framework can serve as a component of river fragmentation and connectivity assessments; as standardized, easily replicable monitoring framework at global and basin scales; and as part of regional dam planning and management strategies.

4.1 Introduction

Dams provide an important source of energy, water for irrigation, protection against floods, and help increase water security, but also have substantial impacts on the ecological integrity of aquatic systems and on the productivity of river systems that provide important resources for rural communities and regional economies (Tockner and Stanford, 2002; Arthington et al., 2010; Poff and Zimmerman, 2010; Richter et al., 2010). Two of the largest consequences of dam construction are river fragmentation and flow regulation, often considered separately in impact assessments despite their known interactions; or they are merged into aggregated impact categories (e.g., Dynesius and Nilsson, 1994; Nilsson et al., 2005). Here, we quantify the degree of river regulation and fragmentation as individual indicators, and we present the results in a matrix framework that allows simultaneous examination of both impacts. Our results can serve as a foundation for future assessments of subsequent environmental impacts resulting from these hydrological alterations, including effects on biodiversity, ecosystem functioning, and fluvial geomorphology.

River fragmentation diminishes the natural connectivity within and among river systems (Tischendorf and Fahrig, 2000; Moilanen and Hanski, 2001). We define connectivity from an ecological viewpoint with a focus on hydrology as "water-mediated transfer of matter, energy or organisms within or between elements of the hydrologic cycle" (Pringle, 2003). Following Ward (1989), connectivity has a longitudinal aspect that connects upstream and downstream ecosystems (Vannote et al., 1980), a lateral dimension by linking riverine systems with wetlands and floodplains (Tockner et al., 1999), and a vertical component that connects surface water with groundwater flows (Stanford and Ward, 1993). Longitudinal connectivity is particularly important for river ecology because of its relation to species migration and dispersal (Fukushima et al., 2007; Cote et al., 2009; Ziv et al., 2012) as well as its role in community structure and biodiversity patterns in river channels (Altermatt, 2013) and riparian zones (Jansson et al., 2000). Longitudinal and lateral connectivity also function as transport pathways for organic and inorganic matter downstream, into riparian zones and floodplains (Vörösmarty et al., 2003; Syvitski et al., 2009; Nilsson et al., 2010). Our analysis focusses on longitudinal connectivity and is based on the assumption of a direct and reciprocal relationship between fragmentation and connectivity.

Dam operation, particularly water storage, is the main contributor to flow regulation, often with the goal to eliminate peak flows, to stabilize low flows, or to impound or divert river flows. These alterations can disrupt ecological functioning (Ward and Stanford, 1995; Pringle et al., 2000; Carlisle et al., 2011), e.g., by reducing sediment delivery to floodplains and deltas (Syvitski et al., 2009), altering thermal regimes (Poole and Berman, 2001), or by disrupting life cycles of freshwater species (Poff et al., 1997). In turn, this may cause the loss of endemic species or the invasion of exotics (Bunn and Arthington, 2002), thereby reducing overall biodiversity (Poff et al., 2007; Reidy Liermann et al., 2012).

Permanent dam disruption of river systems can have effects from species to ecosystem levels and from local to global scales (Rosenberg et al., 1997). Most major global river basins are already impacted by large dams (Nilsson et al., 2005). In the future, dam development is expected to continue, with more than 3700 large hydropower dams alone currently planned or under construction worldwide (Zarfl et al., 2014). As more than one-sixth of the world's population live in glacier- or snowmelt-fed river basins (Kundzewicz et al., 2007), dams are increasingly discussed as an option to buffer against climate-induced fluctuations in water availability (Palmer et al., 2008). However, rapid proliferation of new dams may pose serious impacts on rivers, including those that support high levels of biodiversity or provide important sources of food from fisheries or flood-recession agriculture. Thus, it is of paramount importance to minimize the environmental impacts of new dams.

Recently, advanced strategies to improve the development, distribution and operation of dams by "optimizing" their geospatial location have emerged. These new approaches take into account network structure (Bunn et al., 2000; Erős et al., 2011b) and utilize newly developed hydrographical data (Lehner and Grill, 2013). In this paper, we expand on these proposals and integrate recent methodological approaches to holistically describe the current and future state of dam impacts globally. We address three principal challenges when assessing dam impacts at large scales, which are related to spatial scale, cumulative effects, and impact indicators.

Connectivity has been shown to be scale dependent (Kunin, 1998; Tischendorf and Fahrig, 2000; Fagan et al., 2005), for example, due to different dispersal abilities of species (Wiens, 2002). However, the majority of dam impact studies consider the river basin scale as the fundamental unit of study and may overlook effects at smaller spatial scales (Nilsson et al., 2005; Anderson et al., 2008; Lassalle et al., 2009). Since dispersal ability is highly variable or often unknown, multiple scales should be examined (Calabrese and Fagan, 2004). River networks have a strong hierarchical nesting structure (Fullerton et al., 2010), so advanced dam assessment frameworks should be capable of accommodating nested spatial scales within larger basins. As a step towards addressing this issue, Reidy Liermann et al. (2012) measured the length of the longest undammed stretch of the five largest rivers in each 'freshwater ecoregion' (as defined by Abell et al., 2008) to derive the percentage of free-flowing rivers.

Dams can have cumulative effects many hundred kilometers downstream and upstream of the barrier. Approaches that also take into account adjacent dams within the river system are therefore necessary but rarely performed (Fagan et al., 2005; ICPDR, 2009; Finer and Jenkins, 2012). An emerging method to assess cumulative effects in rivers is provided by graph-theoretic approaches that assess river systems as a network of links and nodes (representing river reaches and confluences, respectively). Network theory, a branch of graph theory focusing on the asymmetrical relations between network objects, can be used to study river networks (Bunn et al., 2000) and to address multiple habitats and hydrological barriers. For example, Schick and Lindley (2007) used graph theory to examine changing patterns in connectivity and the isolation of salmon populations due to dam construction in California's Central Valley. Although such approaches are commonly used in landscape ecology to quantify connectivity, their application has been limited to smaller river systems as computational requirements increase exponentially with the number of network reaches. New concepts such as the Dendritic Connectivity Index (DCI, Cote et al., 2009) and the River Connectivity Index (RCI, Grill et al., 2014) are graph-based models that avoid intense computations through simplified connectivity indices.

The third challenge for dam impact assessments on large scales is that a river constitutes a continuum of habitat types with distinct ecohydrological properties from headwater to lowland rivers (Vannote et al., 1980; Thorp and Delong, 1994; Thorp et al., 2006). Anthropogenic perturbations generally have different impacts depending on the position along a longitudinal gradient (Ward and Stanford, 1983), yet it is difficult to make a definitive statement on where along the gradient perturbations have the most impact. In the absence of such information, current assessments often treat river reaches as equally important, irrespective of their stream order or habitat suitability. For example, the DCI (Cote et al., 2009) measures the proportion of the length of the disconnected network fragments in relation to the entire network, independent of river size. Hence, the same fragmentation value may be obtained if a barrier is placed very high upstream in the network or very low, as long as the disconnected network fragments have the same length. Spatially indiscriminate metrics such as river length or river basin area may therefore be incomplete proxies for habitat in river ecology studies. To address this issue, we here propose using 'river volume' (i.e. reach length *x* width *x* depth at average flow conditions) as the basis for impact calculations in aquatic systems. In freshwater ecology, 'habitat area' and 'habitat volume' are used to consider river channel width and depth as important determinants of species composition (Schlosser, 1982). Habitat volume predict species richness better than habitat area (Angermeier and Schlosser, 1989; Magalhaes et al., 2002) and certain fish species are particularly sensitive to variations of habitat volume as a result of reduced dam releases that diminished habitat availability (Shea and Peterson, 2007).

In the following, we present a novel framework to address these challenges and to evaluate dam impact metrics by emphasising network structure, spatial scale, and incorporation of newly available hydrographical and hydrological information in a holistic connectivity assessment (sensu Fullerton et al., 2010). Our framework combines global high resolution hydrographic data (Lehner et al., 2008) with a graph-based river routing model (HydroROUT; Lehner and Grill, 2013; Grill et al., 2014). Our approach is multi-impact, multi-scale, and indicator-based—intended to compliment, not replace, more traditional local scale impact assessments. Using this framework, we address three questions: a) what are the historical trends and current spatial patterns of dam impacts on river systems resulting from flow regulation and fragmentation?; b) what differences are observed in flow regulation and fragmentation moving from subbasin to global scales?; and c) how will future hydropower dam building impact flow regulation and river fragmentation worldwide?

4.2 Methods

We develop and calculate two new indicators to assess fragmentation and flow regulation at both the basin and subbasin scale based on river volume. We then create a combined matrix of impact scores from our quantitative indicators and apply it to all river basins and subbasins globally. We first examine historic trends in dam impacts on river connectivity and flow regulation and then project future impacts due to planned or proposed dams (additional details are provided in the supplemental information (SI) text).

4.2.1 Data and models (see SI for more information)

River routing model (HydroROUT)

HydroROUT is a river routing model used to conduct tracing, routing and statistical operations in river networks (Lehner and Grill, 2013; Grill et al., 2014) based on a graph-theoretical approach (Bunn et al., 2000). HydroROUT is built on the vector river network of the HydroSHEDS database at 15 arc-second resolution (Lehner et al., 2008). In total, 17.8 million river reaches with an average length of 2.7 km are modeled in HydroROUT, representing a cumulative river length of 48.3 million km. This network includes all global streams and rivers with more than 0.1 m3/s flow (long-term average discharge) or with an upstream area of more than 10 km2. Discharge values from the global hydrological model WaterGAP (Alcamo et al., 2003; Döll et al., 2003) were downscaled to HydroROUT's river network using spatial interpolation methods. Estimates of river volume were derived from mean annual discharge (reflecting the average amount of water available to fish and fauna). According to these simulations, the global river network contains a total of 566.6 km3 of river water.

Dam and reservoir database

We considered 6374 current dams from the GRanD database (Lehner et al., 2011) and 3377 future hydropower dams compiled by Zarfl et al. (2014) in our analysis (Figure 4.1). In the dataset of future dams, 17% are attributed as 'under construction' and the rest are 'planned'. These data do not include dams at a pre-feasibility stage and dams below 1 MW capacity were excluded since information on these is sporadic and often lacks detail due to less onerous licensing requirements (Zarfl *et al.,* 2014). Reservoir storage volumes are available for current dams from the GRanD database, and we added estimates for the future hydropower dams based on a linear regression model.

In the absence of better information, we defined a two-step "future scenario" which assumes that all 'under construction' dams are built by 2020, and that all 'planned' dams are completed by 2030. More information on this scenario and its plausibility is provided in the Discussion section and in the SI (S1.2). For simplicity, the expression "future scenario" will refer to the 2030 horizon from here on, unless stated otherwise.

Uniform spatial units

In addition to river basins, we used a set of subbasin units termed HydroBASINS (Lehner and Grill, 2013) to assess the sensitivity of our index calculations to spatial scale. HydroBASINS is a delineation of global watersheds and was developed to provide nested subdivisions of large river basins to conduct disaggregated spatial analyses in river systems.

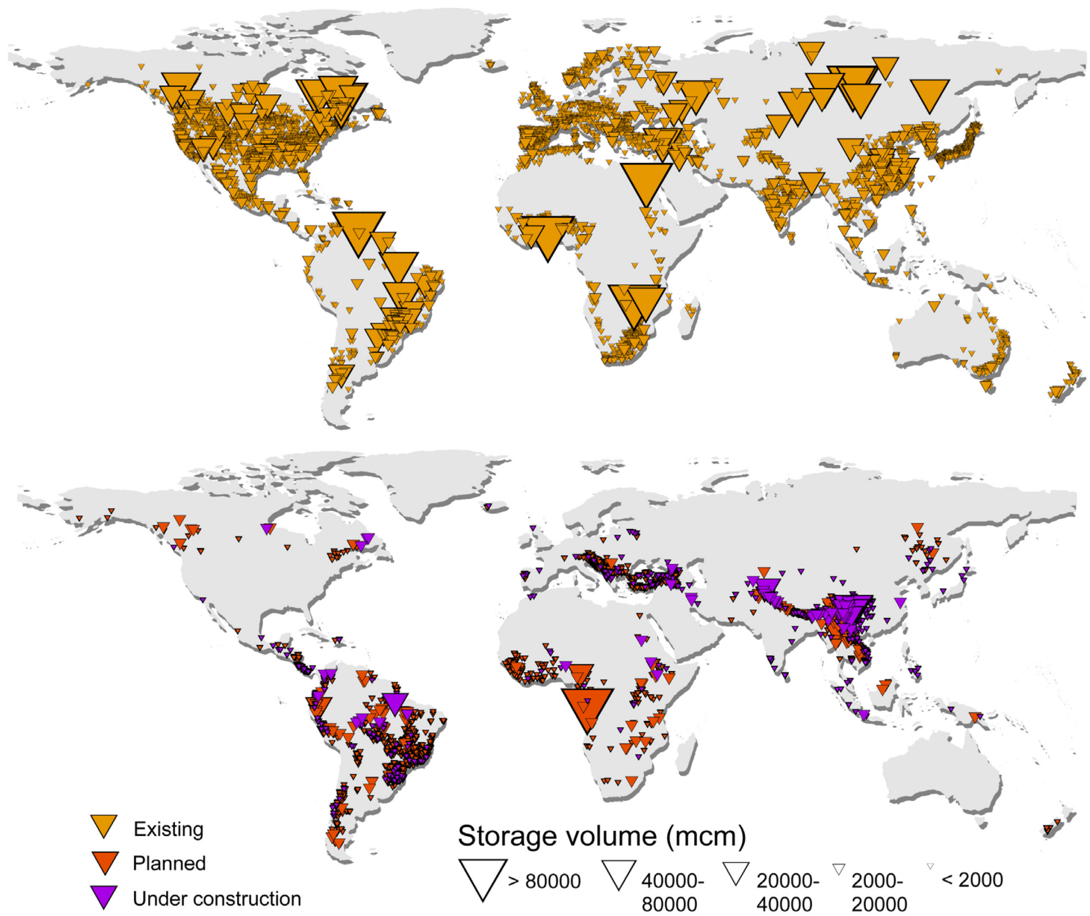


Figure 4.1: Overview of existing dams (GRanD; Lehner et al., 2011) and planned and under construction future dams (Zarfl et al., 2014) by storage volume class (volumes from Lehner et al. (2011) and own estimates).

4.2.2 River Fragmentation Index (RFI)

The River Fragmentation Index (RFI) is a measure of river fragmentation by barriers on structural connectivity (expressed as a percentage of full connectivity) per basin or subbasin; it is equivalent to the River Connectivity Index (RCI) as described in Grill et al. (2014). The RFI is based on the Dendritic Connectivity Index (DCI) by Cote et al. (2009) but substitutes river volume for river length. The RFI of an unfragmented river network is 0%, with each subsequent dam increasing the value to a maximum of 100%. A single dam in a previously undisturbed network leads to greatest fragmentation if it splits the network into two equal volume fragments, in which case the RFI increases to 50%.

4.2.3 River Regulation Index (RRI)

The River Regulation Index (RRI; Grill et al., 2014) is an extension of the Degree of Regulation (DOR) as calculated globally in Lehner et al. (2011) and provides a quantitative proxy of how strongly a river may be affected by alterations to its natural flow regime due to upstream dam operations. The DOR is the proportion of a river's annual flow volume that can be withheld by a reservoir or a cluster of reservoirs upstream of the reach and is calculated for each reach of the network. The DOR has in one form or another been a key component of seminal studies on flow regulation (e.g., Nilsson et al., 2005) or has been analyzed in terms of the hydrologically equivalent 'change in residence time' or 'water aging' (e.g., Vörösmarty et al., 1997). A high DOR indicates an increased probability that substantial discharge volumes can be stored upstream in a given year for future release. We calculate RRI (%) by first weighting the DOR value of each individual reach with its corresponding river volume, and then averaging the results for the entire basin to quantify full-basin impacts in a single index.

4.2.4 Dam Impact Matrix (DIM)

Building upon previous concepts by Dynesius and Nilsson (1994), we assessed fragmentation and regulation effects simultaneously by creating an impact matrix. We first classify each index into four categorical groups based on quartile ranges of occurrence (*weak*: 0-25th; *moderate*: 25-50th; *heavy*: 50-75th; and *severe*: 75-100th percentiles) and then combine the categories from each index to create an integrated four by four matrix. This matrix, based on relative rankings (from low to high), allows comparison across basins worldwide in order to illustrate the large spectrum of possible combinations while identifying four primary groups of impacts at each corner of the matrix (see Figure 4.2). It is not a goal of this study to interpret or compare absolute impact scores or to define ecological thresholds; as such, the assigned class names only represent a statistical ranking and should not be judged as expressing the level of ecological impact (e.g., even the ‘weak’ impact class may include river basins that experience substantial ecological perturbations).

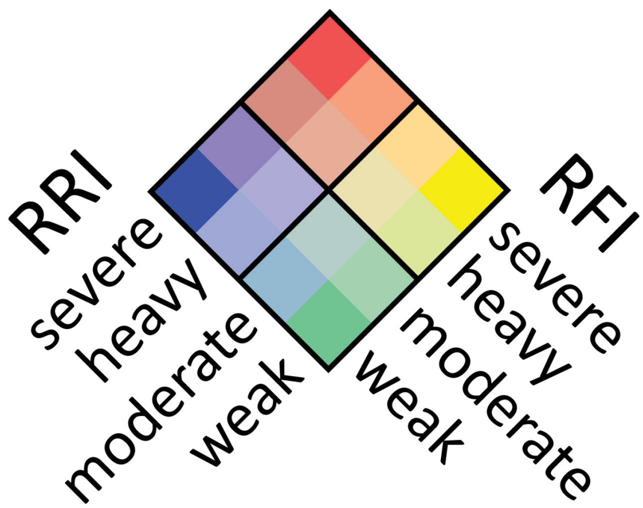


Figure 4.2: Legend for Dam Impact Matrix (DIM) showing qualitative impact categories for the River Fragmentation Index (RFI) and the River Regulation Index (RRI). The 16 possible combinations can be grouped into 4 broader groups representing types of impact: 1) basins with both low fragmentation and low flow regulation (lower quadrant, green colors); 2) basins with high fragmentation but low flow regulation (right quadrant, yellow colors); 3) basins with low fragmentation but high flow regulation (left quadrant, blue colors); and 4) basins with both high fragmentation and high flow regulation (top quadrant, red colors).

4.3 Results

4.3.1 Global trends in fragmentation and flow regulation

Averaged over all basins, both fragmentation (+32%) and flow regulation indices (+43%) deteriorated substantially, with the greatest change between 1950 and 1980 (Figure 4.3). After 1980, the trajectory of both curves indicates that the rate of deterioration due to dam building has slowed considerably, especially after the year 2000. The construction of all future dams by 2030 would further increase fragmentation (+12%) and flow regulation (+10%) at rates similar to the maximum changes of the last century.

Globally, a total of 1293 main (i.e. not subdivided) river basins contain large dams today. These basins represent 59% of global rivers with 28.6 million km of combined length (Figure 4.4 and Table S4.2). An additional 209 basins are affected by at least one future large dam in the 2030 scenario, an increase of 16% in

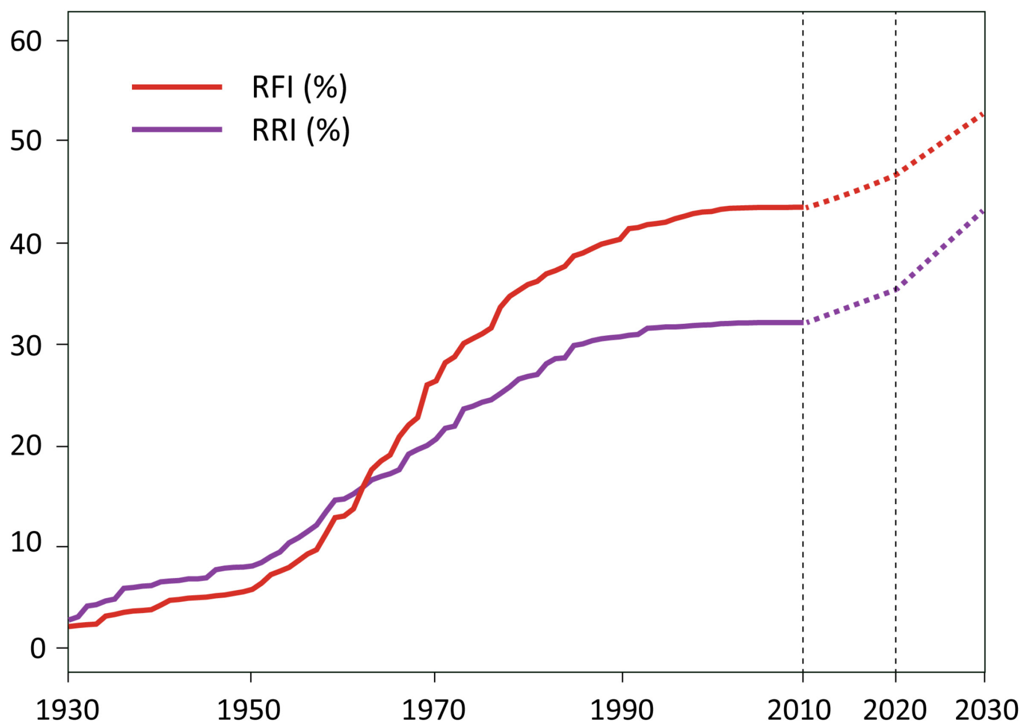


Figure 4.3: Graph showing the trajectory of RFI and RRI indices averaged over all global basins between 1930 and 2010 (based on GRanD) and for a future hydropower scenario based on Zarfl et al. (2014). Values reflect area-weighted means of indices across all global basins.

the number of affected basins; an additional 1.6 million km of rivers would be affected (6% increase). The total storage volume would increase by 39% from 5759 km^3 in 2010 to an estimated 8007 km^3 in 2030.

The total river length in basins unaffected by large dams today amounts to 41%; however, many of these river reaches are in arid or semi-arid regions. If river volume is substituted for river length, we find that basins not impacted by any large dams in our analysis contain just 7% of global river volume, meaning that 93% of the world's river volume lies in basins with at least one dam (Figure 4.4 and Table S4.2). However, this result is highly influenced by the Amazon River which drains roughly one third of global discharge. For a more conservative estimate, we only assessed basins that fall within the fragmentation classes of moderate to severe (i.e. 25th percentile and up). With this alternative estimate, the total river volume of all moderately to severely fragmented basins today amounts to 43% of the global river volume; in the 2030 future scenario, this more than doubles to 89%, suggesting that new large dams will add major pressure to global river basins.

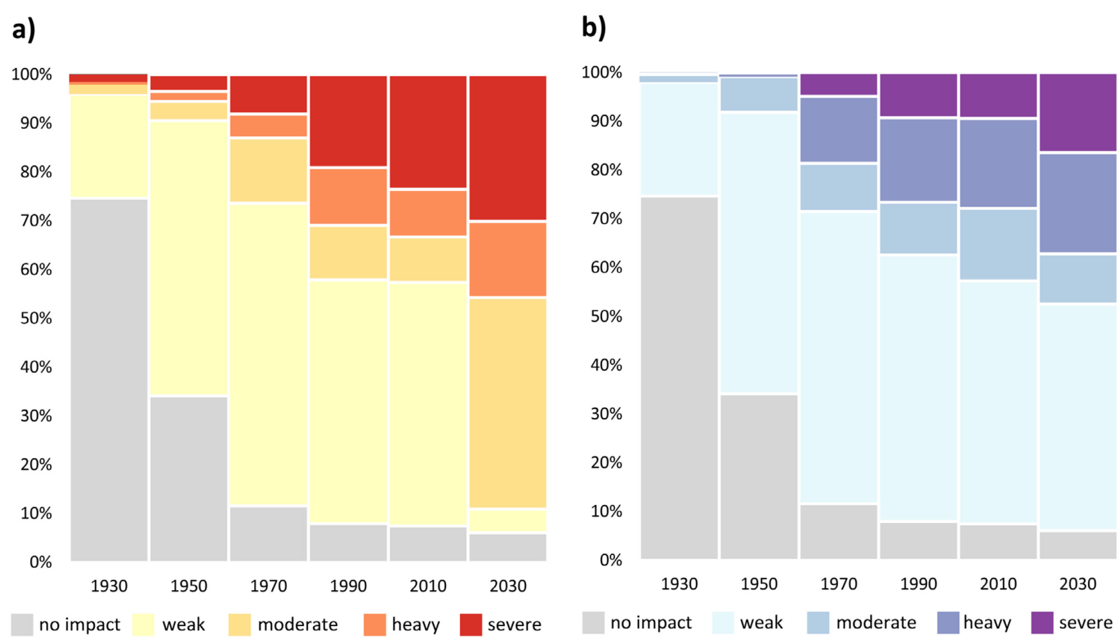


Figure 4.4: Proportion of global river volume impacted by fragmentation (a) and flow regulation (b) for each impact category (see Figure 4.6 for classification criteria). See Table S4.2 for impact values summarized by affected length (km) instead of volume.

Regarding flow regulation, the total river volume of all moderately to severely regulated basins today also amounts to 43% of global river volume, almost the same value as for fragmentation; yet the individual impact classes have a very different distribution: for example, 9% of river volume is severely affected by flow regulation while 24% of river volume is severely affected by fragmentation (Figure 4.4 and Table S4.2). In the 2030 future scenario, the total percentage of moderately to severely impacted rivers for flow regulation increases only slightly to 48%, but many rivers experience shifts from lower to higher impact classes.

In combination, today a total of 48% of river volume is moderately to severely impacted by either flow regulation, fragmentation, or both. In the 2030 future scenario, assuming completion of all dams planned and under construction, this number would rise to 93%, mostly due to large dam construction in the Amazon Basin.

In order to compare our volumetric results to Lehner et al. (2011) who focussed on river length, we calculated the volume of all reaches with DOR values ≥2%, a threshold previously used to distinguish "affected" rivers (Dynesius and Nilsson, 1994; Nilsson et al., 2005; Lehner et al., 2011). The volume of these regulated downstream rivers currently amounts to 34% of global river volume, rising to 65% in the 2030 scenario mostly due to large dam construction along the Amazon. The volume affected is substantially greater than for length (estimated by Lehner et al. (2011) to be ~8% of global river length). The effects of flow regulation are therefore skewed towards high volume rivers, while large dams less directly affect smaller rivers.

While global trends illustrate the general trajectory of worldwide dam developments and their impacts, there is significant regional heterogeneity in both temporal and spatial patterns. For example, some river basins show relatively low impacts by large dams until recently, including the Amazon, Mekong and Salween Rivers (Figure 4.5). On the other hand, river basins such as the Nile, Mississippi, Nelson and Indus were already heavily impacted early in the 20th century. While some river basins have deteriorated in both RFI and RRI in the

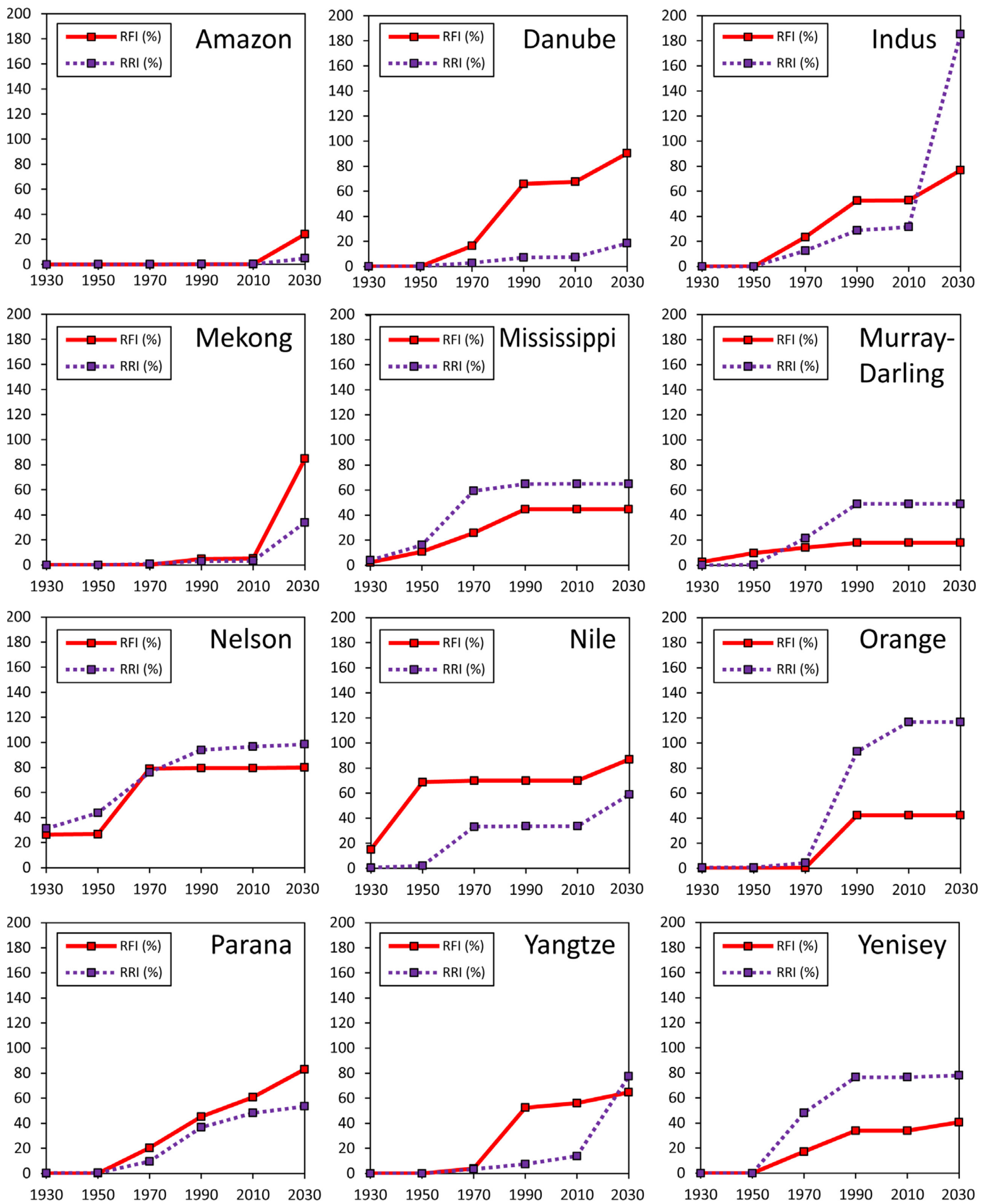


Figure 4.5: Changes in fragmentation and flow regulation for 12 selected large basins.

past, others show impacts only in one of the indicators. For example, the Murray-Darling is only weakly affected by fragmentation, but is heavily impacted by flow regulation (due to large reservoir volumes coinciding with low flow volumes). In contrast, the Danube is severely impacted by fragmentation effects, but only weakly affected by flow regulation. In the future scenario, some basins may not experience much additional change, such as the Yenisei or the Nelson, while others have substantial increases in RFI alone (Yangtze, Danube, Parana), RRI alone (Irrawaddy, Indus), or for both indices (Mekong, Nile, Orange).

4.3.2 Past and current impacts at basin and subbasin scales

River fragmentation

Our fragmentation analysis based on RFI reveals that 43% of the global river volume is moderately to severely impacted today, and that severe impacts affect 24% of the global river volume (Table S4.2). A total of 96 large river basins (defined as >350 m^3/s discharge) are heavily to severely impacted by fragmentation from dams (Figure 4.6 and Table S4.3). However, rivers can appear heavily or severely impacted at the overall basin scale, while the subbasin scale may reveal many less impacted areas, e.g., if most dams are clustered only in certain tributaries. Examples are the Mississippi River in North America, the Parana River in South America, or the Niger, Zambezi and Nile Rivers in Africa, which all appear heavily or severely affected at the basin scale but at the subbasin scale larger proportions or even the majority of reaches are only weakly to moderately affected. In particular, dams at central locations relative to the full basin fragment the entire network resulting in severe degradation at the basin scale; the same dams can result in lower fragmentation scores at the subbasin scale, particularly if they are located at subbasin boundaries.

Flow regulation

Our analysis shows that 9% of global river volume is severely affected by flow regulation from dams (less than by fragmentation) and 18% and 15% are heavily and moderately affected, respectively (more than by fragmentation; Table S4.2). A total of 103 large river basins are heavily and severely affected by flow regulation, in particular extensive regions of North America, Africa, the Iberian Peninsula, Eastern Europe, Eastern Asia, and South-central Asia (Figure 4.6 and

Table S4.3). As with fragmentation, there are noteworthy differences of impact classifications at smaller scales. For example, most flow regulation impacts of the Mackenzie River in northern Canada result from dams in higher tributaries with propagating effects downstream, whereas other parts of the basin are less affected. The Amazon and Congo Rivers have been relatively unaffected by flow regulation, and some of their tributaries show no signs of direct impact by large dams today. Rivers of the Greater Mekong Region (including the Mekong itself) as well as smaller rivers such as the Rhine and Po in Europe have been relatively unaffected by flow regulation from dams as well, due to either fewer dams or lower reservoir capacities.

Dam Impact Matrix of fragmentation and flow regulation

Although many dams have comparable impacts on river flow and fragmentation, some dams cause a bigger impact on one or the other. This translates into numerous basins being more affected in only one of the two impact categories (Figure 4.7).

The top quadrant of the DIM (red colors), i.e. basins heavily to severely affected by both flow regulation and fragmentation, includes a total of 407 basins (21% of global river volume). This category highlights basins that have both dams on mainstem rivers and large reservoir volumes high upstream in the network. These are typically basins with a long history of dam building (e.g., the Nile, Mississippi, or Yangtze).

A total of 221 river basins (right quadrant, yellow colors) are heavily to severely affected by fragmentation, yet only weakly to moderately impacted by flow regulation (12% of global river volume). This category represents basins with a majority of run-of-the-river dams that have high impacts on fragmentation but relatively low storage capacities (e.g., the Danube).

The left quadrant of the DIM (blue colors) combines 234 river basins that are heavily or severely impacted by flow regulation but only show weak to moderate impact from fragmentation (7% of global river volume). Examples in this category include the Nelson, Ob, or Murray-Darling, with large reservoirs located in their headwaters.

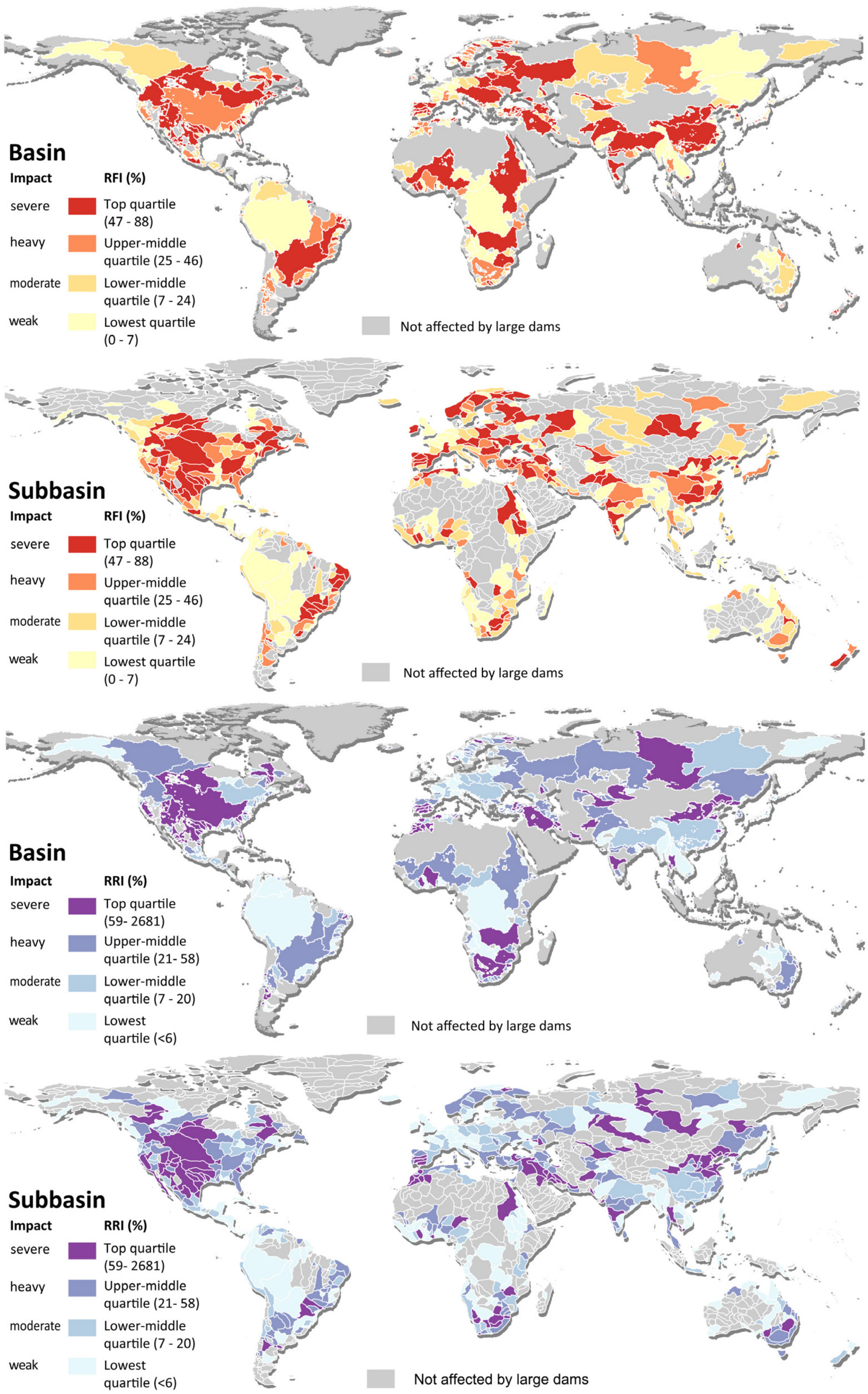


Figure 4.6: Todays fragmentation and flow regulation indices at the basin and subbasin scale. The RFI and RRI values are each classified according to quartiles (0-25th, 25-50th, 50-75th, and 75-100th percentiles).

River basins remaining only weakly to moderately affected by large dams in terms of both fragmentation and flow regulation (bottom quadrant, green colors) amount to 431 (53% of global river volume). Examples of these least impacted basins include certain parts of the Greater Mekong Region, several basins in south-central Asia, the Amazon, Orinoco, Tocantins, as well as large proportions of Western and Eastern Europe.

River basins that remain only moderately to medium affected by large dams in terms of both fragmentation and flow regulation account to 431, or 52.6% of river volume (bottom quadrant, green colors). Examples of these least impacted basins include certain parts of the Greater Mekong Region, several basins in South-central Asia, the Amazon, Orinoco, Tocantins, as well as large proportions of Western and Eastern Europe.

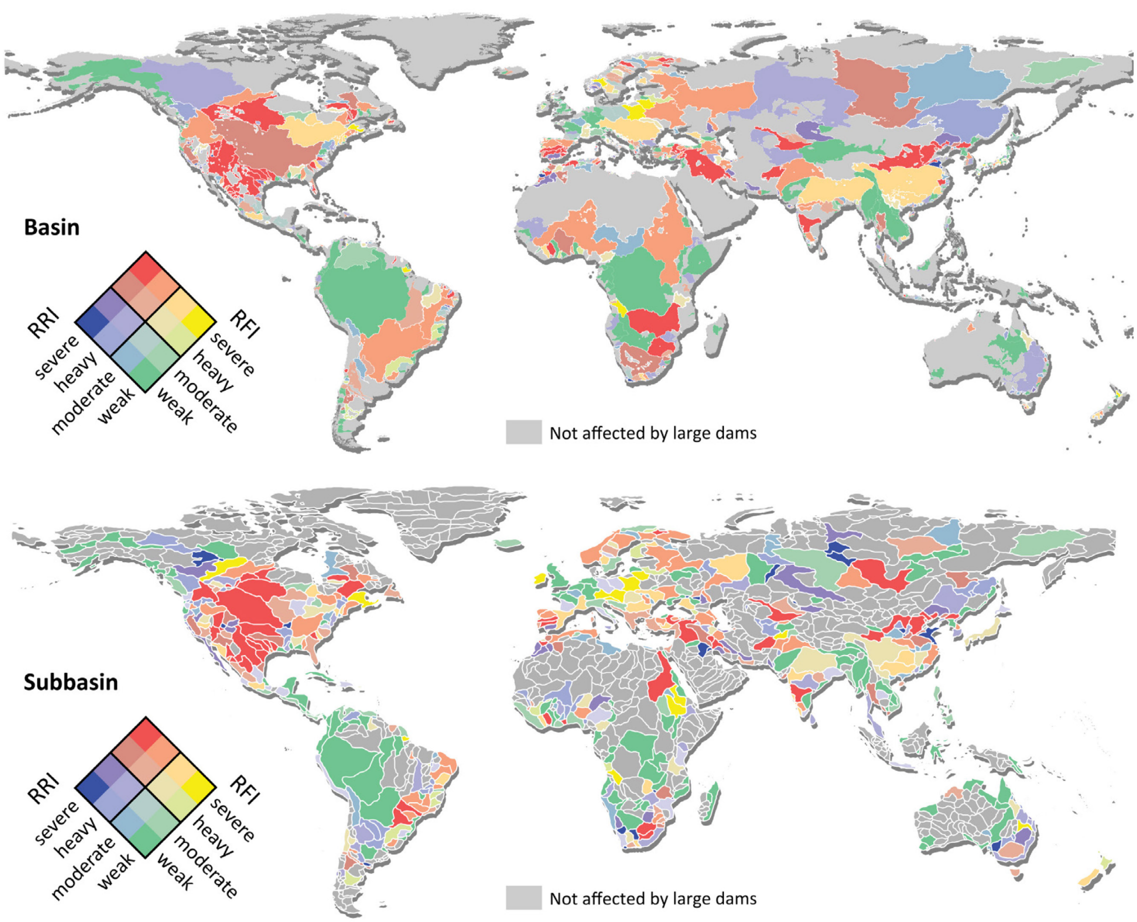


Figure 4.7: Combining fragmentation and flow regulation indices for the basin (top) and subbasin scale (bottom) for today's situation. The RFI and RRI values are given in percent. The class breaks correspond to the percentiles classification in Figure 4.6. This approach allows identification of four primary groups of impacts (see caption of Figure 4.2 for more details on classification scheme).

4.3.3 Future impacts at basin and subbasin scales

River fragmentation

New dams in our 2030 future scenario stress many currently less-affected basins, encompassing a large proportion of global river volume. For example, in the currently weakly affected basins of the Greater Mekong Region, rivers such as the Salween, Irrawaddy and Mekong undergo significant changes, with substantial deterioration of connectivity (see Figure S4.2). These basins are classified in the 2030 scenario as heavily or severely impacted in both RRI and RFI. The Yangtze River shows a similar deterioration, particularly in upstream portions of the basin where a large number of new dams reduce connectivity at the basin scale.

Substantial losses in connectivity are also predicted in the Amazon Basin (RFI +24%). The subbasin scale reveals which reaches contribute to this decline: numerous dams in the middle and lower portions of the Madeira, Tapajos, and Xingu subbasins cause vast increases in fragmentation (RFI +73%, +79%, and +50%; Figure S4.2). A large number of dam projects concentrated in the higher Andes region of the Amazon Basin lead to smaller connectivity losses (RFI +4%) in the subbasin upstream of the Madeira tributary, but these are still considered a substantial change (see Finer and Jenkins, 2012). The Tocantins basin is also highly affected by fragmentation, losing 56% of its connectivity.

A number of planned dams in the upper Nile Basin (Lake Victoria catchment) cause additional impact (+18%) to an already severely affected basin. In North America and Europe, changes in fragmentation are less extreme (e.g., Mackenzie: +12%; Danube: +23%), in part because river systems in these regions are already quite highly fragmented. In Siberia, changes in fragmentation are projected to be mostly smaller, yet with stronger impacts in some subbasins. The Amur in East Asia shows very substantial increases in fragmentation (+46%) due to a new central mainstem dam and numerous large tributary dams.

Flow regulation

Our model results for the 2030 future scenario suggest that an additional 209 currently unaffected river basins will be affected by flow regulation. There are 46 additional basins severely affected as a result of 1156 future dams. This causes the amount of 'severely affected' river volume to rise from 9% to 16%. Smaller increases occur in the 'heavily affected' category (1351 dams), with a total increase of 2% in an additional 28 basins.

A number of hot spots with more drastic changes are shown in Figure S4.2. For example, hundreds of dams are planned in the mountains of the southern slopes of the Himalaya and Pamir Mountains, which would lead to substantial downstream impacts, particularly dominant in the Indus (113 dams; RRI +153%) and across almost all Brahmaputra River subbasins (392 dams; +15%). The Salween River would experience a 132% increase in flow regulation due to 25 planned dams. Another hot spot is the upper Yangtze River where 133 large dams are planned, many of them on the mainstem (+71%).

Dam Impact Matrix of fragmentation and flow regulation

Future dam development will occur mostly in basins that are already affected by dams, so only an additional 2% of river volume from previously unimpacted basins is newly affected by future dam construction in 2030. However, substantial impacts and shifts are found within the individual groups of the DIM (Figure 4.8). The number of basins that become heavily to severely affected by both regulation and fragmentation (northern quadrant of DIM, red colors) increases by 71 (an additional 11% of global river volume). Basins in the southern (green) quadrant also increase by number as formerly unimpacted (and mostly small) basins become impacted. But this group now aggregates less total river volume (-9%) because larger basins have shifted towards the higher impact groups.

