

Towards integrated inundation modelling

Development, testing, and application of a framework for model coupling across scales

Zinergie of onzinerie van het koppelen van hydrologische en hydrodynamische modellen

Ontwikkeling en toepassing van een raamwerk voor nauwkeurigere overstromingssimulaties

(met een samenvatting in het Nederlands)

Synergieeffekte im Hochwasserschutz durch die Kombination von Modellen

Entwicklung und Anwendung eines Rahmenwerkes zur Integrierung von Überflutungsmodellen

(mit einer deutschen Zusammenfassung)

P R O E F S C H R I F T

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties in het openbaar te verdedigen
op vrijdag 5 juli 2019 des middags te 2.30 uur

door

Jannis Michael Hoch

geboren op 19 September 1988 te München (Duitsland)

Promotoren:

Prof. Dr. Ir. M.F.P. Bierkens

Copromotor:

Dr. L.P.H. van Beek

Dr. Ir. H.C. Winsemius

The author would like to acknowledge the funding and support from Climate-KIC in carrying out this research. Climate-KIC is supported by the European Institute of Innovation and Technology (EIT), a body of the European Union.

Towards integrated inundation modelling

Promotoren:

Prof. dr. ir. M.F.P. Bierkens

Copromotor:

Dr. L.P.H. van Beek

Dr. ir. H.C. Winsemius

Examination committee:

Prof. dr. Paul D. Bates

University of Bristol, Bristol, United Kingdom

Prof. dr. Jaap Kwadijk

University of Twente, Enschede, Netherlands

Prof. dr. Ad de Roo

Utrecht University, Utrecht, Netherlands

Joint Research Centre, Ispra, Italy

Prof. dr. Hannah Cloke

University of Reading, Reading, United Kingdom

Uppsala University, Uppsala, Sweden

Prof. dr. ir. Nick van de Giesen

Delft University of Technology, Delft, Netherlands

ISBN 978-90-6266-541-9

Published by Faculty of Geosciences, Universiteit Utrecht, The Netherlands, in:
Utrecht Studies in Earth Sciences (USES), ISSN 2211-4335

Typeset using X_YT_LX

Cover: The inland delta of the Niger River spreads across central Mali – a unique ecosystem in West Africa. A result of the Niger river flooding the sandy Sahelian plains, its vast network of channels, swamps, and lakes mitigates the severity of the arid climate by supplying water during October and November (blue). In contrast the image shows the sparse rain fed vegetation in the surrounding region (dark green); *From ESA Space in Images, released 21-03-2019. More background why the inland delta needs integrated inundation modelling is provided in chapter 6.2.*

Printed by Ipskamp Printing BV, Enschede, The Netherlands

Correspondence to: Jannis M. Hoch, jannishoch@gmail.com



Except where otherwise noted, this work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International Licence, <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

© 2019 by Jannis M. Hoch.

Chapters 2 to 6 are either unpublished submitted articles or final author versions of previously published articles, © by Jannis M. Hoch and co-authors. More information and citation suggestions are provided at the beginning of these chapters.

Utrecht Studies in Earth Sciences 186

Towards integrated inundation modelling

Development, testing, and application of a framework for model coupling across scales

Jannis Michael Hoch

Utrecht 2019

Faculty of Geosciences, Utrecht University

οἶδα οὐκ εἰδώς

Socrates

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 9 |
| 1.1 | Background | 9 |
| 1.2 | General research objectives | 12 |
| 1.3 | Model Coupling in Hydrologic Sciences | 12 |
| 1.4 | Global Flood Modelling | 16 |
| 1.5 | Research questions and Thesis outline | 20 |
| 2 | Assessing the impact of hydrodynamics on large-scale flood wave propagation – a case study for the Amazon Basin | 23 |
| 2.1 | Introduction | 24 |
| 2.2 | Methodology | 26 |
| 2.3 | Results and discussion | 33 |
| 2.4 | Conclusion and recommendations | 38 |
| 3 | GLOFRIM v1.0 – A globally applicable computational framework for integrated hydrological-hydrodynamic modelling | 41 |
| 3.1 | Introduction | 42 |
| 3.2 | Models | 44 |
| 3.3 | The computational framework GLOFRIM | 47 |
| 3.4 | The synthetic test case | 50 |
| 3.5 | Test case: the Amazon River basin | 52 |
| 3.6 | Conclusion and recommendations | 60 |
| 4 | Benchmarking flexible meshes and regular grids for large-scale fluvial inundation modelling | 63 |
| 4.1 | Introduction | 64 |
| 4.2 | Models and methods | 65 |
| 4.3 | Results and Discussion | 70 |
| 4.4 | Conclusion and Recommendations | 78 |
| 5 | Advancing global flood hazard simulations by improving comparability, benchmarking, and integration of global flood models | 83 |
| 5.1 | Introduction | 84 |
| 5.2 | Current global flood models | 85 |
| 5.3 | META study: validation of GFMs | 86 |
| 5.4 | Establishing a Global Flood Model Validation Framework | 88 |
| 5.5 | Opening the black box of model code | 91 |

| | |
|--|------------|
| 5.6 Conclusion and recommendations | 92 |
| 6 On-going developments of GLOFRIM | 95 |
| 6.1 GLOFRIM 2.0: extending the modelling cascade | 97 |
| 6.2 Implementing feedback loops with groundwater | 107 |
| 7 Synthesis | 117 |
| 7.1 Main conclusions | 117 |
| 7.2 Challenges of model coupling | 123 |
| 7.3 Recommendations and outlook | 124 |
| Appendix A Data and code availability | 129 |
| References | 133 |
| Acknowledgements | 147 |
| List of publications | 149 |
| About the author | 151 |

Summary

Flooding due to rivers exceeding their bankfull capacity was, is, and will be a major contributor to global flood risk, imperilling economic and societal development of often vulnerable communities. To fathom the current situation but especially to plan mitigation and adaption measures in the light of climate change and changes in flood risk, computer models simulating flood-triggering processes can provide the basis for decision-making and policy-making. Since riverine flooding is a global issue, yet with locally varying dynamics, trends, and developments, employing models covering the entire terrestrial surface can facilitate detecting hotspots of flood risk, analysing climate change scenarios, and decision-making in data-sparse areas.

Even though global and large-scale inundation modelling as well as the development of relevant modelling tools such as Global Flood Models (GFMs) has been advancing rapidly in the last decade, there still lie challenges ahead. For instance, most GFMs employ either hydrologic or hydrodynamic models as computational core, expediting some applications while impeding others. Hydrologic models are designed to simulate the components of the (terrestrial) water cycle but employ simple approximations of the complex Saint–Venant shallow flow equations. Hydrodynamic models, on the other hand, solve higher-order approximations, but typically depend on observed data as model forcing. Besides, hydrologic and hydrodynamic models applied within GFMs differ in their spatial resolutions with the latter employing finer scales, allowing for a better representation of topographical features. In a nutshell, each GFM has its own distinct strengths depending on the development choices made, making them suitable for some applications and less so for others.

In this thesis, I investigate how those strengths of different GFMs can be integrated in a flexible and efficient way to compensate for model-specific shortcomings of other GFMs, yielding synergetic effects. To that end, spatially explicit and online coupling – that is, on a grid to grid basis and at the (sub)time step level – between models was established, first in a proof-of-concept and later within the computational framework “GLOFRIM”. Currently, the following models can be coupled: the hydrologic models “PCR-GLOBWB” and “wflow” as well as the hydrodynamic models “CaMa-Flood”, “Lisflood-FP”, and “Delft3D Flexible Mesh”.

By means of GLOFRIM, the benefit of coupling hydrologic runoff simulations with hydrodynamic routing for discharge and inundation extent estimates was assessed. Results obtained in the Amazon River basin indicate that using the local inertia equations of hydrodynamic models instead of the kinematic wave approximation of hydrologic models enhances simulated discharge. It is particularly the timing of simulated peak flow – a critical measure for eg. operational flood forecasting – that is greatly improved compared to observations: from a time gap of few months to a few days. Utilizing output from a hydrologic model instead of locally observed discharge furthermore results in a better accuracy of simulated discharge volume. However, results also corroborate that the accuracy of simulated runoff, and therefore the choice of the hydrologic model itself, greatly defines the potential improvement by model coupling. Consequently, this leaves us with the need to carefully assess the quality of simulated runoff by hydrologic models and hence constitutes a major research goal for future

improvements of global hydrologic models. After all, also the potential improvement through model coupling is defined by the eternal modellers' wisdom "garbage in, garbage out".