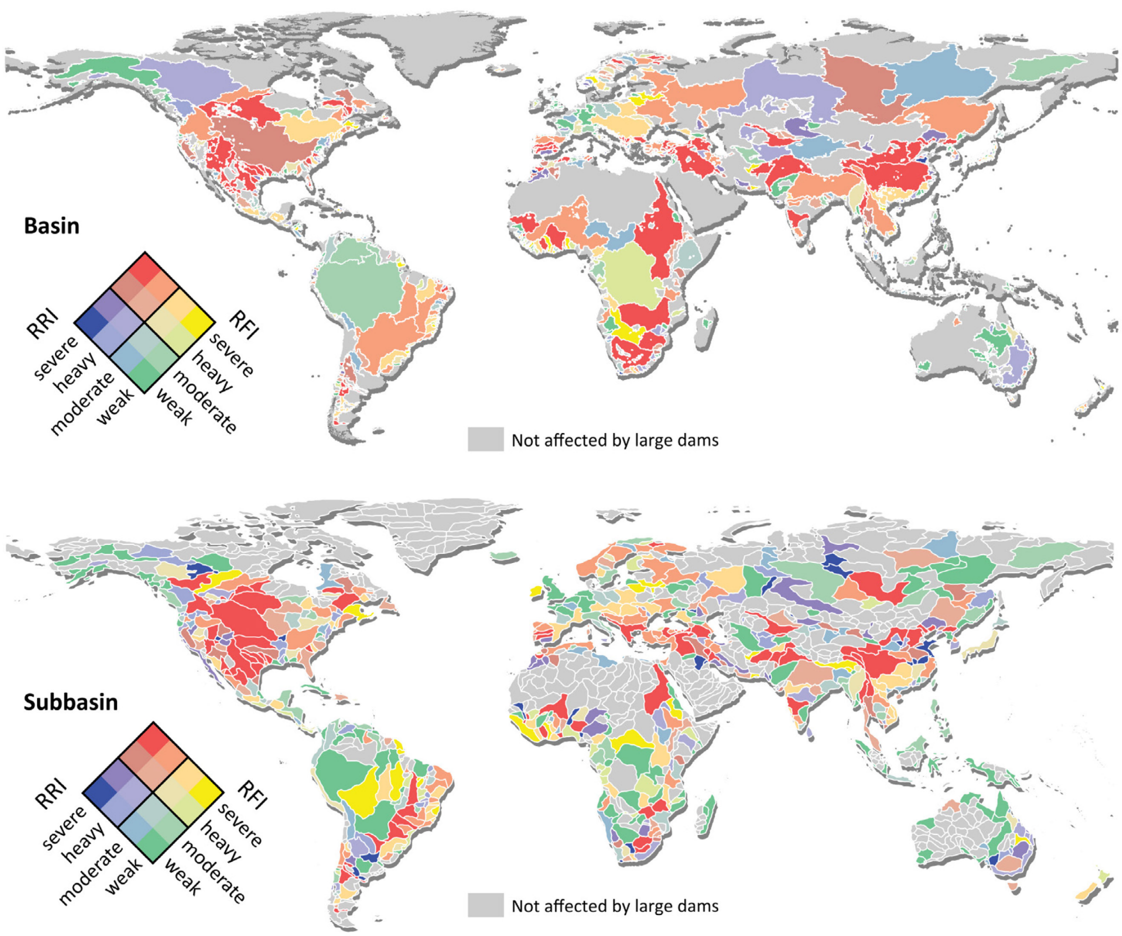


Figure 4.8: Combined impacts for the future scenario of 2030 in which all large hydropower dams currently planned or under construction are built (see caption of Figure 4.2 for more details on classification scheme).

4.4 Discussion

Several studies have reported widespread global effects of dams (e.g., Nilsson et al., 2005; Vörösmarty et al., 2010; Lehner et al., 2011; Reidy Liermann et al., 2012), but mostly at coarse spatial or temporal resolutions. Using the best currently available hydrographic data, our study is the first global analysis where a graph-based river routing model was applied to model past and future impacts of dams at multiple scales and at high spatial resolution. While most frameworks include river length or surface area as an indicator of dam impacts, we developed an indicator based on river volume. River volume may more adequately represent freshwater habitat space and aquatic biodiversity, assuming that larger rivers generally have a greater opportunity for more heterogeneous habitats that fosters greater overall biodiversity (Xenopoulos and Lodge, 2006). Nevertheless, in addition to longitudinal fragmentation, river systems are also affected laterally, with implications for riparian systems and floodplains. As a relatively new measure, more research is needed to determine the strengths and weaknesses of river volume for specific ecohydrological applications.

We also found that river volume improves assessment of cross-scale impacts as it inherently incorporates the concept of stream orders. If river length is used, low level stream orders (first- and second-order streams) can add disproportionately large amounts of river length to the network (>70% of rivers are headwater streams; Lowe and Likens, 2005), yet their contribution to volume is small. By using river volume, network configuration and the range of included (smaller) stream orders becomes less important, enabling greater comparability of indicators between studies with differently detailed river networks.

Our study confirms that examining dam impacts at different scales is critical. Indicators on the basin and subbasin scale can target different applications, each with specific advantages and disadvantages. For example, calculations at the basin scale integrate impacts across the entire river system, allowing for inter-basin comparison. This is particularly useful in cases where connectivity or flow regulation need consideration at the scale of a single basin (e.g., movement patterns of long distance migrating fish, such as large catfish in the Mekong River). However, the wide range of basin sizes—spanning several orders of magnitude globally—confounds comparisons between small and large basins. Our subbasin results reveal higher intra-basin detail and better differentiation among river systems. Nonetheless, achieving a homogenous hierarchical nesting is more difficult in river systems compared to terrestrial systems (see Kunin, 1998). There are several valid ways to partition basins into subbasins, typically based on the size or stream-order of tributaries, each triggering different results. Selecting a useful and homogeneous subbasin breakdown, such as provided by HydroBASINS, is thus a critical task.

Small and medium sized dams (not included in our assessment) can have a significant cumulative effect on flow regulation (Lehner et al., 2011), but similar effects for fragmentation have not been systematically analyzed to date at the global scale. We conducted a sensitivity analysis using the Mississippi Basin and compared the resulting effect of all 25,857 dams included in the National Anthropogenic Barrier Dataset (NABD; Ostroff et al., 2013) to the effect of the set of 704 large dams provided in GRanD. We found substantial changes for both RFI and RRI indicators, with increases in basin-wide flow regulation from 65% to 90% and increases in fragmentation from 45% to 65% (Table S4.1). At the subbasin scale (Figure S4.1), changes vary throughout the basin, with the Arkansas and the Missouri Rivers being most affected by river regulation, and the Upper Mississippi and the Red River showing the greatest fragmentation changes. These findings suggest that global impacts of all dams—large and small—are likely far more severe than illustrated by our results.

Natural barriers, such as waterfalls, could have similar consequences for river network connectivity as dams (Torrente-Vilara et al., 2011; Dias et al., 2013); although they also have unique characteristics such as typically allowing for greater permeability for species in the downstream direction. We conducted an exploratory assessment with a new global dataset of waterfalls (Figure 4.9 and Figure S4.3). When incorporating waterfalls, the natural connectivity of many watersheds is reduced, which provides a different baseline for our fragmentation assessment. The fragmentation values of dams built in the vicinity of waterfalls may therefore be smaller than in a fully connected network; on the other hand, a dam built in the middle of a subbasin upstream of a waterfall may have more significant impacts for this already isolated subbasin than if its downstream part were still connected to the larger system. The inclusion of waterfalls thus has implications for impact assessments of individual dams or groups. Similarly to waterfalls, intermittent rivers lead to seasonally diminished hydrological connectivity and may have comparable effects on natural river connectivity patterns. To include a more differentiated view of natural discontinuities, more data on their location as well as a better understanding of their effect on passage of species up- and downstream is required.

Our future dam scenario is designed to provide an illustration of the potential impacts of plausible new hydropower developments, yet it should not be misinterpreted as a prediction of the 'most likely' future situation. There are typically more dams planned than are actually built, related to funding and other political and economic factors that are difficult to predict, thus determining the likelihood of planned dams being commissioned is afflicted by high uncertainties. In contrast, we consider the completion of dams already under construction rather likely. Our scenario shows that if only the dams that are under construction today (representing 17% of the 3700 dams in Zarfl *et al.*, 2014) were built, fragmentation and regulation indices would increase by 4% each, which constitutes 31% and 38% of the total future increases by 2030 for RFI and RRI, respectively. This confirms that a large proportion of the anticipated impacts will already be caused by rather likely developments in the near future, and a similar trend would then continue if all planned dams were built by 2030 (see Figure 4.3).

However, important shortcomings of our scenario assessment remain. The geographic coverage and completeness of the applied database of future dams is difficult to verify, warranting extra caution when drawing conclusions based on differences between regions. Also, the focus here on hydropower dams may lead to a bias towards fragmentation (rather than flow regulation), yet this may adequately reflect the recent tendency towards building multi-purpose dams while storage-only dams become less common globally. Finally, we acknowledge that our simplified estimation of reservoir volume for future dams based on hydropower capacity has considerable uncertainties. Nevertheless, since our future scenario is based on large hydropower dams only, we likely underestimate the total impacts of other dam types, particularly those related to climate change mitigation and water storage (e.g. irrigation and flood control), as well as small future dams.

An important research challenge remains unaddressed in our study, namely to relate these indicators to actual changes in habitat structure, biotic composition, or biodiversity. As a step towards addressing this challenge, our framework is capable of providing spatially disaggregated changes of impact scores over time and can therefore be used to monitor the fragmentation history of a basin. Such information could be related to past changes in biological indicators to determine if the construction of a specific dam was associated with ecohydrological changes in the river basin.

Given the current capabilities and insights that our framework provides, our approach can also help to identify individual dams or sets of 'hot spot' dams to guide researchers and water resources managers in determining where to conduct more detailed local environmental impact assessments. With an increasing number of dams becoming dysfunctional due to sedimentation, our framework could also inform regional dam removal strategies by prioritizing which dams would potentially provide the biggest benefit if removed (O'Hanley, 2011; Hoenke et al., 2014).

Finally, the framework can be applied in support of conservation planning efforts (Hermoso et al., 2011). Although many basins are currently impacted, we have identified basins that are relatively pristine in terms of dam effects, but under pressure from possible future dam construction. As a large-scale framework, our methodology could be used to minimize further exposure, for example, by identifying free-flowing river sections as part of a strategy to derive conservation targets or to design protected areas (Pringle, 2001; Abell et al., 2007; Thieme et al., 2007).

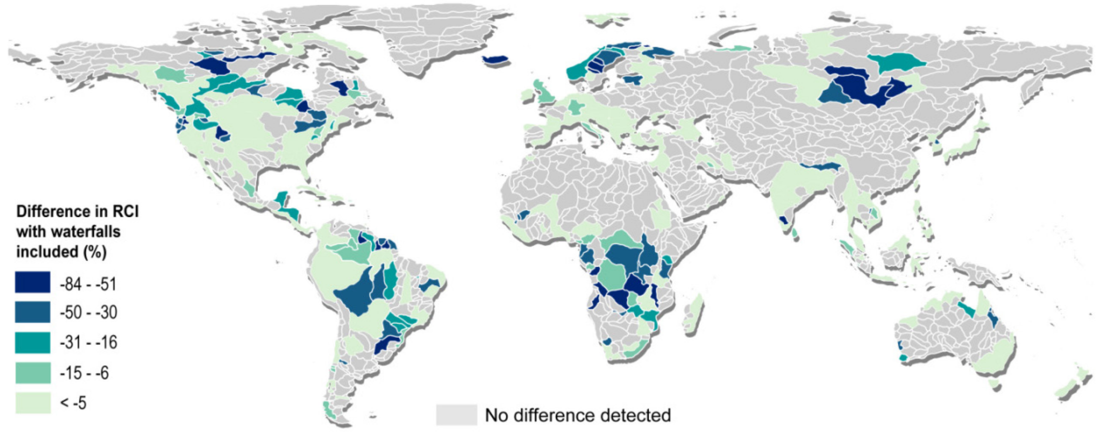


Figure 4.9: The effect of waterfalls on river connectivity at the subbasin scale for the year 2010 calculated as absolute difference in index values ("with waterfalls" versus "without waterfalls").

4.5 Conclusion

We developed a versatile framework to assess river fragmentation and flow regulation by dams based on state-of-the-art global hydrographic data and novel approaches using discharge-based indicators. Almost half of the global river volume is moderately to severely impacted by either flow regulation, fragmentation, or both. Assuming completion of all hydropower dams planned and under construction in our future scenario, this number would increase dramatically.

Assessing the effects of dams on river networks is a complex endeavor, due to the need to account for interacting and cumulative effects of multiple types of flow regulation and fragmentation perturbations. We suggest that multiple indicators should be assessed simultaneously, and that naturally reduced connectivity by waterfalls and intermittent rivers is included. River volume proved to be a more representative and robust metric for assessing river systems across scales compared to commonly used metrics such as river length or basin area. We identified substantial intra-basin heterogeneity of impacts which was previously difficult to assess, suggesting that studies should be conducted at multiple scales.

We found that prolonged and prolific dam building has resulted in large-scale deterioration of the majority of global river basins, with at times heavy to severe impacts. This result is in good agreement with previous studies (Nilsson et al., 2005; Lehner et al., 2011). Yet our new indicators reveal an even higher impact when river volume (rather than length) is used as the basis of assessment. A sensitivity analysis for the Mississippi River in which we added small and medium dams to our global database of large dams suggests that our results are conservative, and that global dam impacts are likely of much greater concern than illustrated here when small dams are considered as well.

Our research offers a consistent framework for assessing large-scale dam impacts over space and time in a world with increasing pressures on water resources. We believe that the proposed method and indicators can be applied in multiple ways: as a standardized, easily replicable monitoring tool that provides comparable global and basin-scale indicators of changes and trends in support of international initiatives such as the Millennium Development Goals (MDG), the Global Biodiversity Outlook (GBO-4) or the Biodiversity Indicator Partnership (BIP); as a complementary set of indicators to support existing methods such as the Indicators of Hydrological Alteration (IHA; Richter et al., 1996) or the Ecological Limits of Hydrologic Alteration (ELOHA; Poff et al., 2010); or as a contribution to more comprehensive assessment strategies that evaluate existing and planned hydropower projects such as the Hydropower Sustainability Assessment Protocol (HSAP). We strongly encourage practitioners and relevant agencies to systematically compile the required information (foremost dam locations, reservoir purpose, and storage volumes) to support these kinds of assessments.

The results of our research emphasize the need for water managers and planners to consider cumulative, large-scale impacts of multiple dams as part of an integrated 'river systems' mindset. In this regard, our framework can be implemented in strategic dam planning efforts and regional scenario developments to help identify the most critical sets of dams or alternative options in efforts to minimize social and environmental tradeoffs associated with dam development while maintaining their socio-economic benefits.

Acknowledgements

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada; World Wildlife Fund (WWF Greater Mekong and WWF US); the EU FP7 BioFresh Project, Berlin, Germany; the Global Environmental and Climate Change Centre (GEC3), Quebec, Canada; and McGill University, Montreal, Canada. G.K. MacDonald received additional research support from the Gordon and Betty Moore Foundation. The authors wish to thank Peter Meier, International Centre for Environmental Management (ICEM) for advice on calculating reservoir storage volume, and Jeff Opperman, The Nature Conservancy (TNC) for contributions to improve the manuscript. We are grateful to three reviewers who provided valuable comments and constructive criticism and to many other colleagues who contributed to this research and the manuscript.

Supplementary Information

S4.1 Supplementary methods

S4.1.1 River routing model (HydroROUT)

The main functions of the HydroROUT model (Lehner and Grill, 2013; Grill et al., 2014) include traversing and analyzing a set of reaches as a connected river network, and to calculate spatial statistics on, over and across river networks, for example by calculating the degree of regulation downstream of a dam, by deriving statistics for the whole basin, or by conducting comparisons between river basins (sensu Peterson et al., 2013). HydroROUT is implemented in C# programming language and uses an ArcGIS geodatabase as the container for all geospatial data.

The underlying network graph of the HydroROUT model is based on a global set of vector river lines, derived from the HydroSHEDS dataset (Lehner et al., 2008) at 15 arc-second spatial resolution (approx. 500m x 500m pixel size at the equator). In total, a set of 17.8 million river reaches with an average length of 2.7 km has been integrated, representing a total global river length of 48.3 million km (this network comprises all rivers and streams with an average flow equal or larger than 0.1 m^3/s, or a catchment area of at least 10 km^2).

River water volume is used as part of the fragmentation and regulation indices, reflecting the amount of water available to fish and fauna. Each river reach has a simulated long-term average discharge value assigned, which has been derived through GIS downscaling techniques (spatial interpolation and re-allocation) from runoff estimates of the global hydrological model WaterGAP for the time period 1961-90 (Alcamo et al., 2003; Döll et al., 2003). River volume, i.e. the water volume (length x width x depth) within the river channel at average discharge is then estimated based on an approximation of channel width and depth following Allen et al. (1994). The global river network as defined above contains a total of 566.6 km^3 of river water according to our simulations.

S4.1.2 Dam and reservoir database

We used information on the location, year of construction, and storage volume of current dams from the Global Reservoir and Dam (GRanD) database (Lehner et al., 2011), yet removed some dams from the database due to missing information about the year of construction or reservoir volumes. We included a total of 6374 out of the 6862 dams available (Figure 4.1). The regulation dam at Lake Victoria was also excluded in our calculations as this dam is not operated at full capacity and would introduce a large bias in the Nile basin due to its central location and extremely large potential storage capacity.

The data for our future scenario assessment was compiled by Zarfl et al. (2014), comprising more than 3700 hydropower dam projects currently under construction (17%) or planned (83%). It should be noted that this database focuses on large hydropower projects only. Dams currently under construction or planned and with a capacity larger than 1 MW were included in the database. They were annotated as 'planned' if they were described as such in the original data source, or if they were reported as being at a feasibility stage where social, cost-benefit and environmental aspects were under evaluation. The level of licensing requirements in terms of hydropower regulation is inconsistent among different sizes of schemes and also differs between countries, thus there remains an inherent uncertainty about the likelihood of the progression of a given project. To reduce this uncertainty, records were cross-validated with multiple data sources, where possible, to confirm the status of the project, or to provide attributes that were missing in the original data source. For consistency with GRanD, a total of 323 future dams were excluded, namely those located at a river with a discharge smaller than 1 m^3/s, as almost all GRanD dams conform to this criteria, leaving 3377 to be included in our analysis.

We estimated storage volumes for these future dams based on a simple linear model that has been derived for observed storage capacity and power generation capacity of planned dams in Asia. We acknowledge the fact that some run-of-the-river dams have little storage volume, whereas some storage dams are not designed to produce much power. In lack of better data, we used 251 dams planned in Asia with available data on live storage (million cubic meters) and installed hydropower capacity (MW) (ICEM, 2010), calculated the ratio between the two variables, and averaged the resulting ratios (mean = 3.19; sd = 8.5). Missing storage volume of future dams were then calculated as: $S_i = Pcap_i * 3.19$, where S_i is the storage volume and $Pcap_i$ is the reported power generation capacity of dam *i*.

S4.1.3 Uniform spatial units

The spatial reference in fragmentation analyses has commonly been the basin scale. However, as the size of the world's river basins span over several orders of magnitude, comparisons between small and large basins are difficult. In addition, if the dams are distributed unequally within the river basin, a smaller scale can produce substantially different spatial patterns of impact indicators for individual subbasins.

To address these issues, we used a consistent set of subbasin units, termed HydroBASINS (Lehner and Grill, 2013), in our framework. HydroBASINS is a delineation of global watersheds and was developed to provide nested subdivisions of large river basins to conduct disaggregated spatial analyses. HydroBASINS consists of 12 hierarchical levels of increasingly smaller subbasin units. For this study, we chose an intermediate level (level 4) where all of the world's largest basins are partitioned into sets of smaller tributaries.

S4.1.4 River Fragmentation Index (RFI)

The River Fragmentation Index (RFI; see also Grill et al., 2014) is a measure of structural connectivity (expressed as the percentage of full connectivity) per river basin or subbasin. It is based on the Dendritic Connectivity Index (DCI) by Cote et al. (2009) but has been modified to better take into account the size of rivers by introducing the concept of river volume. While the DCI considers rivers as lines of a particular length, thus representing habitat space in one dimension, our use of river volume adds a discharge-weighting component. This approach provides a freshwater analogy to the measurements of habitat area commonly applied in terrestrial landscape ecology.

The addition of river volume generally increases the relative impact of dams located on larger or mainstem rivers by shifting the location of largest impact (i.e. the “middle” of the network in terms of river volume) further downstream, towards locations of larger discharge and channel dimensions. The RFI (%) is defined as (Grill et al., 2014):

$$RFI = 100 - \left(\sum_{i=1}^{n} \frac{v_i^2}{V^2} \times 100 \right) \qquad \text{S4.1}$$

where n is the number of fragments (i.e. distinct network sections disconnected by dams); v_i is the total river volume of fragment i; and V is the total river volume of the entire river network in the basin. The RFI of an unfragmented river network is 0%, and each subsequent dam increases the RFI depending on the size distribution of the remaining fragments. A single dam in a previously undisturbed network leads to the maximum fragmentation at the ‘location of largest impact’, i.e. if it splits the network into two equally sized fragments (defined by river volume), and the RFI falls to 50%. If a subsequent dam is built close to a previous one, the added impact is smaller because of the pre-existing dam.

The methodology for calculating the RFI for both river basin and subbasin scale is the same in principal; i.e., the fragmentation effect of dams outside the boundaries of the subbasin is not considered.

S4.1.5 River Regulation Index (RRI)

The River Regulation Index (RRI; Grill et al., 2014) is an extension of the ‘Degree of Regulation’ Index (DOR) as calculated globally in Lehner et al. (2011) and provides a quantitative proxy on how heavily a basin is affected by alterations to the natural flow regime by dam operations. The DOR captures the potential impact of dams on downstream flows by calculating the proportion of a river’s annual discharge volume that can be withheld by a reservoir or a cluster of reservoirs upstream. Spatially explicit DOR values (%) are calculated for each river reach as:

$$DOR = \frac{\sum_{i=1}^{n} s_i}{D} \times 100 \qquad \text{S4.2}$$

where n is the number of all dams upstream of the reach, s_i is the storage capacity of dam i located upstream the river reach, and D is the total annual discharge volume at the river reach. A high DOR value indicates an increased probability that substantial discharge volumes can be stored throughout a given year and released at later times. For example, 10% is used as a threshold in Lehner et al. (2011) to mark the possibility of substantial changes in the natural flow regime to occur. In particular, multi-year reservoirs (DOR > 100%) have the ability to release water in accordance with an artificial, demand-driven regime, often aiming to increase dry-season flows or to eliminate flood peaks.

In order to quantify the overall impact on the basin in a single index value, we applied the RRI (%) (Grill et al., 2014) which is calculated by first weighting the DOR value of each individual reach with its corresponding river volume, and then averaging the results for the entire basin:

$$RRI = \sum_{i=1}^{n} DOR_i \frac{rv_i}{V} \qquad \text{S4.3}$$

where n is the number of reaches in the network; DOR_i is the DOR value of river reach i; rv_i is the river volume of reach i; and V is the total river volume of the entire river network in the basin. Note that RRI values can exceed 100% if the river network is affected by multi-year storage.

For calculations at the subbasin level, the initial DOR values are calculated for the entire river network (as this is the inherent definition of the DOR index), i.e. all upstream dams and their storage capacities are considered, even those of dams that are outside (upstream) of the subbasin. However, the RRI for a subbasin is then calculated by only using the DOR values and river volumes found within the subbasin.

S4.1.6 Dam Impact Matrix (DIM)

Although a dam typically has an effect on both fragmentation and flow regime, one might be greater than the other, depending on the properties and location of the dam. For example, a large storage dam in the upper river network may have heavy impacts on the downstream flow regime over long distances yet little effect on fragmentation, whereas a run-of-the-river type dam with a small reservoir capacity located on a downstream mainstem reach will primarily have implications on fragmentation and little effect on flow regulation. Both effects can be assessed simultaneously by plotting them on an impact matrix. We created this matrix by first classifying each index into four categorical groups based on quartile ranges of occurrence (*weak*: 0-25th; *moderate*: 25-50th; *heavy*: 50-75th; and *severe*: 75-100th percentiles), and then combining the categories from each index to create an integrated four by four matrix. This approach illustrates the large spectrum of possible combinations while identifying four primary groups of impacts at the four corners of the matrix (see Figure 4.2 and Figure 4.4).

S4.2 Additional limitations

S4.2.1 Hydrology

Our downscaled discharge values only represent long-term average conditions, hence special systems with complex seasonal flow regimes are not properly represented. The same applies to variations due to seasonality, or longer-term, natural changes of the river flow. In consequence, our resulting index values are currently only relating to average conditions.

S4.2.2 Other anthropogenic perturbations

Our flow regulation indices do not capture changes to river flows due to anthropogenic perturbations other than dams. However, many rivers have been altered structurally which could affect the river flow; e.g., channelization or levees can affect flow quantities and qualities of rivers downstream of the modification. The same applies to water transfers, which may occur at scales large enough to affect river flow substantially. As a result, some river basins appear as weakly affected due to low numbers of large dams, but may in fact be more impacted by widespread modifications of the river channel and the floodplain, or by water transfers.

S4.2.3 Passability

We assume that any dam, regardless of its design or operation, compromises connectivity, and we assume zero passability for each dam in our model. However, the permeability of barriers for migrating fish and other species depends on the type of dam. For example, a run-of-the-river dam may prevent migration of species in the upstream direction, but can still allow for downstream passage. Smaller dams (albeit not explicitly considered in this study), might be passable by jumping fish in both directions. If reasonable passability estimates were available for each dam, they could be incorporated into the RFI index. But even if passability measures were provided, these are often biased towards the upstream migration of salmonid fish, but neglect other fish species, downstream and lateral migration, and other factors (Kemp and O'Hanley, 2010).

S4.3 Assessment of small dams

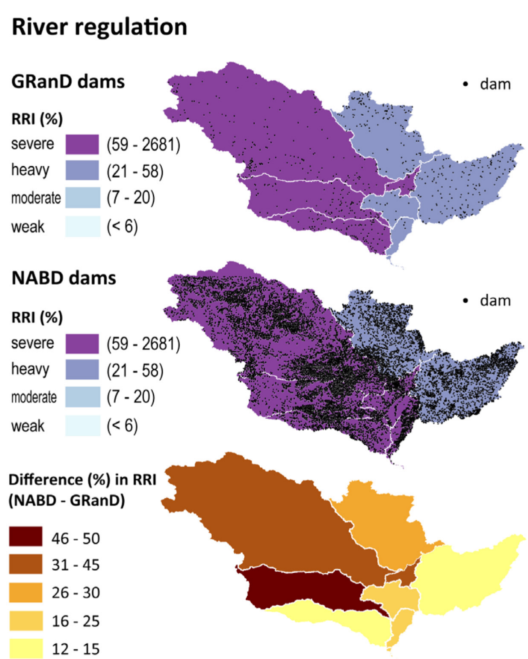


Figure S4.1: Comparing RFI and RRI in the Mississippi basin between two sets of dams: GRanD dams (n = 704) and NABD dams which also include small and medium sized dams in addition to large dams (n = 25857). The lower third of the graph displays absolute differences of RFI and RRI values (%) between the two sets.

Table S4.1: Assessing the effect of small dams on fragmentation and flow regulation indices in the Mississippi Basin (current conditions, 2010). Note that the similar results for impacted river volumes (km^3) suggest that only few (or small) additional river reaches are added to the impacted category (DOR ≥2%), yet the significant increase in RRI indicates that many reaches that are already above the 2% threshold substantially increase in their individual DOR values.

	Dams	RFI	RRI	Volume impacted (DOR >= 2%)		Length impacted (DOR >= 2%)	
	#	%	%	km^3	%(1)	km	%(2)
GRanD (Lehner et al., 2011)	704	44.7	65.0	17.5	95.1	59141	6.2
NABD (Ostroff et al., 2013)	25857	65.4	90.1	17.8	96.7	156084	16.4

(1) ∑ volume network 18.4 km^3

(2) ∑ length network 954464 km

S4.4 Supplementary tables and figures

Table S4.2: Summary table of global impacts by category to compare differences between affected length and volume (see also Figure 4.4).

	1930		1950		1970		1990		2010		2030	
	km	vol	km	vol	km	vol	km	vol	km	vol	km	vol
Fragmentation												
no impact	80.1%	74.7%	63.9%	34.2%	46.1%	11.6%	41.9%	7.9%	40.9%	7.5%	37.6%	6.1%
weak	15.3%	21.3%	25.5%	56.4%	27.3%	62.0%	20.1%	50.0%	20.1%	50.0%	6.8%	4.8%
moderate	2.4%	1.9%	3.4%	3.9%	11.2%	13.4%	9.8%	11.1%	8.6%	9.3%	17.3%	43.4%
heavy	0.9%	0.5%	3.5%	2.1%	5.0%	4.9%	10.1%	11.8%	9.5%	9.8%	11.0%	15.6%
severe	1.2%	1.5%	3.7%	3.4%	10.4%	8.1%	18.1%	19.1%	20.9%	23.5%	27.3%	30.1%
Regulation												
no impact	80.1%	74.7%	63.9%	34.2%	46.1%	11.6%	41.9%	7.9%	40.9%	7.5%	37.6%	6.1%
weak	17.3%	23.3%	28.0%	57.7%	25.5%	59.9%	21.9%	54.7%	18.6%	49.8%	17.8%	46.5%
moderate	1.7%	1.6%	5.9%	7.3%	8.7%	9.9%	7.9%	10.8%	11.1%	14.9%	10.1%	10.3%
heavy	0.7%	0.4%	1.5%	0.7%	13.7%	13.8%	17.6%	17.3%	18.1%	18.4%	18.4%	20.8%
severe	0.3%	0.0%	0.8%	0.1%	5.9%	4.9%	10.7%	9.3%	11.3%	9.4%	16.1%	16.4%

Table S4.3: River Fragmentation Index (RFI), River Regulation Index (RRI), and Dam Impact Matrix (DIM) for selected river basins (top 5 largest basins of each continent). Color scheme for RRI, RFI, and DIM corresponds to Figures 2, 6 and 7.

	RFI		RRI		DIM		Number of dams		Avg. discharge (m^3/s)	Basin size (km^2)
Africa	2010	2030	2010	2030	2010	2030	2010	2030		
Congo	0.2	25.7	0.5	1.9			10	38	44210	3705302
Zambezi	68.7	83.4	87.3	96.4			90	118	5525	1378121
Ogooué	0	20.7	0	0.4			0	4	4694	215217
Niger	60.8	65.3	23.3	54.4			92	134	4203	2098664
Nile	70	88.4	33.6	65			16	92	3702	2916802
Asia										
Ganges	47.6	57	6.5	21.6			134	926	31657	1574223
Yangtze	56.1	64.7	13.8	84.8			588	864	28323	1909199
Irrawaddy	0.7	43.7	1	16.6			10	44	13834	376349
Mekong	5.1	84.8	3.2	50.2			32	278	13096	774281
Amur	3.4	49.2	31.5	38.1			12	66	10126	1998203
Australia										
Murray	18	18	49	49			82	82	791	775219
Clutha	63.2	81.5	10.8	15.1			22	30	739	20630
Mitchell	0	0	0	0			0	0	577	63659
Waiau	51	51	8.5	8.5			4	4	468	8200
Waitaki	74.8	75.1	50.8	52.9			12	14	441	11928
Central America										
Usumacinta	23.2	60.4	13.3	16.2			12	22	3382	123639
San Juan	0.4	2.1	5.3	7.2			2	32	1250	41214
Papaloapan	8.4	8.4	16.7	16.7			2	2	950	39661
Balsas	60.4	60.4	17.9	17.9			28	28	779	111464
Motagua	0	0	0	0			0	0	662	16329
Europe										
Don	78.4	78.4	40.4	40.4			36	36	7799	1404107
Danube	67.5	90.3	7.4	19.4			360	970	6257	786749
Rhine	1.8	4.4	3.9	4			84	88	2390	163008
Dnieper	82.4	83.7	22.8	22.8			12	16	1616	509822
Rhône	38.4	38.4	8.2	8.3			74	76	1594	95894
North America										
Mississippi	44.7	44.7	65	65			704	704	19178	3179496
St Lawrence	63.9	63.9	17.8	17.9			348	356	11724	1053294
Columbia	55.7	56	51.4	52.2			250	254	6830	651407
Fraser	5.7	6.7	34.1	35.5			22	24	3467	230824
Nelson	79.6	80	96.6	98.2			178	184	2130	1004347
South America										
Amazon	0.1	24.1	0.2	4.8			16	386	202212	5912761
Orinoco	10	10.1	3.7	4.7			32	40	34537	938384
Parana	60.8	83	48.2	53.5			140	788	19063	2594295
Tocantins	27.1	82.7	21.9	32			8	178	11443	772470
Magdalena	3.4	23.3	4	8.5			24	74	7466	260739
Arctic										
Yenisei	33.9	33.9	76.6	76.9			14	16	18382	2489751
Lena	2.4	2.5	7.3	9.3			6	8	14210	2453574
Ob	11.1	11.1	24.5	24.5			6	6	12647	2467589
Mackenzie	9.6	21.4	48	55.8			24	32	6627	1795841
Yukon	0	0	0.4	0.4			2	2	2194	832689
Neva	62.2	62.2	27.5	27.5			16	16	2387	285331

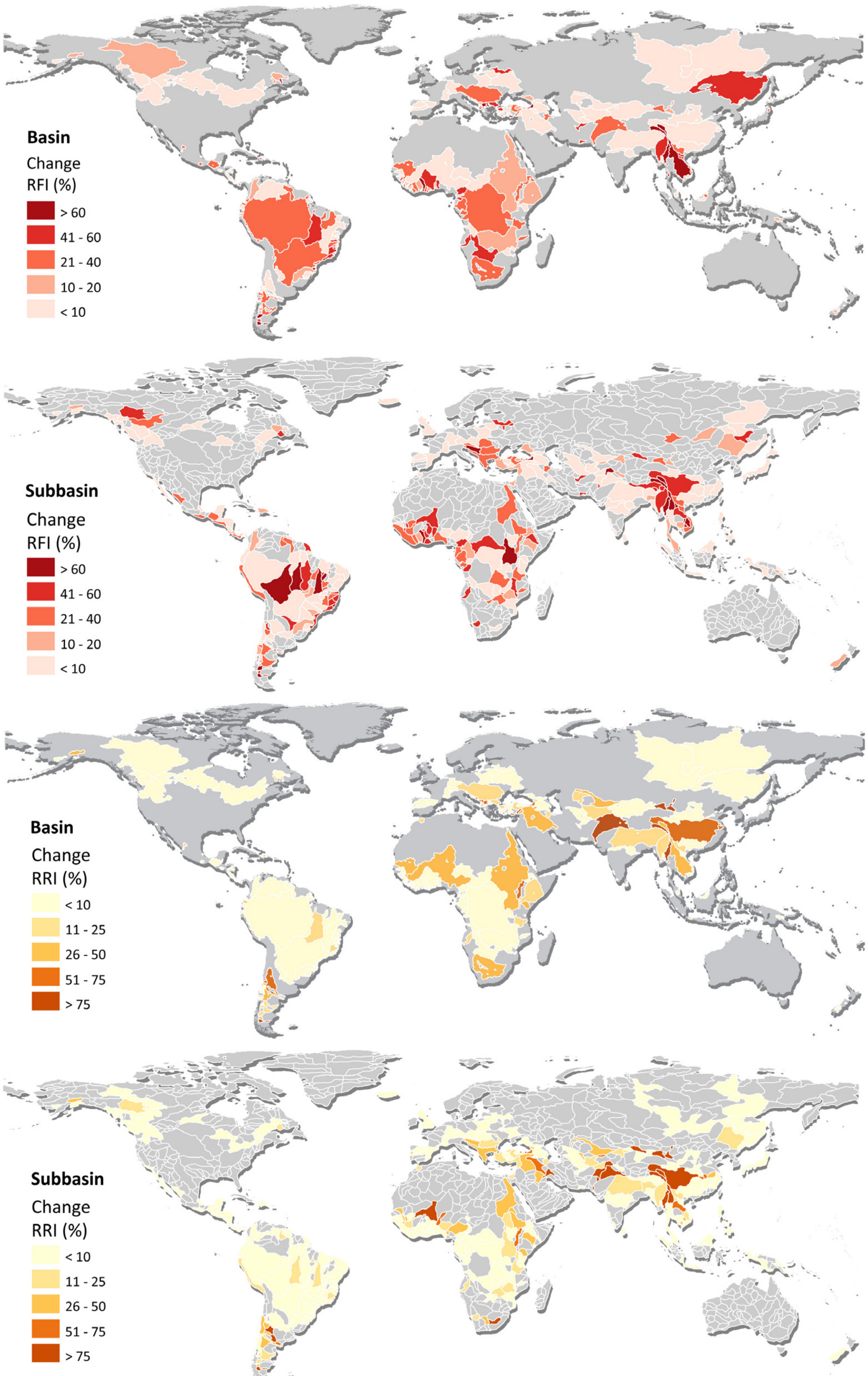


Figure S4.2: Changes in flow regulation and river fragmentation (absolute difference of RRI and RFI in %) between current conditions (2010) and the future scenario of 2030 in which all dams currently planned or under construction are built.

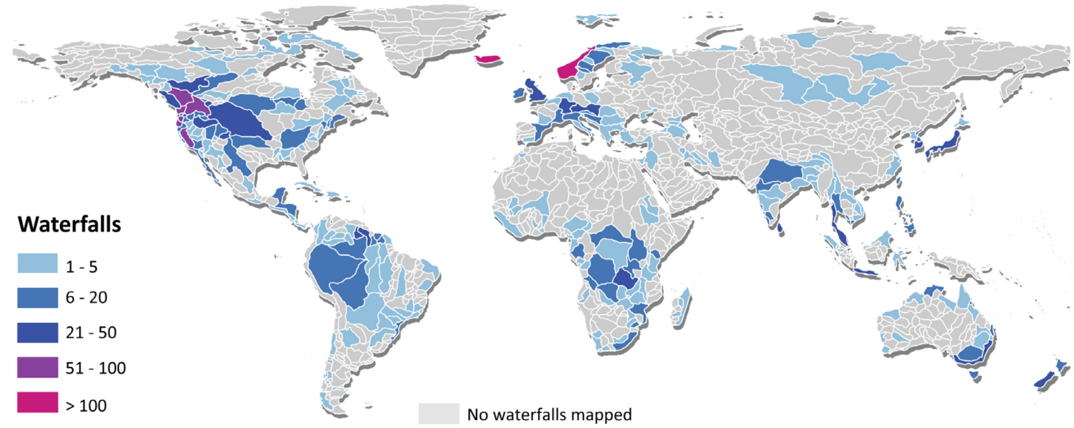


Figure S4.3: Number of waterfalls per subbasin (HydroFALLS; Lehner *et al.,* unpublished data).

Connecting statement (Ch. 5)

Chapters 3 and 4 demonstrated how HydroROUT can be used at the scale of a large river basin and at the global scale to assess human alterations of river connectivity and flow regulation. In chapter 5, I extend HydroROUT in a new direction to serve as an integrated chemical fate model, in order to assess anthropogenic impacts from river pollution from diffuse and point sources. This allows me to focus on a more advanced implementation of connectivity, especially looking at cumulative effects and in-river processes.

Existing approaches to estimate environmental risk from emerging contaminants in Canadian river systems currently do not consider spatially distributed emission sources or local hydrology. Hence, this chapter serves as the foundation for creating a Canada-wide contaminant fate model that could support environmental risk assessment as part of regulatory frameworks.

Furthermore, this chapter includes a regional validation of HydroROUT's representation of flow quantities to ensure the suitability of the applied average and low flow hydrology for the intended contaminant risk assessment. This is a vital step of overall model validation because environmental risk is to a large extent sensitive to variations of discharge.

5 Geospatial contaminant risk assessment in the Saint Lawrence River basin

Günther Grill[1], Usman Khan[2], Bernhard Lehner[1], Jim Nicell[2], and Joseph Ariwi[2]

[1] Department of Geography, McGill University, 805 Sherbrooke Street West, Montreal, Quebec H3A 0B9, Canada

[2] Department of Civil Engineering & Applied Mechanics, McGill University, 817 Sherbrooke Street West, H3A 0C3, Montreal, Canada

In preparation for *Environmental Science & Technology*

Abstract

Chemicals released into freshwater systems threaten ecological functioning and may put aquatic and human health at risk. As part of a pilot study in the Saint Lawrence River Basin, we developed a new contaminant fate model that follows simple, well-established methodologies and is unique in its cross-border, seamless hydrological and geospatial framework, including lake routing, a critical component in northern environments. We validated the model using the pharmaceutical Carbamazepine and predicted environmental risk for 15 chemicals in the Saint-Lawrence River Basin, Canada. The results showed that the majority of tested chemicals are unlikely to pose major environmental risk; however, two pharmaceuticals showed elevated risk in up to 17 percent of rivers affected by municipal effluents. This pilot study provided the foundation to further develop the model towards a Pan-Canadian scale and to wider its scope. As an integrated model, our CFM can facilitate the prioritization of actions, such as reducing contamination sources, protecting drinking water resources or upgrading treatment plants to achieve targeted pollutant removal. In regulatory frameworks, it can help screen new chemicals entering the North American market regarding their potential impact on human and environmental health, for example as part of the Canadian Chemical Management Plan.

5.1 Introduction

The Millennium Ecosystem Assessment identified anthropogenic contaminants released into the environment as a major threat to freshwater ecosystems (Millennium Ecosystem Assessment, 2005). Increasingly, a multitude of household toxicants enter the environment through sewage treatment systems, which include substances and residues from industrial products (additives, lubricants, flame retardants), consumer products (detergents, pharmaceuticals, personal care products), synthetic and natural hormones (e.g., estrogens), and emerging contaminants, such as nanomaterials (Schwarzenbach et al., 2006; Kümmerer, 2011).

This situation creates an urgent need to limit aquatic life and human exposure to these chemicals. In most countries, regulatory frameworks require new chemicals entering the market to undergo an environmental risk assessment. In general, risk is assessed by exposure and effect modelling, whereby exposure modelling predicts environmental concentrations (Predicted Environmental Concentration; PEC), and effect modelling defines levels through which these environmental concentrations have no effect on living organisms (Predicted No-Effect Concentration; PNEC). PEC values higher than PNEC indicate unacceptable risk.

Many traditional methodologies assume that substances are emitted in a standard environment with pre-defined environmental characteristics and use a constant per-capita discharge rate and dilution factors based on country-wide averages. For example, this approach is currently applied in the European Union System for the Evaluation of Substances (EUSES; Vermeire et al., 1997). This simplified method, however, does not account for spatial variation in consumption, population distribution, physio-geographic characteristics (such as hydrology and seasonality), wastewater treatment capabilities, and environmental removal mechanisms. Any of these factors can result in environmental concentrations ranging over several orders of magnitude, creating “contaminant hot-spots” in the environment. Researchers proposed in-situ substance measurements to refine simple non-spatial risk assessments; but measurements are generally complicated by high cost, time-consuming implementation, and lack of robust analytical methods.

Contaminant fate models (CFMs), which combine Geographic Information Systems and water quality models, have emerged as an additional research tool to estimate environmental concentrations of chemicals and their environmental exposure (Johnson et al., 2008; Cowan-Ellsberry et al., 2009). CFMs mitigate some shortcomings of current exposure and risk assessment methods (i.e., measurements, non-spatial models), which can be summarized as follows:

1. CFMs have a strong spatial component, i.e. they take into account population size and distribution, river network complexity and its accumulating effects, and local river flow dynamics that can differ substantially from the emission point and across the stream network.
2. CFMs are particularly useful for measuring analytically challenging chemicals (e.g. nanoparticles), whose environmental quality standard is below current analytical capabilities (e.g. EE2), and those known to act as mixtures (e.g. estrogens).
3. For large spatial assessments, logistical and financial challenges often prohibit comprehensive sampling across watersheds, making modelling the only viable option. Once set up, the CFM can be repeatedly used to address a wide range of contaminants without much additional model adjustment.
4. Besides descriptive modelling, where the model outcome simulates concentrations under current flow conditions, CFMs can be used for predictive modelling, for example to assess a chemical under different climate or emission scenarios. Furthermore, CFMs can aid in normative modelling ("optimizing"), to answer policy questions, e.g. to determine where interventions should be targeted in river systems or to highlight wastewater treatment plants where upgrading would considerably reduce in-river concentrations.

Generally two different approaches exist - multimedia or single-media models. Multimedia models (e.g., EUSES; Vermeire et al., 1997) simulate chemical sources and fate across different environmental compartments (air, water, soil), but are generally difficult to parameterize, spatially coarse, and do not allow for site-specific predictions. This limits their use in environmental risk assessment. Single-media approaches have been used to predict single substance concentrations in river networks at high spatial resolution, such as GREAT-ER (Feijtel et al., 1998), ISTREEM (Wang et al., 2000), LF2000-WQX (Williams et al., 2012), P*h*ATE (Anderson et al., 2004), the MAPPE model (Pistocchi et al., 2012) and the GWAVA model (Johnson et al., 2013). Such models are particularly well-designed for household chemicals and have been applied to chemical exposure assessments in various studies (Atkinson et al., 2009; Cunningham et al., 2009; Hannah et al., 2009; Ort et al., 2009; Cunningham et al., 2012; Hosseini et al., 2012; Johnson et al., 2013). These models share common assumptions and similar key mechanisms: reasonable per capita emission estimates of contaminant mass entering individual wastewater treatment plants (WWTP) can be gained using the average per-capita consumption of a compound of interest and adjusting for human metabolism. The contaminant mass released by individual WWTPs into specific river reaches can be estimated using the average per-capita emissions, knowledge of the local population served by a WWTP, and adjusting for removal during treatment, where relevant (Keller et al., 2006). As for chemical routing in the hydrological system, advection is assumed to be the dominant dilution mechanism, which can be modeled effectively using stream length, velocity, discharge, and a decay function (Pistocchi et al., 2010). Predicted environmental concentrations (PECs) are subsequently based on accumulated load and discharge for each river reach.

The prediction error from these models have consistently fallen within one order of magnitude when compared to actual measurements and, in most cases, range within 2-4 times of the measured environmental concentration (Johnson et al., 2008). As such, and with these limitations in mind, they are deemed suitable for eco-toxicological exposure and risk assessments of low to medium complexity. However, CFM set-up involves a number of complexities, particularly the need for quality cross-boundary hydrographic base maps, including realistic hydrological predictions throughout the river network, consumption variability information, and the chemical’s fate in the system, such as wastewater treatment efficiency and environmental decay. These requirements lead to limited model deployment to regions where such high quality information is not readily available. In Canada, no nationwide, large-scale tool exists that is geospatially explicit (i.e., based on a river network to account for upstream contributions from other wastewater systems) to predict aquatic concentrations of substances released from wastewater systems at sufficient spatial resolution to support risk assessment activities for regulatory frameworks, such as the Canadian Chemical Management Plan.