Besides discharge, the influence of employing models differing in the complexity of the represented physical processes on simulated flood extent was assessed in the Ganges-Brahmaputra basin. Results show that for accurate representation of floodplain inundation, explicit two-dimensional flow simulations outperform more static post-processing approaches such as downscaling of flood volume or inundation depth. What has to be kept in mind, however, is that the spatial resolution needs to be sufficiently fine to make full use of the advantages of two-dimensional flow simulations. As this goes hand in hand with longer run times, users should think carefully beforehand whether they really need this level of model complexity. Either way, the modular design of GLOFRIM supports the representation of different physical processes as the coupled models, and thus the total model complexity, can be defined depending on the envisaged application.

In addition, GLOFRIM was applied to benchmark different hydrodynamic models by aligning all external model settings such as forcing and input data as well as boundary conditions. Results show that internal model factors – such as gridding technique, routing dimension, and channel-floodplain interaction – still have a marked impact on results even if all external factors are aligned. Further research is needed to unravel how those internal processes are responsible for differences in model results.

Since the gridding technique seems to have a marked impact on both simulated discharge and simulated flood extent, various regular grids and flexible meshes were compared for the Elbe River basin. Obtained results confirm previous results at smaller scales, indicating that employing unstructured grids can be computationally efficient alternatives to regular grids, but also require more *a priori* considerations due to the larger number of possible mesh designs. It is especially their application for large-scale studies requiring fast and data-inexpensive grid generation techniques which poses a challenge to accurate yet efficient flexible mesh design.

Due to the wide range of GFM characteristics, recent studies show that there is only little agreement if inundation maps are benchmarked. Hence, a better and especially more systematic analysis of the impact of differences of both internal and external factors is needed to make output from global models more "actionable" for local end-users and reduce uncertainty still associated with model output. As a solution, a cloud-based framework was conceptualized to support the needed systematic validation and benchmarking of GFM output. By providing standard model forcing data and performing validation and benchmarking in a standardized manner, insight can be gained where and why which model performs better or worse than others. Once it is understood which part of a GFM could be improved, it could be replaced with the corresponding module of a better performing GFM via model coupling as supported by GLOFRIM or other comparable frameworks. Eventually, this would allow for the creation of GFMs depending on envisaged application and required model complexity.

Riverine flooding is a highly complex process driven by a wide range of factors. By first better understanding each model's strengths and shortcomings and subsequently integrating those models required to reflect the dominating processes, the simulation of inundation events can be advanced. This thesis provides a first framework specifically designed for benchmarking various models as well as the integration of models representing different levels of complexity. The modular design of the framework supports the integration of even more physical processes (eg. waves and tides) but also non-physical dimensions (eg. human response to floodings) which is essential to eventually perform the holistic flood simulations needed for tackling the ramifications of current and future flood risk challenges world-wide.

Samenvatting

Overstromingen van rivieren contribuert tot het mondiale overstromingsrisico, waardoor de socioculturele en economische ontwikkeling van vaak kwetsbare gemeenschappen in gevaar kan raken. Om zowel de huidige situatie beter te kunnen begrijpen als wel mitigatie en adaptatie maatregelen te kunnen plannen voor de toekomst, kunnen computermodellen inzicht bieden. Aangezien overstromingen een mondiaal fenomeen zijn, alhoewel met regionale verschillen in dynamiek, trend en ontwikkeling, zijn modellen die de hele aardoppervlakte omspannen een goed middel om overstromings-hotspots te detecteren, klimaatscenario's te analyseren en het nemen van beslissingen in data-arme gebieden te ondersteunen.

Hoewel er op het gebied van het simuleren van overstromingen op globale schaal en het ontwikkelen van tools zoals “Global Flood Models” (GFMs) grote stappen zijn gemaakt, zijn er nog een hoop uitdagingen aan te gaan. De meeste GFMs gebruiken bijvoorbeeld óf hydrologische óf hydrodynamische modellen, elk met hun voor- en nadelen voor bepaalde toepassingen. Hydrologische modellen zijn ontwikkeld om de componenten van de waterkringloop te simuleren, maar gebruiken vrij simpele benaderingen van de complexe Saint-Venant stromingsvergelijking. Hydrodynamische modellen maken gebruik van complexere derivaties van de stromingsvergelijking, maar hun forcering (hun aandrijving) is vaak afhankelijk van de beschikbaarheid en kwaliteit van lokale observaties. Daarnaast verschillen hydrologische en hydrodynamische GFMs qua ruimtelijke resolutie, waarbij hydrodynamische modellen meestal met een hogere resolutie worden toegepast, wat het mogelijk maakt om gedetailleerde topografische kenmerken weer te geven. Al met al heeft elk GFM, afhankelijk van de keuzes in model-opzet, zijn specifieke voor- en nadelen en is daardoor voor sommige applicaties beter geschikt dan voor andere.

In dit proefschrift onderzoek ik hoe de sterke kanten van verschillende GFMs kunnen worden geïntegreerd op een flexibele en efficiënte manier, waardoor model-specifieke nadelen van andere GFMs kunnen worden gecompenseerd. Daarvoor werd een “spatially explicit” (rekening houden met ruimtelijke variabiliteit) en “online” (synchrone uitvoering van de modellen) koppeling tussen modellen opgezet; eerst als haalbaarheidsonderzoek en aansluitend in vorm van het raamwerk “GLOFRIM”. Op dit moment kunnen de volgende modellen met elkaar worden gekoppeld: de hydrologische modellen PCR-GLOBWB en wflow alsook de hydrodynamische modellen CaMa-Flood, Lisflood-FP en Delft3D Flexible Mesh.

Met gebruik van GLOFRIM werd de toegevoegde waarde van het koppelen van hydrologische oppervlaktewater simulaties met de afvoerberekeningen van hydrodynamische modellen onderzocht. Resultaten laten zien dat de nauwkeurigheid van de gesimuleerde afvoer is verbeterd door het gebruiken van hydrodynamische modellen. Vooral de timing van gesimuleerde piekafvoer in vergelijking met observaties gaat vooruit: van meerde maanden verschil naar enkele dagen. Dit betekent een essentiële verbetering waarvan bijvoorbeeld de vroegtijdige onderkenning van overstroming kan profiteren. De resultaten laten ook zien dat het gesimuleerde oppervlaktewater, en dus de keuze van het hydrologisch model, bepalend is voor de verbeteringen die bereikt kunnen worden door het koppelen van modellen. Het is dus noodzakelijk om de kwaliteit van oppervlaktewatersimulaties kritisch te analyseren en globale hydrologische modellen te verbeteren waar nodig en mogelijk.

Naast rivier-afvoer werd ook de invloed van modelcomplexiteit op gesimuleerd overstromingsgebied bestudeerd. Resultaten tonen dat tweedimensionale stromingsberekeningen overstromingen nauwkeuriger simuleren dan methodes die gebruik maken van statische nabewerking van resultaten. Desondanks blijft het verstandig om voor elke toepassing een afweging te maken tussen extra complexiteit en de daarmee gepaard gaande langere rekentijden of dat een simpelere aanpak ook voldoende kan zijn. Het modulaire design van GLOFRIM ondersteunt beide opties omdat de gekoppelde modellen, en daardoor de algehele modelcomplexiteit, afhankelijk van de geambieerde toepassing gedefinieerd kunnen worden.

GLOFRIM werd verder toegepast om verschillende hydrodynamische modellen met elkaar te “benchmarken” (vergelijken) door het aanpassen van alle externe modelinstellingen zoals forcering en randvoorwaarden. Resultaten geven aan dat interne factoren – zoals grid type, stromingsberekeningen en rivier-overloopgebied interactie – invloed blijven houden op modelresultaten, zelfs als alle externe factoren zijn bijgesteld. Meer onderzoek is daarom nodig om beter te kunnen begrijpen hoe de verschillen tussen interne factoren kunnen worden verbonden aan verschillen in modelresultaten.

Omdat het grid type een sterke invloed blijkt te hebben op zowel afvoer als overstromingsgebied, hebben wij meerdere “regular grids” (grids met overal dezelfde vorm) en “flexible meshes” (grids met verschillende vormen en grootten) met elkaar vergeleken. Resultaten bevestigen de conclusies van eerdere studies die op kleine schaal waren toegepast, en laten verder zien dat het gebruik van “flexible meshes” rekentijd kan besparen, maar vooraf meer beslissingen vraagt in verband met de grotere mogelijkheden in grid design. Vooral de noodzaak om zowel snelle als accurate en efficiënte designmethoden voor grids te gebruiken is tegenwoordig nog een uitdaging voor de toepassing van “flexible meshes” op de grote schaal en gebieden waar maar weinig lokale data beschikbaar is.