To address this need, in this study, we present a new contaminant fate model that is based on the integration of three components: the HydroSHEDS database (Lehner et al., 2008) which provides the hydrographic backbone of the model; results from the global hydrological model WaterGAP (Alcamo et al., 2003; Döll et al., 2003) which provide the hydrological input data; and the HydroROUT river routing model (Lehner and Grill, 2013) which includes the routing component. We designed our model to predict environmental concentrations for a wide range of chemicals within a complex aquatic network. After initial setup and validation, we applied the model to assess 15 contemporary chemicals in regard to their environmental concentrations and eco-toxicological risk for the Saint Lawrence River Basin.

5.2 Study design and methodology

The Canadian provinces of Quebec and Ontario were to core study area, with a focus on the Saint Lawrence River (Figure 5.1). We also included all river reaches ('contributing areas') that enter the two provinces, as defined by the watershed delineation from the HydroSHEDS database in order to generate a hydrologically complete river network. Cross-boundary watersheds include parts of the United States from Minnesota, Wisconsin, Michigan, Ohio, New York, and Vermont, but most of these US watersheds drain into the Great Lakes before their flow arrives in Canada. We assumed complete dilution of contaminants in large water bodies, hence the US tributaries to the Great Lakes were not considered to pose an immediate risk for Ontario and Quebec. Therefore, in terms of chemical mass balance calculations, we excluded effluents from wastewater treatment plants located in contributing areas south of the Great Lakes. There are, however, parts of New York and Vermont that drain directly into the lower St. Lawrence River that are likely more relevant contaminant sources from the US, so we included these regions in the pilot model version. In total, 1198 WWTPs were part of our study (Figure 5.1).

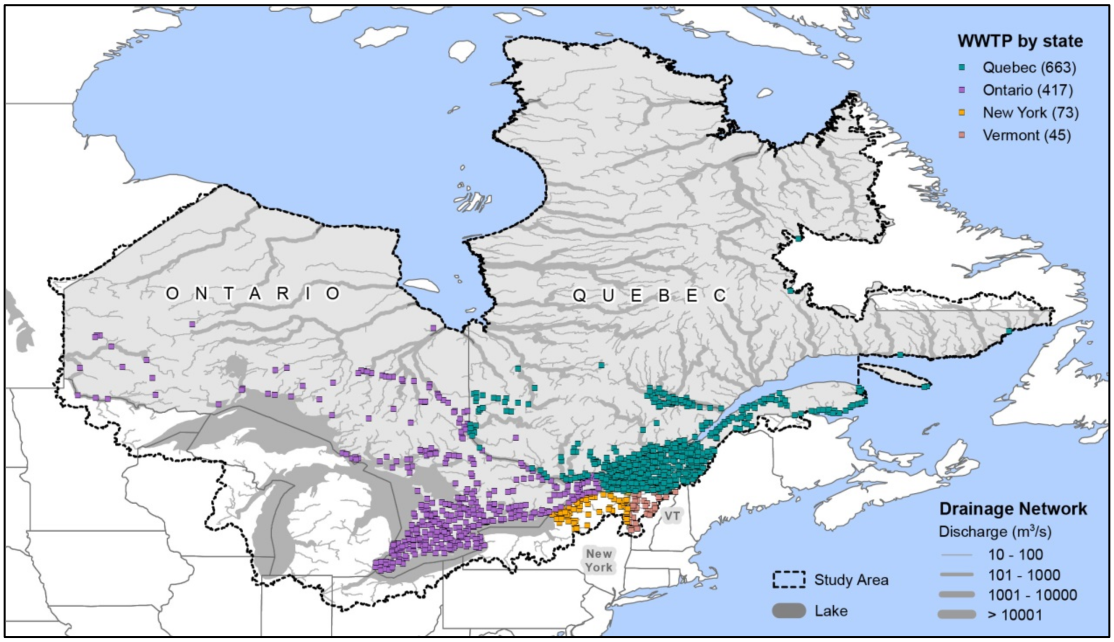


Figure 5.1: The study area included areas of Quebec, Ontario as well as some parts of the United States (hydrologically connected areas draining into the Saint Lawrence River below Kingston). Color markers represent distribution of wastewater treatment plants by state.

Figure 5.2 provides a conceptual model of the study. A geospatial exposure and fate model embedded in a Geographic Information System calculated PEC, based on a cascade of input and removal mechanisms. To determine the PNEC, we used an effect model (Khan, 2014). This enabled us to assess PEC against PNEC to locate river network areas with elevated environmental risk.

We used the pharmaceutical carbamazepine to validate model performance by comparing it with measurements from a literature review. Based on eco-toxicological benchmarks provided by Khan (2014), we evaluated a total of 15 additional substances.

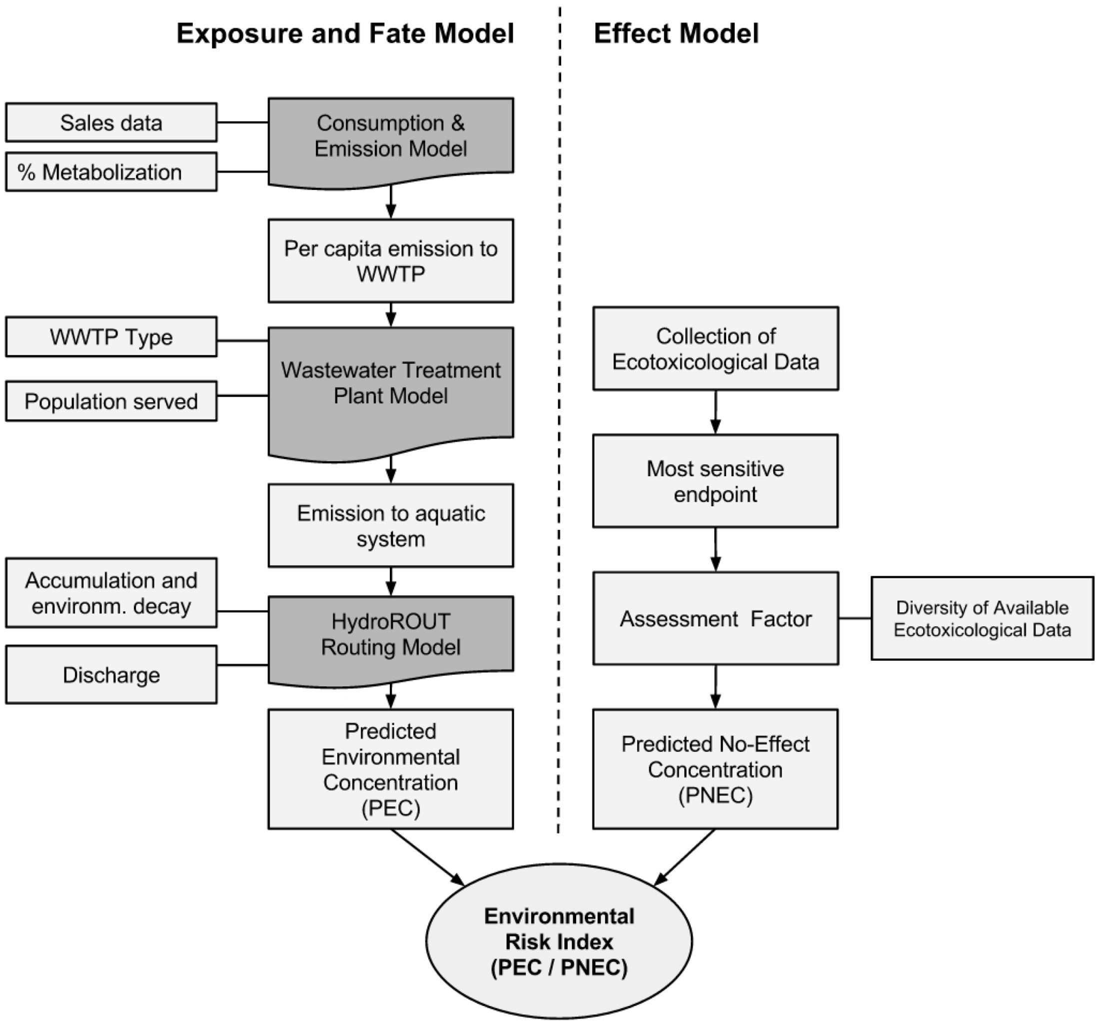


Figure 5.2: Conceptual model of the study. The exposure and fate model is described in the methods section of this paper. The effect model follows established procedures as described in Khan (2014).

5.2.1 Discharge downscaling and validation

To simulate environmental contaminant concentrations, an adequate characterization of discharge of Canadian rivers that receive wastewater effluents was necessary. This ensures that the dilutive capacity of the receiving waters is appropriately parameterized to generate predicted environmental concentrations suitable for contaminant risk assessments.

We used the HydroSHEDS database (Lehner et al., 2008), a publicly available global suite of data layers representing river network topology and watershed boundaries, to provide the baseline hydrographic data. HydroSHEDS defines flow directions at 500m pixel resolution that are used for water and substance transport simulation in a routing model called HydroROUT, currently under development at McGill University (Lehner and Grill, 2013).

The current version of HydroROUT derives discharge by accumulating land surface runoff along the river network; yet the underlying runoff generation simulation (i.e., the vertical water balance) was not performed within the model itself. Instead, we employed decoupled, external runoff estimates provided by the global integrated water balance model WaterGAP (Alcamo et al., 2003; Döll et al., 2003; model version 2.1 as of 2012). WaterGAP provides runoff estimates of long-term monthly averages from 1961-1990 (i.e., the 'climate-normal' period as defined by the World Meteorological Organization) at 0.5 degree grid resolution. We used these runoff estimates with spatial downscaling methods to disaggregate the large grid cells into 500m pixels and then accumulate the runoff along the HydroSHEDS river network.

To evaluate the HydroROUT model accuracy and uncertainty, we compared our downscaled discharge estimates with the reported values of HYDAT gauging stations (Environment Canada, 2012) in three different ways:

1) To evaluate the downscaled discharge estimates applied in HydroROUT, we compared long-term average flows (AVG-YEAR) across stations using a linear regression analysis as recorded at the 57 "most reliable" HYDAT stream gauges (for selection criteria see section S4.5).

2) The risk from chemical substances in surface waters is usually assessed under low flow conditions for which the daily Q90 flow (flow exceeded 90 percent of the time) is a frequently used indicator. Typically, Q90 is calculated from daily discharge measurements, but as the WaterGAP runoff estimates are given as monthly time series, we approximated (daily) Q90 with a substitute, namely Q90-MONTH (for the definition see S4.5). To explore the validity of this approach, first, we assessed the relationship between Q90 and Q90-MONTH for observed HYDAT flows by performing a linear regression analysis. Then, we applied a second linear regression to test the Q90-MONTH correlation between observed HYDAT and simulated HydroROUT values.

3) Finally, we tested the ability of the HydroROUT model to simulate the annual flow regime using the Nash-Sutcliffe Efficiency (NSE) as the quality indicator (Nash and Sutcliffe, 1970).

5.2.2 Exposure model

Pharmaceutical emission model

The pharmaceutical emission model calculates the emission from annual per capita consumption and the number of inhabitants served by the treatment plant. We calculated domestic emission into WWTPs by:

$$L(S)_{WWTP} = [CN(S)_{TP} \times (1 - m_s)] \times P_{TP} \tag{5.1}$$

where $L(S)_{WWTP}$ is the total load arriving at the WWTP, $CN(S)_{TP}$ is the per capita consumption of the chemical s of the connected population, P_{TP} is the total population connected to the WWTP, and m_s is the fraction metabolized in the human body of compound s. However, at the time of this study, the data for consumption on the individual treatment plant level $CN(S)_{TP}$ was not available. Therefore, we created Canada-wide per capita consumption averages for all 15 chemicals generated from total Canadian market sales data divided by the population totals for Canada in 2012. We give further details on the boundary conditions for each chemical in Table 5.1.

Wastewater treatment plant model

Using federal and provincial sources, we assembled the locations and characteristics of 1198 WWTPs (Figure 5.1). We did not target industrial, mining or agricultural sources in this study. Location, name, average daily flow, treatment technology and population served were given as attributes. We manually located ("snapped") the wastewater treatment plants to the appropriate HydroSHEDS river reach using satellite imagery and other mapping sources. Utilizing expert judgment, we classified the treatment plants into the following categories: no treatment, primary treatment, lagoon treatment, and secondary treatment.

Depending on the type of treatment plant and compound simulated, the incoming load is reduced:

$$L(S)_{EFF} = L(S)_{WWTP} \times R(T)_{TP} \qquad 5.2$$

where $L(S)_{EFF}$ is the substance mass after removal, $L(S)_{WWTP}$ is the incoming chemical load from the household emission model, and $R(T)_{TP}$ is the removal efficiency in the wastewater treatment plant of the compound depending on the type of removal T.

5.2.3 River and lake routing model

In this project, we used the HydroSHEDS database (Lehner et al., 2008), a publicly-available global suite of hydrographic data layers. HydroSHEDS defines river flow directions at 500m pixel resolution that create a network of river reaches for water and substances transport simulation in a routing model called HydroROUT (Lehner and Grill, 2013). At 500m resolution, more than 400,000 river reaches with an average length of 2.8 km exist within the study area, and a total of 11,426 river reaches, or 30,506 km are affected by upstream effluent from wastewater treatment plants.

We used HydroROUT's processing engine for mass balance calculations in the river and lake network. For river reaches, we followed a 'plug-flow' approach (Pistocchi et al., 2010), i.e., a 'plug' of substance mass (the amount of contaminant released from the treatment plant) is accumulated downstream as the sum of the input from the current and all upstream reaches flowing into the current reach. The river network was processed iteratively in the hydrological order from source to sink. The outflow mass balance for each river reach was calculated as:

$$PEC(S)_{REACH} = \frac{\left(\sum_i L(S)_{EFF} + \sum_j L(S)_{REACH}\right) \times d}{Q_{REACH}} \qquad 5.3$$

where $PEC(S)_{REACH}$ represents the environmental concentration at the end of the river reach for compound s, as the mass influx sum $L(S)_{EFF}$ from all wastewater treatment plants i located anywhere on the river reach, and the total mass $L(S)_{REACH}$ from all upstream reaches j. Chemical substance degradation in the river body, if applicable, is expected to decrease at a rate proportional to its value, and is represented as the environmental decay factor d. We calculated it based on first-order decay $d = e^{-kt}$ where t is the time a plug of water needs to travel through the river reach, and k is a positive number called the first order rate constant, which determines the environmental decay speed.

Travel time t is derived by dividing river reach length by the average velocity within the river reach. This corresponds to the average retention time in each individual river reach, i.e. the time a plug of fluid needs to travel from the beginning to the end of the river segment. We approximated velocity following an empirically derived formula (Allen et al., 1994):

$$v = 1.07 \times Q^{0.1035} \tag{5.4}$$

where v is the velocity in m/s^{-1} within the river reach and Q is the discharge in m^3/s. Note that for simplicity we used average discharge for Q instead of the required 'bankfull discharge'. The differences resulting from this limitation should not result in substantially different velocities, and is deemed acceptable given the general limitations of the approach by Allen and considering the goal of the study.

We integrated a comprehensive lake dataset, the Surface Water Body Database (SWBD; NASA/NGA, 2003) into HydroROUT to represent lakes larger than 1 km^2 in surface area. SWBD includes vectorized lake polygons that were digitized as part of the Shuttle Radar Topographic Mission (SRTM) at 30m resolution. We also estimated and validated lake volumes using GIS methods, based on correlations with lake surface area and surrounding topography (Pistocchi and Pennington, 2006; Hollister and Milstead, 2010).

As a preliminary measure in our pilot study, we modeled each lake as one single completely stirred tank reactor (Mihelcic et al., 2010):

$$PEC(S)_{LAKE,OUT} = PEC(S)_{LAKE,IN} \times \frac{Q_{LAKE}}{Q_{LAKE} + kV_{LAKE}} \tag{5.5}$$

where $PEC(S)_{LAKE,OUT}$ is the concentration of substance s at the lake outlet, $PEC(S)_{LAKE,IN}$ represents the inflowing concentration from all river sources,

Q_{LAKE} is the discharge in m3/s entering the lake, k is the lake's decay constant, and V_{LAKE} is lake volume, estimated by a geospatial volume estimation model (sensu Hollister and Milstead, 2010). We used the same constant k for both lake and river decay. Future models may improve on modelling lakes by incorporating a "tank-in-series" approach (Shanahan and Harleman, 1984; Anderson et al., 2004).

5.2.4 Dilution factors and percent wastewater in river

We calculated dilution factors DF for each WWTP as the ratio between the downscaled river flow at the WWTP location and the reported discharge leaving each WWTP:

$$DF_{WWTP} = \frac{V_{River} + V_{WWTP}}{V_{WWTP}} \tag{5.6}$$

where V_{River} is the flow volume in the river reach where the treatment plant is located, and V_{WWTP} is the sewage treatment plant effluent volume.

Wastewater flow, expressed as a percentage of river flow, is an indicator that provides general information on a river's status with regard to cumulative WWTP effluents and can help identify critical locations with a high effluent proportion in river water. Based on reported river flow and accumulated effluent volume from all upstream WWTPs, we estimated the wastewater percentage in the watercourse as PW:

$$PW = 100 \times \frac{\sum_i V_{WWTP}}{\sum_i V_{WWTP} + V_{River,j}} \tag{5.7}$$

where $V_{River,j}$ is the river reach' i flow volume, and $\sum_i V_{WWTP}$ is the accumulated effluent volume from all wastewater treatment plants upstream of the river reach. We accumulated the effluent flow from WWTPs along the stream network and applied the above formula for each river reach.

5.2.5 Model uncertainty, sensitivity and validation

Model uncertainty and sensitivity

Contaminant fate models often evaluate multiple substances with differing environmental behaviors. Due to the wide range of parameter input, at times ranging over various orders of magnitude, a calibration for each individual contaminant is complicated, time consuming, and limited by sparse measurement availability. Even if such calibration is achieved, an improvement over uncalibrated models is not always guaranteed (Webster and Mackay, 2003). For these reasons, CFMs are typically not calibrated but, rather, the PECs are based solely on the input parameters (i.e., substance use, metabolism, removal, in-stream decay).

However, we acknowledge that model uncertainty results from disregarding important natural variability of our system. For example, inter-annual and seasonal flow changes increase the range of concentrations in our river system and concentrations may be exceptionally high during extreme low flow events. Another example is added uncertainty from consumption variability, e.g., cross-country or between-state differences in consumption patterns, or temporal changes in consumption (e.g. flu medication). As a third example, the removal efficiency of a treatment plant may be variable, e.g., can be substantially lower for combined sewer overflow systems, which allow some of the wastewater to be discharged untreated in case of high storm water volumes.

Furthermore, we were unable to quantify uncertainty due to inherent modelling and measurement errors. Such uncertainty can result from modelling errors in the Global Hydrological Model, and from the downscaling to higher resolution river networks. The same applies to flow velocities, which affect in-stream removal rates, because these are correlated to discharge in our model.

In most cases, uncertainty is difficult to estimate, especially considering the fact that multiple, relatively unknown pharmaceuticals are modeled. Nevertheless, variability can be accounted for, to some degree, by performing Monte-Carlo simulations (Callahan, 1996) in which random values are selected for key variables from pre-defined probability density functions for a large number of model runs. Generally, each variable uses a specific probability density function, and the variation boundaries for each pharmaceutical are determined individually based on the mean and deviation (e.g. +/- 0.5 std). However, due to the lack of information on the variability of these variables and each pharmaceutical, we used a simpler approach, by which we fixed minimum and maximum variability using a single relative threshold for all variables and pharmaceuticals (we used a 50% deviation from the mean), and drew random values from a triangular probability distribution within the defined variability boundaries. For example, if the discharge at a specific river reach was 5 m^3/s, the Monte-Carlo module randomly selected a value between 2.5 m^3/s and 7.5 m^3/s from a triangular probability distribution. While this does not allow for confidence interval calculations, it does provide a sense of the model boundaries. This preliminary sensitivity analysis should be replaced by a more standard approach once better information about parameter variability becomes available.

Validation of mass balance model

Experimental validation of the simulated PECs with field measurements would be unreasonable for a contaminant fate model that assesses multiple chemicals, mainly due to the logistics and cost of sampling multiple different chemicals. Furthermore, in the case of many emerging contaminants, laboratory methods are not developed enough to detect these chemicals in the environment. In previous studies, contaminant fate models that share similar characteristics have performed reasonably well without calibration for a wide range of substances.

In this study, we use common approaches for contaminant fate model validation (Feijtel et al., 1998; Anderson et al., 2004; Aronson et al., 2012), i.e. semi-quantitative methods, specifically cumulative frequency plot comparisons, to evaluate mass balance module performance for one chemical: Carbamazepine (CBZ). CBZ, a commonly prescribed drug in Canada, shows low removal in wastewater treatment plants and is characterized by surface water persistence (half-life ~70 days), making it a good candidate for the validation of our mass balance model. CBZ is commonly detected in Canadian surface waters. The initial conditions that drive the mass balance model for CBZ are shown in Table 5.1.

We used cumulative probability distributions to evaluate measured concentrations in rivers (e.g. Cunningham et al., 2009). We calculated cumulative distributions for numerous CBZ measurements as found in the literature. In total, we analyzed 373 reported samples from 19 different studies, including: Metcalfe et al. (2003); Miao and Metcalfe (2003); Brun et al. (2006); Hua et al. (2006); Lissemore et al. (2006); Lajeunesse and Gagnon (2007); MacLeod et al. (2007); Yargeau et al. (2007); Viglino et al. (2008); Garcia-Ac et al. (2009); Viglino et al. (2009); Li et al. (2010); Rahman et al. (2010); Tabe et al. (2010); Waiser et al. (2011). We plotted the concentrations as cumulative frequency curves that included non-detects and followed the same procedure for our model simulations to allow direct comparison.

Finally, we used point-by-point validation (e.g., Feijtel et al., 1998) to compare the simulated concentrations from our model to the observed concentrations measured at specific point locations in the river network. Ideally, the on-the-ground observations should be collected following a specific and structured monitoring campaign. In this study, however, we used only observed concentrations from two CBZ studies, one peer-reviewed study by Lajeunesse and Gagnon (2007) and one Master's Thesis from the University of Waterloo (Kormos, 2007).

5.2.6 Environmental concentrations and risk assessment

In total, we selected 15 pharmaceuticals for environmental risk assessment. The selected set comprised those pharmaceuticals for which sufficient eco-toxicological data was available in literature. With the baring exception of antibiotics, we considered a pharmaceutical to have sufficient ecotoxicological data if a chronic No Observed Effect Concentration (NOEC) for fish was available for its exposure. For antibiotics, the availability of an endpoint for cyanobacteria was required before it could be deemed to have sufficient data. The reason for this data availability requirement was that fish are known to be the most sensitive taxon to non-antibiotic pharmaceuticals (Fent et al., 2006), while cyanobacteria are the most sensitive taxon to antobiotics (Ebert et al., 2011).

Subsequently, we predicted river concentrations for each chemical using the parameterization given in Table 5.1 for both low flow and average flow conditions. In addition, we calculated risk indices for river reaches by calculating the ratio between predicted environmental concentrations (PEC) and the predicted no effect concentration (PNEC). Except for Ethinylestradiol (EE2), PNECs were based on relevant toxicity benchmarks developed by Khan (2014). For EE2, we adopted the environmental quality standard proposed by the European Union of 0.035 ng/L (SCHER, 2011). In general, PEC/PNEC > 1 is indicative of unacceptable risk. Finally, we plotted the risk indices as density curves for each chemical to further explore the potential risk of the chemical on environmental health.

Table 5.1: Parameter values used for the calculation of the source and fate of pharmaceuticals.

	Tot. con.	Per cap. con.	Metabolization[1]	Removal efficiency (R_1)	Removal efficiency (R_{1+2})	Instream degradation constant	Predicted No-Effect Concentration
	CN_T	CN_{TP}	m_s	$R1_{TP}$	$R2_{TP}$	k	$PNEC$
Pharmaceutical	kg/ yr	µg/ cap·d	Fraction	Fraction	Fraction	day-1	ng/L
5-Fluorouracil	537	45	0.221	0.07	0.41	0.3465	200
Anastrozole	7.4	1	0.24	0.04	0.06	0	1000
Atenolol	11436	960	0.88	0.11	0.40	0.024	148000
Azithromycin	3061	257	1	0	0	0	9.4
Bicalutamide	232	19	0.48	0.04	0.05	0	1000
Carbamazepine	22443	1885	0.17	0	0.09	0.0088	500
Diazepam	285	24	0.11	0	0	0	100
Diclofenac	9057	761	0.24	0.01	0.42	0.28875	100
Ethinylestradiol (EE2)	9.7-13.1	0.81-1.10	0.43	0.21	0.84	0.07	0.035
Fulvestrant	0.15	0.013	0.19	0.65	0.72	0	0.57
Hydrochlorothiazide	13778	1157	1	0	0	0	1000000
Quetiapine	10182	855	0.029	0.09	0.12	0.0385	10000
Sulfamethoxazole	16804	1411	0.30	0.003	0.50	0.0364	590000
Tamoxifen	264	22	0.09	0.26	0.26	0.3465	22
Trimethoprim	3651	307	0.79	0.10	0.25	0	240000

[1] $U_{ex} + F_{ex}$
[2] 7-22 if injected, and 100 if topical

5.3 Results

5.3.1 Discharge validation

Long-term average discharge

Figure 5.3 shows the results of long-term annual discharge (AVG-YEAR) compared to and downscaled discharge used by the HydroROUT model at 57 stations. The linear regression model suggests a high correlation and indicates that the downscaled discharge accounts for 98.2% of observed data variation (coefficient: 1.05; residual standard error: 0.244). Additional validation sets and results incorporating a larger number of HYDAT gauging stations, as well as validation of monthly flows are available in S4.5.

Low flow discharge validation

We proposed to use the Q90-MONTH (i.e. the lowest long-term average monthly flow) as a substitute for the commonly used daily Q90 low flow index.

Figure 5.3b shows that a very good correspondence exists between the two indicators, and Q90-MONTH accounts for 96.7% of (daily) Q90 value variations (coefficient: 1.123; residual standard error: 0.4461; statistically significant results at the 95% confidence level). However, Q90-MONTH tends to systematically overestimate Q90, in particular for very small streams. This means, with respect to fate modelling, that low flow assessments based on Q90-MONTH are likely to underestimate substance concentrations and contamination risk as compared to analyses using daily Q90.

Figure 5.3c shows the comparison results between the observed and modeled low flow index Q90-MONTH for the "most reliable" 57 HYDAT stations. The linear regression model confirms a high correlation and indicates that the HydroROUT model accounts for 90.2% of the variation in the observed data (coefficient: 1.043; residual standard error: 0.6743; statistically significant results at the 95% confidence level). As expected, the associated error for low flow conditions is higher than for long-term average discharges. Differences between modeled and reported low flows can vary by up to one order of magnitude, especially for smaller rivers.

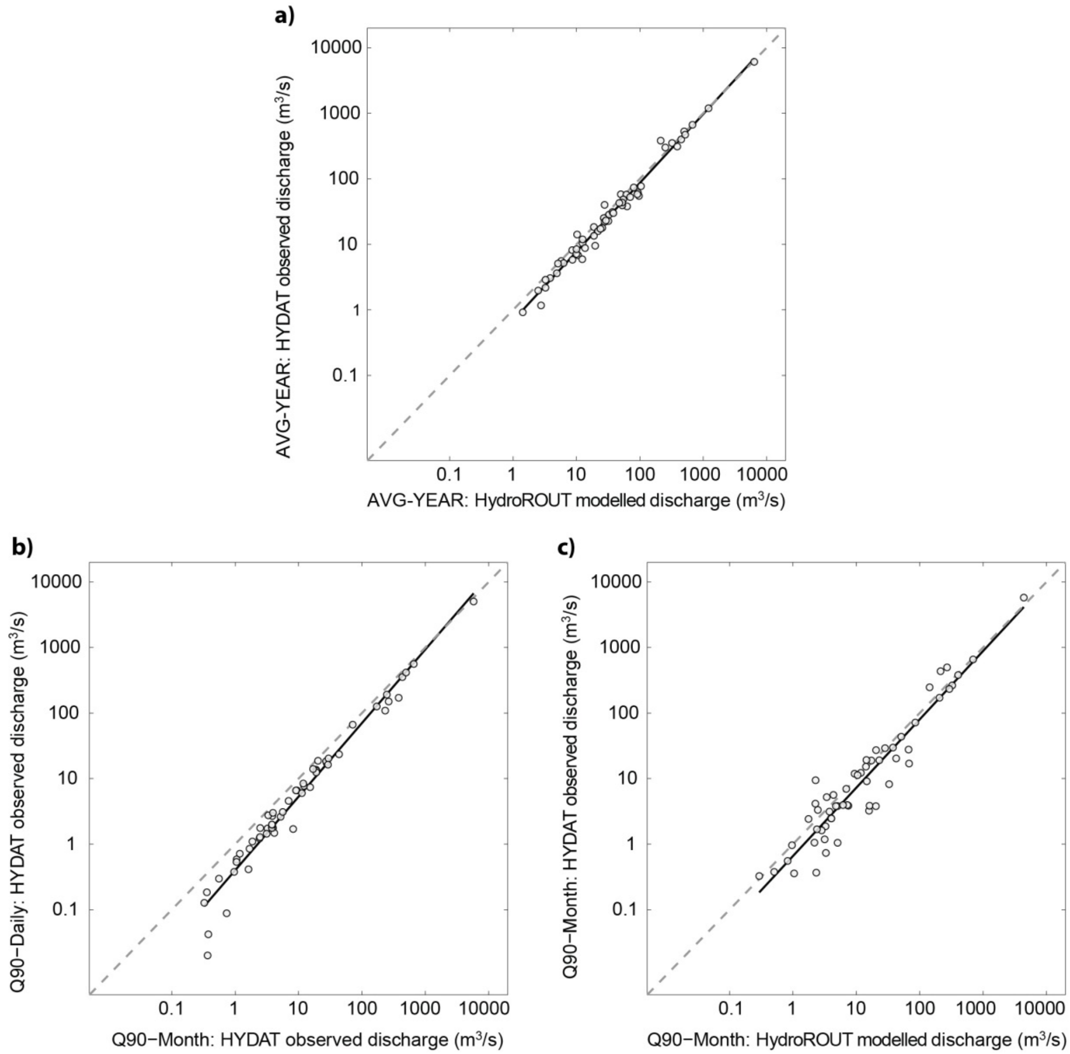


Figure 5.3: A) Scatterplot of observed and simulated long-term average flows (AVG-YEAR). B) Comparison between daily Q90 and monthly proxy (Q90-MONTH). C) Observed (re-calculated from daily data) and modeled Q90-MONTH low flow values.

This is likely caused by the more extreme nature of low flow conditions as well as the potential for significant anthropogenic influences in terms of flow regulation that are not adequately represented in the hydrological model. Accordingly, the uncertainty for predicting contaminant concentrations for low flow conditions is increasing. Nevertheless, in terms of contamination modelling this range of error is still tolerable for general screening and risk assessment analysis, assuming the goal of such assessments is to derive environmental concentrations within a factor of 10 (USEPA, 1996; Anderson et al., 2004).

5.3.2 Mass balance model validation

Figure 5.4 shows a quantitative comparison between simulated and observed CBZ concentrations. Roughly 52% of samples across 19 studies detected CBZ in surface waters (lakes and bays were excluded). The observed and simulated concentrations plotted against each other showed relatively good agreement. The observed concentrations fall within the range of both average and low flow simulated concentrations. The 95th percentile concentrations for CBZ were 13 ng/L (min 6 - max 29), 31 ng/L (11-65), and 85 ng/L under simulated average and low flow conditions, and for observed concentrations, respectively.

Due to limited availability of geo-referenced measurements, we could only conduct point-by-point validations at a few locations, reported by Lajeunesse and Gagnon (2007) and Kormos (2007). Lajeunesse and Gagnon (2007) measured upstream and at distances of up to 8 km downstream of the Montreal wastewater treatment plant in the St. Lawrence River. The measurements of Lajeunesse and Gagnon (2007) represent low flow conditions and indicate that the Montreal WWTP contributes strongly to the surface water concentrations of Carbamazepine. This observation is replicated (albeit at lower magnitude) in our model by the sharp increase of concentrations in the St. Lawrence River below the location of Montreal's WWTP (Table 5.2).

The study by Kormos (2007) measured and analysed raw surface water concentrations of Carbamazepine at two drinking water plants in the Grand River Basin, Ontario (see Figure S5.6) and included detailed river discharge at the time of measurement (see Table S5.8). The comparison between observed and simulated flow showed good overall agreement between HYDAT's reported low and average flow values with our model (Table 5.3) although the average flow is simulated notably higher than observed at Facility B. We then compared simulated concentrations under low flow conditions with predicted environmental concentrations from our model (Table S5.9). Despite the differences in modeled discharge for one of the stations, we still observed a good agreement between observed and simulated concentrations. We give a further detail of the evaluation of the two studies in S5.5.

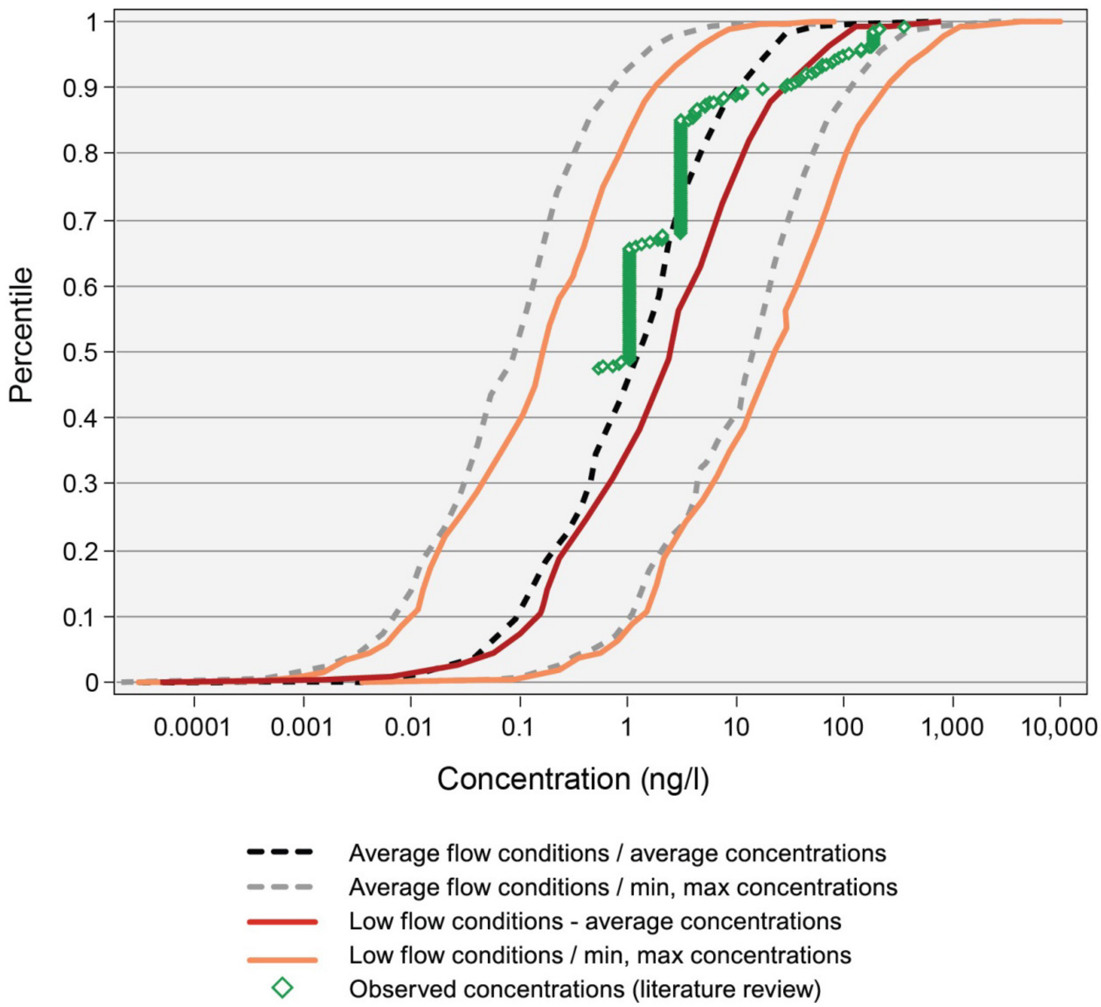


Figure 5.4 Cumulative frequency plot with minimum, maximum, and average concentrations of Carbamazepine for all reaches of the river network. Black and grey colors represent average flow conditions; red and orange represent low flow conditions. Ranges are computed based on 500 Monte-Carlo simulations with random variation of parameters up to 50% for substance usage, river discharge, and in-stream decay. Also plotted are observed concentrations for Carbamezapine in Canadian surface waters, compiled from 19 studies which analyzed 373 samples for the presence of Carbamezapine. Cumulatively, Carbamezapine was detected in 52% of the samples analysed (non-detects are included in the graph).

Table 5.2: Simulated and observed Carbamazepine concentrations (ng/L) in the St. Lawrence River near Montreal.

	0.5 km upstream WWTP	0.5 km downstream WWTP	2.5 km downstream WWTP	4.5 km downstream WWTP	8 km downstream WWTP*
Lajeunesse and Gagnon (2007)	0.8	7.4	5	4	3.5
HydroROUT simulated	0.77	1.7	1.7	1.7	2.0

* Note: This location includes the load from two additional treatment plants and from the Miles Iles River

Table 5.3: Comparison between observed and simulated long-term flow (m^3/s) and between observed and simulated Carbamazepine concentrations (ng/L) at two drinking water stations at the Grand River, Ontario. The concentration for the “observed Q90-MONTH low flow” was approximated by taking the average of the shaded cells in Table S5.8 (representing months of low flow) across samples of both facilities.

Location	Average Flow (m^3/s)		Q90-MONTH (m^3/s)		CBZ conc. (ng/L)	
	Obs.	Sim.	Obs.	Sim.	Obs.	Sim.
Facility A	37.9	42.3	14.4	11.4	21.6	25.8
Facility B	58.1	91	20.1	25.1	49.6	48

5.3.3 Dilution Factors

Many screening-level risk assessment models are built on a hypothetical “representative” scenario, based on apparently conservative assumptions regarding environmental dilution in multiple water bodies. Typically, such models for Canada assume that a conservative risk assessment can be performed by using a default dilution factor (DF) of 10 (HC/EC, 2010). The suitability of such an assumption remains to be evaluated. The model we developed here is particularly suited to perform such an evaluation.

Under low flow conditions, our results indeed show that 90% percent of DFs of the analysed WWTPs are higher than 10 (Figure 5.5), suggesting that this value is a reasonable assumption for screening level assessments where risk management prioritization is the goal. However, our results also show that the actual risk is strongly dependent on hydrological conditions: 118 WWTPs showed low (1 to 5) to medium (5 to 10) DF values, suggesting potentially elevated risk downstream of these locations (Figure 5.6).

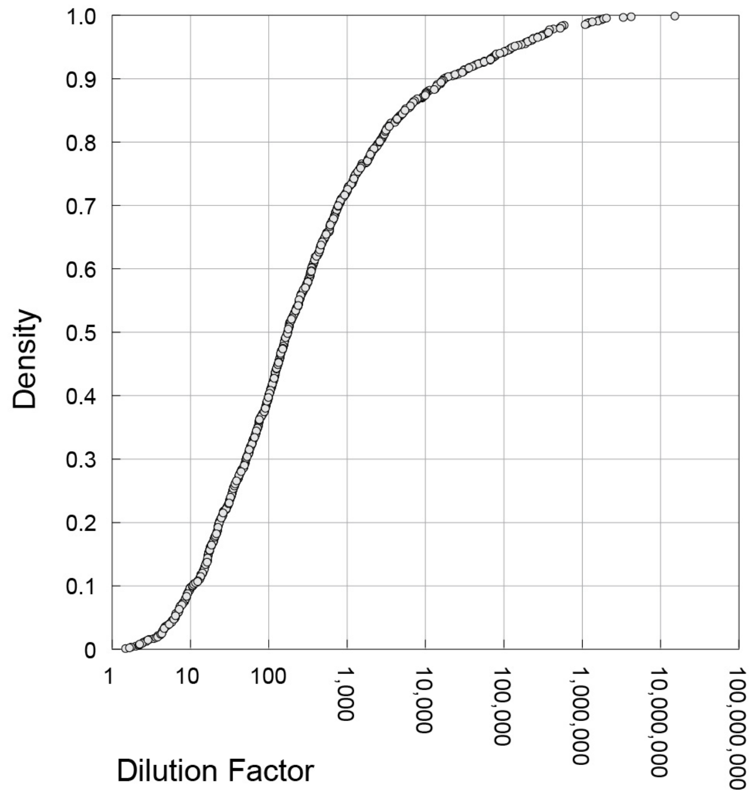


Figure 5.5: Cumulative frequency plot of dilution factors. A total of 888 WWTPs were included; WWTPs that discharge into lakes, or discharge seasonally were excluded.

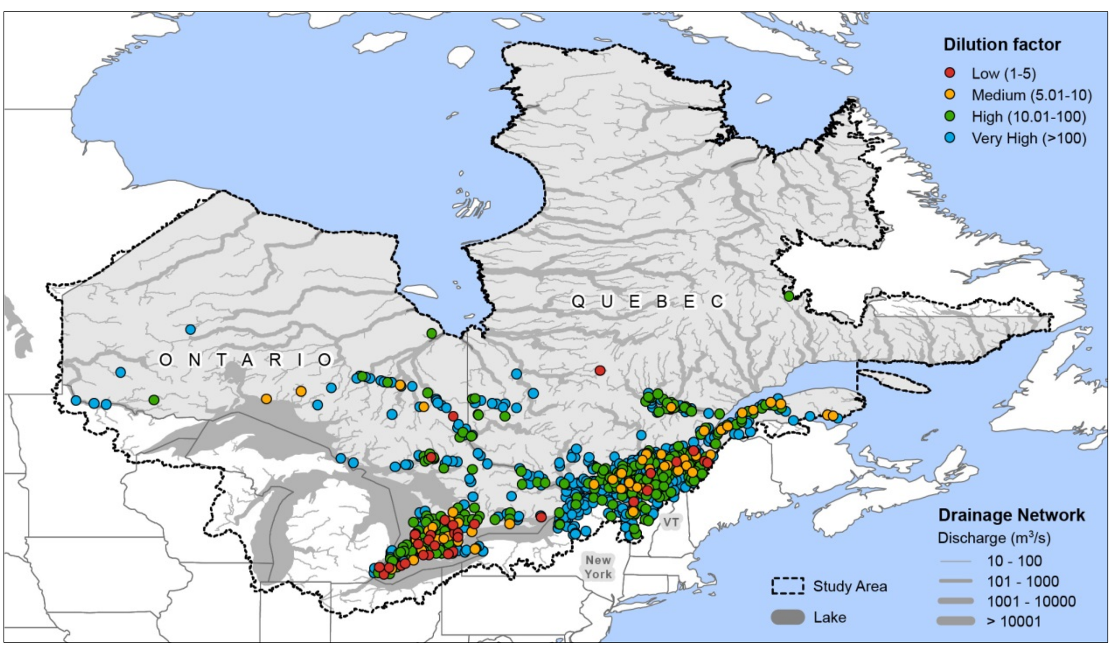


Figure 5.6: Map of dilution factors under low flow conditions (Q90-MONTH).

5.3.4 Percentage wastewater in rivers

Even if dilution is high for individual treatment plants, cumulative effects from other treatment plants may increase the potential for environmental risk. Our results suggest that numerous important rivers in our study area are potentially at risk from wastewater treatment plant effluents under low flow conditions, due to substantial accumulated effluents volumes (Figure 5.7). For example, under low flow conditions, the effluent discharge percentage of total water flow can reach mean values of 14%, 16%, and 17% for the Thames, Credit, and Don Rivers respectively. Maximum concentrations in shorter stretches of rivers can yield up to 29%, 38%, and 61% of wastewater in the Maitland, Blanche, and Credit Rivers, respectively (Table 5.4).

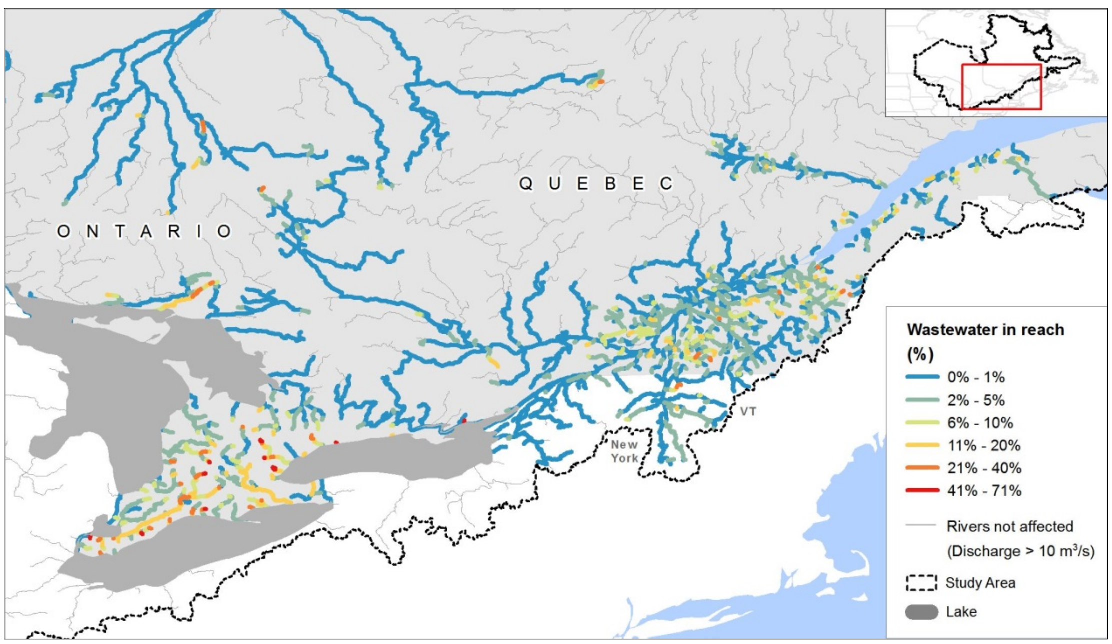


Figure 5.7: Percentage of wastewater in river course under low flow conditions (Q90-MONTH).