Recente studies hebben aangetoond dat modelresultaten van verschillende GFMs maar deels overeenstemmen, waarschijnlijk door verschillen in de vele modelkarakteristieken van GFMs. Het is daarom nodig om de invloed van externe en interne factoren van GFMs op een systematische manier te onderzoeken. Zowel om de onzekerheid omtrent modelresultaten te verminderen als om globale output beter begrijpelijk en toepasbaar te maken voor lokale gebruikers. Als mogelijke oplossing werd een conceptueel raamwerk ontworpen waardoor de nodige validatie en vergelijking van GMF output gefaciliteerd wordt. Door het beschikbaar maken van forceringdata en uitvoering van gestandaardiseerde validatie en benchmarking, kan nieuwe kennis worden opgedaan waar en waarom welk model beter is dan een ander model. Het zal dan uiteindelijk niet alleen mogelijk zijn om minder presterende GFM modules door de corresponderende module van een ander GFM te vervangen, maar ook om overstromingsmodellen specifiek als functie van de beoogde toepassing te ontwerpen.

Overstromingen zijn bijzonder complex en gedreven door verschillende factoren. Door eerst beter te begrijpen wat zowel de sterke als de zwakke kanten van modellen zijn en aansluitend de modellen te integreren die de complexiteit van overstromingen het best reflecteren, kunnen wij overstromingssimulaties verbeteren. Dit proefschrift biedt een raamwerk aan dat specifiek is ontworpen om zowel modellen te vergelijken als wel modellen met verschillende complexiteiten te integreren. Het design van het raamwerk maakt het mogelijk niet alleen andere fysieke processen (zoals golven en getij) maar ook niet-fysieke processen (zoals het menselijk gedrag naar overstromingen) toe te voegen aan de modelstructuur. Op die manier zal het uiteindelijk mogelijk zijn de holistische overstromingssimulaties uit te voeren die noodzakelijk zijn om de complexiteit van het hedendaagse en toekomstige overstromingsrisico op mondiale schaal aan te kunnen pakken.

Zusammenfassung

Überschwemmungen von Flüssen (“fluviale Überflutungen”) sind maßgeblicher Bestandteil des heutigen und zukünftigen globalen Überflutungsrisikos, da sie sowohl die ökonomische als auch sozio-kulturelle Entwicklung von Gesellschaften gefährden können. Um nicht nur die gegenwärtige Gefahr von Überflutungen besser begreifen, sondern auch um zukünftig notwendige Schutz- und Anpassungsmassnahmen treffen zu können, bieten Computer-Modelle eine adäquate Diskussionsgrundlage. Obwohl Überflutungen weltweit ein Risiko darstellen, können Dynamiken, Trends und Entwicklungen lokal variieren. Im Gegensatz zu regional begrenzten Modellen können globale Überflutungsmodelle dabei helfen Gefahrenherde zu identifizieren, räumlich zusammenhängende Auswirkungen verschiedener Klimaszenarien zu analysieren und Risikoabwägungen in Gebieten mit wenig verfügbaren Daten zu erleichtern.

Obwohl im letzten Jahrzehnt globale fluviale Überflutungssimulationen und auch die Entwicklung von entsprechenden “Global Flood Models” (GFMs; *Globale Überflutungsmodelle*) grosse Fortschritte verbucht haben, gibt es noch eine Vielzahl an Herausforderungen. So benutzen die meisten GFMs entweder hydrologische oder hydrodynamische Modelle, was für manche Anwendungen förderlich, für andere wiederum nachteilig sein kann. Hydrologische Modelle wurden entworfen um die Komponenten des Wasserkreislaufes, wie zum Beispiel Oberflächenabfluss, zu simulieren. Diese verwenden jedoch nur simple Approximationen der komplexen Saint-Venant Strömungsgleichungen. Hydrodynamische Modelle können wiederum komplexere Strömungsgleichungen lösen, sind aber in der Regel auf gemessene Daten zur Forcierung (zum Antreiben) angewiesen. Ausserdem unterscheiden sich die von GFMs verwendeten hydrologischen und hydrodynamischen Modelle in Bezug auf ihre räumliche Auflösung, wobei letztere eine feinere Auflösung haben und dadurch in der Lage sind topographische Merkmale besser darzustellen. Man kann also feststellen, dass jedes GFM, je nach Entwicklungsentscheidungen, nicht für alle Anwendungen in gleichem Maße geeignet ist.

In dieser These untersuche ich, wie die Stärken verschiedener GFMs auf eine flexible und effiziente Art und Weise miteinander kombiniert werden können, um die Nachteile anderer GFMs durch Synergieeffekte auszugleichen. Zu diesem Zweck wurden Modelle “spatially-explicit” (durch eindeutige räumliche Zuordnung zusammengehörender Raster) und “online” (bei zeitgleicher Ausführung der Modelle) gekoppelt; zuerst in einer Machbarkeitsstudie und anschließend im Rahmenwerk GLOFRIM. Derzeit können mit GLOFRIM die folgenden Modelle integriert werden: die hydrologischen Modelle PCR-GLOBWB und wflow sowie die hydrodynamischen Modelle CaMa-Flood, Lisflood-FP und Delft3D Flexible Mesh.