Table 5.4: Percentage of wastewater in river courses under low flow conditions (Q90-MONTH)

No.	Name	Wastewater % (max)	Wastewater % (mean)	Province
1	Credit River	62	14	ON
2	Blanche River	38	4	ON
3	Maitland River	29	7	ON
4	Thames (North Thames)	21	16	ON
5	Grand River	19	10	ON
6	Don River	18	17	ON
7	Mississippi River	17	9	ON
8	Ausable River	13	3	ON
9	Yamaska	9	5	QC
10	Rivière du Nord	9	6	QC
11	Rivière Bécancour	8	4	QC
12	Rivière des Envies	7	2	QC
13	Rivière Doncaster	7	2	QC
14	Rivière L'Assomption	4	2	QC
15	Rivière Champlain	4	2	QC

5.3.5 Environmental risk assessment

We evaluated environmental risk assessment for the 15 chemicals selected for the study area and present the results in Figure 5.8 and Table 5.5. Of the 15 chemicals analyzed, the release of the antibiotic Azithromycin and the exogenous estrogen used in birth control pills, Ethinylestradiol (EE2), were found to present risk to the receiving aquatic environment.

The geospatially disaggregated concentrations of the antibiotic Azithromycin are shown in Figure 5.9 (see also Figure S5.1 and Figure S5.2 for average flow conditions). Substantial river courses show elevated concentrations over extensive river lengths. On average, the Monte-Carlo runs indicated that the substance showed PEC concentrations higher than PNEC in 17.2% of the affected river reaches with a minimum of 8.6% and a maximum of 35.3% affected. Under average flow conditions, Azithromycin still triggered risk in 7.6% (min: 1.7%; max: 16.9%) of the river courses.

Geospatial concentrations of EE2 are shown in Figure 5.10 (see also Figure S5.1 and Figure S5.2 for average flow conditions). EE2 showed average PEC concentrations higher than PNEC in 3.0% (min: 0.8%; max; 7.8%) of the affected river reaches (Table 5.5). Under average flow conditions, EE2 triggered risk in 0.6% (min: 0.1%; max: 2.7%) of the river courses. We assumed PNEC for EE2 to be equal to the environmental quality standard developed by the EU (SCHER, 2011).

The risk predicted for EE2 release is lower than that predicted by Johnson et al. (2013) for a number of countries in Europe. The reason for this may be higher per capita EE2 consumption in Europe compared to Canada (Johnson et al., 2013) and lower dilution levels that are available in Europe compared to Canada for sewages upon their release to the environment (Keller et al., 2014).

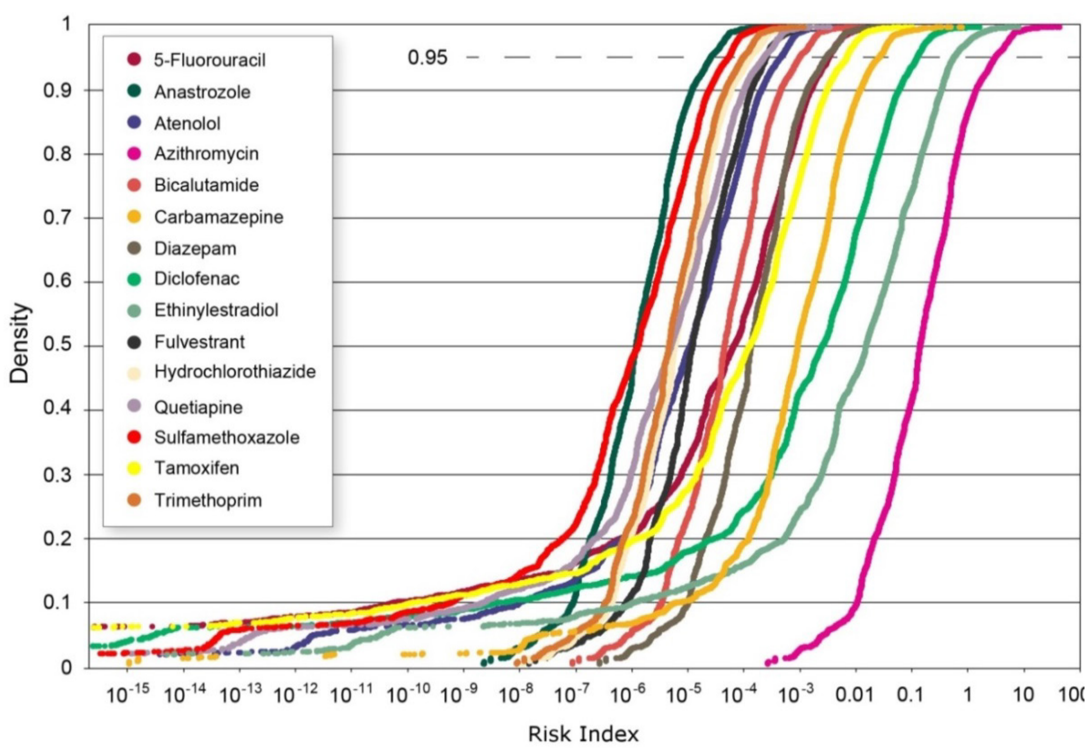


Figure 5.8: Risk indices (PEC/PNEC) for simulated chemicals under low flow conditions. Risk Index is calculated as the ration of Predicted Environmental Concentration (PEC) over the Predicted No-Effect Concentration (PNEC).

Table 5.5: Minimum, mean and maximum 95th percentile concentrations of pharmaceuticals (ng/L) in rivers under average flow conditions (AVG-YEAR) and low flow conditions (Q90-MONTH). 'Min' and 'Max' percentile concentration refer to the results of the Monte-Carlo runs, representing boundary concentrations, whereas 'mean' refers to the average of Monte-Carlo runs. Also shown are the percentages of river length affected downstream of WWTPs (total length of rivers 36,419 km) that exceed PNEC concentrations (Risk Index > 1). Substances that trigger risk in any form are emphasized in bold.

		Average flow scenario (AVG-YEAR)		**Low flow scenario (Q90-MONTH)**	
Pharmaceutical	**PNEC (ng/L)**	**Mean 95th percentile conc. (ng/L)**	**Risk Index >1 (% river length)**	**Mean 95th percentile conc. (ng/L)**	**Risk Index >1 (% river length)**
5-Fluorouracil	200	0.24 (min: 0.11; max: 0.55)	0	0.64 (min: 0.29; max: 1.41)	0
Anastrozole	1000	0.01 (min: 0.005; max: 0.025)	0	0.025 (min: 0.01; max: 0.06)	0
Atenolol	1.48 x 10^5	25.7 (min: 11.4; max: 58.4)	0	64.2 (min: 29.1; max: 142.1)	0
Azithromycin	**9.4**	**12.1 (min: 5.4; max: 27.7)**	**7.6 (min: 1.7; max: 16.9)**	**27.5 (min: 12.3; max: 61.2)**	**17.2 (min: 8.6; max: 35.3)**
Bicalutamide	1000	0.42 (min: 0.19; max: 0.95)	0	0.95 (min: 0.42; max: 2.11)	0
Carbamazepine	500	12.2 (min: 5.4; max: 27.5)	0	30 (min: 13.6; max: 66.4)	0
Diazepam	100	0.13 (min: 0.05; max: 0.28)	0	0.28 (min: 0.12; max: 0.63)	0
Diclofenac	**100**	**4.5 (min: 2.1; max: 10.3)**	**0.00006 (min: 0; max: 0.1)**	**12 (min: 5.6; max: 26.5)**	**0.0004 (min: 0; max: 0.3)**
Ethinylestradiol (EE2)	**0.035**	**0.0075 (min: 0.0033; max: 0.0172)**	**0.6 (min: 0.1; max: 2.7)**	**0.0188 (min: 0.0086; max: 0.042)**	**3.0 (min: 0.8; max: 7.8)**
Fulvestrant	0.57	0.00005 (min: 0.000022; max: 0.000115)	0	0.000139 (min: 0.000063; max: 0.0003)	0
Hydrochloro-thiazide	10^6	54.7 (min: 24.2; max: 124.6)	0	123.7 (min: 55.4; max: 275.3)	0
Quetiapine	10^4	0.9 (min: 0.4; max: 2)	0	2.2 (min: 1; max: 4.8)	0
Sulfamethoxa-zole	5.9 x 10^5	11.3 (min: 5.1; max: 25.5)	0	30.2 (min: 13.6; max: 66.2)	0
Tamoxifen	22	0.06 (min: 0.03; max: 0.13)	0	0.15 (min: 0.07; max: 0.33)	0
Trimethoprim	2.4 x 10^5	9.5 (min: 4.2; max: 21.7)	0	22.17 (min: 9.9; max: 49.18)	0

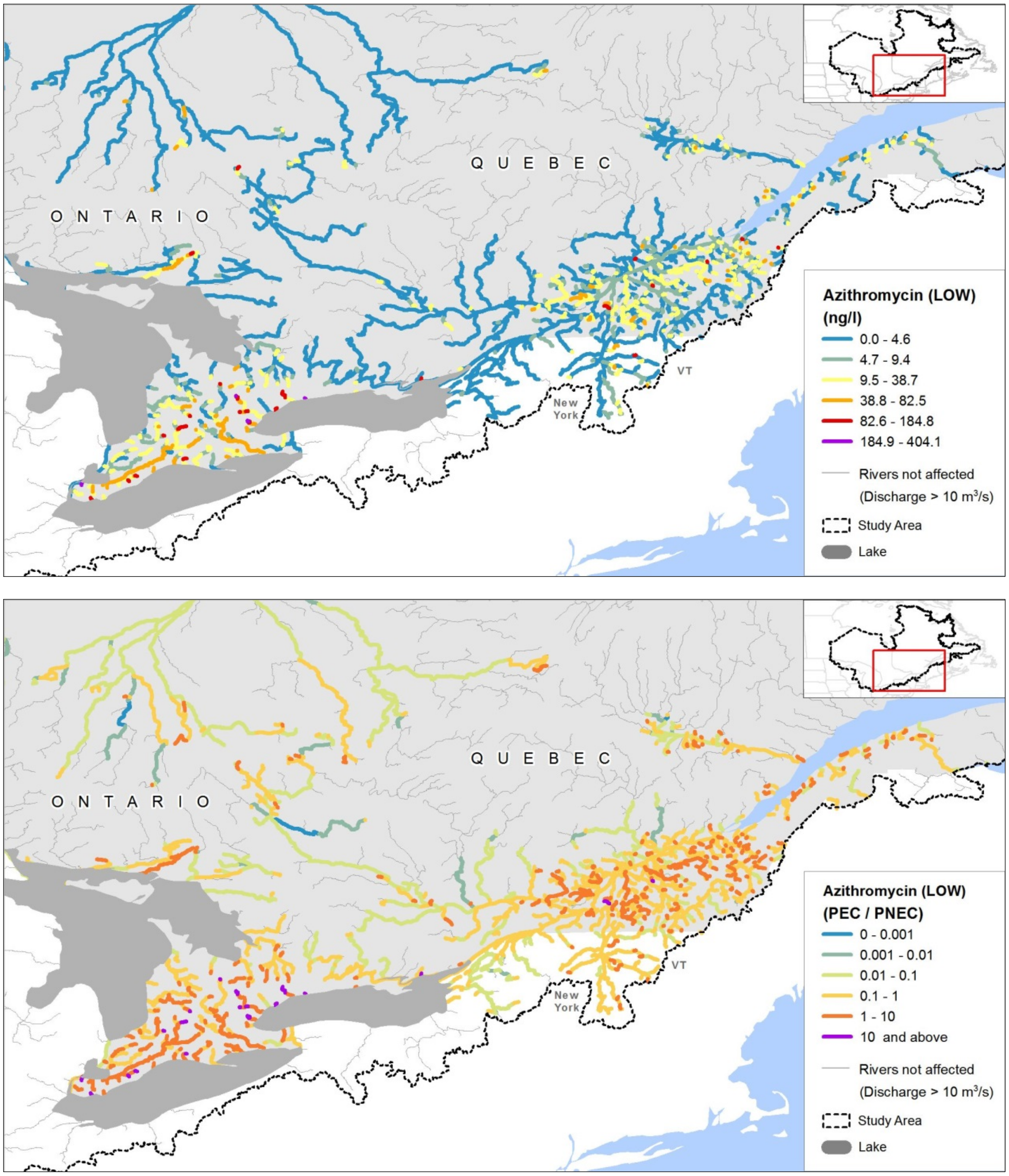


Figure 5.9: Simulated Azithromycin concentrations in rivers under low flow conditions and associated risk indices assuming PNEC of 9.4ng/L (see SI for maps under average flow conditions).

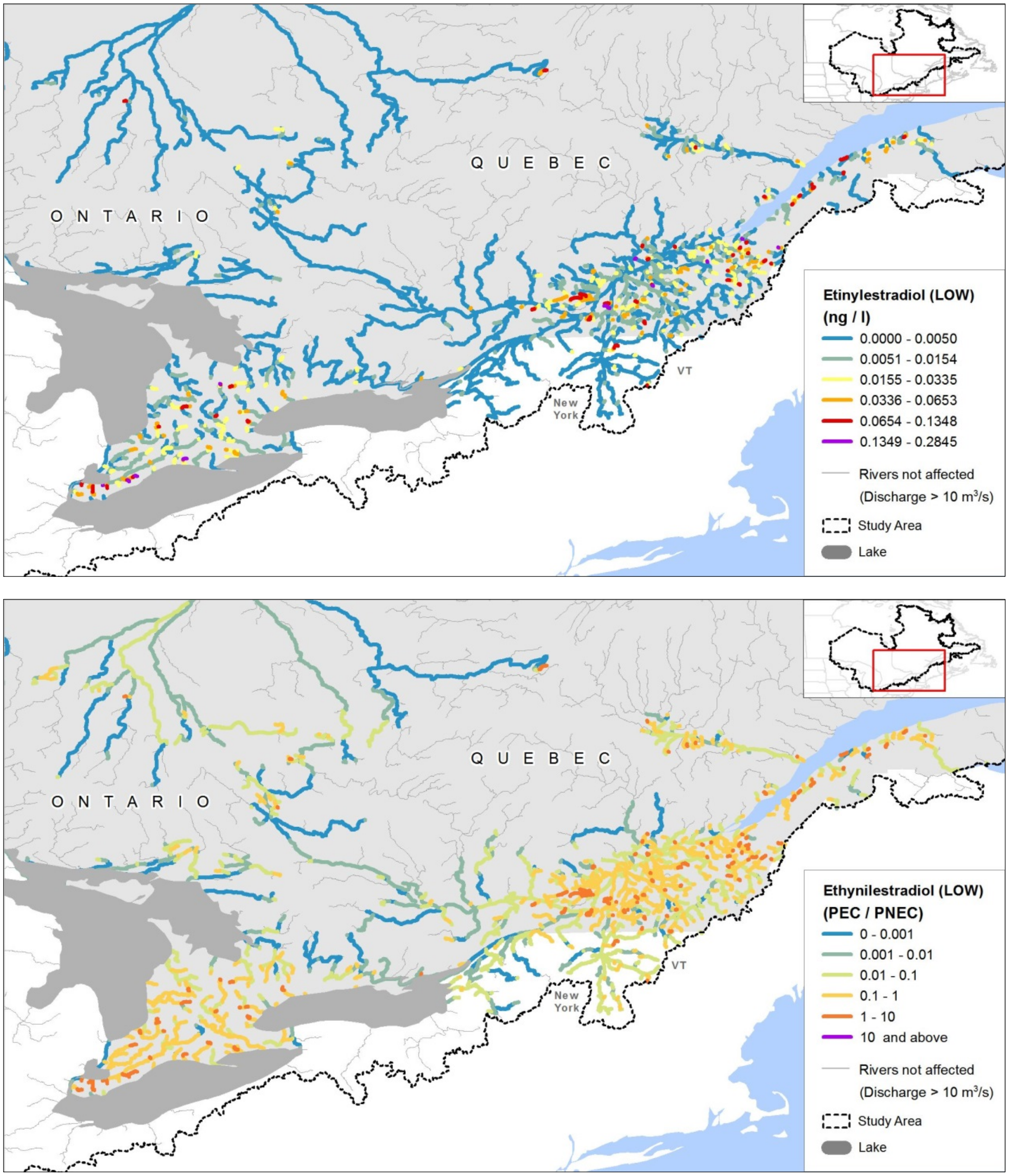


Figure 5.10: Simulated Ethinylestradiol concentrations in rivers under low flow conditions and associated risk indices assuming PNEC of 500ng/L (see SI for maps under average flow conditions).

5.4 Discussion

We first calculated the percentage of wastewater in river flow indicator. As a separate indicator which can be derived relatively easy, the results pinpoint geographic hotspots of elevated generic risk and may be targeted in management plans to improve the wastewater treatment infrastructure or reduce consumption and disposal of certain chemicals upstream.

We then estimated risk for 15 pharmaceutical and found quite high risk for Azithromycin. To better appraise risk posed by the aquatic release of Azithromycin, it is worth discussing the manner in which its PNEC was estimated. We estimated Azithromycin's PNEC by dividing an acute endpoint for cyanobacteria with an application factor (AF) of 100 (Khan, 2014). We deemed this necessary, since the available endpoint for the critical taxon was an acute endpoint. Apparent Azithromycin risk may reside partially in the high AF required to evaluate its PNEC. Therefore, a chronic endpoint for the exposure of Azithromycin to cyanobacteria needs to be developed. Nevertheless, currently available data indicates that Azithromycin release presents an unacceptable risk to the study area.

The risk of EE2 was also found elevated. PNEC for EE2 was developed with an AF of 2, which is indicative of the fact that there is good confidence in the ecotoxicological data available for fish exposure to EE2. Therefore, the predications made here can be taken to suggest that EE2 may be presenting an unacceptable risk to a small number of river reaches in the study area. Yet, the risk posed by EE2 when assessed on its own also has limitations, because EE2 is known to act with other endogenous estrogens to result in fish developing intersex characteristics (Williams et al., 2009).

However a number of shortcomings and unaddressed uncertainties should be mentioned and steps should be taken to address uncertainties more completely in the future. Regarding discharge, we found a good agreement between our downscaled model values and the discharge measured at gauging stations. Nevertheless, for large-scale runoff modelling, important factors affecting routed discharge accuracy include, but are not limited to, errors in the routing process (e.g., related to flow velocity, channel geometry, flow attenuation) and errors due to inadequately represented flow regulation structures. Human controlled flow regulation features, such as dams and reservoirs, are common in the study area, possibly causing a misrepresentation of low flow conditions in particular. Also, errors in routing diverging (bifurcating) river channels (which may converge again further downstream) can cause significant inaccuracies in HydroROUT as the model is currently unable to represent this situation. Another source of uncertainty is the simplified estimation of flow velocity. In the current model version, we use variable flow velocity based on Allen et al. (1994). This approach is very simplistic and affects, among other things, the time available for in-stream decay. This method is subject to further verification and uncertainty analysis; however we do not consider it a main contributor to uncertainty in this case.

The validation using Carbamazepine revealed differences between observed and simulated 95th percentile concentrations, which can be attributed, in part, to the limited data point availability across different river sizes, but more likely to the fact that measurements are generally biased towards collecting samples from surface waters with high concentrations, such as downstream of wastewater treatment plants (Hannah et al., 2009). Furthermore, almost no study records river flow at the time of measurement, which further complicates the comparison. A possible improvement to the validation methodology could be achieved by comparing cumulative probability plots for specific regions with simulated and reported concentrations from similar river reaches.

To assess model sensitivity related to parameter uncertainty, the Monte-Carlo simulation module is currently based on simplified assumptions to generate combinations of parameter sets. We believe, however, that better uncertainty calculation implementations can further improve model output confidence.

We have not addressed the temporal variation in product consumption and associated chemical emissions. Consumption may vary inter-annually or seasonally, change between different days of the week, or follow diurnal fluctuations (Kormos, 2007). These patterns in product consumption are currently disregarded since such information is typically not available. More insight into specific chemical variations could improve sensitivity analysis settings and allow for more realistic results. Sales data at individual province or city levels or even at finer scales is readily available. Therefore, in cases where substance consumption differs significantly by region, the accuracy of predictions could be much improved. Such data is available (e.g., from the firm IMS Health); however, the associated costs have prohibited their inclusion in this study.

Additionally, the contaminant fate model does not currently include input from atmospheric deposition, erosion, surface runoff, or from non-treated wastewater. The latter could be a significant contributor since the connection of population to sewage treatment plants is incomplete.

5.5 Conclusions

We developed a first version of a new geospatial chemical fate model that combines a global, high-resolution hydrographic framework HydroSHEDS with a river routing model, HydroROUT, to estimate sources and fate of chemicals, and validated the performance in a case study focusing on the Saint Lawrence River basin. We simulated environmental concentrations for 15 pharmaceuticals and identified potentially elevated risk for the estrogen Ethinylestradiol (EE2) and the antibiotic Azithromycin in a moderate and large number of river reaches, respectively.

Based on the results presented, we believe that our model is capable of realistically estimating in-river contaminant concentrations for the purpose of screening for problematic chemicals, especially considering that a typical goal of such models is to simulate concentrations that are within an order of magnitude of measured concentrations. Yet, a number of improvements can be made to gain greater confidence in our model simulation:

The discharge used in the HydroROUT chemical fate model is in good agreement with data from gauging stations, although further improvement can be made to model low flows more accurately, which tend to be simulated too high compared with commonly used low flow metrics such as Q90. As such our model may underestimate risk under low flow conditions.

Our model is unique in linking to high-resolution databases of lakes, and by taking into account lake removal processes, which are particularly important in

Canada's lake-rich environments. Yet future development should include more validation of lake volumes estimation and possibly also test other lake removal models.

A future goal is to extend this model to the Canadian scale to support the implementation of the Canadian Chemical Management Plan. Based on the experiences with this pilot study, we see this goal within our reach, if substantial work is conducted to gain greater confidence in our discharge model, especially focussing on low flow metrics, and if enough resources are available to invest in the time to link the contaminant sources such as WWTP, which took a substantial share of the model development time. The seamless hydrographic database facilitated modelling of cross-border contributions from sources outside Canada, which is otherwise difficult, and typically a major roadblock in developing models across national boundaries. This is especially true for the availability of discharge data, which is seamlessly available from HydroSHEDS, two of the major components of a Canadian-wide contaminant fate model.

We anticipate that this model will serve other functions as well. For example, it could be used in conjunction with on-going research to evaluate the importance of a number of emerging contaminants; to evaluate the cumulative effects of pollutant loadings originating from urban communities and industries (e.g., hospitals, mining, petroleum, pulp and paper) across entire watersheds at large scales; and to simulate the potential impacts of specific events that can result in significant particular pollutant emissions across watersheds for relatively short periods of time (e.g., during the extensive use of pharmaceuticals during pandemics).

Our model aims to enhance our scientific understanding of the legacy of contaminants at large and to create a leading, large-scale contaminant fate model that can serve as a unique tool to study risks and future threats related to cumulative effluents of substances from domestic, agricultural, or industrial sources; screen for critically impacted areas; and test or optimize mitigation scenarios. If extended to the national scale, the model can support the implementation of the Canadian Chemical Management Plan by adding a geospatial risk assessment component.

Supplementary information

S5.1 Maps of chemical concentration

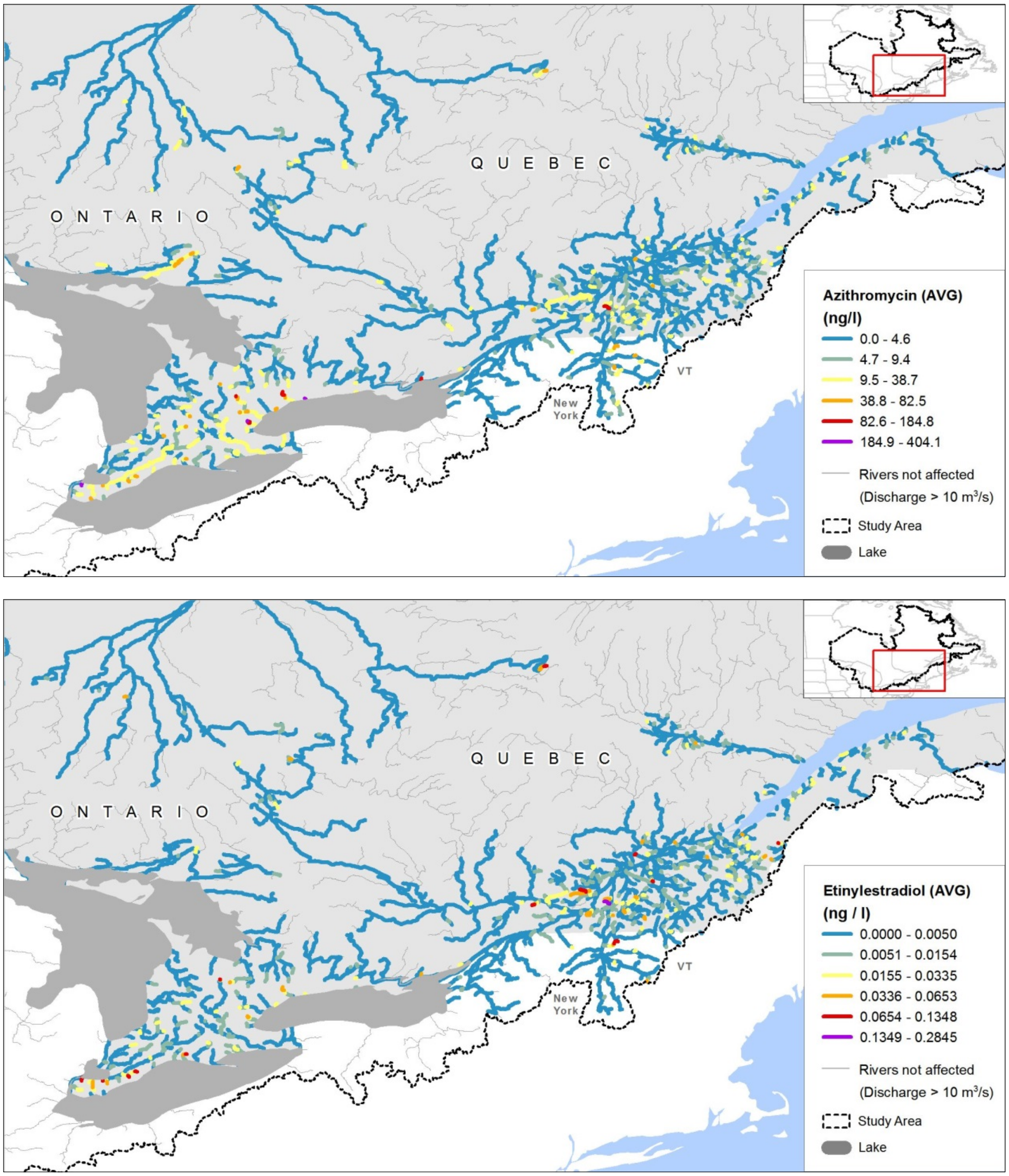
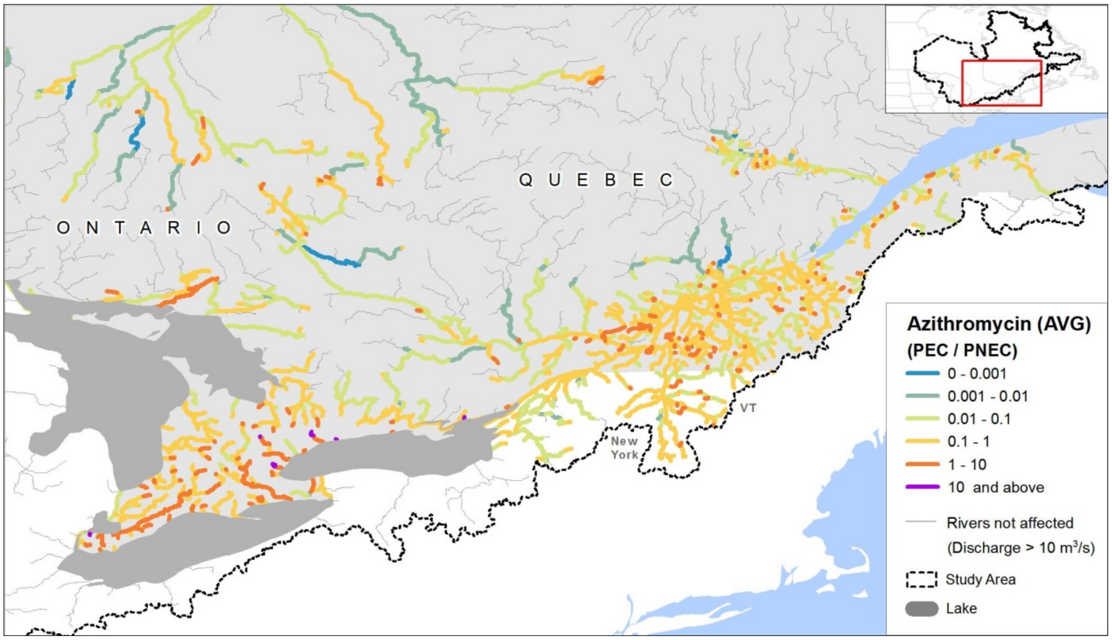


Figure S5.1: Simulated concentrations for Azithromycin (top panel) and Ethinylestradiol (bottom panel) under average flow conditions (AVG-YEAR).

S5.2 Maps of risk index



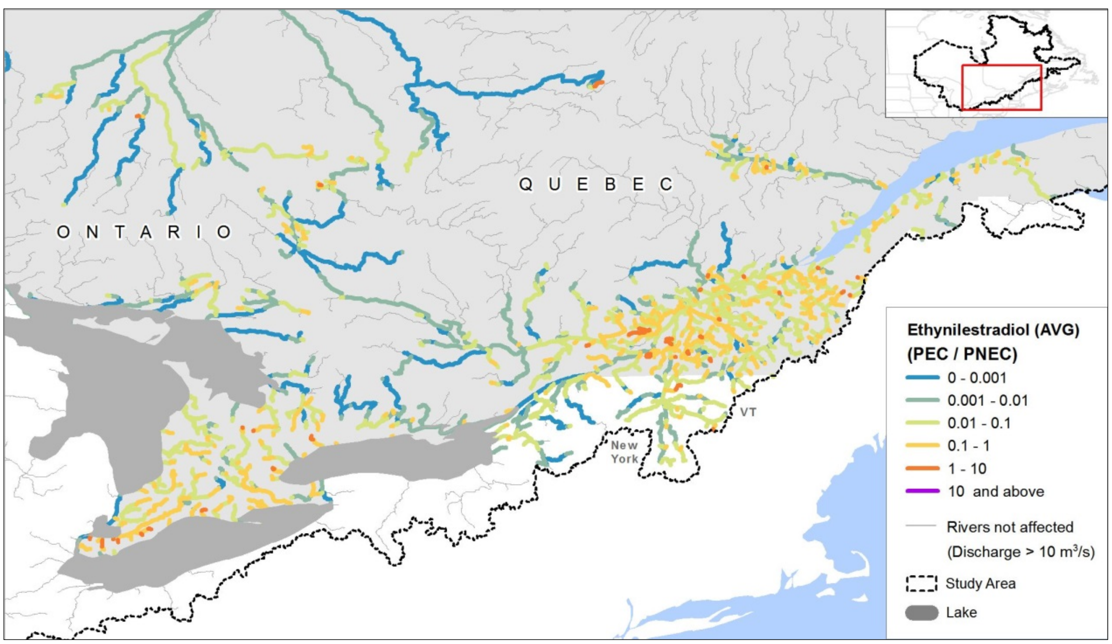


Figure S5.2: Simulated risk indices for Azithromycin (PNEC = 9.4 ng/L; top panel) and for Carbamazepine (PNEC = 500 ng/L; bottom panel) under average flow conditions.

S5.3 Detailed hydrological model validation

In order to evaluate the accuracy and uncertainty associated with the HydroROUT model, we compared our downscaled discharge estimates with the reported values of HYDAT gauging stations (Environment Canada, 2012).

To evaluate the general quality of the downscaled discharge estimates used in HydroROUT, long-term average flows and low flows were compared across stations using a linear regression analysis. In a second comparison, the ability of the HydroROUT model to simulate the annual flow regime was tested using the Nash-Sutcliffe Efficiency (NSE) as the quality indicator.

By means of the Environment Canada Data Explorer, average daily stream flows of gauges in the provinces of Quebec and Ontario were extracted from HYDAT (Environment Canada, 2012) in a time series format. Location information is provided for all gauging stations as coordinates in decimal degrees. The HYDAT gauging stations had to be co-registered (i.e., snapped) to the stream network of HydroROUT to allow for the assignment of simulated (downscaled) long-term average monthly flows and upland area from the river network model to each station. With the use of provincial dam and reservoir data, gauging stations downstream in close proximity to potential sources of flow regulation were flagged accordingly; yet it is important to note that this information was not readily available for Ontario, thus only Quebec rivers were attributed for the presence of dam structures in near vicinity. Finally, the difference in upland watershed area between the reported values of the HYDAT stations and the simulated values based on HydroSHEDS was calculated. This difference serves as a measure of accuracy for the co-registration of the HYDAT gauging station to the HydroROUT stream network.

For the subsequent validation purposes, different subsets of stations in the provinces of Quebec and Ontario were selected by extracting them from the total of 522 available stations. In particular, a set of 57 “most reliable” stations was derived based on the following, consecutive criteria:

1) Stations need to have 30 years of observed, uninterrupted records in the ‘climate normal’ period (1961-90) - number of stations reduced to 235

2) Stations need to show a difference of maximum 10% between reported HYDAT and estimated HydroSHEDS watershed area - number of stations reduced to 156.

3) Stations cannot have a known major dam or reservoir downstream in close proximity as the effect of flow regulations are not included in our downscaled discharge estimations - number of stations reduced to 57.

Additional validation sets with more HYDAT stations and different criteria have also been evaluated (see Table S5.1 and Table S5.2, as well as Figure S5.3 and Figure S5.4 for results). The R statistical software and its HydroTSM package (Zambrano-Bigiarini, 2011) were used to calculate statistics relevant to the validation of the modeled discharge data. Daily and monthly statistics were calculated for each HYDAT station from daily flow data during the study period. For the HydroROUT model, only monthly statistics were calculated from the long-term average monthly flow data. The following hydrological indicators were assessed for observed flows (HYDAT) and simulated flows (HydroROUT):

- MQ: long-term average flow for time series 1961-90
- Annual flow regime: series of 12 representative monthly flow values (Jan-Dec) representing the average flow for each month calculated from the period 1961-90
- Q90: daily flow that is exceeded at 90% of time (only calculated for HYDAT data)
- Q90-MONTH : average flow of the lowest month in the annual flow regime (see above)

The risk from chemical substances in surface waters is usually assessed under low flow conditions for which Q90 is a frequently used indicator. Typically, Q90 is calculated from daily discharge measurements, but as the WaterGAP runoff estimates are given as monthly time series, we approximated (daily) Q90 with a substitute, namely Q90-MONTH. To explore the validity of this approach, we first assessed the relationship between Q90 and Q90-MONTH for observed HYDAT flows by performing a linear regression analysis, and then applied a second linear regression to test the correlation of Q90-MONTH between observed HYDAT and simulated HydroROUT values.

Table S5.1: Statistics for low flow (Q90-MONTH) validation using different sets of HYDAT gauging stations with different characteristics (for corresponding scatterplots see Figure 5.3).

Set	Max. upland error (%)	Years of data	R^2	Number of stations
A	10	10	0.89	307
B	10	30	0.92	97
C	25	10	0.88	367
D	25	30	0.91	111
E	100	10	0.43	458
F	100	30	0.35	138

Table S5.2: Statistics for long-term average flow validation using different sets of HYDAT gauging stations with different characteristics (for corresponding scatterplots see Figure S5.4).

Set	Maximum upland error (%)	Years of data	R^2	Number of stations
G	10	10	0.97	307
H	10	30	0.97	97
I	25	10	0.97	367
J	25	30	0.97	111
K	100	10	0.40	458
L	100	30	0.31	138

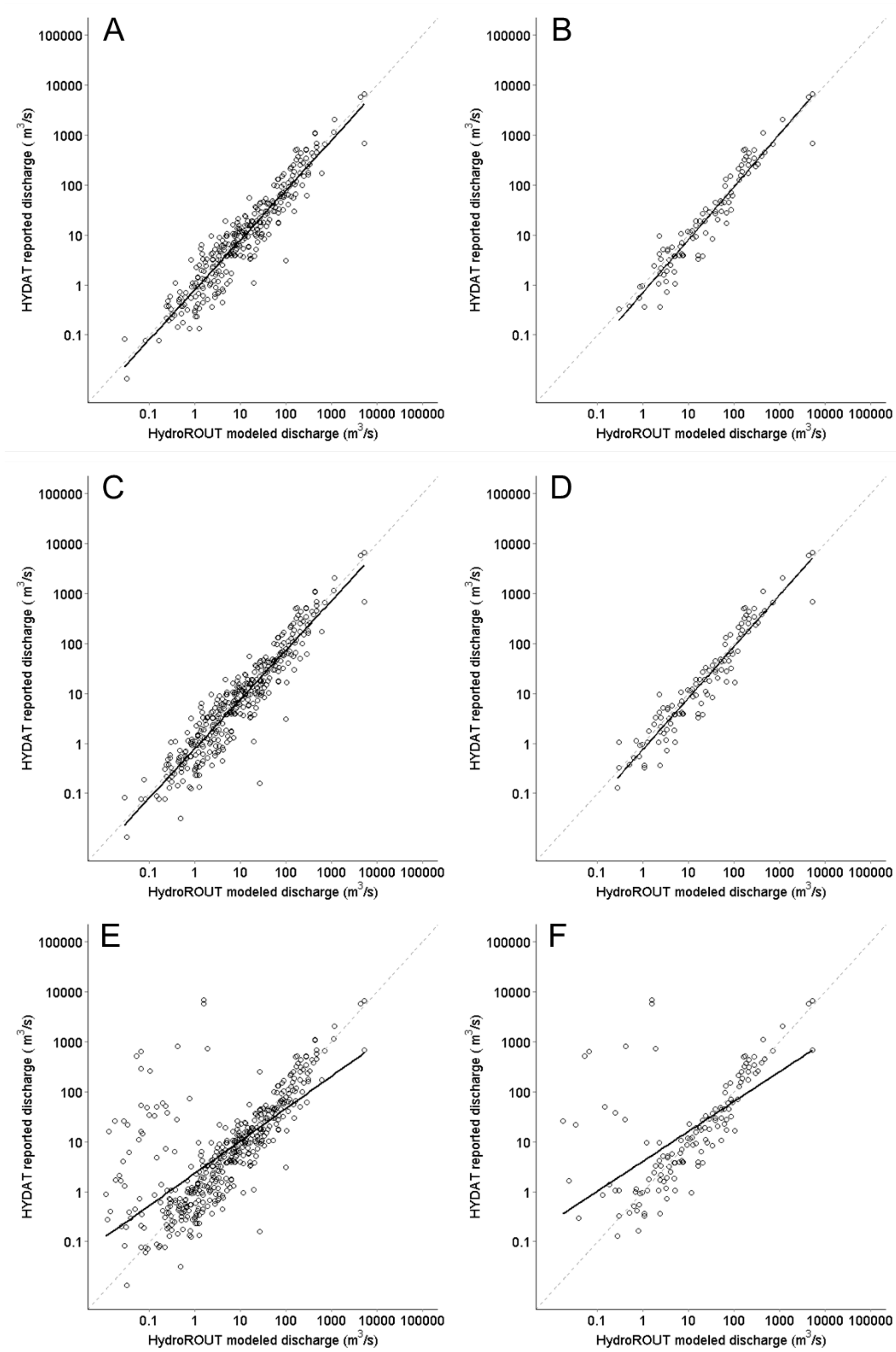


Figure S5.3: Scatterplots for observed and modeled low flow (Q90-MONTH) values (see Table S5.1 for more explanation). Plots E and F include stations with large discrepancies in reported versus modeled watershed areas; these HYDAT stations may either be co-registered to incorrect tributaries on the HydroSHEDS river network; or may be located on braided stream channels (or canals) which are not properly represented in HydroSHEDS; or their watershed may be incorrectly delineated in HydroSHEDS.

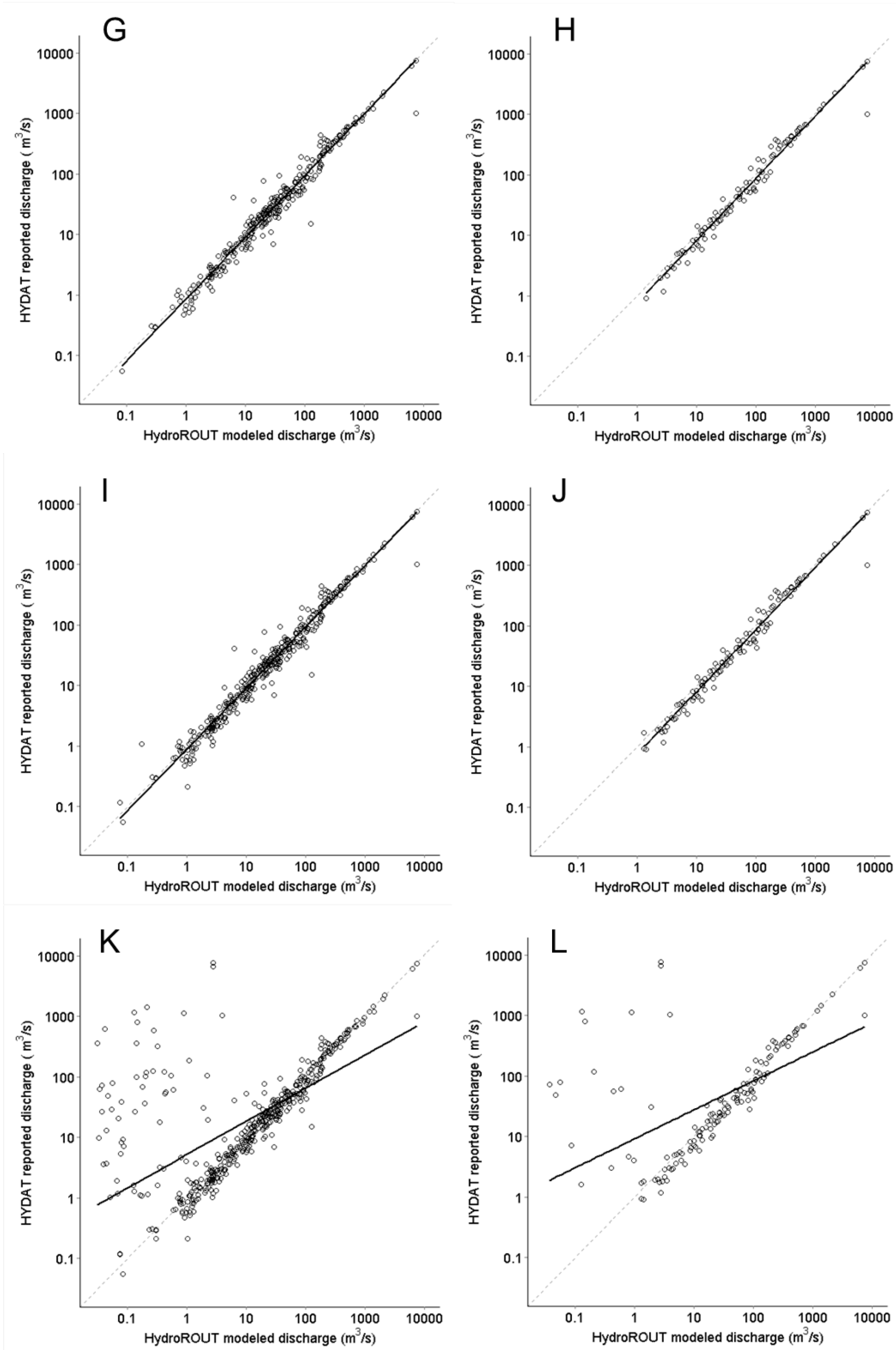


Figure S5.4: Scatterplots for observed and modeled long-term average flow values (see Table S5.2 for more explanation). Plots K and L include stations with large discrepancies in reported versus modeled watershed areas; these HYDAT stations may either be co-registered to incorrect tributaries on the HydroSHEDS river network; or may be located on braided stream channels (or canals) which are not properly represented in HydroSHEDS; or their watershed may be incorrectly delineated in HydroSHEDS.

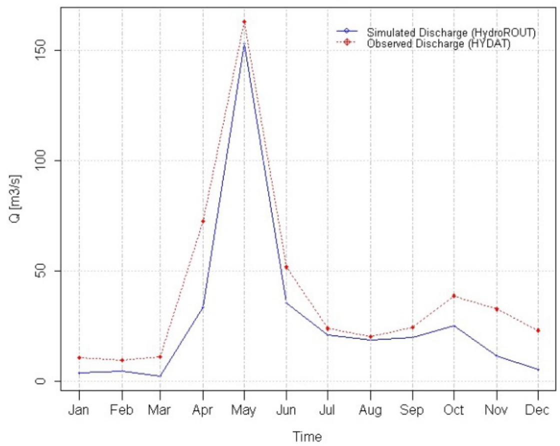
Comparison of long-term average monthly flows

Figure S5.5 depicts four graphs with comparisons of flow regimes for long-term average monthly flows derived from observed HYDAT station data and downscaled for use in the HydroROUT model. The Nash-Sutcliffe model efficiency (NSE) was used as a measure for the strength of the correlation. This analysis was executed using the HydroGOF package (Zambrano-Bigiarini, 2011) in R. NSE values can range from minus infinity to 1, with 1 being the best value, and negative values indicating a less than random fit.

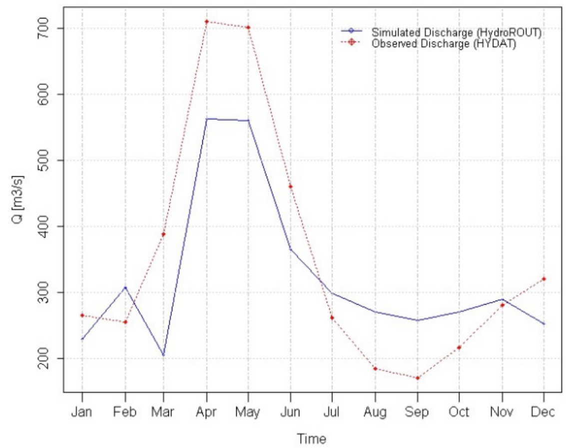
Based on these NSE values, we prescribed ratings following (with slight modifications) suggestions by Moriasi and Arnold (2007) for examining river flow data with a monthly time step (Table S5.3). Although the resulting NSE coefficients are low or unsatisfactory in many cases, these results need careful interpretation. In particular, the NSE coefficient is not very suitable for capturing situations where a temporal shift in discharge regimes exists while the overall shape of the curve is similar (see Figure S5.5, lower right panel). Such shifts are often observed in large-scale models for high latitudes, where snowmelt process are predominant, and temporal shifts of one or two months due to inadequate representation of daily temperature fluctuations (which induce snow melt) are common. The most important measure for our contamination risk assessment model, i.e. the magnitude of the low flow index Q90-MONTH, may still be represented reasonably well despite this temporal shift in the flow regime.

Table S5.3: NSE rating summary for 57 selected HYDAT stations

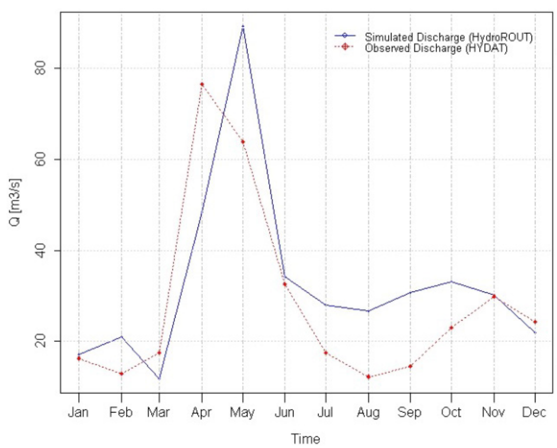
NSE rating	Range	Count
Very Good	1.00 ≥ NSE >0.75	2
Good	0.75 ≥ NSE > 0.65	3
Satisfactory	0.65 ≥ NSE > 0.50	5
Inconclusive	0.50 ≥ NSE ≥ 0	14
Unsatisfactory	NSE < 0	33



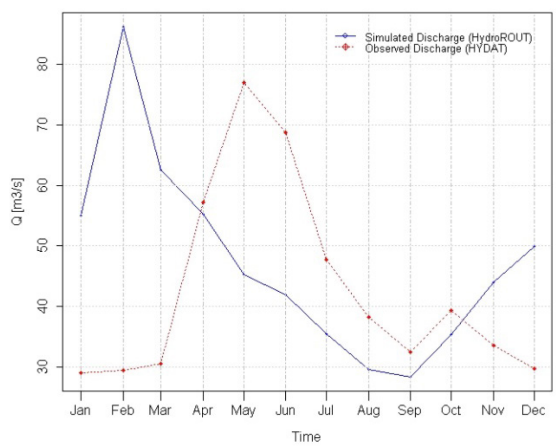
NSE rating example: very good fit (NSE=0.85; Station ID: 02QB001)



NSE rating example: good fit (NSE=0.7; Station ID: 02OJ007)



NSE rating example: satisfactory fit (NSE=0.53; Station ID: 02CC005)



NSE rating example: Unsatisfactory fit (NSE=-1.49; Station ID: 05QD006)

Figure S5.5: Flow regime examples and Nash-Sutcliffe Efficiency (NSE) ratings.

S5.4 Wastewater treatment plants

Data was available for a total of 1283 WWTPs within the study area. At the time of data collection in 2013, Quebec had the most treatment plants with a total of 692, followed by Ontario with 473. The states of Vermont and New York in the contributing basin areas had 45 and 73 WWTPs, respectively. All 1283 WWTPs discharge a total effluent volume of 13.8 million m^3 per day and serve a total population of 15.3 million people (see Table S5.4). From the total of 1283 given treatment plants, we included 1198 in the current model and excluded 85 due to lack of information (treatment type, population, location).

The distinction of direct discharge into lakes is an important characteristic of the model implementation. According to the available data, 221 WWTPs discharge a total volume of 4.5 million m^3 per day into 89 lakes across the four provinces/states. Further information, a breakdown of WWTPs that discharge into lakes by region, and statistics for all lakes with WWTPs are included in Table S5.5 and Table S5.6.

Table S5.4: Summary statistics of discharge volume, population served, and number of WWTPs by province/state

Treatment type	Ontario	Quebec	New York	Vermont	Total
No treatment					
Discharge (m^3/day)	4,195	39,238			43,433
Population served	1,708	44,002			45,710
Number of WWTP	5	63			68
Primary treatment					
Discharge (m^3/day)	535,395	3,628,418	14,896	42	4,178,750
Population served	533,873	2,734,625	24,805	366	3,293,669
Number of WWTP	21	35	25	1	82
Lagoon treatment					
Discharge (m^3/day)	224,181	1,473,220	9,539	8,487	1,715,427
Population served	299,551	1,915,737	10,213	18,171	2,243,672
Number of WWTP	171	516	9	12	708
Activated sludge					
Discharge (m^3/day)	6,019,814	626,299	29,583	4,471	6,680,167
Population served	7,414,725	735,791	34,759	9,077	8,194,352
Number of WWTP	232	37	23	2	294
Trickling filter					
Discharge (m^3/day)	103,983	651,841	41,185	9,615	806,624
Population served	115,852	885,881	35,121	7,831	1,044,685
Number of WWTP	11	30	6	1	48
Tertiary treatment					
Discharge (m^3/day)	39,582	101,210	66,699	97,016	304,507
Population served	41,012	138,846	67,435	183,279	430,572
Number of WWTP	17	11	10	29	67
Non-surface discharge or treatment type unknown					
Discharge (m^3/day)	82,612				82,612
Population served	57,570				57,570
Number of WWTP	16				16
Total effluent (m^3/day)	**7,009,762**	**6,520,226**	**161,902**	**119,630**	**13,811,520**
Total pop. served	**8,464,291**	**6,454,882**	**172,333**	**218,724**	**15,310,230**
Total # of WWTP	**473**	**692**	**73**	**45**	**1,283**

Table S5.5: Volume of wastewater directly discharged into lakes by region.

Region	# WWTPs discharging into lakes	Effluent discharged (m^3/day)	Population served
Quebec	44	71,153	108,717
Ontario	149	4,405,978	5,468,433
Vermont	7	25,764	55,903
New York	21	45,512	45,512
Total	221	4,548,407	5,679,571

Table S5.6: Statistics of lakes that receive direct discharge from WWTPs

Lake Name	# of WWTPs discharging into lake	Total population served	Effluent Received (m^3/day)	Lake surface area (sq. km)
Lake Ontario	46	4,576,034	3,509,777	19,347
Lake Michigan	31	190,515	180,729	117,120
Lake Champlain	12	75,793	50,605	1,141
Lake Erie	12	57,685	88,072	25,767
Lake Simcoe	7	107,065	100,901	759
Lac St. Jean	5	15,455	11,896	1,066
Lake Superior	5	110,388	116,869	81,843
Lake Timiskaming	4	7,055	6,520	204
Lac Aylmer	3	5,480	4,213	32
Lake Muskoka	3	6,408	4,135	115
Buckhorn Lake	2	2,452	3,373	109
Lac Magog	2	3,655	1,195	11
Lac Megantic	2	440	110	27
Sturgeon Lake	2	16,993	18,547	43
Other Lakes (total 78)	85	504,153	451,464	4,904
Total	221	5,679,571	4,548,407	252,488

S5.5 Mass balance validation: point-by-point

Case 1: St. Lawrence River at Montreal

In order to conduct a point-by-point validation, we used point locations provided by Lajeunesse and Gagnon (2007), who measured upstream and at distances of up to 8 km downstream of the Montreal wastewater treatment plant in the St. Lawrence River. Montreal's WWTP is only equipped with primary treatment technology. As a result, little pharmaceuticals are removed due to the treatment process (see also Gagnon and Lajeunesse, 2012). The measurements of Lajeunesse and Gagnon (2007) represent low flow conditions and indicate that the Montreal WWTP contributes strongly to the surface water concentrations of Carbamazepine. This observation is replicated (albeit at lower magnitude) in our model by the sharp increase of concentrations in the St. Lawrence River below the location of Montreal's WWTP (Table S5.7).

Table S5.7: Simulated and observed Carbamazepine concentrations (ng/L) in the St. Lawrence River near Montreal

	0.5 km upstream WWTP	**0.5 km downstream WWTP**	**2.5 km downstream WWTP**	**4.5 km downstream WWTP**	**8 km downstream WWTP***
Lajeunesse and Gagnon (2007)	0.8	7.4	5	4	3.5
HydroROUT simulated	0.77	1.7	1.7	1.7	2.0

* Note: This location includes the load from two additional treatment plants and from the Miles Iles River

The comparison with the study by Lajeunesse and Gagnon (2007) needs careful interpretation. The measured river flow at the day of observation was 8340 m^3/s according to HYDAT at station St. Lawrence/Lasalle (Station ID: 020A016), whereas HydroROUT's Q90-MONTH index for low flow conditions indicates 7776 m^3/s, i.e. nearly 10% lower than the measured value. Higher concentrations are thus expected from HydroROUT, yet simulated concentrations are roughly 40% lower than observed. On the other hand, Lajeunesse and Gagnon (2007) do not provide measured effluent concentrations from the WWTP for the observation date. Although little inter-annual variation may be expected, the same study references two effluent measurements at the same time of the year that differ by a factor of seven (656 ng/L on 27 April 2005 compared to 91 ng/L on 26 April 2006). Since we do not know the effluent concentrations on the day of observation, we cannot be sure whether the given in-river concentrations are on the high or low end of the spectrum.

Finally, it should be noted that literature studies in general tend to rely on measurements made directly in the sewage effluent or immediately downstream of a treatment plant, likely within the discharge plume of the WWTP. Our current model cannot represent concentrations at specific points within the mixing zone since it assumes full mixing of the wastewater parcel upon release. Therefore, simulated concentrations might be smaller than those reported in literature.

Case 2: Grand River at Kitchener

A study by Kormos (2007) measured and analysed raw surface water concentrations of Carbamazepine at two drinking water plants in the Grand River Basin, Ontario (see Figure S5.6) and included detailed river discharge at the time of measurement (Table S5.8). The comparison between observed and simulated flow showed good overall agreement between HYDAT's reported low and average flow values with our model (Table S5.9) although the average flow is simulated notably higher than observed at Facility B. We then compared simulated concentrations under low flow conditions with predicted environmental concentrations from our model (Table S5.9). Despite the differences in modeled discharge for one of the stations, we still observed a good agreement between observed and simulated concentrations. Note that the variability in monthly loading was quite high, ranging from 441g to 855g for Facility A, and 1205g to 2757g (excluding the outliers in August 2005) for Facility B.

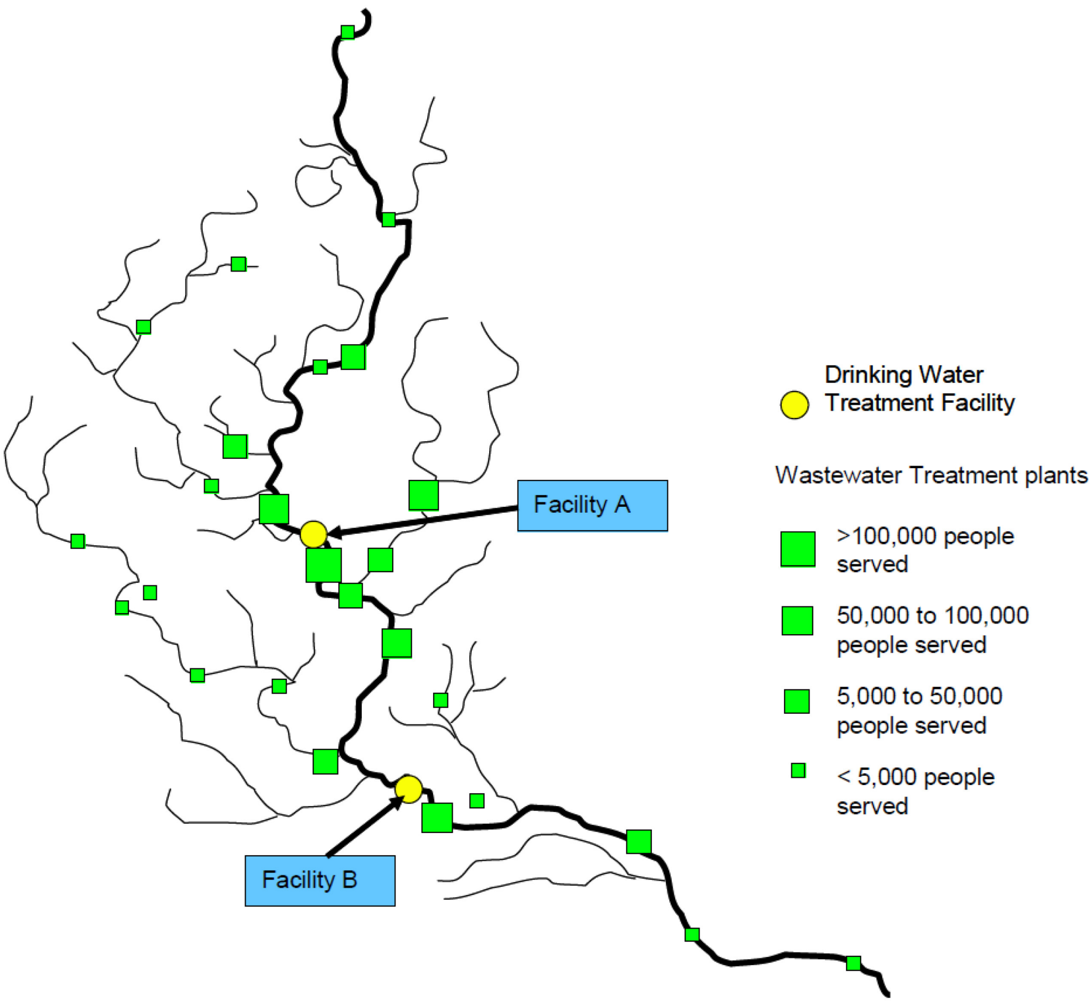


Figure S5.6: Location of measurement stations (Facility A and B). Map from Kormos (2007).

Table S5.8: Raw surface water concentrations of Carbamazepine (ng/L) and mean daily flows (m^3/s) at two drinking water stations at the Grand River, Ontario (Kormos, 2007; see p. 123 and 152). Flow measurements from HYDAT stations; grey shading indicates measurements in low flow periods (as used in Table S5.9). Numbers in parenthesis are considered outliers.

Year	**'05**	**'05**	**'05**	**'05**	**'05**	**'05**	**'05**	**'05**	**'05**	**'06**	**'06**	**'06**
Day/Month	**5/4**	**3/5**	**7/6**	**5/7**	**2/8**	**6/9**	**4/10**	**15/11**	**6/12**	**3/1**	**7/2**	**7/3**
Facility A												
Con. Sample 1	2.8	7.6	25	27	22	16	19	24	11	7.1	3.2	12
Con. Sample 2	2.7	7.1	22	28	20	17	18	22	11	8.5	2.8	12
Observed flow	85	30	15	10	11	10	13	10	25	30	65	15
Facility B												
Con. Sample 1	7.9	14	52	72	(1015)	51	33	42	18	19	8.2	27
Con. Sample 2	7.6	14	53	67	(961)	52	31	43	21	16	7.9	29
Observed flow	140	50	20	15	13	14	15	15	30	48	95	25

Table S5.9: Comparison between observed and simulated long-term flow (m^3/s) and between observed and simulated Carbamazepine concentrations (ng/L) at two drinking water stations at the Grand River, Ontario. The concentration for the "observed Q90-MONTH low flow" was approximated by taking the average of the shaded cells in Table S5.8 (representing months of low flow) across samples of both facilities.

Location	**Average Flow (m^3/s)**		**Q90-MONTH (m^3/s)**		**CBZ conc. (ng/L)**	
	Obs.	Sim.	Obs.	Sim.	Obs.	Sim.
Facility A	37.9	42.3	14.4	11.4	21.6	25.8
Facility B	58.1	91	20.1	25.1	49.6	48

Connecting statement (Ch. 6)

Chapter 5 demonstrated the extension of HydroROUT as a chemical fate model in a major Canadian river basin, and estimated risk for 15 commonly used pharmaceuticals in Canada. In chapter 6, HydroROUT is adapted further to perform at an even larger scale as the first high-resolution chemical fate model applied to modelling the risk of hormones as endocrine disruptors in continental China.

China was chosen due to its pressing water quality problems, but also because current models designed for North America or Europe cannot easily be transferred to China (and other developing or transitional countries) due to data limitations. Current models (including my own adaptation developed in chapter 5) neglect sources from populations not connected to wastewater treatment plants, but these sources release contaminants directly into the environment, which may in fact present greater ecological risk than local releases from wastewater treatment plants which serve larger populations.

In this chapter, I developed a pilot model where the sources of non-connected households are specifically included as a distributed source of contamination. I integrated raster data and developed downscaling methods to estimate contaminant releases from these distributed sources. I then explore the relative importance of distributed (non-point) versus point sources (i.e., wastewater treatment plants), and estimate the combined risk from hormones for Chinese freshwater systems.

6 A high-resolution contaminant fate model for continental China

Günther Grill[1], Jing Li[2], Usman Khan[2], Zhong Yan[3], Bernhard Lehner[1], Jim Nicell[2], Joseph Ariwi[1]

[1] Department of Geography, McGill University, 805 Sherbrooke Street West, H3A 0B9, Montreal, Canada

[2] Department of Civil Engineering & Applied Mechanics, McGill University, 817 Sherbrooke Street West, H3A 0C3, Montreal, Canada

[3] Beijing Municipal Research Institute of Environmental Protection, Beijing, China

In preparation for the journal *Environment International*

Abstract

The contamination of freshwater systems is increasing in many river basins due to industrialization and population growth, posing risks to ecosystems and human health. However, the sources and fate of many chemicals are currently unknown. We developed a first version of a new spatially-explicit contaminant fate model as an extension of the river routing model HydroROUT, a vector-based framework for river routing. HydroROUT operates at very high spatial resolution (~500m), simulates river and stream chemical transport with in-stream removal, and contains links to a set of lakes larger than 1 km^2 which act as a partial sink during the transport. The chemical fate model includes a consumption and release model, considers point-source contributions from wastewater treatment plants, and accumulates contributions of rural and urban populations not connected to sewage treatment plants. As a case study, we modeled the sources and fate of the estrogens estrone (E1), 17β-estradiol (E2), estriol (E3), as well as the synthetic estrogenic steroid hormone 17α-ethinylestradiol (EE2) in Chinese surface water bodies. Preliminary validation of the results revealed that our predictions were within the range of concentrations reported in the literature. Our 'most probable' scenario results in unacceptable risks, i.e. higher than the predicted no effect concentrations in 27% of rivers of China that have an average discharge > 1 m^3/s. Larger rivers also showed substantial risk with 28%, 14%, and 3% of rivers larger than 10, 100, and 1000 m^3/s, respectively, affected by unacceptable levels of hormone concentrations, whereas no risk was predicted in the largest rivers (>10,000m^3/s) of China. Wastewater treatment plants play an important role in water quality control by reducing the risk in substantial portions of the river network, which would otherwise show unacceptable risk. Nevertheless, releases from untreated population dominate by far the overall contribution to risk. Our estimates of contaminant concentrations should be interpreted as incomplete because estrogen sources from animal husbandry are not included in the current model, but are likely substantial sources of these contaminants in parts of China.

6.1 Introduction

Widespread concerns have arisen that environmental quality in China is deteriorating due to increasing urbanization, population growth and industrial, agricultural and economic intensification (Wang et al., 2012; Liu et al., 2013; SEPA, 2013; Varis et al., 2014; Zhou et al., 2014). Extensive contamination of freshwater ecosystems due to releases from municipal and industrial sewage plants, untreated wastewater, and leakage of chemicals from agriculture production and animal husbandry operations has been reported (Shao et al., 2006). An emerging issue that exacerbates freshwater degradation are releases of pharmaceutically active compounds into rivers and lakes (Richardson et al., 2005). The extent to which these sources contribute to overall freshwater degradation is unclear, because the identification of these compounds in the environment is complicated by limited detection capacities of current technology and by the high cost associated with extensive measurement campaigns. In addition to conducting measurements, the modelling of the sources and risks of exposure has become a commonly-used strategy to provide reasonable results at low cost for a variety of compounds (Johnson et al., 2008; Cowan-Ellsberry et al., 2009). Spatially distributed contaminant fate models such as GREAT-ER (Feijtel et al., 1998), ISTREEM (Wang et al., 2000), LF2000-WQX (Williams et al., 2012), P*h*ATE (Anderson et al., 2004), MAPPE (Pistocchi et al., 2012), and GWAVA (Johnson et al., 2013) have been widely used in Europe and North America (Atkinson et al., 2009; Cunningham et al., 2009; Hannah et al., 2009; Ort et al., 2009; Cunningham et al., 2012; Hosseini et al., 2012; Johnson et al., 2013). These models share common assumptions and similar key mechanisms; i.e., reasonable estimates of per capita emissions of the mass of contaminants released into individual wastewater treatment plants (WWTPs) are calculated using the average per capita consumption of a compound of interest and then adjusted for human metabolism, removal of contaminants in wastewater treatment plants are estimated (Keller et al., 2006), and dilution through advection after discharged into natural waters are modeled using stream length, velocity, discharge and a decay function (Pistocchi et al., 2010). Predicted environmental concentrations (PECs) are subsequently based on the accumulated load and the discharge at the river reach scale. These types of models have not been extensively applied in less “data-rich” countries. In particular, China poses a significant challenge to large-scale contaminant fate models because of a large rural population not connected to wastewater treatment plants, the country's large variability of hydrological regimes, and the lack of publicly available datasets regarding the location of WWTPs.

Several studies have attempted to assess the risks associated with chemicals in China using non-spatial screening techniques (see review in Wang et al., 2012). Other studies used spatially-limited models that either focussed on individual watersheds (Liu et al., 2014) or assessed risk on very large spatial scales (Zhang et al., 2014). These models were based on a coarse scale (e.g., entire watersheds) and only estimated risks in large rivers (> $1000 m^3/s$). (Whelan et al., 2011) use a coarse scale (0.5 x 0.5 degree) raster based hydrological and routing model to estimate concentrations of "down-the-drain" products in China based on a global hydrological model. The estimated levels of risk arising from these coarse models are generally low, mainly due to high dilution in these large rivers and environmental decay during the long travel time between headwater sources and downstream locations. However, the distribution of risks in smaller headwater streams, which were ignored in the previous models for China, remains unknown, but could potentially be high because sources in headwaters are generally concentrated releases from small urban and rural congregations.

The identification of risk in a spatially more accurate way in the river system is a prerequisite to the development of strategies to reduce environmental contamination and exposure. Releases from WWTPs are generally considered the major source of pollutants of surface waters. Recent studies took into account the percentage of populations served by WWTPs and account for removal (Zhang et al., 2014). A model to predict steroid inflow and effluent from sewage treatment plants in China was developed by Johnson and Williams (2004) based on detailed population assemblages. Such a model can contribute to better parameterize and validate WWTP emissions into rivers. But to our knowledge, there is currently no risk assessment model that precisely geolocates the WWTPs to the discharge point at individual river reaches in order to account for more localized impacts.

Almost half of the population in China still lives in rural areas (China Population Census Office, 2012) and the vast majority of these populations are not connected to WWTPs (MEP, 2011). Hormone releases into small and medium sized rivers from these non-connected rural populations can be reasonably presumed to be major sources. To address the challenges posed by sources from rural area population and to spatially model the fate of estrogens in China, we used high resolution population density maps to derive spatially-distributed release estimates similar to Hodges et al. (2011). These release estimates were combined with a high-resolution river routing model to accumulate the sources and determine the fate of chemicals and the risks posed by them in the river network downstream.

In order to test whether the model predictions were reasonable, as case study for China, we calculated the environmental exposure and risk of four hormones: estrone (E1), 17β-estradiol (E2), estriol (E3), as well as the synthetic estrogenic steroid hormone 17α-ethinylestradiol (EE2). We selected estrogens because every human releases certain quantities of hormones, which allowed us to research spatially distributed ("diffuse") sources from residents not connected to WWTPs, versus localized ("point") sources from populations connected to WWTPs. Furthermore, hormones are very potent endocrine disruptors (Crisp et al., 1998) and act in an additive way (Caldwell et al., 2012) and thus, may pose risk to the wildlife and humans even at low concentrations. For example, the feminization of fish leading to the collapse of fish populations caused by estrogen pollution has been reported at concentrations as low as a few nanograms per litre (Kidd et al.). Concentrations in excess of predicted no effect concentrations (PNECs) have been reported at several sites in rivers and lakes across China (Chang et al., 2009; Zhao et al., 2009; Chang et al., 2011; Jiang et al., 2012). Jiang et al. (2012) found alarmingly high levels of source water contamination in Chinese rivers, lakes and reservoirs. Yet, the assessment of risk of hormones in rivers and lakes based on contaminant fate models have been mainly conducted in North America and Europe (Hannah et al., 2009; Caldwell et al., 2010; Writer et al., 2010; Anderson et al., 2012) and very limited research has been conducted in China on risk assessment for hormones.

The share of risk contributed by communities not connected to WWTPs is likely to be high in China for two reasons: First, a significant proportion is not connected to any sewage treatment facility, especially in rural areas. Approximately 26% of the population in large cities, and 46% in small and medium cities, lack wastewater treatment, whereas the coverage in urban areas is practically non-existent (MOHURD, 2012). Our own calculations indicate an average of 52% connectivity across urban and rural areas. Second, the removal efficiency of hormones via sewage treatment is generally very high (Caldwell et al., 2012). This creates a situation where even a small number of people that are not connected to wastewater treatment facilities may produce significant risk, which makes the assessment of the small scale contributions of local residents even more important.

Importantly, little work has been done to account for animal sources of estrogens. The main sources of natural and synthetic estrogens are humans and animal husbandry operations (Hanselman et al., 2003) and, to a minor extent, sources from wildlife. An exposure model was created for Shanghai (Liu et al., 2012), which attributed human and livestock contributions of hormones. A comprehensive multimedia model for China was recently developed by Zhang et al. (2014), which included steroid sources from humans, animals, land application of livestock waste and sewage irrigation, with an assessment of concentrations in rivers, sediments and soil for 58 river basins in China. The model estimates predicted high emission densities of steroidal hormones in East China in general alignment with GDP densities. High risks were predicted in 12 of 58 basins, mainly located in Northeastern China. The predicted environmental concentrations (PECs) were within the range of 0.08 - 5.7 ng/L. However, it appears that average flows were used to calculate PECs, which would result in relatively low predicted risks. Furthermore, these concentrations were almost all estimated for rivers with large dilutive capacity (>500m^3/s average flow).

To address the arguments made above, a realistic estimation of risk for smaller rivers must account for sources and pathways of contaminants at high spatial resolution, must take into account the locations, capacity and discharge volumes of WWTPs, must include reasonable estimates of hydrological flow, and must account for populations not connected to wastewater treatment facilities. The first goal of this study is, therefore, to develop a high-resolution spatial contaminant fate model for China, which considers local estimates of stream flow at the river reach scale and takes into account point-sources of WWTPs as well as sources of populations not connected to WWTPs. The second goal of this study is to apply our contaminant fate model to study the spatial distribution and magnitude of estrogen emission sources, predict their fate in the environment, and to determine the environmental risk of these compounds in Chinese surface water bodies. To better understand the spatial distribution of chemical sources, our high-resolution approach is essential to determine the risk from spatially distributed sources in headwaters, which will allow us to better assess the risk for smaller rivers. Due to substantial uncertainty regarding these sources and their fate in the environment, we use several scenarios to understand our model's sensitivity with respect to key parameters. Overall, this study contributes the first comprehensive, high-resolution multi-contaminant fate model for continental China and sets the stage for the application of the model to other areas of the world.

6.2 Methodology

6.2.1 General outline

Based on county- and city-level administrative units, we first calculated per-capita contributions of residents based on the demographics of each administrative unit by separating sources into urban and rural. We then assess pathways either via WWTPs or directly into the environment (Figure 6.1). Using high-resolution river hydrography and a routing model, we accumulate the load downstream in the river system and then account for the decay, transformation and removal occurring in lakes. The four assessed natural and synthetic estrogens are then weighted by their potency and then total concentrations in surface waters are calculated based on low-flow discharge estimates generated from long-term model simulations. PECs are then assessed against the PNEC to assess risk. We ran the model for 11 scenarios (i.e., scenario 1, scenario 2, etc.) with different settings of parameters reflecting WWTP removals, direct discharge coefficients (*ddc*) of untreated wastewater, river decay, and lake removal. A description of the methodology used is given in the sections below with further details provided in the Supplementary Information (SI).

6.2.2 Data sources

1. River network: The study area comprises of continental China. However, a number of rivers in the northern part of the country are leaving China, and enter the country at a later point downstream. These river reaches were also included, to create a complete hydrologically connected network. The baseline hydrographic data used in this study is provided by the HydroSHEDS database (Lehner et al., 2008), a publicly-available global suite of data layers representing river network topology and watershed boundaries (Figure 6.2). HydroSHEDS defines river flow directions at 500 m pixel resolution, which are then used to create a network of river reaches for transport simulation of water and substances in a routing model

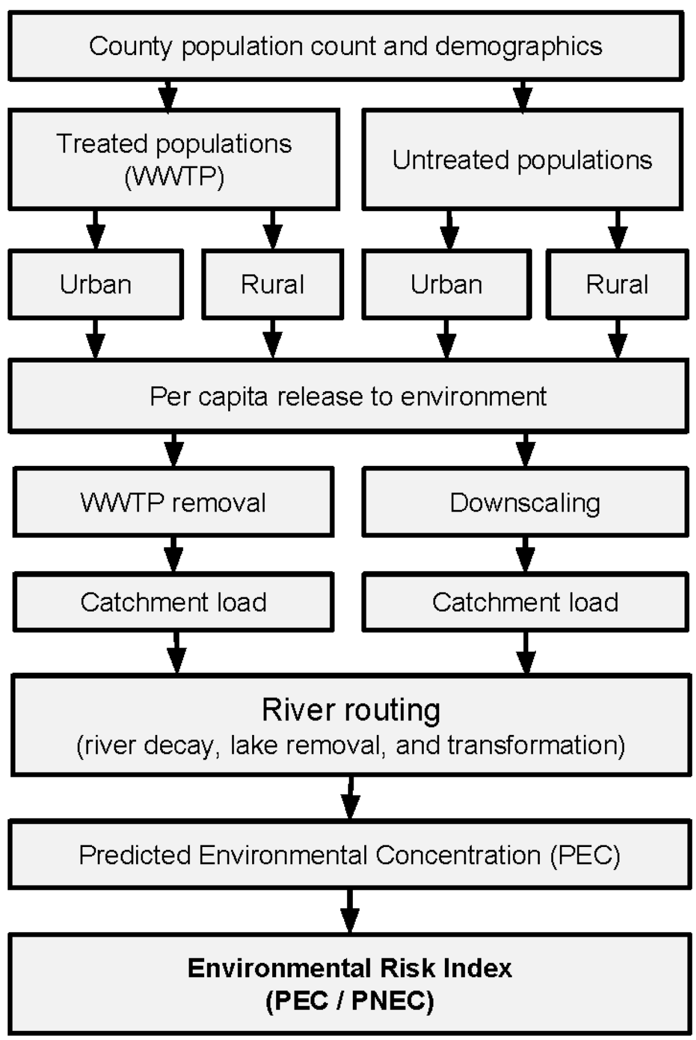


Figure 6.1: Conceptual representation of the contaminant fate model.

called HydroROUT (Lehner and Grill, 2013). At the 500 m resolution, there are more than 1 million river reaches with an average length of 2.8 km within the study area. Each river reach has a contributing catchment, which was used to aggregate fine-scale spatial information. A set of 25,583 lakes with a surface area larger than 1 km^2 (NASA/NGA, 2003) were also integrated into HydroROUT.

2. Discharge data: Decoupled, external runoff estimates, provided by the global integrated water balance model WaterGAP (Alcamo et al., 2003; Döll et al., 2003), are employed. WaterGAP provides runoff estimates of long-term monthly averages for the period of 1961-90 at a 0.5 degree grid resolution. These runoff estimates are spatially downscaled by disaggregating the large grid cells into 500 m

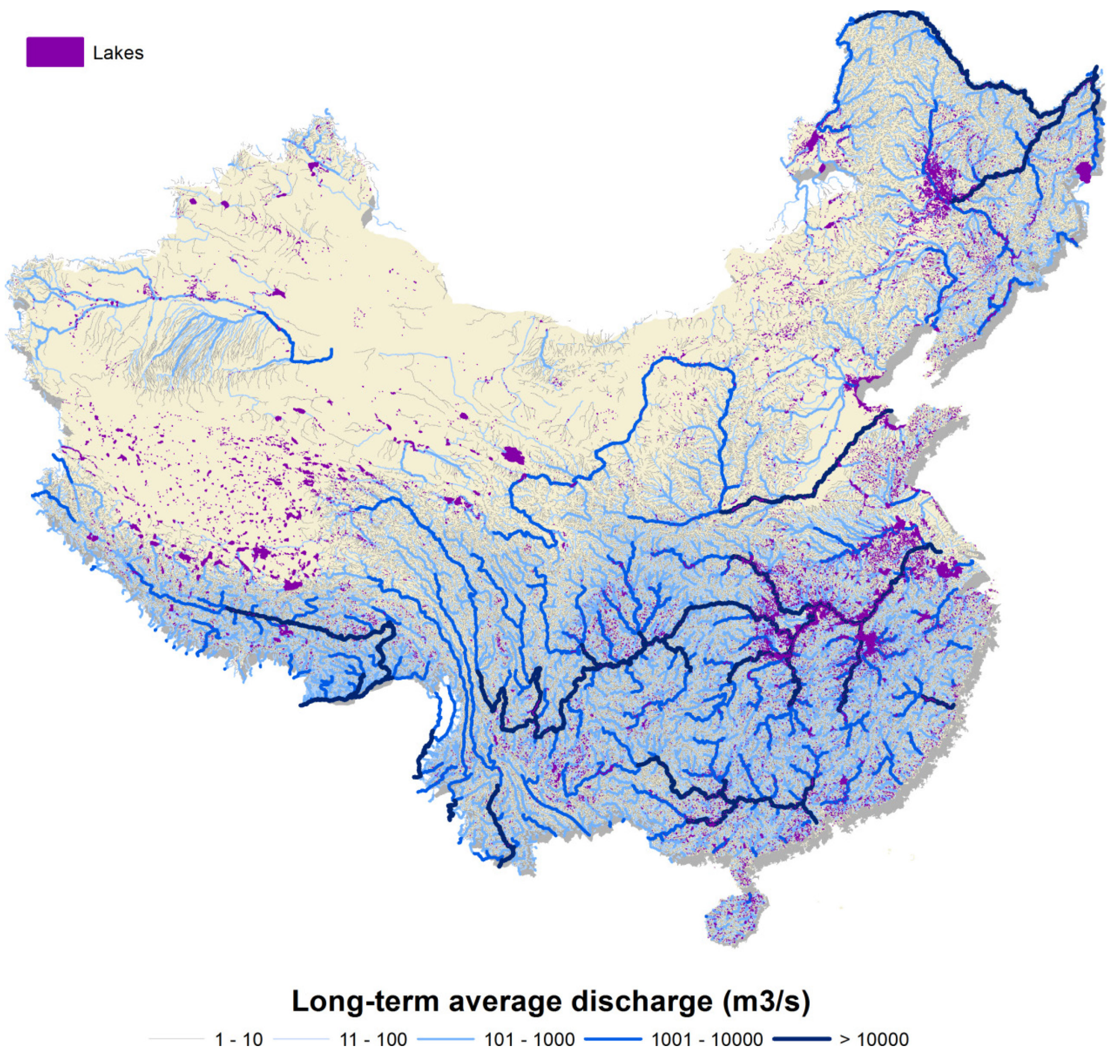


Figure 6.2: River network (Lehner et al., 2008) and long-term average discharge of rivers > 1 m^3/s in China (Alcamo et al., 2003; Döll et al., 2003).

pixels and then accumulating them along the HydroSHEDS river network. Monthly model low flows were used to substitute more traditional low flow indicators, such as daily Q_{90} flows (see chapter 5).

3. Wastewater treatment plants: The locations and characteristics of 2739 WWTPs for China were obtained from official sources in tabular form (MEP, 2011), corresponding to August 2010. A total of 192 WWTPs with flow rates less than 0.001 million m^3/d were excluded (due to difficulties to determine information on 'population served' for small plants), resulting in 2547 WWTPs considered in this study (Figure 6.3). The total designed treatment capacity of all WWTPs is 125.0 million m^3/d with a total average treatment capacity of 96.1 million m^3/d. All WWTPs considered used secondary treatment technologies (i.e., biological treatment). The population served by each WWTP was assigned based on data of the Beijing Municipal Research Institute of Environmental Protection compiled by Zhong Yan (unpublished data). The specific longitudes and latitudes of the WWTPs were determined by using geolocation techniques, which translate addresses into geographic coordinates. Each data point was manually verified for each individual plant and corrected using satellite imagery. Finally, the locations were co-registered to the river network of HydroROUT.

4. Administrative Units (AUs): The calculations of hormone contributions were conducted at the county administrative level. Large cities such as Beijing are composed of several smaller counties, and these were grouped into homogenous "city" units to avoid inconsistencies in calculating treatment coverage in these areas due to the peripheral location of WWTPs in these large conglomerates. We calculated population demographics for a total of 2345 AUs based on the latest China population census conducted in 2010 (China Population Census Office, 2012) (see Figure 6.4). Population density maps (Gaughan et al., 2013) and spatial classifications of rural and urban areas (Schneider et al., 2009, 2010) were used to calculate hormone releases at the river reach scale.

Figure 6.3: Wastewater treatment plants and population served in China (MEP, 2011). Locations are based on geocoding and manual adjustments using satellite images.

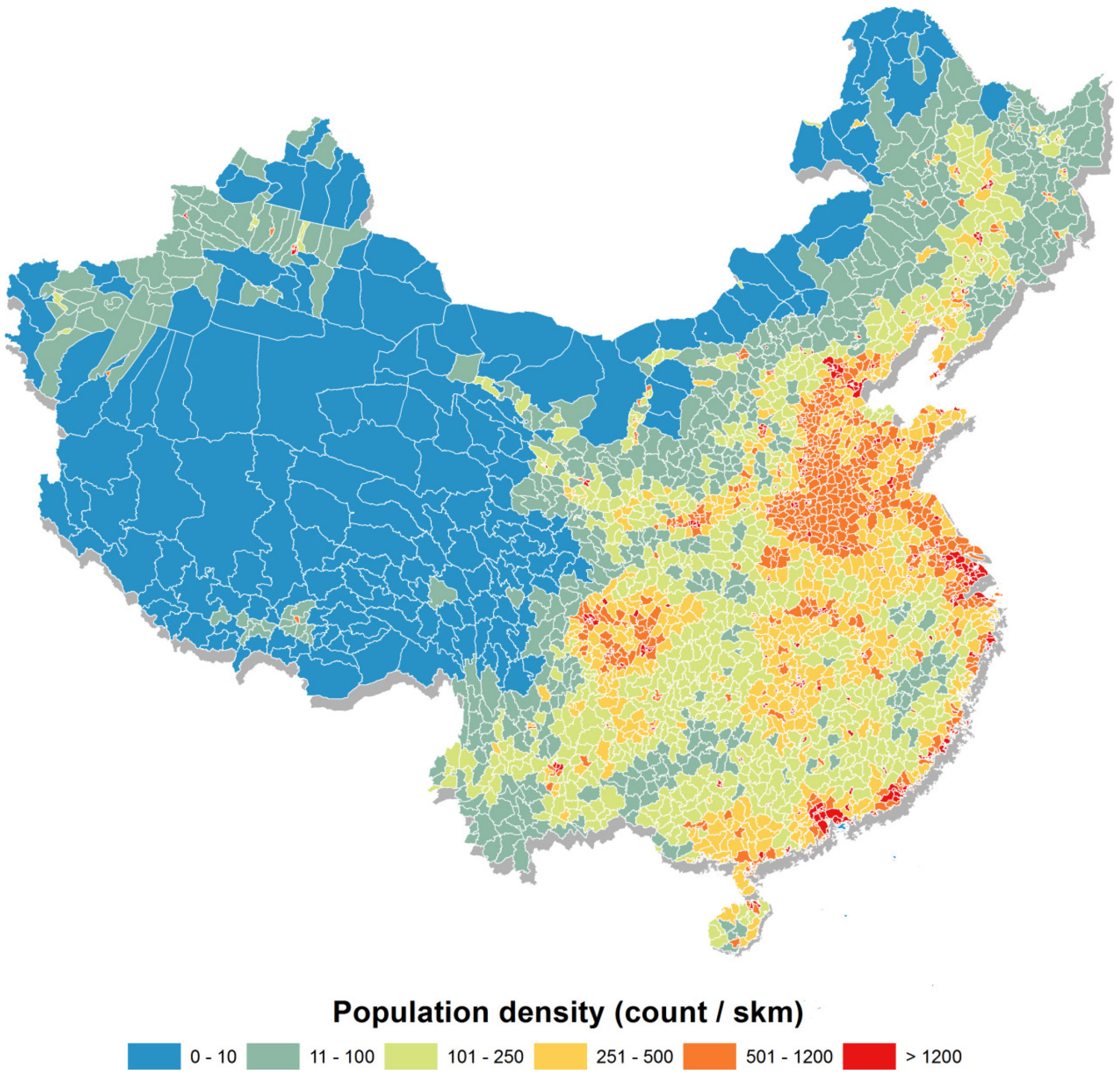


Figure 6.4: County level administrative units and population density in 2010 (China Population Census Office, 2012).

6.2.3 Per capita hormone release

Per-capita estrogen hormone releases (E1, E2, E3, and EE2) were calculated from natural and synthetic sources for each AU, based on the demographic characteristics. Residents of each AU were grouped based on gender and age group, and per capita contributions of natural hormone releases were calculated using average excretion rates (Khan, 2014; Khan and Nicell, 2014). Synthetic hormone contributions from contraceptives and hormone replacement therapy (HRT) were calculated using total national consumption (CMSN, 2012) and adjusted for elimination and metabolic conversion to other forms of estrogen (Khan, 2014; Khan and Nicell, 2014). The total consumption was distributed spatially based on relative consumption rates reported for each province (CMSN, 2012). See section S6.1 in SI for details on these calculations.

6.2.4 Source separation and pathways

Two types of loading were incorporated into the model: (1) hormone sources from populations connected to a WWTP; and (2) populations not connected to any WWTP. We first determined the population count of treated versus untreated populations in each AU based on the sum of populations treated by all WWTPs versus the total population. We then calculated loadings from treated and untreated populations in each AU separately for rural and urban populations, based on land use classifications (Schneider et al., 2009, 2010).

We calculated the loading from rural and urban populations connected to WWTPs using information about the capacity associated with each treatment plant (see S6.2) and the per capita hormone excretion rates calculated previously (see 0). After removal in the WWTP, the load was linked to the river reach or lake into which the plant discharges its effluent.

The contribution of estrogens from the population not connected to WWTPs were calculated by taking the estimated total hormone contributions of the entire population minus the contributions estimated for population connected to WWTPs. We made the assumption that not all human releases enter directly into surface water bodies; however, the factors affecting this proportion are currently not well understood. We therefore incorporated a new variable into the model representing the proportion of hormone load from untreated wastewater reaching the water body, termed *direct discharge coefficient (ddc).* In our 'most probable' scenario, we parameterized *ddc* as 40% for untreated wastewater discharged into nearby rivers in rural areas as reported by Wang et al. (2011) and used a higher *ddc* of 80% for urban areas (Dr. Wang Kaijun, School of Environment, Tsinghua University, personal communication) due to the presence of impervious surfaces. We conducted modelling experiments with a range of other *ddc* values to evaluate the sensitivity of model outputs to this variable.

We used a population density map from WorldPop (Gaughan et al., 2013)—aggregated to match our model's 500 m spatial resolution—to distribute the contributions from untreated populations spatially according to population density in the AUs - separately for rural and urban areas. Rural and urban areas were classified according to data of a MODIS-based land use classification (Schneider et al., 2009, 2010). The downscaling procedure essentially resulted in a high-resolution surface of hormone input (see S6.2 for details).

6.2.5 River and lake routing model

We used an extension of the global routing model HydroROUT (Lehner and Grill, 2013) for mass balance calculations in the river and lake network. For river reaches, we followed a 'plug-flow' approach (Pistocchi et al., 2010), i.e., a 'plug' of substance mass (the amount of contaminant released from the treatment plant) is accumulated downstream as the sum of the input from the current and all upstream reaches flowing into the current reach. We currently model lakes as either passing the load fully to the downstream river or as eliminating the load entirely (Ort et al., 2009), depending on user input. In future model iterations, once the necessary lake volumes have been calculated, a more appropriate method based on a completely stirred reactor model will be incorporated. The river network is processed iteratively in the hydrological order from source to sink (see S6.3 for details). We also accounted for the assumption that an estimated 50% of E2 is converted into E1 during treatment or in the environment (Zhang et al., 2014).

6.2.6 Environmental risk assessment

After accumulation and degradation, we calculated PECs per river reach by dividing the accumulated chemical mass by the discharge of the reach (see also section S6.4). Due to the seasonal variability of river flows, we used the Q90-month, which is a low flow indicator equivalent to the daily 90th percentile flow; i.e. a flow that is exceeded at 90% of the time and, which represents our 'worst-case' scenario with minimal dilution and maximal concentrations.