GLOFRIM wurde angewandt um den Mehrwert des Kombinierens von hydrologischem Oberflächenabfluss (“runoff”) mit hydrodynamischen Strömungsberechnungen zu bewerten, um Durchfluss (“discharge”) und Überflutungsausmass (“inundation extent”) von Flüssen genauer simulieren zu können. Resultate zeigen, dass bei Verwendung von hydrodynamischen Modellen die Genauigkeit des simulierten Durchflusses profitiert. Vor allem das Timing des simulierten Spitzendurchflusses wird im Vergleich zu gemessenen Werten verbessert: von mehreren Monaten Differenz zu einigen Tagen. Diese Verbesserung kann entscheidend sein,

zum Beispiel für das rechtzeitige Auslösen von Frühwarnsystemen. Gleichzeitig zeigen die Resultate jedoch auch, dass die eben genannte Verbesserung der Genauigkeit, die durch Integration hydrodynamischer Modelle möglich ist, von der Genauigkeit des Oberflächenabflusses, und damit der Wahl des hydrologischen Modells, abhängt. Es ist daher notwendig die Qualität der Oberflächenabflussberechnungen von hydrologischen Modellen sorgfältig zu analysieren und gegebenenfalls die Modelle zu verbessern.

Es wurde ebenfalls untersucht wie das Verwenden verschiedener Modelle, die sich bezüglich der Komplexität der simulierten physikalischen Prozesse unterscheiden, das simulierte Ausmaß von Überflutungen beeinflusst. Ergebnisse zeigen, dass für eine genaue Wiedergabe von Überflutungen explizite zweidimensionale Simulationen vorteilhaft sind. Daraus fließt fort, dass es wichtig ist im Voraus ab zu wägen ob die längere Simulationszeit, die für diese zweidimensionalen Berechnungen nötig ist, vertretbar ist oder ob das benötigte Prozessverständnis auch mit einfacheren Mitteln zu beschaffen ist. GLOFRIM kann ob seines modulären Designs in beiden Fällen eingesetzt werden. Indem Modelle in Abhängigkeit von der geplanten Anwendung und der maßgebenden physikalischen Prozesse miteinander integriert werden können, ist es möglich verschiedene Komplexitätsgrade abzubilden.

GLOFRIM wurde zudem angewandt um verschiedene hydrodynamische Modelle zu "benchmarks" (zu vergleichen). Zu diesem Zweck wurden alle externen Faktoren, zum Beispiel Forcierung und Randbedingungen wie etwa der Meeresspiegel, in allen Modellen angeglichen. Resultate zeigen, dass interne Faktoren wie die Generierung des Modelrasters, numerische Prozesse oder die simulierte Interaktion zwischen Fluss und Auen starken Einfluss auf Modelresultate haben, selbst wenn externe Faktoren angepasst sind. Um die bestehenden Unterschiede zwischen Ergebnissen verschiedener Modelle an die Unterschiede zwischen den internen Faktoren dieser Modelle zu knüpfen ist daher die Notwendigkeit für weitergehende Forschung angezeigt.

Da die Art wie Modelraster generiert werden simulierten Durchfluss und Überflutungsgebiet beeinflusst, wurden diverse "regular grids" (gleichförmige Raster) und "flexible meshes" (in Größe und Form varierende Raster) miteinander verglichen. Resultate bestätigen frühere Studien und indizieren außerdem, dass "flexible meshes" zwar recheneffizient sind, aber ob der größeren Möglichkeiten auch mehr vorhergehende Abwägungen benötigen. Vor allem bei der Anwendung für großflächige Studien in womöglich datenarmen Gebieten ist eine einerseits schnelle und datenunaufwändige und andererseits akkurate als auch recheneffiziente Generation von "flexible meshes" derzeit noch herausfordernd.

Ob der verschiedenen GFM Charakteristiken stimmt das Ausmaß simulierter Überflutungsgebiete nur teilweise zwischen Modellen überein. Eine systematischere Analyse der Unterschiede von externen und internen Faktoren ist daher notwendig, um die Ergebnisunsicherheiten zu reduzieren und dadurch die Ergebnisse globaler Modelle besser für lokale Benutzer begreifbar und benutzbar zu machen. Als potentielle Lösung wurde ein Rahmenwerk konzeptualisiert, das das notwendige systematische Validieren und Vergleichen von GFM Resultaten ermöglicht. Durch das Bereitstellen von Standardforcierung sowie das Durchführen von standardisierter Validierung und Vergleich von Modelresultaten können Einsichten gewonnen werden wo und warum welches Model besser abschneidet als andere. Sobald man einschätzen kann welcher Teil eines GFMs Verbesserungspotential hat, kann er mit dem korrespondierenden Modul eines besser abschneidenden GFMs ersetzt werden, zum Beispiel durch das Kombinieren verschiedener Modelle oder Module.

Fluviale Überflutungen sind äußerst komplexe Prozesse und werden von einer Vielzahl Faktoren beeinflusst. In dem man erst die Stärken und Schwächen jedes Modells analysiert und anschließend die Modelle integriert, die nötig sind um die Komplexität der maßgebenden

physikalischen Prozesse darzustellen, kann die Simulation von Überflutungen verbessert werden. Die vorliegende Arbeit stellt ein Rahmenwerk bereit, das spezifisch entworfen wurde um unterschiedlich entworfene Modelle zu vergleichen und zu integrieren. Es ist dabei so strukturiert worden, dass auch weitere physikalische (zum Beispiel Wellen und Gezeiten) aber auch nicht-physikalische Prozesse (zum Beispiel die menschliche Reaktion auf Überflutungen) integriert werden können. Auf diese Art und Weise können die holistischeren Überflutungssimulationen durchgeführt werden, die nötig sind um die Herausforderungen von heutigem und zukünftigem Überflutungsrisiko weltweit angehen zu können.

Chapter 1

Introduction

1.1 Background

Flood events were, are, and will be a major hydrologic hazard. There have been reports about large areas of land being inundated already in ancient times, even without taking the Noachian deluge (Eusebius, 405AD) into account. For a number of rivers, such as the Nile and the Yellow river, such ancient events nowadays termed paleofloods, have been dated back for millennia using tree rings or sediment cores (cf. Baker (1987), Baker (2006), Ballesteros-Cánovas et al. (2015), Bell (1970), Greenbaum et al. (2014), and Huang et al. (2010)).

Affected cultures learned to live with the river and kept their damages manageable by not placing valuable assets in the regularly inundated floodplain. In time, however, construction along rivers and oceans increased as these locations were needed for trade, livelihood, and military power to guarantee further socio-economic development. Consequently, protection against riverine and coastal flood events was needed and dikes were built to avoid major economic and cultural setbacks or even complete societal collapse (Zhang et al., 2005). Despite the protection of the already present assets, the resulting so-called “levee effect”, “safe development paradox” or “safety dilemma” (Di Baldassarre et al., 2018) often also leads to intensified settlement along water bodies, further increasing the value at risk (cf. Di Baldassarre et al. (2009), Collenteur et al. (2015), and Mård et al. (2018)).

In recent decades, casualties and economic damage due to riverine flooding world-wide increased greatly although varying in the numbers reported (Ward et al., 2013; Hirabayashi et al., 2013; Jongman et al., 2012). This development cannot solely be attributed to continued urbanization on flood-prone areas along water bodies (Mård et al., 2018; Ceola et al., 2014), but also the impact of climate change and the associated changes in meteorological patterns are causing shifts in flood damage (Winsemius et al., 2016; Hirabayashi et al., 2013).

As a result, flood hazard, associated flood risk as well as mitigation and adaption measures are hot topics and therefore acknowledged in several global frameworks. These frameworks, such as in the United Nations Framework Convention on Climate Change (UNFCCC) or the Intergovernmental Panel on Climate Change reports (IPCC, 2014), not only provide an overview and address challenges of climate change in general but also specifically of disaster risk and their reduction, for example in the Sendai Framework (UNISDR, 2015b).

While the global picture shows an overall increase in future flood risk, there are marked regional differences. For instance, higher income countries even show a decrease in vulnerability and thus are less susceptible to flood events, leading to reduced flood risk if normalized with the Gross Domestic Product (GDP) (Jongman et al., 2015). Research by Winsemius et al. (2016) furthermore shows that for some regions, flood risk may decrease as consequence of climate change (eg. in the Volta basin) or socio-economic development (eg. in the Amazon basin), meaning that economic activities concentrate to a lesser extent in floodplain areas in the future. For most regions, however, both increased urbanization and climate change in-

crease predicted future flood risk, such as the Mekong basin and the Lena basin, respectively (Figure 1.1).