Since estrogens act in an additive manner (Thorpe et al., 2001) and the ecotoxicological potency of all estrogens is not equal (SCHER, 2011; Anderson et al., 2012; Caldwell et al., 2012), the total amount of estrogen released from residents was expressed in terms of E2 equivalents, the most potent natural estrogen. Specifically, EE2, E1 and E3 were assumed to be 10, 0.33 and 0.04 times as potent as E2, respectively (Khan and Nicell, 2014). After the weighting, we calculated risk quotients (RQs), based on the ration of PECs to PNEC for all estrogens. Combined PNEC was assumed to be 1 ng of E2-eq/L (Williams et al., 2009).

6.3 Results

6.3.1 Total emissions

The total excretion of hormones from residents in China was estimated at 29,548 kg/yr. This apportions to 6033, 3018, 20454, and 43 kg/yr of E1, E2, E3, and EE2, respectively, as shown in Table 6.1. A total of 11,115 kg/yr was discharged into rivers and lakes, partitioned as 3267, 582, 7249, and 16 kg/yr of E1, E2, E3, and EE2 respectively. This represents a total of 62.4% estrogen removal from WWTPs and includes the reduced proportion of the load from untreated wastewater reaching the water body (*ddc*) from urban and rural populations who are not connected to WWTPs. Figure 6.5 shows the emission density of estrogens per square kilometer. Spatial patterns of estrogen emission densities are generally consistent with the spatial pattern of population density shown in Figure 6.4. After routing, first order decay processes and lake removal eliminated another 33.7% of the total estrogen excretions, leaving 3.9% of the total estrogen excretions into the environment being exported to sinks (i.e., either to the sea or to endorheic sinks), or to neighboring countries, with total of 12.6, 4.2, 1.3, and 19.7% of E1, E2, E3, and EE2 being exported.

Table 6.1: Sources and fate of estrogen compounds in China (Data based on CMSN (2012) and own calculations*).

Process	**E1 (kg)**	**%**	**E2 (kg)**	**%**	**E3 (kg)**	**%**	**EE2 (kg)**	**%**	**Total (kg)**	**%**
Total Excretion	6033	100	3018	100	20,454	100	43	100	29,548	100
Removal in WWTPs or not directly discharged	2766	45.8	2436	80.7	13,205	64.6	27	62.8	18,433	62.4
Discharge into rivers and lakes	3267	54.2	582	19.3	7249	35.4	16	37.2	11,115	37.6
Decay and removal in rivers and lakes	2510	41.6	456	15.1	6979	34.1	7.5	17.4	9952	33.7
Export to sinks or neighboring countries	757	12.6	127	4.2	270	1.3	8.5	19.7	1163	3.9

*Totals may not match or account to 100% due to rounding.

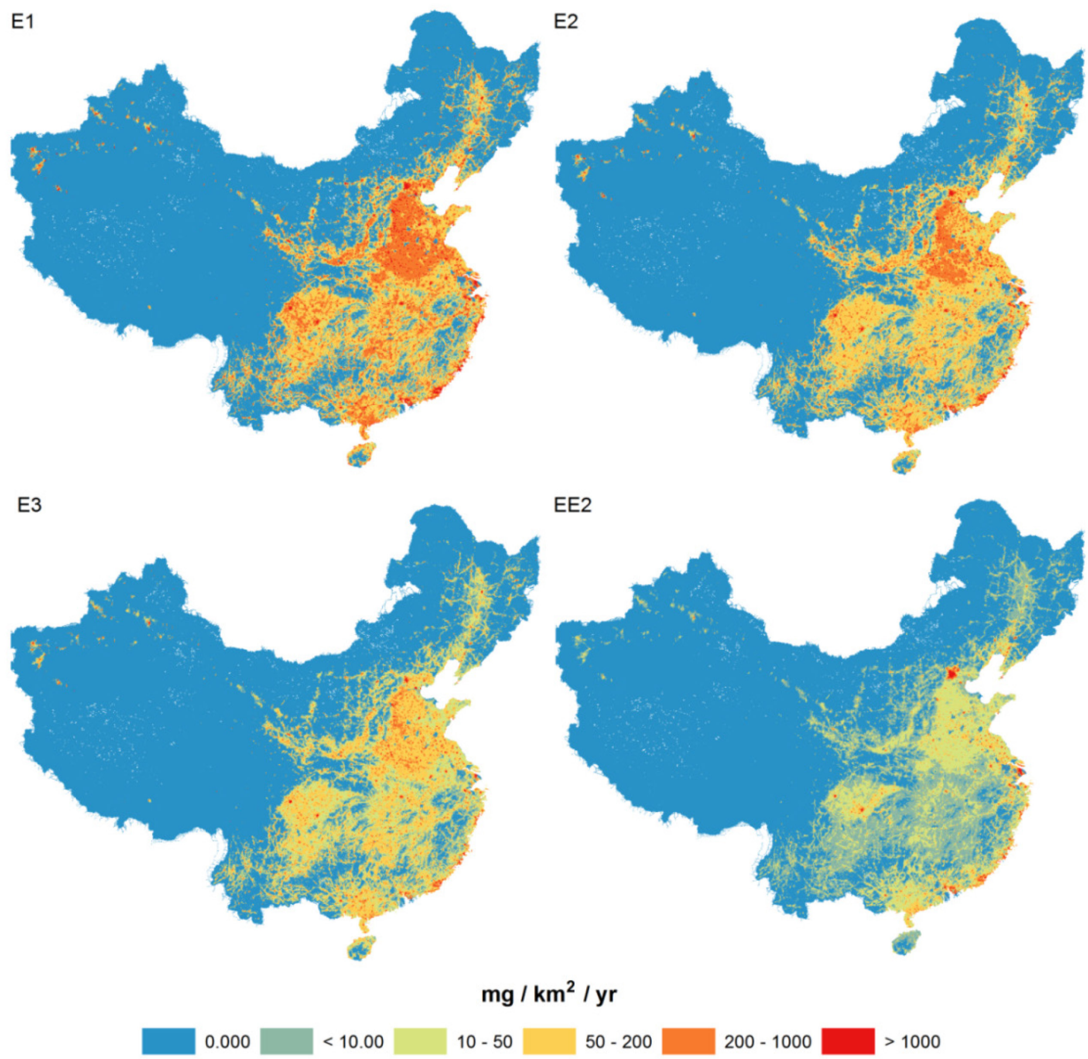


Figure 6.5: Potency adjusted emission density (reach catchment scale) for simulated estrogens (after treatment in WWTP and accounted for direct discharge coefficient *ddc*; before environmental decay and lake removal). EE2, E1 and E3 were assumed to be 10, 0.33 and 0.04 times as potent as E2, respectively (Khan and Nicell, 2014).

6.3.2 Environmental risk and sensitivity

We detected a high sensitivity with respect to the contributions of rural populations unserved by WWTPs, combined with a high parameter uncertainty. We therefore ran the model for 11 scenarios (i.e., scenario 1, scenario 2, etc.) with different settings of parameters reflecting WWTP removals, direct discharge coefficients (*ddc*) of untreated wastewater, river decay, and lake removal as illustrated in Table 6.2. The scenario that was currently deemed as most probable (scenario 7 in Table 6.2) based on parameters from a literature review resulted in a predicted substantial risk (PEC > PNEC) for widespread parts of China.

The resultant spatial patterns of risk—based on potency-adjusted estrogen concentrations assessment and expressed as risk quotient (RQ)—are shown in Figure 6.6. Risk patterns loosely follow the population density patterns (as shown in Figure 6.4), modified by regional variations of precipitation and discharge and by local flow magnitude. Large parts of the North China Plains appear with very high concentrations, due to low river flow and high population density.

Overall, this scenario indicates risk for 19.8% of all rivers (considering river length) in China. By examining different river sizes, we detect more risk if the smallest rivers ($0.1 m^3/s$-$1 m^3/s$) are no longer accounted for, with 27.1% of rivers >1 m^3/s and 28.3% of rivers >10 m^3/s at risk. This may be due to the fact that numerous small streams do not receive releases according to our population

Table 6.2: Percentage of river length at risk for different model scenarios by river class. River classes are defined by discharge size thresholds, for example the river class '>1000' also includes the rivers in the next larger category, '> 10,000'.

Scenario Number	Description	WWTP removal	ddc % (urb)	ddc % (rur)	Lake rem.	River decay	Percent of river length at risk (%) by river class (m^3/s)					
							>0.1 (complete network)	>1	>10	>100	>1000	>10,000
1	Downstream of WWTP only	Yes	0	0	Yes	Yes	1.3	4.7	4.8	0.8	0	0
2	low ddc	Yes	20	20	Yes	Yes	12.8	16.2	15.4	4.7	0	0
3	med. ddc	Yes	40	40	Yes	Yes	19.5	25.7	25.4	11.4	2.1	0
4	high ddc	Yes	60	60	Yes	Yes	23.3	32	32.5	17.6	4.4	0
5	v. high ddc	Yes	80	80	Yes	Yes	25.4	36.4	37.8	23.5	7.6	0
6	full ddc	Yes	100	100	Yes	Yes	26.9	39.6	41.7	28.4	10.4	0
7 ('most probable')	**variable ddc**	**Yes**	**80**	**40**	**Yes**	**Yes**	**19.8**	**27.1**	**28.3**	**14.1**	**2.9**	**0**
8	high ddc, no riv. rem.; no lake rem.	Yes	80	40	No	No	22.4	38.3	50.8	54.8	54.2	87.3
9	S8: high ddc, no lake rem.	Yes	80	40	No	Yes	20.9	32.5	38.3	28.2	17.2	0
10	high ddc, no river rem.	Yes	80	40	Yes	No	21.0	31.0	35.8	26.5	16.8	13.7
11	high ddc, no WWTP removal	No	80	40	Yes	Yes	20.1	28.6	32.8	22.7	8.3	0

distribution map, and due to the observation that WWTPs are typically located at larger rivers. There were also numerous small river reaches that fall dry for one or more months per year, as indicated by the color grey in Figure 6.6, in which case the low flow values become difficult to derive with our discharge model. In these cases, we assumed no river flow and assigned zero concentration. If only medium to large sized rivers (>100 m^3/s) are assessed, 13.9% of rivers showed risk, whereas, only 2.9% of large to very large rivers (>1000 m^3/s) showed risk. Very large rivers (>10,000 m^3/s) did not show any risk due to their high dilution capabilities.

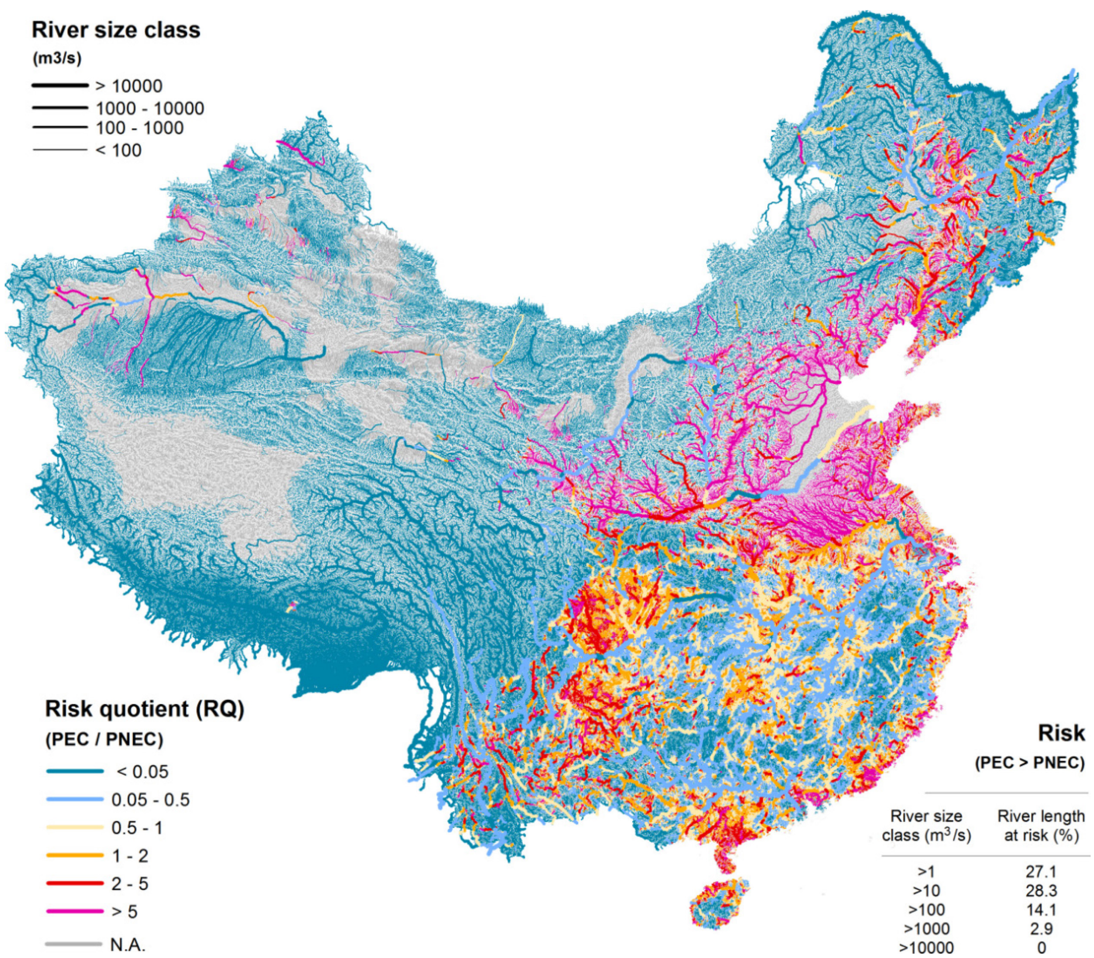


Figure 6.6: Risk quotient (RQ) based on the predicted environmental concentration (PEC) to predicted no-effect concentration (PNEC; assumed 1 ng/L based on Williams et al. (2009)) for each river reach. The colors orange, red and purple indicate RQ above 1. The small table at the lower right indicates the percentage of river reaches in each size class for which the combined estrogen concentrations exceed PNEC. All calculations were conducted under low flow conditions for the scenario 7 of Table 6.2. Rivers marked grey showed zero flow over one or several months and could not be assessed due to lack of data.

WWTPs alone produce risk in 4.8% of rivers > 10 m^3/s, as illustrated in scenario 1 and scenario 11 in Table 6.2. This highlights the large contribution of populations who are not served with WWPTs, even if *ddc* is set low (i.e., 20% for both urban and rural areas) as demonstrated in scenario 2. However, removal by WWTPs is increasingly important for larger rivers. When comparing scenario 11 to scenario 7, risk without WWTPs removal increased by 4.5% and 8.6% in medium (>10 m^3/s) and large rivers (> 100 m^3/s) respectively, and affects 5.4% more of the very large rivers (> 1000 m^3/s). This indicates that WWTPs play an important role in water quality control of increasingly large rivers. Without WWTP removal, the risk in many river reaches would be increased to unacceptable levels (PEC>PNEC). Nevertheless, releases from untreated population dominate, by far the overall contribution to risk.

The removal of chemical load due to processes in lakes and rivers is very important. A comparison of scenario 7 and scenario 8 shows that both lake and in-stream removal reduced the risk in all river classes substantially. Lake removal by itself has a very substantial influence on results, as by the results of scenario 10. Another important observation is that lake removal becomes increasingly important for larger rivers: if the model passes the load directly through lakes (without any removal), as shown in scenario 9, the risk is twice as high for rivers > 100 m^3/s and 5 times as high for very large rivers compared to the scenario with lake removal (i.e., scenario 7).

A similarly large effect is triggered by the in-stream removal in rivers, which substantially reduces risk compared to no in-stream removal (i.e., scenario 10). It appears that in-stream removal processes are largely responsible for reducing risk in large and very large rivers, possibly due to the longer travel time with opportunities for environmental decay.

6.3.3 Validation

Modelling results were validated by comparing predicted estrogen concentrations with available measurements in selected river reaches, as shown in Table 3. The results revealed our predictions to be within the range of concentrations reported in the literature (Chang et al., 2009; Zhao et al., 2009; Chang et al., 2011; Jiang et al., 2012). Some explorations for individual sites at the Yangtze River

(Table 6.3) showed that our predictions were within an order of magnitude (or better) from observations, except for one location (S10). However, a more extensive set of point-by-point comparisons have not been conducted yet. These comparisons are generally complicated by the lack of available data on stream flow quantities at the time when the concentrations in the rivers were measured. Some more observational data is available (Chang et al., 2009; Zhao et al., 2009; Chang et al., 2011; Jiang et al., 2012) but the studies did not include coordinates for the measurements, making the matching between site and river network difficult, or lack discharge data at the time of measurement. Additional validation including the application of other validation techniques, such as cumulative frequency plots, requires extensive efforts that are beyond the scope of the present study.

Table 6.3: Observed (Jiang et al., 2012) versus predicted PEC (this study) of estrogens in the Yangtze River. From the 23 locations examined in Jiang et al. (2012), only 16 locations provided measurements. Out of these, nine measurements were taken in a lake or reservoir, and were thus not considered.

	E1 (ng/L)			**E2 (ng/L)**			**E3 (ng/L)**		
Site	**obs.**	**pred.**	**factor diff.**	**obs.**	**pred.**	**factor diff.**	**obs.**	**pred.**	**factor diff.**
S6	1.53	1.67	1.1	0.25	0.32	1.3	n.a.	n.a.	n.a.
S7	0.96	2.0	2.1	0.34	0.38	1.1	n.a.	n.a.	n.a.
S8	0.87	0.5	1.7	0.31	0.1	3.1	n.a.	n.a.	n.a.
S9	1.08	10.05	9.3	0.55	1.5	2.7	4.4	2.78	1.6
S10	1.93	0.15	12.9	0.71	0.02	35.5	3.9	0.014	281.4
S11	2.37	10.09	4.3	0.58	1.53	2.6	4.2	3.42	1.2
S14	2.98	6.34	2.1	1.51	0.75	2.0	2.6	0.38	6.8

6.4 Discussion

A high-resolution contaminant fate model was developed for China as a basis for assessing risk due to contaminants in rivers of China. The model was applied to the case of estrogens, was tested for its sensitivity to model parameter values, and was validated in a preliminary way against available measurement data. This model is of an exploratory nature and, thus, a number of further assessments and improvements should be conducted before greater confidence in the results can be achieved. In particular, the issues discussed below point to the need for further exploration and refinement of the model.

First, we identified a high sensitivity of model output to the direct discharge coefficient (*ddc*), which reflects the contaminant load discharged directly into rivers. Research on this issue is sparse and *ddc* estimates rely on sparsely available evidence. Chang et al. (2009) attributed almost 63% of the hormone loads in 45 Chinese rivers to discharges of untreated wastewater. In rural areas, this proportion could be lower; for example, Wang et al. (2011) report 47% of untreated wastewater is discharged into trenches or ditches directly, whereas, 7% is discharged into nearby rivers. In urban areas, the proportion of untreated wastewater reaching a water body is deemed even higher due to a larger proportion of impervious surfaces and can reach levels of up to 80% (Dr. Wang Kaijun, School of Environment, Tsinghua University, personal communication). Although *ddc* is larger in urban areas, our model sensitivity was much higher for rural populations, because on the one hand, a low proportion of the population is connected to wastewater treatment facilities, and on the other hand the proportion of untreated urban populations is generally much lower than for rural populations. Furthermore, this issue varies spatially, related to hydrological stream density and flow quantities. Particularly large areas of the North China Plains, where our model indicates increased risk, is mostly characterized by “artificial” hydrology, consisting of a network of canals with the aim to control and retain as much water as possible. This is in contrast to North American and European landscapes, where the goal is often to drain the landscape from excess water in case of flooding (Ongley et al., 2010). As a consequence, the proportion of wastewater reaching a significant surface water body and routed downstream may be exaggerated in our model and could lead to overestimations of risk in these areas.

Our model includes high-resolution pathways of estrogens from residents to the river system previously not considered by other models at this spatial scale. There is a high sensitivity associated with releases of contaminants and routing in these headwater streams. For example the range of concentrations quickly rises when smaller rivers are assessed given that extremely high concentrations of estrogens are observed in many small river reaches. Uncertainties with respect to the underlying hydrography and hydrology in small river reaches contribute to uncertainty of in-river concentrations and may explain those high concentrations. A thorough validation of the hydrological model could not yet be performed for Chinese river courses; however, our hydrological data has been validated in previous studies (see chapter 5) with reasonable success. Since the hydrological data stems from the same global model, we assume comparable accuracy and uncertainty for China; however, a region-specific validation should be conducted before future implementations of this model.

Population density maps (PDMs) can greatly affect the overall risk, especially for smaller rivers. PDMs are modeled densities based either on nighttime light distribution or on modeled spatial distributions based on roads, river networks and major population centers. Population data stems from administrative level population statistics. The level of detail of these administrative units can play a role for accurately distributing the population spatially. To test this, we conducted additional sensitivity analysis with three different population distribution maps. The Global Population of the World data product (GPW; CIESIN, 2005) appears to distribute populations rather evenly across rural areas, and this triggers risk is numerous very small headwater stream. The population distribution of LandScan (LandScan, 2006), on the other hand, appears to be more unevenly distributed, resulting in higher spatial variation. Compared to GPW, this causes some catchments to show very high population counts, but also results in a large amount of basins with very little contributions from residents. The third dataset termed WorldPop (Gaughan et al., 2013) used advanced spatial regression models, which included a variety of data sources such as street-level geodata. The spatial distribution appeared to be more realistic than for the others. Three test runs with the different data sets produced significant differences in risk—GPW produced the most risk, followed by Landscan, and finally WorldPOP. However, due to the notable sensitivity, a more thorough comparison in the future is warranted.

A limitation of the model is that contributions of animal husbandry operations are not currently accounted for because their locations and loads are difficult to determine, which is a common problem when assessing risk from estrogens using CFMs (Caldwell et al., 2010; Anderson et al., 2012). Animals in livestock farms excrete several orders of magnitudes higher levels of estrogen per head than humans (Zhang et al., 2014) and can contribute to substantial hormone activity in rivers (Matthiessen et al., 2006; Burkholder et al., 2007; Chen et al., 2010). Depending on the ratio of animals to humans and depending on the mechanisms of how these sources are released into the environment, the contribution could vary substantially over space (Wint and Robinson, 2007); for example, animal husbandry operations are spatially concentrated in the mid-west of the United States. In China, an average of 62% of E1, E2, and E3 may stem from animals (Zhang et al., 2014). Due to the lack of data to account for these sources, our model results should be interpreted as incomplete, and the real risk is likely substantially higher. Comprehensive data should be collected to include these sources in future studies.

Due to limitations in the underlying global hydrological runoff model WaterGAP temporal resolution, we proposed Q90-Month (i.e. the lowest long-term average monthly flow) as a substitute for the commonly used daily Q90 low flow index. There is generally very good correspondence exists between the two indicators (see chapter 5 of this thesis), yet Q90-Month tends to be systematically overestimated compared to Q90, in particular for very small streams. This means, with respect to fate modelling, that low flow assessments based on Q90-Month are likely to underestimate substance concentrations and contamination risk as compared to analyses using daily Q90. Nevertheless, in terms of contamination modelling this range of error is still tolerable for general screening and risk assessment analysis, assuming the goal of such assessments is to derive environmental concentrations within a factor of 10 (USEPA, 1996; Anderson et al., 2004).

The representation of lakes in our model is currently simplified. If we would model lakes as tank reactors or tank-in-series reactors (Anderson et al., 2004), we can expect partial removal, resulting in risk percentages falling between the results shown in scenario 7 and scenario 9. However, the extent to which this increases risk downstream is currently unknown.

Despite the current limitations of the model, and pending further improvements and validation, we anticipate a variety of future application of this contaminant fate model. Contributions from livestock per county to model the additional risk from animal feeding operations in China should be included in this model to more comprehensively model the risk from estrogens. Following this, other contaminants from animal husbandry operations, such as antibiotics, nitrogen and phosphorus could be modeled as well. There are a number of substances for which the magnitude of their release to the environment has been linked to the purchasing power of the population such as GDP (Hodges et al., 2011). We thus suggest using county-level GDP distributions to generate a spatially explicit general indicator of risk from such substances, or to utilize GDP to spatially downscale the use of certain chemicals.

6.5 Conclusion

In this work, the first high-resolution contaminant fate model for China has been developed with the goal of providing a tool for assessing sources, fate, and environmental risk of a variety of chemicals. The model operates at an unprecedented spatial resolution of 500m for continental China and estimates emissions from human populations using even higher resolution population density maps and from co-registered wastewater treatment plants, which were previously not considered as georeferenced point sources. An important source of dilution, lakes and reservoirs are explicitly represented as well. As a test case, we explored the fate of hormones in freshwater systems, in an environment driven by high hydrological variability and large contaminant contributions from populations not connected to wastewater treatment plants. Under low flow conditions, our model was used to predict high-resolution patterns of environmental concentrations. The results indicate that predicted environmental concentrations are expected to exceed the predicted no-effect concentration in approximately one quarter of all rivers larger than 1 m^3/s when accounting for river length. We found that wastewater treatment plants contribute substantially to risks by themselves, but releases from untreated rural populations in headwaters dominated by far the contribution to overall risk. Regarding overall performance, a preliminary validation indicated good agreement between predicted spatial patterns and measured local contaminant concentrations. Despite these promising results, the extent to which our model can produce accurate predictions needs further investigation. When specifically considering the case of China, uncertainty around the direct discharge of untreated wastewater from populations should be reduced and further validation of the hydrological database is necessary, with a specific focus on China. Nevertheless, once these improvements have been made and greater confidence in the model results can be gained, the model could support environmental risk assessment for a variety of chemicals in China or elsewhere.

Supplementary Information

S6.1 Calculation of natural and synthetic estrogen contributions

Hormone releases from age and gender groups (see Table S6.1) were calculated for each administrative unit (AU) based on its population demographics.

The total amount of synthetic estrogen, 17β-estradiol (E2), estrone (E1), and ethinylestradiol (EE2) used in 2010 in China was 439, 2177, and 114 kg respectively (CMSN, 2012). Furthermore, we accounted for the fact that E1 and E2 are transformed into other forms of estrogen. A total of 37.8% of consumed EE2 is eliminated through excretion (see Table 6.2).

We were unable to find consumption data at the county level and instead used provincial level data provided by CMSN (2012) regarding the total consumption of contraception and hormone replacement therapy (HRT) pills. Consumption data was available for the entire country and for the provinces of Beijing, Shanghai and Chongqing, Guangdong, Jiangsu, Zhejiang, Liaoning, Fujian, Shandong, Guangxi, and Henan (Table 6.3). These AUs combine 81.6% and 92.8% of contraceptive and HRT pill use of China, respectively. The levels of consumption for the missing provinces were estimated by distributing the proportion of pill usage unaccounted for to the remaining provinces relative to the provinces' population count.

Table S6.1: Natural steroid excretion per person (Khan, 2014; Khan and Nicell, 2014).

		Excretion (μg/d)		
Population group		Estrone (E1)	17β-estradiol (E2)	Estriol (E3)
female	prepubescent	0.08	0.06	0.2
	cycling-menstrual	8.69	4.42	9.40
	menopausal	1.7	0.8	1.9
	pregnant	248	126	1337
male	prepubescent	0.08	0.06	0.2
	adult male	3.6	2.13	2.4

Table S6.2: Use, excretion and conversion of synthetic forms of estrogens.

Estro-gen	Use in kg/yr (CMSN, 2012)	Conversion to other estrogens (%)	Reference	excreted load (kg/yr)			
				EE2	E1	E2	E3
EE2	114	37.8 to EE2	(Khan and Nicell, 2014)	43			
E2	439	25.0 to E1 8.5 to E2 6.0 to E3	(Khan, 2014)		110	37	22
E1	2177	11.0 to E1 2.0 to E2	(Khan, 2014)		240	44	
Total				43	350	81	22

Table S6.3: Distribution of contraceptive and hormone replacement therapy (HRT) pill use among the provinces and municipalities in China (CMSN, 2012).

Province or municipality	Contraceptive pill use (%)	HRT pill use (%)
Anhui	1.60	2.30
Beijing	18.2	3.90
Chongqing	4.37	3.90
Fujian	4.29	2.80
Gansu*	0.69	0.50
Guangdong	14.66	7.10
Guangxi*	3.43	0.92
Guizhou*	1.00	0.72
Hainan*	0.23	0.16
Hebei	1.84	3.00
Heilongjiang*	1.00	0.73
Henan	3.00	2.90
Hubei	1.49	4.50
Hunan	1.67	2.00
Inner Mongolia*	0.63	0.46
Jiangsu	7.93	11.1
Jiangxi*	1.16	0.84
Jilin	0.72	2.40
Liaoning	5.50	4.70
Ningxia*	0.16	0.12
Qinghai*	0.15	0.11
Shaanxi*	0.99	0.72
Shandong	4.49	8.80
Shanghai	10.25	7.60
Shanxi*	0.90	0.65
Sichuan	2.14	2.80
Tianjin	0.32	3.40
Tibet	n.a	n.a
Xinjiang*	0.55	0.40
Yunnan*	1.19	0.87
Zhejiang	5.45	19.60

* estimated

S6.2 Source separation and pathways to environment

Contributions of populations connected to wastewater treatment plants

We assume that WWTPs primarily serve urban populations. The total of treated population is therefore subtracted from the total urban population first, irrespective of the ratio between urban and rural population:

$$d = Pop_{urb} - Pop_{treated} \tag{6.1}$$

Depending on the result, we proceed with calculations as follows:

(a) If d is a positive number, d is the estimate of the number of urban people not served by a WWTP ($d = Pop_{urb,untreated}$). It also follows that the urban population served by treatment systems is equal to the total population served by treatment ($Pop_{urb,treated} = Pop_{treated}$). In this case, none of the rural population is served by WWTPs ($Pop_{rur,treated} = 0$) and the rural population not served by WWTPs is equals to the total rural population ($Pop_{rur,untreated} = Pop_{rur}$).

b) If d is zero, all urban population is served by WWTPs and none of the rural population is.

c) If d is a negative number, all urban people are served by treatment facilities

$$Pop_{urb,treated} = Pop_{urb} | \, Pop_{urb,untreated} = 0 \tag{6.2}$$

and the absolute value of d is the estimated number of rural people provided access to treatment in addition to the urban served.

$$Pop_{rur,treated} = |d| \tag{6.3}$$

The rural population not served by treatment is equal to the total rural population minus the rural population served by treatment.

The individual contributions from any given WWTP is the sum of the population served $Pop_{urb,treated} + Pop_{rur,treated}$ multiplied with the daily capita excretion rate ($ex_{cap/day}$; see Table 6.1) multiplied by a removal factor depending on the type of treatment. There were no primary (i.e., physical/chemical treatment) WWTPs found in our study area—only secondary WWTPs (i.e., biological treatment). We assumed average removal rates (rem_{type}) of 66.8%, 85%, 97%, and 84% for E1, E2, E3, EE2 based on reported literature values (Caldwell et al., 2010). The total input of hormones from wastewater treatment plants per river reach (Ld_{reach}) is the sum of the contributions from all WWTPs releasing waster into that reach:

$$Ld_{reach,treated} = \sum_{WWTP} \left(\left(Pop_{urb,treated} + Pop_{rur,treated} \right) \times ex_{\frac{cap}{day}} \times rem_{type} \right) \qquad 6.4$$

Contributions of populations not connected to wastewater treatment and spatial downscaling procedure

We upscaled the WorldPop population density raster (Gaughan et al., 2013) from its native 100 m resolution to our 500 m pixel resolution and created a per-pixel population ratio grid:

$$ratio_i = \left(\frac{Pop_i}{Pop_{Admin}} \right) \qquad 6.5$$

where $ratio_i$ is the fraction of the total population Pop_{Admin} of the AU in each grid cell i, and Pop_i is the population in that grid cell. This ratio allowed us to spatially distribute and downscale estrogen contributions from populations calculated at the AU level to a finer (500 m × 500 m) resolution. We separated urban and rural ratios based on land use classifications (Schneider et al., 2009, 2010). Note that this method may underestimate the contributions from current urban populations whereas rural population may be overestimated because a population shift into urban environments has recently occurred in most counties (Hodges et al., 2011). The loading of an individual pixel within the AU in an area is then calculated as:

$$Export_i = Pop_{Admin,untreated} \times lF(S)_{cap/day} \times ddc \times ratio_i \times 365 \qquad 6.6$$

where $Export_i$ is the annual mass load of substance s, $Pop_{Admin,untreated}$ is the total count of population that is not served by WWTPs in the AU (rural or urban), ddc is the proportion of hormone load from untreated wastewaters reaching the water body.

In our ‘most probable’ scenario, we parameterized *ddc* as 40% for untreated wastewater discharged into nearby rivers in rural areas, as reported by Wang et al. (2011), and used a higher *ddc* of 80% for urban areas due to the presence of more impervious surfaces (Dr. Wang Kaijun, School of Environment, Tsinghua University, personal communication).

Finally, after the contribution of untreated populations was downscaled to the pixel resolution, we used zonal statistics with reach catchments as zones to aggregate the hormone load from the individual pixels to the respective river reach.

$$L(S)_{reach,i} = \sum_{WWTP} Input_{pix,untreated,\mathrm{urban|rural}} \tag{6.7}$$

S6.3 River and lake routing of mass balances

The outflow mass balance for each river reach is calculated as:

$$TL_{reach,i} = \left(\sum_{i} Ld_{wwtp} + \sum_{j,n} Ld_{reach} \right) \times d_s \tag{6.8}$$

where $TL_{reach,i}$ represents the total accumulated mass at the end of the river reach *i*, calculated as the sum of mass influx from all wastewater treatment plants located anywhere on the river reach Ld_{wwtp}, and the total mass Ld_{reach} from all upstream reaches *j,n*. The degradation of a chemical substance in the river body, if applicable, is expected to decrease at a rate proportional to its mass, and is represented as the factor *d* is the environmental decay factor of estrogen *s*. The factor *d* is calculated based on first-order decay $d = e^{-kt}$ where *t* is the time a plug of water needs to travel through the river reach (based on average flow velocity), and *k* is a positive number called the first-order rate constant, which determines the time speed of environmental decay. Based on averages given in Caldwell et al. (2010), we used 0.3, 0.3, 5.7 and 0.07 for E1, E2, E3 and EE2, respectively, as *k* (day^{-1}).

In the river network, *t* is derived by dividing river reach length by the average velocity within the river reach. This corresponds to the average retention time in each individual river reach (i.e., the time a plug of fluid needs to travel from the beginning to the end of the river segment). Velocity was estimated from long-term average discharge following the empirically derived formula (Allen et al., 1994):

$$v = 1.07 \times (Q^{0.1035}) \qquad 6.9$$

where v is the velocity in m/s within the river reach and Q is the long-term average discharge in m^3/s.

Lakes are currently modeled as either passing the load fully to the downstream river or eliminating the load entirely (Ort et al., 2009), but will be included as a completely stirred tank reactor (Anderson et al., 2004) in future upgrades of the model.

S6.4 Environmental Risk Assessment

After accumulation and degradation, we calculated the predicted environmental concentrations, PEC_i , for each river reach i by dividing the accumulated chemical mass $TL_{reach,i}$ by the discharge Q_i of the corresponding reach (which includes the accumulated amount of water volume from the wastewater treatment plants upstream):

$$PEC_i = \frac{TL_{reach,i}}{Q_i} \qquad 6.10$$

Due to the seasonal variability of river flows, we used the Q90-month, a low flow indicator equivalent to the daily 90th percentile flow (i.e. flow that is exceeded 90% of the time and which represents the worst-case scenario with minimal dilution and maximal concentration). See chapter 5 of this thesis for a discussion on Q90-month.

Since estrogens act in an additive manner (Thorpe et al., 2001) and since the eco-toxicological potency of all estrogens is not equal (SCHER, 2011; Anderson et al., 2012; Caldwell et al., 2012), the amount of estrogen release from residents was estimated in terms of E2 equivalents, the most potent natural estrogen. Specifically, EE2, E1 and E3 were assumed to be 10, 0.33 and 0.04 times as potent as E2, respectively (Khan and Nicell, 2014). After the weighting, we compared the E2-equivalent concentrations (PECs) to the combined Predicted No Effect Concentration (PNEC) for all estrogens, which is assumed to be 1 ng of E2-eq/L (Williams et al., 2009).

7 Synthesis and concluding remarks

This thesis contributes to advancing the field of global hydrological modelling by providing a comprehensive, new framework to conduct eco-hydrological research at high spatial resolution. The application of this modelling platform presents opportunities to bridge the divide between large-scale modelling and more local, applied research conducted by ecologists or water managers who require higher-resolution data and/or specific analyses not typically available in global hydrological and routing models. I conducted analyses in regional (Saint Lawrence River; Mekong River), national (China), and global research settings, which would not have been possible without the presented innovations in global river routing. My research has helped to alleviate the challenges that many large-scale hydrological and routing models face related to inadequate spatial resolution, inflexible data structures, missing multi-scale support, limited support for connectivity, and the lack of integrated modelling.

7.1 Scientific contributions

This thesis and related research introduced the first global integrated hyper-resolution river routing model designed for eco-hydrological assessments of anthropogenic effects on rivers. It made substantial contributions to advance scientific understanding and methodologies in the field of global hydrology, as follows:

Contributions to knowledge:

i. In chapter 3, I developed and compared several new indicators of river fragmentation and flow regulation. While most river assessments and applications use river length as a fundamental measure to represent river habitat for species, I find that indicators based on river volume may be more adequate in this application for ecological modelling. I showed that an indicator that uses river volume (i.e. size) provides a better representation of ecological information such as species migratory ranges or river classifications. I also found that indicators based on river volume are less scale dependent as they are less susceptible to the spatial extent of the underlying river network (i.e. additional small streams have little volume). The development of new volume-based indicators (chapter 3) led me to their application at the global scale in chapter 4. These results further supported my hypothesis that river volume is more representative as a proxy or indicator in ecological assessments than river length. This hypothesis warrants further investigation as part of a more systematic and directed assessment. As a broader opportunity, the use of river volume as demonstrated in HydroROUT could provide an alternative or additional measure for ecohydrological research in general.

ii. In chapter 4, I conducted a multi-scale assessment in which I showed that river fragmentation is highly scale dependent both within and across river basins. At the sub-basin scale, I identified substantial intra-basin heterogeneity of impacts which was previously difficult to assess due to model scale and resolution. The results suggest that studies on river network connectivity should be conducted at multiple scales.

iii. In chapter 4, I also demonstrate that small dams significantly contribute to fragmentation when compared to scenarios where only large dams are assessed. Specifically, the relative magnitude of small versus large dam impacts was quantified for the first time in detail for the Mississippi Basin.

iv. Finally, chapter 4 revealed that natural barriers significantly affect the fragmentation calculations, in particular waterfalls. This finding suggests that future assessments should take into account both anthropogenic and natural sources of fragmentation, which has implications for environmental impact assessments of individual or groups of dams. For example, a proposed dam at a location with a waterfall nearby may not reduce connectivity as much as would be expected in a fully connected network. Furthermore, the inclusion of waterfalls in eco-hydrological models could enable new assessments altogether, such as predicting the distribution of species throughout river systems.

v. In chapter 5, I show that 90% of all wastewater treatment plants in the study area indicate a dilution factor of 10 or higher, which provides evidence that this commonly applied assumption underlying most current risk assessment methodologies is indeed valid. However, I also show that that assessments based on dilution factors alone cannot replace spatial risk assessments such as the one conducted with HydroROUT. For example, I found that two out of 15 commonly used pharmaceuticals in Canada show elevated environmental risk in extensive reaches of the Saint Lawrence River.

vi. In chapter 6, I determined novel spatial patterns of risk from hormones in continental China, and I show that wastewater treatment plants are critical to reducing risk from hormones in river systems. Yet, a second key finding of this study is that the combined contributions from small-scale populations that are not connected to sewage treatment plants may be far greater pollution sources than treated wastewater from large urban municipalities. This is an important finding for the implementation of contaminant fate models in regions where large parts of the population are not connected to sewage treatment plants, including large parts of the developing world. My findings also suggest that these non-point sources may be equally important for other population-sourced chemicals, such as pharmaceuticals.

Contributions to scientific methodologies:

i. This research provides a global, multi-scale framework in which river routing is feasible from very small scales to increasingly larger scales; i.e., from small river reaches to larger, nested hydrological subdivisions such as sub-basins and basins. This provides the research community with a framework to conduct eco-hydrological research at multiple scales using a common routing approach. The framework is based on graph-theoretical principles, where river networks are analysed using network theory. This type of approach has been shown to support the linkage of landscape ecology and hydrology domains (Bunn et al., 2000), and as such provides new opportunities to support integrated modelling in freshwater systems.

ii. My research led to the design and critical evaluation of indicators of river fragmentation and flow regulation that can be derived rapidly as a first-order proxy even in data-poor settings, and provided a proof-of-concept that such simple indicators are useful to rank, distinguish, or monitor varying levels of impacts from different dams. Combined with stakeholder involvement, such indicators provide opportunities to be used in policy-support and environmental assessments, for example with the goal to find alternative or 'optimal' solutions in large-scale dam planning and management exercises.

iii. This study combined, for the first time, multiple impacts of dams, namely flow regulation and fragmentation of rivers in one framework. While the two effects were previously assessed separately, or in the form of a lumped indicator, they were here combined to form one integrated "Dam Impact Matrix" which helps to improve our understanding of trade-offs between different types of dams in general and in relation to their societal benefit (i.e. energy production; chapter 3).

iv. Furthermore, this research allowed for much more refined and detailed results regarding dam impacts than previously achievable. The higher spatial resolution of the HydroROUT framework supports the distinction of intra-basin variability and trends, i.e. at sub-basin scales, while the computationally efficient routing algorithms enable fast, repetitive scenario calculations to provide temporal trends of river fragmentation that have never been presented before.

v. This research also showed the applicability of HydroROUT as a high-resolution chemical fate model, which is expected to provide novel opportunities to investigate the sources and fate of chemicals in the environment at large scales. As a chemical fate model, HydroROUT's relatively simple, vector-based design supports assessments of a large variety of chemicals. HydroROUT has been shown to produce reasonable predictions of chemical concentrations in rivers to support first-order risk assessments and screening of chemicals.

7.2 Limitations and future directions

HydroROUT provides significant improvements over current approaches in eco-hydrological modelling at large scales, and my research has addressed many of the identified challenges with substantial success. Despite these advancements, a number of challenges remain, which are mainly related to limitations identified in my diverse studies on river networks. These limitations, at the same time, provide avenues for future improvements and research directions.

Limitations in the implementation of hydrological and routing processes

- HydroROUT's routing scheme does not include dynamic runoff routing but instead uses, in a decoupled mode, monthly and yearly long-term average discharge from a Global Hydrological Model that was downscaled to HydroROUT. This currently limits the application of HydroROUT to studies that do not require dynamic interactions between streamflow and other model components. Yet, the non-dynamic integration of streamflow substantially decreases model complexity and model run time, leading to superior performance and allowing, for example, rapid repeated model executions required for Monte-Carlo simulations or in sensitivity analyses. If dynamic flow routing would be included in HydroROUT, such applications may no longer be feasible.

- The omission of an internal runoff generation scheme means that scenarios of land use or climate change and their resulting effects on discharge cannot be calculated. The implementation of a combined land-surface model with a routing scheme would drastically increase model complexity, and negate some of the performance advantages of HydroROUT.

- The currently implemented routing and tracing scheme is limited to upstream and downstream routing, accumulation and first-order decay functions. More complex in-river routing may be required for representing higher levels of complexity in river processes, e.g. dispersal and population-level dynamics. More advanced in-stream decay functions (e.g. nutrient-spiralling) also require modifications to the current routing algorithms.

- The routing component currently does not consider the effects of floodplains including flood wave attenuation of retention processes. Also, their effect on chemical removal processes is not considered, although floodplains are known to act as sinks for river constituents, such as sediments, nutrients, and contaminants. With the recent integration of a global floodplain map (Fluet-Chouinard *et al.,* in prep) there is an opportunity to include a "floodplain module" in HydroROUT. This could provide further research avenues into the effect of floodplains on the distribution of water and substances along the river network.

- The representation of velocity in HydroROUT is currently derived from river discharge using hydraulic geometry laws (Allen et al., 1994). While this approach is inherently prone to large errors, it may be considered an improvement over the use of constant flow velocities implemented in some other models. Yet using this approach in HydroROUT makes the model estimates susceptible to uncertainty and error linked to the accuracy of representing discharge (further discussed below). Future iterations of HydroROUT should consider the use of more advanced methods to estimate velocity similar to Ngo-Duc et al. (2007), Fiedler and Döll (2010), and Verzano et al. (2012). The latter is using discharge, stream gradients, and global estimates of simplified Manning coefficients.

- I was able to show that small dams and waterfalls play an important role in fragmentation assessments, and may significantly alter outcomes of dam impact assessments. Yet I was not able to include them systematically in my large-scale assessments in the Mekong or globally, because of ongoing discussion about how the differences of passability between small dams, waterfalls, and large dams should be addressed by the connectivity indices. Large dams are considered generally non-permeable (assuming no structures to enhance passability), but waterfalls are typically passable in one direction, and smaller barriers are more passable overall. There is an opportunity to conduct research on the effects of waterfalls on river connectivity, and on the combined effect of including waterfalls into dam planning frameworks.

- Most of the river routing calculations in HydroROUT are conducted based on steady-state assumptions using long-term average discharge, or, in the case of applications related to chemical fate routing, based on low flow indicators calculated from steady-state monthly averages. My research indicated that the prediction of chemical concentrations in rivers is very sensitive to the accurate representation of low flow hydrology. Typical low flow indicators (e.g., 7Q10, Q90) in risk assessment are based on daily flows, but this temporal detail is not included in HydroROUT. Steps, including statistical approaches, should be taken in future iterations of the model to improve the low flow hydrology used in HydroROUT.

- The dam impact studies should be interpreted with caution due to a number of shortcomings: The calculation of river regulation disregards important processes that modify the flow, in particular—due to lack of data—the actual operational rules of dams. Furthermore, the calculations of fragmentation do not take into account any form of permeability of barriers, and assumes a complete barrier effect for every dam. However, numerous dams implicitly (by design) or explicitly (by structural modifications) allow the passage of species to pass the dam under certain circumstances. The DOR and RRI calculations should therefore be considered as a proxy for accumulated large-scale dam impacts rather than as a quantitative absolute measure of impact for individual dams.

Limitations in underlying baseline hydrography and the downscaled discharge estimates

The underlying baseline hydrographic data (HydroSHEDS) as well as the downscaled discharge used in HydroROUT display uncertainties that translate into limitations of the HydroROUT framework as follows:

- The HydroSHEDS database was developed originally only for regions below 60 degrees northern latitude, as data beyond this latitude were not available from the underlying digital elevation model (Farr et al., 2007). Substitute data from a different topographic dataset (USGS, 2000) with substantially inferior quality was added for the northernmost regions. There is potential for improvement in these areas, as better elevation data recently became available from Robinson et al. (2014).

- In predominantly flat terrain, the representation of hydrological flow paths is generally less accurate due to elevation 'noise' introduced from vegetation and built-up areas. In these cases, manual corrections were conducted in the production of HydroSHEDS; e.g., stream burning techniques were employed to enforce hydrological accuracy, where high resolution river networks were available (Lehner et al., 2006). However, significant deviations from observed flow paths are noticeable in some regions.

- Because of the underlying algorithm to derive flow directionality does not allow for multiple downstream directions, river bifurcations (i.e. splits into multiple flow channels), braided river systems, or secondary channels in river deltas cannot be represented with HydroROUT. Canals and artificial structures are also not implemented, which becomes relevant in places such as the North China Plain (chapter 6), which is heavily modified due to canalization. These shortcomings cannot easily be circumvented, and the implementation of these features would add substantial complexity to the connectivity and flow routing model. Nevertheless, for many applications these omissions, which are typically at smaller scales, are less critical.

- Highly detailed and complex topographic features such as floodplain channels that regulate local hydrological connectivity are not adequately represented in HydroSHEDS (Yamazaki et al., 2012). These limitations may become relevant in certain situations; i.e., where such features are critical. Even at HydroSHEDS' 90m resolution, such spatial detail is not well represented. Dedicated data models with higher spatial resolution should be considered (e.g. based on LIDAR).

- The preliminary validation of downscaled river discharge that is used in HydroROUT showed overall good agreement with the observed flows at 166 gauging stations in Canada for long-term average discharge values (see chapter 5). However, the monthly flow values, important in environmental risk assessment methodologies where a measure of low flows is needed (chapter 5 and 6), showed substantially larger errors. In cases where we compared low flow indicators derived from HydroROUT's downscaled discharge to low flow indicators at gauging stations (monthly Q90 flows), values of HydroROUT were generally too high. This, in turn, affects the environmental risk assessments conducted in chapter 5 and 6, leading to underestimated risk for low flow conditions. Human controlled flow regulation features, such as dams and reservoirs, can further cause misrepresentation of low flow conditions in particular. Overall, more validation of the discharge input into HydroROUT is needed—ideally a systematic global scale assessment to better understand the limitations in low flow hydrology.

- Regarding the lake database used in HydroROUT, it should be mentioned that there is a severe quality discrepancy between the representation of a lake's surface area and its storage volume. While the first is based on manual, high-resolution digitizing, the latter is estimated with spatial regression models. Overall, the volume estimations have been shown to provide acceptable first-order proxies based on tests against a selection of 5950 lakes across the world (Messager *et al.,* in prep.), yet the values of individual lakes may be greatly over- or underestimated.

7.3 Policy application and cross-cutting themes

The contributions made in this research have so far been applied to two fields of research—dam impact assessments and contaminant fate modelling. As evidence of their usefulness in integrated assessments in a policy-relevant context, the indicators of river fragmentation and flow regulation that I developed in chapters 3 and 4 have already been implemented in a case study prepared for the Asian Development Bank to support an international integrated assessment entitled 'Ensuring Sustainability of Greater Mekong Sub-region Regional Power Development (ADB TA7764-REG), sub-project: Determination of Flow Regulation and River Connectivity for Different Scenarios'.

Furthermore, these indicators will serve as new global scale indicators of fragmentation and flow regulation as part of the Biodiversity Indicator Partnership (BIP), a global initiative to promote and coordinate development and delivery of biodiversity indicators (http://www.bipindicators.net/).

The past and current development of HydroROUT as a chemical fate model (chapter 5) has been supported by Health Canada and Environment Canada with the goal to assist the implementation of the Canadian Chemical Management Plan (Ambrose and Clement, 2006)—a plan for addressing the legacy of approximately 4300 chemical substances prioritized through a categorization process by 2020. Between October 2012 and March 2013, a pilot study was carried out on behalf of Health Canada to test the feasibility of developing a large-scale contaminant fate model for the provinces of Quebec and Ontario (Lehner et al., 2013)[1]. Further initiatives are under way to extend the scope of the model from the current East Canada extent to a Canada-wide model.

The HydroROUT framework was recently used to create linkages between hydrological objects and cities, with the goal to map the water sources of the world's largest cities. This work has resulted in my co-authorship in the publication 'Water on an urban planet: Urbanization and the reach of urban water infrastructure' (McDonald et al., 2014).

The importance of upstream connectivity for improved conservation planning has been widely acknowledged (Hermoso et al., 2012) and HydroROUT, with its graph-theoretical approach, is ideally suited to provide such functionality at a global scale. In particular, there is, to my knowledge, no current framework that includes natural and anthropogenic barriers in large-scale freshwater conservation planning approaches, a gap that could be filled by HydroROUT with its unique capabilities to model connectivity and fragmentation.

[1] A consulting report was delivered to Health Canada, with Bernhard Lehner and Jim Nicell as the Principal Investigators. I was responsible for conducting the research and model development.

7.4 Overarching conclusions

The presented research contributes substantially to the body of knowledge in integrated eco-hydrological modelling. I have developed the first of a new generation of global river routing models that can support local modelling and decision making while providing a large-scale perspective. The resulting model, HydroROUT, has a number of unique characteristics, which I have demonstrated can contribute to novel integrated research applications.

HydroROUT combines preprocessed data, models and tools into a single large-scale modelling framework providing researchers with the opportunity to adapt the modelling scheme to the specific requirements of their geographic location and study system. Despite the current limitations related to the quality of the underlying HydroSHEDS database and implementation of the HydroROUT modelling framework, results have to be judged against large-scale needs of ecologists or water resources managers for which HydroROUT was designed.

For example, even if discharge values are prone to substantial uncertainty and error, ecological changes in the river network are often most pronounced at confluences between small tributaries and large main-stem rivers, where flow magnitudes can differ by one or several orders of magnitude. Many critical characteristics along the river network, such as highly altered conditions, disruptions in connectivity, distances along the flow path, or contributing catchment areas are well represented in the current model version, despite errors in discharge, due to the very high spatial resolution of the hydrographic baseline data. Major changes between river orders are well represented and can be related to species distributions. Detailed spatial objects, such as effluent points of wastewater treatment plants, dams, waterfalls and lakes, can be included as part of the assessments, even if uncertainties in the exact hydrological values are present.

The current version of HydroROUT is not able to conduct its own runoff and discharge generation, which leaves the model unsuitable for scenario calculations related to changes in land use or climate. The decoupled nature of HydroROUT, however, presents other research opportunities in terms of facilitating novel and computationally highly efficient eco-hydrological applications at large scales, previously not possible. The advanced implementation of connectivity and river routing in HydroROUT using a vector framework and a graph-theoretical model are key characteristics of this new generation of routing models.

In this thesis, I have made several technical and theoretical improvements to advance the field of global hydrological modelling and routing, by providing new data, models, methods, knowledge and insights that are useful for large-scale eco-hydrological assessments. Beyond this research domain, there are ample opportunities to apply HydroROUT and its concepts in a variety of related fields including aquatic ecology, biogeochemistry, geo-statistical modelling, and health risk assessments.

Bibliography

Abbott MB, Bathurst JC, Cunge JA, O'Connell PE, Rasmussen J (1986). An introduction to the European Hydrological System — Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system. *Journal of Hydrology* **87**, 61-77. doi:10.1016/0022-1694(86)90115-0

Abbott MB, Refsgaard JC (1996). Distributed hydrological modelling, first ed. Springer: Dordrecht, Netherlands.

Abell R, Allan JD, Lehner B (2007). Unlocking the potential of protected areas for freshwaters. *Biological Conservation* **134**, 48-63. doi:10.1016/j.biocon.2006.08.017

Abell R, Thieme ML, Revenga C, Bryer M, Kottelat M, Bogutskaya N, Coad B, Mandrak N, Balderas SC, Bussing W, Stiassny MLJ, Skelton P, Allen GR, Unmack P, Naseka A, Ng R, Sindorf N, Robertson J, Armijo E, Higgins JV, Heibel TJ, Wikramanayake E, Olson D, Lopez HL, Reis RE, Lundberg JG, Perez MHS, Petry P (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *Bioscience* **58**, 403-414. doi:10.1641/B580507

Ahuja RK, Magnanti TL, Orlin JB (1993). Network flows: theory, algorithms and applications. Englewood Cliffs, NJ: Prentice Hall.

Alcamo J, Döll P, Henrichs T, Kaspar F, Lehner B, Rösch T, Siebert S (2003). Development and testing of the WaterGAP 2 global model of water use and availability. *Hydrological Sciences Journal* **48**, 317-337. doi:10.1623/hysj.48.3.317.45290

Alexander RB, Bohlke JK, Boyer EW, David MB, Harvey JW, Mulholland PJ, Seitzinger SP, Tobias CR, Tonitto C, Wollheim WM (2009). Dynamic modeling of nitrogen losses in river networks unravels the coupled effects of hydrological and biogeochemical processes. *Biogeochemistry* **93**, 91-116. doi:10.1007/s10533-008-9274-8

Allen PM, Arnold JC, Byars BW (1994). Downstream channel geometry for use in planning-level models. *Journal of the American Water Resources Association* **30**, 663-671

Altermatt F (2013). Diversity in riverine metacommunities: a network perspective. *Aquatic Ecology* **47**, 365-377. doi:10.1007/s10452-013-9450-3

Alvarez-Cobelas M, Sánchez-Carrillo S, Angeler DG, Sánchez-Andrés R (2009). Phosphorus export from catchments: a global view. *Journal of the North American Benthological Society* **28**, 805-820. doi:10.1899/09-073.1

Ambrose R, Clement T (2006). Plan for the assessment and management of certain substances on the Domestic Substances List. Canada Gazette, Part I 140[49]. *Canada Gazette*, 4109–4116

Amoros C, Bornette G (2002). Connectivity and biocomplexity in waterbodies of riverine floodplains. *Freshwater Biology* **47**, 761-776. doi:10.1046/j.1365-2427.2002.00905.x

Amoros C, Roux A (1988). Interaction between water bodies within the floodplains of large rivers: function and development of connectivity, Connectivity in Landscape Ecology. Proceedings of the 2nd International Seminar of the "International Association for Landscape Ecology", Münster, Germany, pp. 125-130.

Andersen J, Refsgaard JC, Jensen KH (2001). Distributed hydrological modelling of the Senegal River Basin — model construction and validation. *Journal of Hydrology* **247**, 200-214. doi:10.1016/s0022-1694(01)00384-5

Anderson EP, Pringle CM, Freeman MC (2008). Quantifying the extent of river fragmentation by hydropower dams in the Sarapiquí River Basin, Costa Rica. *Aquatic Conservation: Marine and Freshwater Ecosystems* **18**, 408-417. doi:10.1002/aqc.882

Anderson MG, Bates PD (2001). Model validation: perspectives in hydrological science, first ed. Wiley: Chichester, UK.

Anderson PD, D'Aco VJ, Shanahan P, Chapra SC, Buzby ME, Cunningham VL, Duplessie BM, Hayes EP, Mastrocco FJ, Parke NJ, Rader JC, Samuelian JH, Schwab BW (2004). Screening analysis of human pharmaceutical compounds in U.S. surface waters. *Environmental Science & Technology* **38**, 838-849

Anderson PD, Johnson AC, Pfeiffer D, Caldwell DJ, Hannah R, Mastrocco F, Sumpter JP, Williams RJ (2012). Endocrine disruption due to estrogens derived from humans predicted to be low in the majority of U.S. surface waters. *Environmental Toxicology and Chemistry* **31**, 1407-1415. doi:10.1002/etc.1824

Andersson E, Nilsson C, Johansson ME (2000). Effects of river fragmentation on plant dispersal and riparian flora. *Regulated Rivers: Research & Management* **16**, 83-89

Angermeier PL, Schlosser IJ (1989). Species-area relationship for stream fishes. *Ecology* **70**, 1450-1462. doi:10.2307/1938204

Arnell NW, Livermore MJL, Kovats S, Levy PE, Nicholls R, Parry ML, Gaffin SR (2004). Climate and socio-economic scenarios for global-scale climate change impacts assessments: characterising the SRES storylines. *Global Environmental Change* **14**, 3-20. doi:10.1016/j.gloenvcha.2003.10.004

Arnold JG (1998). Large area hydrologic modeling and assessment. I. Model development. *Journal of the American Water Resources Association* **34**, 73-89

Aronson D, Weeks J, Meylan B, Guiney PD, Howard PH (2012). Environmental release, environmental concentrations, and ecological risk of N,N-Diethyl-m-toluamide (DEET). *Integrated Environmental Assessment and Management* **8**, 135-166. doi:10.1002/ieam.271

Arthington AH, Bunn SE, Puff (2006). The challenge of providing environmental flow rules to sustain river ecosystems. *Ecological Applications* **16**, 1311-1318

Arthington AH, Naiman RJ, McCLAIN ME, Nilsson C (2010). Preserving the biodiversity and ecological services of rivers: new challenges and research opportunities. *Freshwater Biology* **55**, 1-16. doi:10.1111/j.1365-2427.2009.02340.x

Aspinall R, Pearson D (2000). Integrated geographical assessment of environmental condition in water catchments: Linking landscape ecology, environmental modelling and GIS. *Journal of Environmental Management* **59**, 299-319. doi:10.1006/jema.2000.0372

Atkinson R, Power R, Lemon D, O'Hagan R (2008). The Australian Hydrological Geospatial Fabric – Development methodology and conceptual architecture. CSIRO:Water for a Healthy Country National Research Flagship., Canberra, Australia, pp. 1-55.

Atkinson SF, Johnson DR, Venables BJ, Slye JL, Kennedy JR, Dyer SD, Price BB, Ciarlo M, Stanton K, Sanderson H, Nielsen A (2009). Use of watershed factors to predict consumer surfactant risk, water quality, and habitat quality in the upper Trinity River, Texas. *The Science of The Total Environment* **407**, 4028-4037. doi:10.1016/j.scitotenv.2009.02.029

Aufdenkampe AK, Mayorga E, Raymond PA, Melack JM, Doney SC, Alin SR, Aalto RE, Yoo K (2011). Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere. *Frontiers in Ecology and the Environment* **9**, 53-60. doi:10.1890/100014

Bagstad KJ, Johnson GW, Voigt B, Villa F (2013). Spatial dynamics of ecosystem service flows: A comprehensive approach to quantifying actual services. *Ecosystem Services* **4**, 117-125. doi:10.1016/j.ecoser.2012.07.012

Bahn V, McGill BJ (2007). Can niche-based distribution models outperform spatial interpolation? *Global Ecology and Biogeography* **16**, 733-742. doi:10.1111/j.1466-8238.2007.00331.x

Baird IG, Flaherty MS (2005). Mekong River fish conservation zones in southern Laos: assessing effectiveness using local ecological knowledge. *Environmental Management (New York)* **36**, 439-454. doi:10.1007/s00267-005-3093-7

Baran E (2006). Fish migration triggers in the Lower Mekong Basin and other tropical freshwater systems. MRC technical paper No. 14. Mekong River Commission (MRC): Ventiane, Lao PDR.

Baran E (2010). Mekong fisheries and mainstream dams., in: ICEM (Ed.), Mekong River Commission Strategic Environmental Assessment of hydropower on the Mekong mainstream. International Centre for Environmental Management (ICEM): Hanoi, Viet Nam.

Beighley RE, Eggert KG, Dunne T, He Y, Gummadi V, Verdin KL (2009). Simulating hydrologic and hydraulic processes throughout the Amazon River Basin. *Hydrological Processes* **23**, 1221-1235. doi:10.1002/hyp.7252

Bennett EM, Carpenter SR, Caraco NF (2001). Human impact on erodable phosphorus and eutrophication: a global perspective. *Bioscience* **51**, 227-234

Berkamp G, McCartney M, Dugan P, McNeely J, Acreman MC (2000). Dams, Ecosystem Functions and Environmental Restoration Thematic Review II.1 prepared as an input to the World Commission on Dams (WCD). Secretariat of the World Commission on Dams, Cape Town, Australia.

Beusen AHW, Bouwman AF, Dürr HH, Dekkers ALM, Hartmann J (2009). Global patterns of dissolved silica export to the coastal zone: Results from a spatially explicit global model. *Global Biogeochemical Cycles* **23**, 1-13. doi:10.1029/2008GB003281

Biggs BJF, Nikora VI, Snelder TH (2005). Linking scales of flow variability to lotic ecosystem structure and function. *River Research and Applications* **21**, 283-298. doi:10.1002/rra.847

Birkel C, Soulsby C, Ali G, Tetzlaff D (2014). Assessing the cumulative impacts of hydropower regulation on the flow characteristics of a large atlantic salmon river system. *River Research and Applications* **30**, 456-475. doi:10.1002/rra.2656

Bontemps S, Defourny P, Van Bogaert E, Sophie B, Pierre D, Eric VB (2010). GLOBCOVER 2009 Products Report Description. European Space Agency. Available at: http://ionia1.esrin.esa.int/ (Accessed April 3rd, 2012), Ionia, pp. 1-30.

Brown BL, Swan CM, Auerbach DA, Grant EHC, Hitt NP, Maloney KO, Patrick C (2011). Metacommunity theory as a multispecies, multiscale framework for studying the influence of river network structure on riverine communities and ecosystems. *Journal of the North American Benthological Society* **30**, 310-327

Brun GL, Bernier M, Losier R, Doe K, Jackman P, Lee H-B (2006). Pharmaceutically active compounds in atlantic canadian sewage treatment plant effluents and receiving waters, and potential for environmental effects as measured by acute and chronic aquatic toxicity. *Environmental Toxicology and Chemistry* **25**, 2163-2176

Bunn AG, Urban DL, Keitt TH (2000). Landscape connectivity: A conservation application of graph theory. *Journal of Environmental Management* **59**, 265-278. doi:10.1006/jema.2000.0373

Bunn SE, Arthington AH (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management* **30**, 492-507. doi:10.1007/s00267-002-2737-0

Burkholder J, Libra B, Weyer P, Heathcote S, Kolpin D, Thorne PS, Wichman M (2007). Impacts of waste from concentrated animal feeding operations on water quality. *Environmental Health Perspectives* **115**, 308-312. doi:10.1289/ehp.8839

Buser HR, Poiger T, Muller MD (1998). Occurrence and fate of the pharmaceutical drug diclofenac in surface waters: Rapid photodegradation in a lake. *Environmental Science & Technology* **32**, 3449-3456. doi:10.1021/Es980301x

Butkus SR, Welch EB, Horner RR, Spyridakis DE (1988). Lake response modeling using available biologically phosphorus. *Journal of the Water Pollution Control Federation* **60**, 1663-1669

Butts M, Loinaz M, Bauer-Gottwein P, Unnasch R, Gross D (2012). Mike SHE-ECOLAB–an integrated catchment-scale eco-hydrological modelling tool, 19th International Conference on Computational Methods in Water Resources, University of Illinois at Urbana-Champaign.

Calabrese JM, Fagan WF (2004). A comparison-shopper's guide to connectivity metrics. *Frontiers in Ecology and the Environment* **2**, 529-536. doi:10.2307/3868383

Caldwell DJ, Mastrocco F, Anderson PD, Lange R, Sumpter JP (2012). Predicted-no-effect concentrations for the steroid estrogens estrone, 17beta-estradiol, estriol, and 17alpha-ethinylestradiol. *Environmental Toxicology and Chemistry* **31**, 1396-1406. doi:10.1002/etc.1825

Caldwell DJ, Mastrocco F, Nowak E, Johnston J, Yekel H, Pfeiffer D, Hoyt M, DuPlessie BM, Anderson PD (2010). An assessment of potential exposure and risk from estrogens in drinking water. *Environmental Health Perspectives* **118**, 338-344. doi:10.1289/ehp.0900654

Callahan BG (1996). Commemoration of the 50th anniversary of Monte Carlo. *Human and Ecological Risk Assessment: An International Journal* **2**, 627-627. doi:10.1080/10807039609383638

Campbell Grant EH, Lowe WH, Fagan WF (2007). Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters* **10**, 165-175. doi:10.1111/j.1461-0248.2006.01007.x

Campbell IC (2009). The Mekong - Biophysical Environment of an International River Basin. Aquatic Ecology. Academic Press: New York.

Carlisle DM, Wolock DM, Meador MR (2011). Alteration of streamflow magnitudes and potential ecological consequences: a multiregional assessment. *Frontiers in Ecology and the Environment* **9**, 264-270. doi:10.1890/100053

Carlsson C, Johansson AK, Alvan G, Bergman K, Kuhler T (2006). Are pharmaceuticals potent environmental pollutants? Part I: environmental risk assessments of selected active pharmaceutical ingredients. *Science of the Total Environment* **364**, 67-87. doi:10.1016/j.scitotenv.2005.06.035

Chang H, Wan Y, Hu J (2009). Determination and source apportionment of five classes of steroid hormones in urban rivers. *Environmental Science & Technology* **43**, 7691-7698. doi:10.1021/es803653j

Chang H, Wan Y, Wu S, Fan Z, Hu J (2011). Occurrence of androgens and progestogens in wastewater treatment plants and receiving river waters: comparison to estrogens. *Water Research* **45**, 732-740. doi:10.1016/j.watres.2010.08.046

Chapra SC (1997). Surface water-quality modeling. McGraw–Hill: New York.

Chen TS, Chen TC, Yeh KJ, Chao HR, Liaw ET, Hsieh CY, Chen KC, Hsieh LT, Yeh YL (2010). High estrogen concentrations in receiving river discharge from a concentrated livestock feedlot. *Science of the Total Environment* **408**, 3223-3230. doi:10.1016/j.scitotenv.2010.03.054

Cherkassky BV, Goldberg AV, Radzik T (1996). Shortest paths algorithms: Theory and experimental evaluation. *Mathematical programming* **73**, 129-174. doi:10.1007/BF02592101

China Population Census Office (2012). Tabulation on the 2010 Population Census of the People's Republic of China by county. China Statistics Press, Beijing.

CIESIN (2005). Gridded Population of the World Version 3 (GPWv3). Available at http://sedac.ciesin.columbia.edu/gpw/. Socioeconomic Data and Applications Center (SEDAC), Palisades, New York.

CMSN (2012). Chinese Medical Statistical Yearbook (2010–2011). Chinese Medicine Statistics Net (CMSN), Beijing, China.

Coe MT (1998). A linked global model of terrestrial hydrologie processes: Simulation of modern rivers, lakes, and wetlands. *Journal of Geophysical Research* **103**, 8885-8899. doi:10.1029/98JD00347

Coe MT (2000). Modeling terrestrial hydrological systems at the continental scale: Testing the accuracy of an atmospheric GCM. *Journal of Climate* **13**, 686-704. doi:10.1175/1520-0442(2000)013<0686:MTHSAT>2.0.CO;2

Coe MT, Costa MH, Howard EA (2008). Simulating the surface waters of the Amazon River basin: impacts of new river geomorphic and flow parameterizations. *Hydrological Processes* **22**, 2542-2553. doi:10.1002/hyp.6850

Coe MT, Costa MH, Soares-Filho BS (2009). The influence of historical and potential future deforestation on the stream flow of the Amazon River - Land surface processes and atmospheric feedbacks. *Journal of Hydrology* **369**, 165-174. doi:10.1016/j.jhydrol.2009.02.043

Cote D, Kehler DG, Bourne C, Wiersma YF (2009). A new measure of longitudinal connectivity for stream networks. *Landscape Ecology* **24**, 101-113. doi:10.1007/s10980-008-9283-y

Cowan-Ellsberry CE, McLachlan MS, Arnot JA, Macleod M, McKone TE, Wania F (2009). Modeling exposure to persistent chemicals in hazard and risk assessment. *Integrated Environmental Assessment and Management* **5**, 662-679. doi:10.1897/IEAM_2008-084.1

Crisp TM, Clegg ED, Cooper RL, Wood WP, Anderson DG, Baetcke KP, Hoffmann JL, Morrow MS, Rodier DJ, Schaeffer JE, Touart LW, Zeeman MG, Patel YM (1998). Environmental endocrine disruption: an effects assessment and analysis. *Environmental Health Perspectives* **106 Suppl 1**, 11-56

Cunningham VL, Binks SP, Olson MJ (2009). Human health risk assessment from the presence of human pharmaceuticals in the aquatic environment. *Regulatory Toxicology and Pharmacology* **53**, 39-45. doi:10.1016/j.yrtph.2008.10.006

Cunningham VL, D'Aco VJ, Pfeiffer D, Anderson PD, Buzby ME, Hannah RE, Jahnke J, Parke NJ (2012). Predicting concentrations of trace organic compounds in municipal wastewater treatment plant sludge and biosolids using the PhATE model. *Integrated Environmental Assessment and Management* **8**, 530-542. doi:10.1002/ieam.1274

David CH, Maidment DR, Niu GY, Yang ZL, Habets F, Eijkhout V (2011). River network routing on the NHDPlus dataset. *Journal of Hydrometeorology* **12**, 913-934. doi:10.1175/2011jhm1345.1

David CH, Yang Z-L, Hong S (2013). Regional-scale river flow modeling using off-the-shelf runoff products, thousands of mapped rivers and hundreds of stream flow gauges. *Environmental Modelling & Software* **42**, 116-132. doi:10.1016/j.envsoft.2012.12.011

Dias MS, Cornu J-F, Oberdorff T, Lasso CA, Tedesco PA (2013). Natural fragmentation in river networks as a driver of speciation for freshwater fishes. *Ecography* **36**, 683-689. doi:10.1111/j.1600-0587.2012.07724.x

Döll P, Berkhoff K, Bormann H, Fohrer N, Gerten D (2008). Advances and visions in large-scale hydrological modelling: findings from the 11th workshop on large scale hydrological modelling. *Advances in Geosciences* **18**, 51-61

Döll P, Fiedler K, Zhang J (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrology and Earth System Sciences Discussions* **6**, 4773-4812. doi:10.5194/hessd-6-4773-2009

Döll P, Kaspar F, Lehner B (2003). A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology* **270**, 105-134. doi:10.1016/s0022-1694(02)00283-4

Döll P, Lehner B (2002). Validation of a new global 30-min drainage direction map. *Journal of Hydrology* **258**, 214-231. doi:10.1016/s0022-1694(01)00565-0

Donner SD, Coe MT, Lenters JD, Twine TE, Foley JA (2002). Modeling the impact of hydrological changes on nitrate transport in the Mississippi River Basin from 1955 to 1994. *Global Biogeochemical Cycles* **16**, 1-19. doi:10.1029/2001gb001396

Dudgeon D, Arthington AH, Gessner MO, Kawabata Z, Knowler DJ, Leveque C, Naiman RJ, Prieur-Richard AH, Soto D, Stiassny ML, Sullivan CA (2006). Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological Reviews of the Cambridge Philosophical Society* **81**, 163-182. doi:10.1017/S1464793105006950

Dugan PJ, Barlow C, Agostinho AA, Baran E, Cada GF, Chen D, Cowx IG, Ferguson JW, Jutagate T, Mallen-Cooper M, Marmulla G, Nestler J, Petrere M, Welcomme RL, Winemiller KO (2010). Fish migration, dams, and loss of ecosystem services in the Mekong basin. *Ambio* **39**, 344-348. doi:10.1007/s13280-010-0036-1

Dynesius M, Nilsson C (1994). Fragmentation and flow regulation of river systems in the northern third of the world. *Science* **266**, 753-762. doi:10.1126/science.266.5186.753

Ebert I, Bachmann J, Kuhnen U, Kuster A, Kussatz C, Maletzki D, Schluter C (2011). Toxicity of the fluoroquinolone antibiotics enrofloxacin and ciprofloxacin to photoautotrophic aquatic organisms. *Environmental Toxicology and Chemistry* **30**, 2786-2792. doi:10.1002/etc.678

Elith J, Leathwick JR (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* **40**, 677-697. doi:10.1146/annurev.ecolsys.110308.120159

Ellis EC, Ramankutty N (2008). Putting people in the map: anthropogenic biomes of the world. *Frontiers in Ecology and the Environment* **6**, 439-447. doi:10.1890/070062

EMEA (2006). Guideline on the environmental risk assessment of medicinal products for human use, EMEA/CHMP/SWP/4447/00. European Medicines Agency (EMEA), London, UK.

Environment Canada (2012). Water Survey of Canada's Hydrometric Database (HYDAT). Retrieved from http://www.wsc.ec.gc.ca (September 2012). Environment Canada (EC), Ottawa, Canada.

Erős T, Olden JD, Schick RS, Schmera D, Fortin M-J (2011a). Characterizing connectivity relationships in freshwaters using patch-based graphs. *Landscape Ecology* **27**, 303-317. doi:10.1007/s10980-011-9659-2

Erős T, Schmera D, Schick RS (2011b). Network thinking in riverscape conservation – A graph-based approach. *Biological Conservation* **144**, 184-192. doi:10.1016/j.biocon.2010.08.013

ESRI (2011). ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

Fagan WE, Aumann C, Kennedy CM, Unmack PJ (2005). Rarity, fragmentation, and the scale dependence of extinction risk in desert fishes. *Ecology* **86**, 34-41. doi:10.1890/04-0491

FAO, IIASA, ISRIC, ISS-CAS, and JRC (2012). Harmonized World Soil Database (version 1.2). FAO: Rome, Italy and IIASA: Laxenburg, Austria.

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, Seal D, Shaffer S, Shimada J, Umland J, Werner M, Oskin M, Burbank D, Alsdorf DED (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics* **45**, 1-33. doi:10.1029/2005RG000183

Feijtel T, Boeije G, Matthies M, Young A, Morris G, Gandolfi C, Hansen B, Fox K, Matthijs E, Koch V, Schroder R, Cassani G, Schowanek D, Rosenblom J, Holt M (1998). Development of a geography-referenced regional exposure assessment tool for European rivers—GREAT-ER. *Journal of Hazardous Materials* **61**, 59-65. doi:10.1016/s0304-3894(98)00108-3

Fekete BM, Gibson JJ, Aggarwal P, Vörösmarty CJ (2006). Application of isotope tracers in continental scale hydrological modeling. *Journal of Hydrology* **330**, 444-456. doi:10.1016/j.jhydrol.2006.04.029

Fekete BM, Vörösmarty CJ, Lammers RB (2001). Scaling gridded river networks for macroscale hydrology: Development, analysis, and control of error. *Water Resources Research* **37**, 1955-1967. doi:10.1029/2001wr900024

Feldman A (2000). Hydrologic Modeling System-HEC-HMS. Technical Reference Manual. US Army Corps of Engineers Hydrologic Engineering, Washington, D.C., p. 157.

Fent K, Weston Aa, Caminada D (2006). Ecotoxicology of human pharmaceuticals. *Aquatic toxicology (Amsterdam, Netherlands)* **76**, 122-159. doi:10.1016/j.aquatox.2005.09.009

Fiedler K, Döll P (2010). Monthly and daily variations of continental water storage and flows, in: Flechtner FM, Gruber T, Güntner A, Mandea M, Rothacher M, Schöne T, Wickert J (Eds.), System Earth via Geodetic-Geophysical Space Techniques. Advanced Technologies in Earth Sciences. Springer: Berlin, Heidelberg, pp. 407-414.

Finer M, Jenkins CN (2012). Proliferation of hydroelectric dams in the Andean Amazon and implications for Andes-Amazon connectivity. *PLoS ONE* **7**, e35126. doi:10.1371/journal.pone.0035126

Friedl G, Wüest A (2002). Disrupting biogeochemical cycles-Consequences of damming. *Aquatic Sciences-Research Across Boundaries* **64**, 55-65

Fukushima M, Kameyama S, Kaneko M, Nakao K, Steel EA (2007). Modelling the effects of dams on freshwater fish distributions in Hokkaido, Japan. *Freshwater Biology* **52**, 1511-1524. doi:10.1111/j.1365-2427.2007.01783.x

Fullerton AH, Burnett KM, Steel EA, Flitcroft RL, Pess GR, Feist BE, Torgersen CE, Miller DJ, Sanderson BL (2010). Hydrological connectivity for riverine fish: measurement challenges and research opportunities. *Freshwater Biology* **55**, 2215-2237. doi:10.1111/j.1365-2427.2010.02448.x

Fürst J, Hörhan T (2009). Coding of watershed and river hierarchy to support GIS-based hydrological analyses at different scales. *Computers & Geosciences* **35**, 688-696. doi:10.1016/j.cageo.2008.04.007

Galpern P, Manseau M, Fall A (2011). Patch-based graphs of landscape connectivity: A guide to construction, analysis and application for conservation. *Biological Conservation* **144**, 44-55. doi:10.1016/j.biocon.2010.09.002

Ganio LM, Torgersen CE, Gresswell RE (2005). A geostatistical approach for describing spatial pattern in stream networks. *Frontiers in Ecology and the Environment* **3**, 138-144. doi:10.1890/1540-9295(2005)003[0138:AGAFDS]2.0.CO;2

Garcia-Ac A, Segura PA, Gagnon C, Sauvé S (2009). Determination of bezafibrate, methotrexate, cyclophosphamide, orlistat and enalapril in waste and surface waters using on-line solid-phase extraction liquid chromatography coupled to polarity-switching electrospray tandem mass spectrometry. *Journal of environmental monitoring : JEM* **11**, 830-838. doi:10.1039/b817570e

Gaughan AE, Stevens FR, Linard C, Jia P, Tatem AJ (2013). High resolution population distribution maps for Southeast Asia in 2010 and 2015. *PLoS ONE* **8**, e55882. doi:10.1371/journal.pone.0055882

Goldewijk KK, Beusen A, van Drecht G, de Vos M (2011). The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography* **20**, 73-86. doi:10.1111/j.1466-8238.2010.00587.x

Gong L, Halldin S, Xu C-Y (2011). Global-scale river routing-an efficient time-delay algorithm based on HydroSHEDS high-resolution hydrography. *Hydrological Processes* **25**, 1114-1128. doi:10.1002/Hyp.7795

Gong L, Widén-Nilsson E, Halldin S, Xu C-Y (2009). Large-scale runoff routing with an aggregated network-response function. *Journal of Hydrology* **368**, 237-250. doi:10.1016/j.jhydrol.2009.02.007

Gordon E, Meentemeyer R (2006). Effects of dam operation and land use on stream channel morphology and riparian vegetation. *Geomorphology* **82**, 412-429. doi:10.1016/j.geomorph.2006.06.001

Gordon ND (2004). Stream hydrology: an introduction for ecologists, second ed. John Wiley and Sons: West Sussex.

Gosain A, Rao S (2006). Climate change impact assessment on hydrology of Indian river basins. *Current Science* **90**, 346-353

Gosling SN, Arnell NW (2011). Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. *Hydrological Processes* **25**, 1129-1145. doi:10.1002/hyp.7727

Graham ST, Famiglietti JS, Maidment DR (1999). Five-minute, 1/2°, and 1° data sets of continental watersheds and river networks for use in regional and global hydrologic and climate system modeling studies. *Water Resources Research* **35**, 583-587. doi:10.1029/1998wr900068

Grill G, Ouellet Dallaire C, Fluet Chouinard E, Sindorf N, Lehner B (2014). Development of new indicators to evaluate river fragmentation and flow regulation at large scales: A case study for the Mekong River Basin. *Ecological Indicators* **45**, 148-159. doi:10.1016/j.ecolind.2014.03.026

Gupta A, Liew SC (2007). The Mekong from satellite imagery: A quick look at a large river. *Geomorphology* **85**, 259-274. doi:10.1016/j.geomorph.2006.03.036

Haddeland I, Clark DB, Franssen W, Ludwig F, Voss F, Arnell NW, Bertrand N, Best M, Folwell S, Gerten D, Gomes S, Gosling SN, Hagemann S, Hanasaki N, Harding R, Heinke J, Kabat P, Koirala S, Oki T, Polcher J, Stacke T, Viterbo P, Weedon GP, Yeh P (2011). Multimodel estimate of the global terrestrial water balance: setup and first results. *Journal of Hydrometeorology* **12**, 869-884. doi:10.1175/2011jhm1324.1

Haddeland I, Skaugen T, Lettenmaier DP (2007). Hydrologic effects of land and water management in North America and Asia: 1700-1992. *Hydrology and Earth System Sciences* **11**, 1035-1045. doi:10.5194/hess-11-1035-2007

Halls AS, Kshatriya M (2009). Modelling the cumulative barrier and passage effects of mainstream hydropower dams on migratory fish populations in the Lower Mekong Basin, MRC Technical Report No. 25. Mekong River Commission (MRC), Vientiane, Lao PDR.

Hanasaki N, Inuzuka T, Kanae S, Oki T (2010). An estimation of global virtual water flow and sources of water withdrawal for major crops and livestock products using a global hydrological model. *Journal of Hydrology* **384**, 232-244. doi:10.1016/j.jhydrol.2009.09.028

Hanasaki N, Kanae S, Oki T (2006). A reservoir operation scheme for global river routing models. *Journal of Hydrology* **327**, 22-41. doi:10.1016/j.jhydrol.2005.11.011

Hannah R, D'Aco VJ, Anderson PD, Buzby ME, Caldwell DJ, Cunningham VL, Ericson JF, Johnson AC, Parke NJ, Samuelian JH, Sumpter JP (2009). Exposure assessment of 17alpha-ethinylestradiol in surface waters of the United States and Europe. *Environmental Toxicology and Chemistry* **28**, 2725-2732. doi:10.1897/08-622.1

Hanselman TA, Graetz DA, Wilkie AC (2003). Manure-Borne Estrogens as Potential Environmental Contaminants: A Review. *Environmental Science & Technology* **37**, 5471-5478. doi:10.1021/es034410+

Harrison JA, Bouwman L, Mayorga E, Seitzinger S, Bouwman AF (2010). Magnitudes and sources of dissolved inorganic phosphorus inputs to surface fresh waters and the coastal zone: A new global model. *Global Biogeochemical Cycles* **24**, 1-16. doi:10.1029/2009GB003590

Havnø K, Madsen M, Dørge J, Singh V (1995). MIKE 11-a generalized river modelling package, in: Singh V (Ed.), Computer models of watershed hydrology. Water Resources Publications: Colorado, USA, pp. 733-782.

HC/EC (2010). Priority substances list assessment report. Follow-up on the state of science report, 2000: Ethylene Glycol. Environment Canada / Health Canada, Ottawa, Canada.

He B, Kanae S, Oki T, Hirabayashi Y, Yamashiki Y, Takara K (2011). Assessment of global nitrogen pollution in rivers using an integrated biogeochemical modeling framework. *Water Research* **45**, 2573-2586. doi:10.1016/j.watres.2011.02.011

Helfenstein J, Kienast F (2014). Ecosystem service state and trends at the regional to national level: A rapid assessment. *Ecological Indicators* **36**, 11-18. doi:10.1016/j.ecolind.2013.06.031

Hermoso V, Kennard MJ, Linke S (2012). Integrating multidirectional connectivity requirements in systematic conservation planning for freshwater systems. *Diversity and Distributions* **18**, 448-458. doi:10.1111/j.1472-4642.2011.00879.x

Hermoso V, Linke S, Prenda J, Possingham HP (2011). Addressing longitudinal connectivity in the systematic conservation planning of fresh waters. *Freshwater Biology* **56**, 57-70. doi:10.1111/j.1365-2427.2009.02390.x

Hernando MD, Mezcua M, Fernández-Alba A, Barceló D (2006). Environmental risk assessment of pharmaceutical residues in wastewater effluents, surface waters and sediments. *Talanta* **69**, 334-342

Higgins JV, Bryer MT, Khoury ML, Fitzhugh TW (2005). A freshwater classification approach for biodiversity conservation planning. *Conservation Biology* **19**, 432-445. doi:10.1111/j.1523-1739.2005.00504.x

Hodges JEN, Holmes CM, Vamshi R, Mao D, Price OR (2011). Estimating chemical emissions from home and personal care products in China. *Environmental Pollution* **165**, 199-207. doi:10.1016/j.envpol.2011.11.009

Hoenke KM, Kumar M, Batt L (2014). A GIS based approach for prioritizing dams for potential removal. *Ecological Engineering* **64**, 27-36. doi:10.1016/j.ecoleng.2013.12.009

Hollister J, Milstead WB (2010). Using GIS to estimate lake volume from limited data. *Lake and Reservoir Management* **26**, 194-199. doi:10.1080/07438141.2010.504321

Hortle KG (2009). Fisheries of the Mekong River Basin, in: Campbell IC (Ed.), The Mekong. Biophysical Environment of a Transboundary River. Elsevier: New York (NY), pp. 199-253.

Hosseini NA, Parker WJ, Matott LS (2012). Modelling Concentrations of Pharmaceuticals and Personal Care Products in a Canadian Watershed. *Canadian Water Resources Journal* **37**, 191-208. doi:10.4296/cwrj2012-949

Hua W, Bennett ER, Letcher RJ (2006). Ozone treatment and the depletion of detectable pharmaceuticals and atrazine herbicide in drinking water sourced from the upper Detroit River, Ontario, Canada. *Water Research* **40**, 2259-2266. doi:10.1016/j.watres.2006.04.033

ICEM (2010). MRC Strategic Environmental Assessment (SEA) of Hydropower on the Mekong Mainstream. International Centre for Environmental Management (ICEM), Hanoi, Viet Nam.

ICPDR (2009). Danube River Basin District Management Plan. Part A – Basin-wide overview. Doc.No.1C/151. Available at http://www.icpdr.org/icpdr-pages/danube_rbm_plan_ready.htm. International Commission for the Protection of the Danube River (ICPDR), Vienna, Austria.

James WF, Richardson WB, Soballe DM (2008). Effects of residence time on summer nitrate uptake in mississippi river flow-regulated backwaters. *River Research and Applications* **24**, 1206-1217. doi:10.1002/rra.1150

Jansson R, Nilsson C, Renofalt B (2000). Fragmentation of riparian floras in rivers with multiple dams. *Ecology* **81**, 899-903. doi:10.1890/0012-9658(2000)081[0899:Forfir]2.0.Co;2

Jiang W, Yan Y, Ma M, Wang D, Luo Q, Wang Z, Satyanarayanan SK (2012). Assessment of source water contamination by estrogenic disrupting compounds in China. *Journal of Environmental Sciences* **24**, 320-328. doi:10.1016/s1001-0742(11)60746-8

Johnson AC, Dumont E, Williams RJ, Oldenkamp R, Cisowska I, Sumpter JP (2013). Do concentrations of ethinylestradiol, estradiol, and diclofenac in European rivers exceed proposed EU environmental quality standards? *Environmental Science & Technology* **47**, 12297-12304. doi:10.1021/es4030035

Johnson AC, Ternes T, Williams RJ, Sumpter JP (2008). Assessing the concentrations of polar organic microcontaminants from point sources in the aquatic environment: measure or model? *Environmental Science & Technology* **42**, 5390-5399. doi:10.1021/es703091r

Johnson AC, Williams RJ (2004). A model to estimate influent and effluent concentrations of estradiol, estrone, and ethinylestradiol at sewage treatment works. *Environmental Science & Technology* **38**, 3649-3658. doi:10.1021/Es035342u

Johnson WC (1992). Dams and riparian forests: case study from the upper Missouri River. *Rivers* **3**, 229-242

Kang B, Perrett L, Li Y, He D, Bin K, Lisa P, Yungang LI, Daming HE (2009). Are the fish of the upper and lower Mekong interconnected? *Chinese Journal of Oceanology and Limnology* **27**, 400-407. doi:10.1007/s00343-009-9132-0

Kanno Y, Russ WT, Sutherland CJ, Cook SB (2012). Prioritizing aquatic conservation areas using spatial patterns and partitioning of fish community diversity in a near-natural temperate basin. *Aquatic Conservation-Marine and Freshwater Ecosystems* **22**, 799-812. doi:10.1002/Aqc.2266

Keller V, Fox K, Rees HG, Young AR (2006). Estimating population served by sewage treatment works from readily available GIS data. *Science of the Total Environment* **360**, 319-327. doi:10.1016/j.scitotenv.2005.08.043

Keller VD, Williams RJ, Lofthouse C, Johnson AC (2014). Worldwide estimation of river concentrations of any chemical originating from sewage-treatment plants using dilution factors. *Environmental Toxicology and Chemistry* **33**, 447-452. doi:10.1002/etc.2441

Kemp PS, O'Hanley JR (2010). Procedures for evaluating and prioritising the removal of fish passage barriers: a synthesis. *Fisheries Management and Ecology* **17**, 297-322. doi:10.1111/j.1365-2400.2010.00751.x

Kennard MJ, Pusey BJ, Olden JD, Mackay SJ, Stein JL, Marsh N (2010). Classification of natural flow regimes in Australia to support environmental flow management. *Freshwater Biology* **55**, 171-193. doi:10.1111/j.1365-2427.2009.02307.x

Khan U (2014). Pharamceutically Active Compounds in the Environment: Risks, Trade-offs and Sewer Epidemology, Department of Civil Engineering and Applied Mechanics. McGill University, Montreal, QC.

Khan U, Nicell JA (2014). Contraceptive options and their associated estrogenic environmental loads: relationships and trade-offs. *PLoS ONE* **9**, e92630. doi:10.1371/journal.pone.0092630

Kidd KA, Blanchfield PJ, Mills KH, Palace VP, Evans RE, Lazorchak JM, Flick RW (2007). Collapse of a fish population after exposure to a synthetic estrogen. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8897-8901. doi:10.1073/pnas.0609568104

Kindlmann P, Burel F (2008). Connectivity measures: a review. *Landscape Ecology* **23**, 879-890. doi:10.1007/s10980-008-9245-4

Kormos JL (2007). Occurrence and seasonal variability of selected pharmaceuticals in southern Ontario drinking water supplies. University of Waterloo, Waterloo, Ontario.

Kümmerer K (2011). Emerging Contaminants, in: Wilderer P (Ed.), Treatise on Water Science. Elsevier: Oxford, pp. 69-87.

Kummu M, Lu XX, Wang JJ, Varis O (2010). Basin-wide sediment trapping efficiency of emerging reservoirs along the Mekong. *Geomorphology* **119**, 181-197. doi:10.1016/j.geomorph.2010.03.018

Kundzewicz Z, Mata L, NW A, Döll P, Kabat P, Jiménez B, Miller KA, Oki T, Sen Z, Shiklomanov I (2007). Freshwater resources and their management, in: Parry ML, Canziani OF, Palutikof JP, Linden PJvd, Hanson CE (Eds.), Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press: Cambridge, UK, pp. 173-210.

Kunin WE (1998). Extrapolating species abundance across spatial scales. *Science* **281**, 1513-1515. doi:10.1126/science.281.5382.1513

Lajeunesse A, Gagnon C (2007). Determination of acidic pharmaceutical products and carbamazepine in roughly primary-treated wastewater by solid-phase extraction and gas chromatography-tandem mass spectrometry. *International Journal of Environmental Analytical Chemistry* **87**, 565-578. doi:10.1080/03067310701189083

LandScan (2006). Landscan global population database. Available at: http://www.ornl.gov/landscan (accessed June 2008). Oak Ridge National Laboratory, Oak Ridge, TN.

Lassalle G, Crouzet P, Rochard E (2009). Modelling the current distribution of European diadromous fishes: an approach integrating regional anthropogenic pressures. *Freshwater Biology* **54**, 587-606. doi:10.1111/j.1365-2427.2008.02135.x

Lazar AN, Butterfield D, Futter MN, Rankinen K, Thouvenot-Korppoo M, Jarritt N, Lawrence DS, Wade AJ, Whitehead PG (2010). An assessment of the fine sediment dynamics in an upland river system: INCA-Sed modifications and implications for fisheries. *Science of the Total Environment* **408**, 2555-2566

Lehner B, Döll P (2004). Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology* **296**, 1-22. doi:10.1016/j.jhydrol.2004.03.028

Lehner B, Grill G (2013). Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes* **27**, 2171-2186. doi:10.1002/hyp.9740

Lehner B, Liermann CR, Revenga C, Vörösmarty C, Fekete B, Crouzet P, Döll P, Endejan M, Frenken K, Magome J, Nilsson C, Robertson JC, Rödel R, Sindorf N, Wisser D (2011). High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment* **9**, 494-502. doi:10.1890/100125

Lehner B, Nicell J, Grill G, Khan U, Ariwi J (2013). Down-the-drain geospatial fate model for substances in consumer products: a pilot study for the provinces of Quebec and Ontario. McGill University, Montreal, Canada.

Lehner B, Verdin KL, Jarvis A (2006). HydroSHEDS Technical Documentation v1.0. World Wildlife Fund US, Washington, DC, pp. 1-27.

Lehner B, Verdin KL, Jarvis A (2008). New global hydrography derived from spaceborne elevation data. *EOS, Transactions of the American Geophysical Union* **89**, 93-104. doi:10.1029/2008EO100001

Li H, Helm PA, Metcalfe CD (2010). Sampling in the Great Lakes for pharmaceuticals, personal care products, and endocrine-disrupting substances using the passive polar organic chemical integrative sampler. *Environmental Toxicology and Chemistry* **29**, 751-762. doi:10.1002/etc.104

Liang X, Lettenmaier DP, Wood EF, Burges SJ (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research* **99**, 415-414

Lissemore L, Hao C, Yang P, Sibley PK, Mabury S, Solomon KR (2006). An exposure assessment for selected pharmaceuticals within a watershed in Southern Ontario. *Chemosphere* **64**, 717-729. doi:10.1016/j.chemosphere.2005.11.015

Liu J, Zang C, Tian S, Liu J, Yang H, Jia S, You L, Liu B, Zhang M (2013). Water conservancy projects in China: Achievements, challenges and way forward. *Global Environmental Change* **23**, 633-643. doi:10.1016/j.gloenvcha.2013.02.002

Liu X, Shi J, Zhang H, Zhan X, Shen G, Hu S (2014). Estimating estrogen release and load from humans and livestock in Shanghai, China. *Journal of Environment Quality* **43**, 568-577. doi:10.2134/jeq2013.08.0328

Liu Z, Wang H, Carmichael PL, Deag EJ, Duarte-Davidson R, Li H, Howe P, Meng W, Price OR, Shen Y, Shore RF, Sweetman AJ, Yan ZG, Zang WC, Undeman E, Jones KC (2012). China begins to position for leadership on responsible risk-based global chemicals management. *Environmental Pollution* **165**, 170-173. doi:10.1016/j.envpol.2012.02.024

Llovel W, Becker M, Cazenave A, Crétaux J-F, Ramillien G (2010). Global land water storage change from GRACE over 2002–2009; Inference on sea level. *Comptes Rendus Geoscience* **342**, 179-188. doi:10.1016/j.crte.2009.12.004

Loveland TR, Reed BC, Brown JF, Ohlen DO, Zhu Z, Yang L, Merchant JW (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *International Journal of Remote Sensing* **21**, 1303-1330. doi:10.1080/014311600210191

Lowe WH, Likens GE (2005). Moving headwater streams to the head of the class. *Bioscience* **55**, 196-197

Lowe WH, Likens GE, Power ME (2006). Linking scales in stream ecology. *Bioscience* **56**, 591-597. doi:10.1641/0006-3568(2006)56[591:Lsise]2.0.Co;2

MacDonald GK, Bennett EM, Potter PA, Ramankutty N (2011). Agronomic phosphorus imbalances across the world's croplands. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 3086-3091. doi:10.1073/pnas.1010808108

MacLeod SL, McClure EL, Wong CS (2007). Laboratory calibration and field deployment of the polar organic chemical integrative sampler for pharmaceuticals and personal care products in wastewater and surface water. *Environmental toxicology and chemistry / SETAC* **26**, 2517-2529. doi:10.1897/07-238.1

Magalhaes MF, Beja P, Canas C, Collares-Pereira MJ (2002). Functional heterogeneity of dry-season fish refugia across a Mediterranean catchment: the role of habitat and predation. *Freshwater Biology* **47**, 1919-1934. doi:10.1046/j.1365-2427.2002.00941.x

Maidment DR (1993). Handbook of hydrology. McGraw-Hill Professional: New York.

Maidment DR (2002). Arc Hydro: GIS for water resources. http://www.crwr.utexas.edu/gis/archydrobook/ArcHydro.htm. ESRI Press: Redlands, California.

Mari L, Casagrandi R, Bertuzzo E, Rinaldo A, Gatto M (2014). Metapopulation persistence and species spread in river networks. *Ecology Letters* **17**, 426-434. doi:10.1111/ele.12242

Marinov D, Pistocchi A, Trombetti M, Bidoglio G (2014). Assessment of riverine load of contaminants to European seas under policy implementation scenarios: an example with 3 pilot substances. *Integr. Environ. Assess. Manag.* **10**, 48-59. doi:10.1002/ieam.1459

Matthiessen P, Arnold D, Johnson AC, Pepper TJ, Pottinger TG, Pulman KG (2006). Contamination of headwater streams in the United Kingdom by oestrogenic hormones from livestock farms. *Science of the Total Environment* **367**, 616-630. doi:10.1016/j.scitotenv.2006.02.007

Mayorga E, Seitzinger SP, Harrison JA, Dumont E, Beusen AHW, Bouwman AF, Fekete B, Kroeze C, Van Drecht G (2010). Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation. *Environmental Modelling & Software* **25**, 837-853. doi:10.1016/j.envsoft.2010.01.007

McDonald RI, Weber K, Padowski J, Flörke M, Schneider C, Green PA, Gleeson T, Eckman S, Lehner B, Balk D, Boucher T, Grill G, Montgomery M (2014). Water on an urban planet: Urbanization and the reach of urban water infrastructure. *Global Environmental Change* **27**, 96-105. doi:10.1016/j.gloenvcha.2014.04.022

McLaughlin F, Lapointe M, Bourque G, Boisclair D (2014). Using regional flow classes as references to analyse flow regime anomalies across a set of regulated Canadian rivers. *Journal of Hydrology* **519**, 307-328. doi:10.1016/j.jhydrol.2014.07.015

McManamay RA (2014). Quantifying and Generalizing Hydrologic Responses to Dam Regulation using a Statistical Modeling Approach. *Journal of Hydrology* **519 A**. doi:10.1016/j.jhydrol.2014.08.053

MEP (2011). National list of urban sewage treatment facilities in operation. Downloaded at http://www.mep.gov.cn/gkml/hbb/bgg/201104/ W020110420407642318640.pdf (April, 2013). Ministry of Environmental Protection of China (MEP).

Metcalfe CD, Koenig BG, Bennie DT, Servos M, Ternes TA, Hirsch R (2003). Occurrence of neutral and acidic drugs in the effluents of Canadian sewage treatment plants. *Environmental Toxicology and Chemistry* **22**, 2872-2880. doi:10.1897/02-469

Miao X-S, Metcalfe CD (2003). Determination of carbamazepine and its metabolites in aqueous samples using liquid chromatography-electrospray tandem mass spectrometry. *Analytical Chemistry* **75**, 3731-3738

Mihelcic JR, Zimmerman JB, Auer MT (2010). Environmental engineering: Fundamentals, sustainability, design. Wiley Hoboken: NJ, USA.

Millennium Ecosystem Assessment (2005). Ecosystems and human well-being: wetlands and water synthesis. www.MAweb.org. World Resources Institute, Washington, DC.

MOHURD (2012). State of China's urban drainage and sewage treatment 2006-2010. Online at: http://www.mohurd.gov.cn/xytj/tjzljsxytjgb/ w020120515197618301527 01250.docx (April, 2013). Ministry of Housing and Urban-Rural Development (MOHURD), Beijing, China.

Moilanen A, Hanski I (2001). On the use of connectivity measures in spatial ecology. *Oikos* **95**, 147-151. doi:10.1034/j.1600-0706.2001.950116.x

Monfreda C, Ramankutty N, Foley JA (2008). Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year. *Global Biogeochemical Cycles* **22**, 1-19. doi:10.1029/2007gb002947

Moriasi D, Arnold J (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers* **50**, 885-900

MRC (2003). Fish Migration in the Mekong River Basin. MRC Technical Report. Mekong River Commission (MRC), Vientiane, Lao PDR.

MRC (2009). Hydropower Project Database. Mekong River Commission (MRC), Vientiane, Lao PDR.

MRC (2010). Initiative on Sustainable Hydropower (ISH) for 2011-2015. Available at (Aug. 2012) http://www.mrcmekong.org/assets/Publications/Programme-Documents/2011-2015-Initiative-Final-Version-30-OctISH.pdf. Mekong River Commission (MRC), Vientiane, Lao PDR.

Muneepeerakul R, Weitz JS, Levin SA, Rinaldo A, Rodriguez-Iturbe I (2007). A neutral metapopulation model of biodiversity in river networks. *Journal of Theoretical Biology* **245**, 351-363. doi:10.1016/j.jtbi.2006.10.005

Naiman R, Latterell J, Pettit N, Olden J (2008). Flow variability and the biophysical vitality of river systems. *Comptes Rendus Geosciences* **340**, 629-643. doi:10.1016/j.crte.2008.01.002

NASA/NGA (2003). SRTM Water Body Data. v2.1. Digital Media. Available at: http://dds.cr.usgs.gov/srtm/version2_1/SWBD/, in: NASA/NGA (Ed.).

Nash J, Sutcliffe J (1970). River flow forecasting through conceptual models part I: A discussion of principles. *Journal of Hydrology* **10**, 282-290

Ngo-Duc T, Oki T, Kanae S (2007). A variable streamflow velocity method for global river routing model: model description and preliminary results. *Hydrology and Earth System Sciences Discussions* **4**, 4389-4414. doi:10.5194/hessd-4-4389-2007

Nilsson C, Brown RL, Jansson R, Merritt DM (2010). The role of hydrochory in structuring riparian and wetland vegetation. *Biological Reviews of the Cambridge Philosophical Society* **85**, 837-858. doi:10.1111/j.1469-185X.2010.00129.x

Nilsson C, Dynesius M (1994). Ecological effects of river regulation on mammals and birds: a review. *Regulated Rivers-Research & Management* **9**, 45-53. doi:10.1002/rrr.3450090105

Nilsson C, Reidy CA, Dynesius M, Revenga C (2005). Fragmentation and flow regulation of the world's large river systems. *Science* **308**, 405-408. doi:10.1126/science.1107887

O'Hanley JR (2011). Open rivers: barrier removal planning and the restoration of free-flowing rivers. *Journal of Environmental Management* **92**, 3112-3120. doi:10.1016/j.jenvman.2011.07.027

Oki T, Agata Y, Kanae S, Saruhashi T, Yang D, Musiake K (2001). Global assessment of current water resources using total runoff integrating pathways. *Hydrological Sciences Journal* **46**, 983-995. doi:10.1080/02626660109492890

Oki T, Kanae S (2006). Global hydrological cycles and world water resources. *Science* **313**, 1068-1072. doi:10.1126/science.1128845

Oki T, Sud YC (1998). Design of Total Runoff Integrating Pathways (TRIP)—A Global River Channel Network. *Earth Interactions* **2**, 1-37

Olden JD, Kennard MJ, Pusey BJ (2012). A framework for hydrologic classification with a review of methodologies and applications in ecohydrology. *Ecohydrology* **5**, 503-518. doi:10.1002/eco.251

Olivera F, Famiglietti J (2000). Global-scale flow routing using a source-to-sink algorithm. *Water Resources Research* **36**, 2197-2207

Olivera F, Maidment DR (1999). Geographic information systems (GIS)-based spatially distributed model for runoff routing. *Water Resources Research* **35**, 1155-1164

Ongley ED, Xiaolan Z, Tao Y (2010). Current status of agricultural and rural non-point source pollution assessment in China. *Environmental Pollution* **158**, 1159-1168. doi:10.1016/j.envpol.2009.10.047

Orr S, Pittock J, Chapagain A, Dumaresq D (2012). Dams on the Mekong River: lost fish protein and the implications for land and water resources. *Global Environmental Change* **22**, 925-932. doi:10.1016/j.gloenvcha.2012.06.002

Ort C, Hollender J, Schaerer M, Siegrist H (2009). Model-based evaluation of reduction strategies for micropollutants from wastewater treatment plants in complex river networks. *Environmental Science & Technology* **43**, 3214-3220. doi:10.1021/es802286v

Ostroff A, Wieferich D, Cooper A, Infante DM (2013). National Anthropogenic Barrier Dataset (NABD). United States Geological Service (USGS) Aquatic GAP Program, Denver, Colorado.

Padgham M, Webb JA (2010). Multiple structural modifications to dendritic ecological networks produce simple responses. *Ecological Modelling* **221**, 2537-2545. doi:10.1016/j.ecolmodel.2010.04.022

Paiva RCD, Buarque DC, Collischonn W, Bonnet M-P, Frappart F, Calmant S, Bulhões Mendes CA (2013a). Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin. *Water Resources Research* **49**, 1226-1243. doi:10.1002/wrcr.20067

Paiva RCD, Collischonn W, Buarque DC (2013b). Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon. *Hydrological Processes* **27**, 333-346. doi:10.1002/hyp.8425

Paiva RCD, Collischonn W, Tucci CEM (2011). Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach. *Journal of Hydrology* **406**, 170-181. doi:10.1016/j.jhydrol.2011.06.007

Palmer MA, Reidy Liermann CA, Nilsson C, Flörke M, Alcamo JM, Lake PS, Bond N (2008). Climate change and the world's river basins: anticipating management options. *Frontiers in Ecology and the Environment* **6**, 81-89. doi:10.1890/060148

Pascual-Hortal L, Saura S (2006). Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. *Landscape Ecology* **21**, 959-967. doi:10.1007/s10980-006-0013-z

Peterson EE, Ver Hoef JM, Isaak DJ, Falke JA, Fortin MJ, Jordan CE, McNyset K, Monestiez P, Ruesch AS, Sengupta A, Som N, Steel EA, Theobald DM, Torgersen CE, Wenger SJ (2013). Modelling dendritic ecological networks in space: an integrated network perspective. *Ecology Letters* **16**, 707-719. doi:10.1111/ele.12084

Petts GE, Gurnell AM (2005). Dams and geomorphology: research progress and future directions. *Geomorphology* **71**, 27-47

Phillips RW, Spence C, Pomeroy JW (2011). Connectivity and runoff dynamics in heterogeneous basins. *Hydrological Processes* **3075**, 3061-3075. doi:10.1002/hyp.8123

Pinto-Coelho RM (1998). Effects of eutrophication on seasonal patterns of mesozooplankton in a tropical reservoir: a 4-year study in Pampulha Lake, Brazil. *Freshwater Biology* **40**, 159-173. doi:10.1046/j.1365-2427.1998.00327.x

Pistocchi A, Groenwold J, Lahr J, Loos M, Mujica M, Ragas AMJ, Rallo R, Sala S, Schlink U, Strebel K, Vighi M, Vizcaino P (2011a). Mapping Cumulative Environmental Risks: Examples from the EU NoMiracle Project. *Environmental Modeling and Assessment* **16**, 119-133. doi:10.1007/s10666-010-9230-6

Pistocchi A, Marinov D, Pontes S, Gawlik BM (2012). Continental scale inverse modeling of common organic water contaminants in European rivers. *Environmental Pollution* **162**, 159-167. doi:10.1016/j.envpol.2011.10.031

Pistocchi A, Marinov D, Pontes S, Zulian G, Trombetti M (2011b). Multimedia assessment of pollutant pathways in the environment: a Global scale model. Joint Research Center / Institute for Environment and Sustainability (JRC-IES), Ispra, Italy.

Pistocchi A, Pennington D (2006). European hydraulic geometries for continental scale environmental modelling. *Journal of Hydrology* **329**, 553-567. doi:10.1016/j.jhydrol.2006.03.009

Pistocchi A, Sarigiannis DA, Vizcaino P (2010). Spatially explicit multimedia fate models for pollutants in Europe: state of the art and perspectives. *The Science of The Total Environment* **408**, 3817-3830. doi:10.1016/j.scitotenv.2009.10.046

Pistocchi A, Vizcaino P, Hauck M (2009). A GIS model-based screening of potential contamination of soil and water by pyrethroids in Europe. *Journal of Environmental Management* **90**, 3410-3421. doi:10.1016/j.jenvman.2009.05.020

Poff NL, Allen JD, Bain MB, Karr JR, Prestegaard KL, Richter BD, Sparks RE, Stromberg JC (1997). The natural flow regime – a paradigm for river conservation and restoration. *Bioscience* **47**, 769-784

Poff NL, Olden JD, Merritt DM, Pepin DM (2007). Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5732-5737. doi:10.1073/pnas.0609812104

Poff NL, Richter BD, Arthington AH, Bunn SE, Naiman RJ, Kendy E, Acreman M, Apse C, Bledsoe BP, Freeman MC, Henriksen J, Jacobson RB, Kennen JG, Merritt DM, O'Keeffe JH, Olden JD, Rogers K, Tharme RE, Warner A (2010). The ecological limits of hydrologic alteration

(ELOHA): a new framework for developing regional environmental flow standards. *Freshwater Biology* **55**, 147-170. doi:10.1111/j.1365-2427.2009.02204.x

Poff NL, Zimmerman JKH (2010). Ecological responses to altered flow regimes: a literature review to inform the science and management of environmental flows. *Freshwater Biology* **55**, 194-205. doi:10.1111/j.1365-2427.2009.02272.x

Poole GC, Berman CH (2001). An ecological perspective on in-stream temperature: natural heat dynamics and mechanisms of human-caused thermal degradation. *Environmental Management* **27**, 787-802. doi:10.1007/s002670010188

Portmann FT, Siebert S, Döll P (2010). MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles* **24**, 1-24. doi:10.1029/2008GB003435

Poulsen AF, Poeu O, Viravong S, Suntornratana U, Tung NTT (2002). Deep pools as dry season fish habitats in the Mekong River basin, MRC technical paper. Mekong River Commission (MRC), Phnom Penh.

Prigent C, Papa F, Aires F, Rossow WB, Matthews E (2007). Global inundation dynamics inferred from multiple satellite observations, 1993-2000. *Journal of Geophysical Research* **112**, 1-13. doi:10.1029/2006JD007847

Pringle CM (2001). Hydrologic connectivity and the management of biological reserves: a global perspective. *Ecological Applications* **11**, 981-998

Pringle CM (2003). What is hydrologic connectivity and why is it ecologically important? *Hydrological Processes* **17**, 2685-2689. doi:10.1002/hyp.5145

Pringle CM, Freeman MC, Freeman BJ (2000). Regional effects of hydrologic alterations on riverine macrobiota in the New World: tropical-temperate comparisons. *Bioscience* **50**, 807-823. doi:10.1641/0006-3568(2000)050[0807:Reohao]2.0.Co;2

Rahman MF, Yanful EK, Jasim SY, Bragg LM, Servos MR, Ndiongue S, Borikar D (2010). Advanced oxidation treatment of drinking water: Part I. Occurrence and removal of pharmaceuticals and endocrine-disrupting compounds from Lake Huron water. *Ozone: Science & Engineering* **32**, 217-229. doi:10.1080/01919512.2010.489185

Ramankutty N, Foley JA (1999). Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles* **13**, 997-1027. doi:10.1029/1999gb900046

Rayfield B, Fortin M-J, Fall A (2011). Connectivity for conservation: a framework to classify network measures. *Ecology* **92**, 847-858. doi:10.1890/09-2190.1

Reidy Liermann CR, Nilsson C, Robertson J, Ng RY (2012). Implications of dam obstruction for global freshwater fish diversity. *Bioscience* **62**, 539-548. doi:10.1525/bio.2012.62.6.5

Renard P, Allard D (2013). Connectivity metrics for subsurface flow and transport. *Advances in Water Resources* **51**, 168-196. doi:10.1016/j.advwatres.2011.12.001

Renssen H, Knoop JM (2000). A global river routing network for use in hydrological modeling. *Journal of Hydrology* **230**, 230-243. doi:10.1016/S0022-1694(00)00178-5

Richardson BJ, Lam PK, Martin M (2005). Emerging chemicals of concern: pharmaceuticals and personal care products (PPCPs) in Asia, with particular reference to Southern China. *Marine Pollution Bulletin* **50**, 913-920. doi:10.1016/j.marpolbul.2005.06.034

Richter BD, Baumgartner JV, Powell J, Braun DP (1996). A method for assessing hydrologic alteration within ecosystems. *Conservation Biology* **10**, 1163-1174. doi:DOI 10.1046/j.1523-1739.1996.10041163.x

Richter BD, Postel S, Revenga C, Scudder T, Lehner B, Churchill A, Chow M (2010). Lost in development's shadow: the downstream human consequences of dams. *Water Alternatives* **3**, 14-42

Robinson N, Regetz J, Guralnick RP (2014). EarthEnv-DEM90: A nearly-global, void-free, multi-scale smoothed, 90m digital elevation model from fused ASTER and SRTM data. *ISPRS Journal of Photogrammetry and Remote Sensing* **87**, 57-67. doi:10.1016/j.isprsjprs.2013.11.002

Rosenberg DM, Berkes F, Bodaly RA, Hecky RE, Kelly CA, Rudd JWM (1997). Large-scale impacts of hydroelectric development. *Environmental Reviews* **54**, 27-54

Rost S, Gerten D, Bondeau A, Lucht W, Rohwer J, Schaphoff S (2008). Agricultural green and blue water consumption and its influence on the global water system. *Water Resources Research* **44**, 1-17. doi:10.1029/2007WR006331

Russell GL, Miller JR (1990). Global river runoff calculated from a global atmospheric general circulation model. *Journal of Hydrology* **117**, 241-254. doi:10.1016/0022-1694(90)90095-F

Sahoo AK, Pan M, Troy TJ, Vinukollu RK, Sheffield J, Wood EF (2011). Reconciling the global terrestrial water budget using satellite remote sensing. *Remote Sensing of Environment* **115**, 1850-1865. doi:10.1016/j.rse.2011.03.009

SCHER (2011). Opinion on draft environmental quality standards under the Water Framework Directive –Ethinylestradiol, 30 March 2011. Scientific Committee on Health and Environmental Risks (SCHER).

Schick RS, Lindley ST (2007). Directed connectivity among fish populations in a riverine network. *Journal of Applied Ecology* **44**, 1116-1126. doi:10.1111/j.1365-2664.2007.01383.x

Schlosser IJ (1982). Fish community structure and function along two habitat gradients in a headwater stream. *Ecological Monographs* **52**, 395-414. doi:10.2307/2937352

Schneider A, Friedl MA, Potere D (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters* **4**, 044003. doi:10.1088/1748-9326/4/4/044003

Schneider A, Friedl MA, Potere D (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sensing of Environment* **114**, 1733-1746

Schneider C, Flörke M, Eisner S, Voss F (2011). Large scale modelling of bankfull flow: An example for Europe. *Journal of Hydrology* **408**, 235-245. doi:10.1016/j.jhydrol.2011.08.004

Schuol J, Abbaspour KC, Srinivasan R, Yang H (2008). Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model. *Journal of Hydrology* **352**, 30-49. doi:10.1016/j.jhydrol.2007.12.025

Schwarzenbach RP, Egli T, Hofstetter TB, von Gunten U, Wehrli B (2010). Global water pollution and human health. *Annual Review of Environment and Resources* **35**, 109-136. doi:10.1146/annurev-environ-100809-125342

Schwarzenbach RP, Escher BI, Fenner K, Hofstetter TB, Johnson CA, von Gunten U, Wehrli B (2006). The challenge of micropollutants in aquatic systems. *Science* **313**, 1072-1077. doi:10.1126/science.1127291

Segurado P, Branco P, Ferreira MT (2013). Prioritizing restoration of structural connectivity in rivers: a graph based approach. *Landscape Ecology* **28**, 1231-1238. doi:10.1007/s10980-013-9883-z

SEPA (2013). State of the environment China. SEPA (State Environmental Protection Administration), Beijing, China.

Serna-Chavez HM, Schulp CJE, van Bodegom PM, Bouten W, Verburg PH, Davidson MD (2014). A quantitative framework for assessing spatial flows of ecosystem services. *Ecological Indicators* **39**, 24-33. doi:10.1016/j.ecolind.2013.11.024

Seyler F, Calmant S, da Silva JS, Moreira DM, Mercier F, Shum CK (2013). From TOPEX/Poseidon to Jason-2/OSTM in the Amazon basin. *Advances in Space Research* **51**, 1542-1550. doi:10.1016/j.asr.2012.11.002

Shanahan P, Harleman DRF (1984). Transport in lake water quality modeling. *Journal of Environmental Engineering* **110**, 42-57. doi:10.1061/(asce)0733-9372(1984)110:1(42)

Shao M, Tang XY, Zhang YH, Li WJ (2006). City clusters in China: air and surface water pollution. *Frontiers in Ecology and the Environment* **4**, 353-361. doi:10.1890/1540-9295(2006)004[0353:Ccicaa]2.0.Co;2

Shea CP, Peterson JT (2007). An evaluation of the relative influence of habitat complexity and habitat stability on fish assemblage structure in unregulated and regulated reaches of a large southeastern warmwater stream. *Transactions of the American Fisheries Society* **136**, 943-958. doi:10.1577/t06-165.1

Shields Jr FD, Simon A, Steffen LJ (2000). Reservoir effects on downstream river channel migration. *Environmental Conservation* **27**, 54-66. doi:10.1017/s0376892900000072

Shiklomanov IA (2000). Appraisal and assessment of world water resources. *Water International* **25**, 11-32. doi:10.1080/02508060008686794

Siebert S, Döll P (2010). Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation. *Journal of Hydrology* **384**, 198-217. doi:10.1016/j.jhydrol.2009.07.031

Siebert S, Döll P, Hoogeveen J, Faures JM, Frenken K, Feick S (2005). Development and validation of the global map of irrigation areas. *Hydrology and Earth System Sciences* **9**, 535-547. doi:10.5194/hessd-2-1299-2005

Sindorf N, Wickel B (2011). Connectivity and Fragmentation – Hydrospatial Analysis of Dam Development in the Mekong River Basin. World Wildlife Fund, Conservation Science Program, Washington DC.

Snelder TH, Biggs BJF (2002). Multiscale river environment classification for water resources management. *Journal of the American Water Resources Association* **38**, 1225-1239. doi:10.1111/j.1752-1688.2002.tb04344.x

Snelder TH, Biggs BJF, Woods RA (2005). Improved eco-hydrological classification of rivers. *River Research and Applications* **21**, 609-628. doi:10.1002/rra.826

Stanford JA, Ward JV (1993). An ecosystem perspective of alluvial rivers: connectivity and the hyporheic corridor. *Journal of the North American Benthological Society* **12**, 48-60. doi:10.2307/1467685

Stein JL, Hutchinson MF, Stein JA (2014). A new stream and nested catchment framework for Australia. *Hydrology and Earth System Sciences* **18**, 1917-1933. doi:10.5194/hess-18-1917-2014

Sverdrup-Jensen S (2002). Fisheries in the Lower Mekong Basin: status and perspectives, MRC technical paper No. 6. Mekong River Commission (MRC), Phnom Penh.

Syvitski JPM, Kettner AJ, Overeem I, Hutton EWH, Hannon MT, Brakenridge GR, Day J, Vörösmarty C, Saito Y, Giosan L, Nicholls RJ (2009). Sinking deltas due to human activities. *Nature Geoscience* **2**, 681-686. doi:10.1038/Ngeo629

Tabe S, Yang P, Zhao X, Hao C, Seth R (2010). Occurrence and Removal of PPCPs and EDCs in the Detroit River Watershed. *Water Practise & Technology* **5**, 1-8. doi:10.2166/WPT.2010.015

Thieme M, Lehner B, Abell R, Hamilton SK, Kellndorfer J, Powell G, Riveros JC (2007). Freshwater conservation planning in data-poor areas: An example from a remote Amazonian basin (Madre de Dios River, Peru and Bolivia). *Biological Conservation* **135**, 484-501. doi:10.1016/j.biocon.2006.10.054

Thorp JH, Delong MD (1994). The Riverine Productivity Model: An Heuristic View of Carbon Sources and Organic Processing in Large River Ecosystems. *Oikos* **70**, 305. doi:10.2307/3545642

Thorp JH, Thoms MC, Delong MD (2006). The riverine ecosystem synthesis: Biocomplexity in river networks across space and time. *River Research and Applications* **22**, 123-147. doi:10.1002/Rra.901

Thorpe KL, Hutchinson TH, Hetheridge MJ, Scholze M, Sumpter JP, Tyler CR (2001). Assessing the biological potency of binary mixtures of environmental estrogens using vitellogenin induction in juvenile rainbow trout (Oncorhynchus mykiss). *Environmental Science & Technology* **35**, 2476-2481. doi:10.1021/es001767u

Tischendorf L, Fahrig L (2000). On the usage and measurement of landscape connectivity. *Oikos* **90**, 7-19. doi:10.1034/j.1600-0706.2000.900102.x

Tockner K, Pennetzdorfer D, Reiner N, Schiemer F, Ward JV (1999). Hydrological connectivity, and the exchange of organic matter and nutrients in a dynamic river–floodplain system (Danube, Austria). *Freshwater Biology* **41**, 521-535

Tockner K, Stanford JA (2002). Riverine flood plains: present state and future trends. *Environmental Conservation* **29**, 308-330. doi:10.1017/S037689290200022X

Torrente-Vilara G, Zuanon J, Leprieur F, Oberdorff T, Tedesco PA (2011). Effects of natural rapids and waterfalls on fish assemblage structure in the Madeira River (Amazon Basin). *Ecology of Freshwater Fish* **20**, 588-597. doi:10.1111/j.1600-0633.2011.00508.x

Urban D, Keitt T (2001). Landscape connectivity: a graph-theoretic perspective. *Ecology* **82**, 1205-1218. doi:10.1890/0012-9658(2001)082[1205:lcagtp]2.0.co;2

USEPA (1996). Environmental indicators of water quality in the United States. EPA Number: 841R96002. US Environmental Protection Agency (USEPA), Washington, DC.

USEPA and USGS (2008). NHDPlus, NHDPlus User Guide, online: http://horizon-systems.com/NHDPlus/index.php. United States Environmental Protection Agency and United States Geological Survey, Washington, DC.

USGS (2000). HYDRO1k Elevation Derivative Database. *U.S. Geological Survey Center for Earth Resources Observation and Science (EROS): Sioux Falls, S. D. (Available at http://eros.usgs.gov/#/Find_Data/Products_and_Data_Available/gtopo30/hydro)*

Vannote RL, Minshall GW, Cummins KW, Sedell JR, Cushing CE (1980). The river continuum concept. *Canadian Journal of Fsheries and Aquatic Sciences* **37**, 130-137

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**, 105-116. doi:10.1016/j.apgeog.2014.05.012

Verdin KL, Verdin JP (1999). A topological system for delineation and codification of the Earth's river basins. *Journal of Hydrology* **218**, 1-12. doi:10.1016/S0022-1694(99)00011-6

Vermeire TG, Jager DT, Bussian B, Devillers J, den Haan K, Hansen B, Lundberg I, Niessen H, Robertson S, Tyle H, vander Zandt PTJ (1997). European Union system for the evaluation of substances (EUSES). Principles and structure. *Chemosphere* **34**, 1823–1836

Verzano K, Barlund I, Flörke M, Lehner B, Kynast E, Voss F, Alcamo J (2012). Modeling variable river flow velocity on continental scale: Current situation and climate change impacts in Europe. *Journal of Hydrology* **424**, 238-251. doi:10.1016/j.jhydrol.2012.01.005

Vieux BE, Needham S (1993). Nonpoint-pollution model sensitivity to grid-cell size. *Journal of Water Resources Planning and Management* **119**, 141-157. doi:10.1061/(Asce)0733-9496(1993)119:2(141)

Viglino L, Aboulfadl K, Mahvelat AD, Prévost M, Sauvé S (2008). On-line solid phase extraction and liquid chromatography/tandem mass spectrometry to quantify pharmaceuticals, pesticides and some metabolites in wastewaters, drinking, and surface waters. *Journal of Environmental Monitoring* **10**, 482-489. doi:10.1039/b800861b

Viglino L, Aboulfald K, Prévost M, Sauvé S (2009). Detection of Pharmaceuticals, Personal Care Products, and Pesticides in Water Resources by APCI-LC-MS / MS, Application Note. Thermo Fisher Scientific Inc., San Jose, USA.

Visser TAM, Valbo-Jørgensen J, Ratanachookmanee T (2003). Mekong Fish Database. Mekong River Commission (MRC), Phnom Penh.

Vogt JV, Soille P, De Jager A, Rimavičiūtė E, Mehl W, Haastrup P, Paracchini MML, Dusart J, Bódis K, Foisneau S, Bamps C (2007). Developing a pan-European database of drainage networks and catchment boundaries from a 100 metre DEM, in: Wachowicz M, Bodum L (Eds.), Proceedings 10th AGILE Int. Conference on Geographic Information Science, Aalborg, Denmark, pp. 8-11.

Vörösmarty CJ, Fekete BM, Meybeck M, Lammers RB (2000a). Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution. *Journal of Hydrology* **237**, 17-39. doi:10.1016/s0022-1694(00)00282-1

Vörösmarty CJ, Green P, Salisbury J, Lammers RB (2000b). Global water resources: vulnerability from climate change and population growth. *Science* **289**, 284-288

Vörösmarty CJ, McIntyre PB, Gessner MO, Dudgeon D, Prusevich A, Green P, Glidden S, Bunn SE, Sullivan CA, Reidy Liermann C, Davies PM (2010). Global threats to human water security and river biodiversity. *Nature* **467**, 555-561. doi:10.1038/nature09440

Vörösmarty CJ, Meybeck M, Fekete B, Sharma K, Green P, Syvitski JPM (2003). Anthropogenic sediment retention: major global impact from registered river impoundments. *Global and Planetary Change* **39**, 169-190. doi:10.1016/S0921-8181(03)00023-7

Vörösmarty CJ, Moore III B, Grace AL, Gildea MP, Melillo JM, Peterson BJ, Rastetter EB, Steudler PA (1989). Continental scale models of water balance and fluvial transport: an application to South America. *Global Biogeochemical Cycles* **3**, 241-265

Vörösmarty CJ, Sharma KP, Fekete BM, Copeland AH, Holden J, Marble J, Lough JA (1997). The storage and aging of continental runoff in large reservoir systems of the world. *AMBIO: Journal of the Human Environment* **26**, 210-219

Wade A, Whitehead P, Butterfield D (2002a). The Integrated Catchments model of Phosphorus dynamics (INCA-P), a new approach for multiple source assessment in heterogeneous river systems: model structure and equations. *Hydrology and Earth System Sciences Discussions* **6**, 583-606

Wade AJ, Durand P, Beaujouan V, Wessel WW, Raat KJ, Whitehead PG, Butterfield D, Rankinen K, Lepisto A (2002b). A nitrogen model for European catchments: INCA, new model structure and equations. *Hydrology and Earth System Sciences* **6**, 559-582. doi:10.5194/hess-6-559-2002

Wainwright J, Turnbull L, Ibrahim TG, Lexartza-Artza I, Thornton SF, Brazier RE (2011). Linking environmental régimes, space and time: Interpretations of structural and functional connectivity. *Geomorphology* **126**, 387-404. doi:10.1016/j.geomorph.2010.07.027

Waiser MJ, Humphries D, Tumber V, Holm J (2011). Effluent-dominated streams. Part 2: Presence and possible effects of pharmaceuticals and personal care products in Wascana Creek, Saskatchewan, Canada. *Environmental Toxicology and Chemistry* **30**, 508-519. doi:10.1002/etc.398

Walker RS, Novaro AJ, Branch LC (2007). Functional connectivity defined through cost-distance and genetic analyses: a case study for the rock-dwelling mountain vizcacha (Lagidium viscacia) in Patagonia, Argentina. *Landscape Ecology* **22**, 1303-1314. doi:10.1007/s10980-007-9118-2

Wang H, Yan ZG, Li H, Yang NY, Leung KM, Wang YZ, Yu RZ, Zhang L, Wang WH, Jiao CY, Liu ZT (2012). Progress of environmental management and risk assessment of industrial chemicals in China. *Environmental Pollution* **165**, 174-181. doi:10.1016/j.envpol.2011.12.008

Wang J, Zhang L, Qiu H, Bai J (2011). Rural sewage treatment choices and policy impact analysis. *Agro-Environment and Development* **28**, 11-17

Wang X, White-Hull C, Dyer S, Yang Y (2000). GIS-ROUT: a river model for watershed planning. *Environment and Planning B: Planning and Design* **27**, 231-246. doi:10.1068/b2624

Ward JV (1989). The Four-Dimensional Nature of Lotic Ecosystems. *Journal of the North American Benthological Society* **8**, 2. doi:10.2307/1467397

Ward JV, Stanford J (1983). The serial discontinuity concept of lotic ecosystems, in: Fontaine III TD, Bartell SM (Eds.), Dynamics of lotic ecosystems. Ann Arbor Scientific Publishers: Ann Arbor: Michigan, pp. 229-242.

Ward JV, Stanford J (1995). Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation. *Regulated Rivers: Research & Management* **11**, 105-119

Ward JV, Tockner K, Arscott DB, Claret C (2002). Riverine landscape diversity. *Freshwater Biology* **47**, 517-539. doi:10.1046/j.1365-2427.2002.00893.x

Watts K, Handley P (2010). Developing a functional connectivity indicator to detect change in fragmented landscapes. *Ecological Indicators* **10**, 552-557. doi:10.1016/j.ecolind.2009.07.009

Webster E, Mackay D (2003). Defining uncertainty and variability in environmental fate models, CEMC Report No. 200301. Canadian Environmental Modelling Centre, Trent University, Peterborough, Ontario.

Whelan MJ, Hodges JEN, Williams RJ, Keller VDJ, Price OR, Li M (2011). Estimating surface water concentrations of "down-the-drain" chemicals in China using a global model. *Environmental pollution (Barking, Essex : 1987)* **165**, 233-240. doi:10.1016/j.envpol.2011.10.035

Whiteaker TL, Maidment DR, Goodall JL, Takamatsu M (2006). Integrating Arc Hydro features with a schematic network. *Transactions in GIS* **10**, 219-237. doi:10.1111/j.1467-9671.2006.00254.x

Whitehead PG, Wilson EJ, Butterfield D (1998). A semi-distributed Integrated Nitrogen model for multiple source assessment in Catchments (INCA): Part I - model structure and process equations. *Science of the Total Environment* **210**, 547-558. doi:10.1016/S0048-9697(98)00037-0

Wiens JA (2002). Riverine landscapes: taking landscape ecology into the water. *Freshwater Biology* **47**, 501-515. doi:10.1046/j.1365-2427.2002.00887.x

Williams PW (2008). World Heritage Caves and Karst. IUCN: Gland, Switzerland.

Williams PW, Ford DC (2006). Global distribution of carbonate rocks. *Zeitschrift fuer Geomorphologie* **147**, 1-2

Williams RJ, Churchley JH, Kanda R, Johnson AC (2012). Comparing predicted against measured steroid estrogen concentrations and the associated risk in two United Kingdom river catchments. *Environmental Toxicology and Chemistry* **31**, 892-898. doi:10.1002/etc.1756

Williams RJ, Keller VD, Johnson AC, Young AR, Holmes MG, Wells C, Gross-Sorokin M, Benstead R (2009). A national risk assessment for intersex in fish arising from steroid estrogens. *Environmental Toxicology and Chemistry* **28**, 220-230. doi:10.1897/08-047.1

Wint W, Robinson T (2007). Gridded Livestock of the World 2007. Food and Agriculture Organisation of the United Nations, Animal Production and Health Division, Rome, Italy, pp. 1-131.

Wisser D, Fekete BM, Vörösmarty CJ, Schumann AH (2010). Reconstructing 20th century global hydrography: a contribution to the Global Terrestrial Network- Hydrology (GTN-H). *Hydrology and Earth System Sciences* **14**, 1-24

Wisser D, Frolking S, Douglas EM, Fekete BM, Vörösmarty CJ, Schumann AH (2008). Global irrigation water demand: Variability and uncertainties arising from agricultural and climate data sets. *Geophysical Research Letters* **35**, L24408. doi:10.1029/2008GL035296

Wood EF, Roundy JK, Troy TJ, van Beek LPH, Bierkens MFP, Blyth E, de Roo A, Döll P, Ek M, Famiglietti J, Gochis D, van de Giesen N, Houser P, Jaffé PR, Kollet S, Lehner B, Lettenmaier DP, Peters-Lidard C, Sivapalan M, Sheffield J, Wade A, Whitehead P (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research* **47**, 1-10. doi:10.1029/2010WR010090

World Comission On Dams (2000). Dams and development: a new framework for decision-making. The report of the World Commission on dams. Earthscan Publications Ltd, London and Sterling, VA.

Writer JH, Barber LB, Brown GK, Taylor HE, Kiesling RL, Ferrey ML, Jahns ND, Bartell SE, Schoenfuss HL (2010). Anthropogenic tracers, endocrine disrupting chemicals, and endocrine disruption in Minnesota lakes. *Science of the Total Environment* **409**, 100-111. doi:10.1016/j.scitotenv.2010.07.018

Xenopoulos MA, Lodge DM (2006). Going with the flow: using species-discharge relationships to forecast losses in fish biodiversity. *Ecology* **87**, 1907-1914. doi:10.1890/0012-9658(2006)87[1907:GWTFUS]2.0.CO;2

Xenopoulos MA, Lodge DM, Alcamo JM, Marker M, Schulze K, Van Vuuren DP (2005). Scenarios of freshwater fish extinctions from climate change and water withdrawal. *Global Change Biology* **11**, 1557-1564. doi:10.1111/j.1365-2486.2005.01008.x

Yamazaki D, Baugh CA, Bates PD, Kanae S, Alsdorf DE, Oki T (2012). Adjustment of a spaceborne DEM for use in floodplain hydrodynamic modeling. *Journal of Hydrology* **436-437**, 81-91. doi:10.1016/j.jhydrol.2012.02.045

Yamazaki D, de Almeida GAM, Bates PD (2013). Improving computational efficiency in global river models by implementing the local inertial flow equation and a vector-based river network map. *Water Resources Research* **49**, 7221-7235. doi:10.1002/wrcr.20552

Yamazaki D, Kanae S, Kim H, Oki T (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research* **47**, 1-21. doi:10.1029/2010WR009726

Yamazaki D, Oki T, Kanae S (2009). Deriving a global river network map and its sub-grid topographic characteristics from a fine-resolution flow direction map. *Hydrology and Earth System Sciences* **13**, 2241-2251. doi:10.5194/hess-13-2241-2009

Yargeau V, Lopata A, Metcalfe C (2007). Pharmaceuticals in the Yamaska River, Quebec, Canada. *Water Quality Research Journal of Canada* **42**, 231-239

Zambrano-Bigiarini M (2011). hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series. *R package version 0.3-2. [Available at http://CRAN.R-project.org/package=hydroGOF]*

Zarfl C, Lumsdon AE, Berlekamp J, Tydecks L, Tockner K (2014). A global boom in hydropower dam construction. *Aquatic Sciences*. doi:10.1007/s00027-014-0377-0

Zhang QQ, Zhao JL, Ying GG, Liu YS, Pan CG (2014). Emission estimation and multimedia fate modelling of seven steroids at the river basin scale in China. *Environmental Science & Technology* **48**, 7982–7992. doi:10.1021/es501226h

Zhao JL, Ying GG, Wang L, Yang JF, Yang XB, Yang LH, Li X (2009). Determination of phenolic endocrine disrupting chemicals and acidic pharmaceuticals in surface water of the Pearl Rivers in South China by gas chromatography-negative chemical ionization-mass spectrometry. *Science of the Total Environment* **407**, 962-974. doi:10.1016/j.scitotenv.2008.09.048

Zhou Y, Khu ST, Xi BD, Su J, Hao FH, Wu JY, Huo SL (2014). Status and challenges of water pollution problems in China: learning from the European experience. *Environmental Earth Sciences* **72**, 1243-1254. doi:10.1007/s12665-013-3042-3

Ziv G, Baran E, Nam S, Rodríguez-Iturbe I, Levin SA (2012). Trading-off fish biodiversity, food security, and hydropower in the Mekong River Basin. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 5609-5614. doi:10.1073/pnas.1201423109