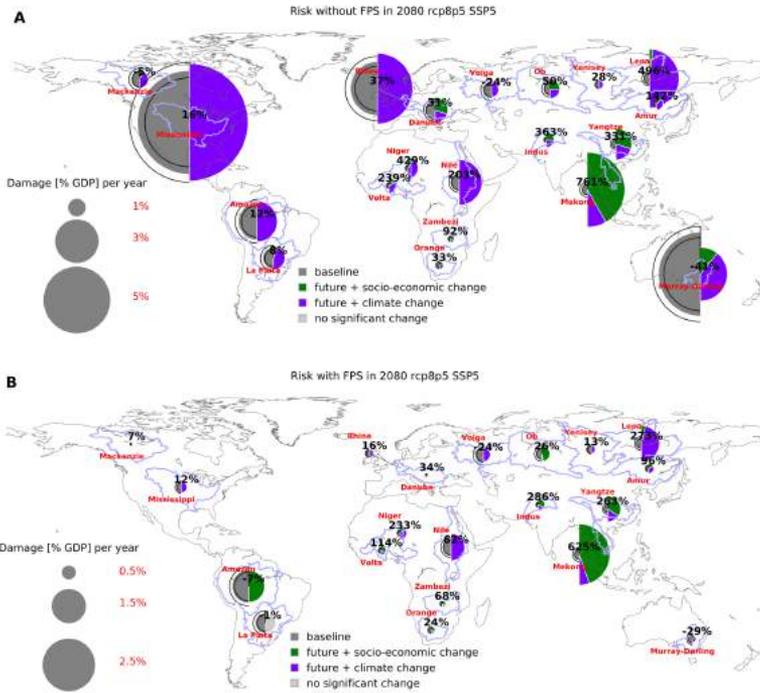


Figure 1.1: Projected change in economic risk until 2080 in the Fossil fuel-based development projection. **A**, The ratio of annual urban damage over the basins total GDP per year under No flood protection standards (FPS) conditions. **B**, Same as a but for Partial FPS conditions. Note that the scales of the circular diagrams of A,B are different due to the large difference between No FPS and Partial FPS conditions. The grey left halves of the circles represent the current risk, with estimated uncertainty bounds in black lines. The right half of the circles represents future risk. The relative sizes of the two different colours represent the relative contributions of climate change and socio-economic change to risk increases or decreases. The percentage for each basin indicates the increase in the risk metric displayed from the present day (2010) to 2080. From Winsemius et al. (2016).

Also, research showed that larger flood events cannot be seen as mere national or catchment issues but have to be treated in an international context as they often affect multiple countries simultaneously due to their origin in large-scale atmospheric circulations (Jongman et al., 2014). Consequently, fluvial flood risk should be considered as a multi-lateral problem to establish optimal management practices involving all countries along the entire river. Examples are the International Commission for the Protection of the Rhine (IPCR), the Inter-

national Commission for the Protection of the Danube River (ICPDR) or the Mekong River Commission (MRC).

To allow comparison of flood risks between climatic and economic regions as well to predict the impact of large flood events across borders, models covering the entire earth surface are developed, so-called Global Flood Models (GFMs). By now, there is a range of GFMs publicly available. The most common GFMs are according to Trigg et al. (2016): GLOFRIS (Winsemius et al., 2013), JRC (Dottori et al., 2016), CIMA-UNEP (Rudari et al., 2015), Fathom (Sampson et al., 2015), ECMWF (Pappenberger et al., 2012), and CaMa-Flood (Yamazaki et al., 2011). See section 1.4.2 for more details about the individual model set-ups and properties.

As the added value of GFMs is acknowledged, they have found application in policy and decision-making. For instance, CIMA-UNEP was used for the Global Assessment Report of the United Nations Office for Disaster Risk Reduction (UNISDR, 2015a), the JRC model is linked with the operational Global Flood Awareness System (GloFAS, 2018), and GLOFRIS is the computational basis of the “Aqueduct Floods” tool developed by the World Resources Institute (*Aqueduct Global Flood Analyzer* 2019). GFMs thus not only provide scientific explorations of what is technically possible, but also are highly application-oriented products aiming at assessing the consequences of flood events for societies now and in the future.

Hence, the results of GFMs need to fulfil high requirements of model output accuracy. However, each GFM inherently has its own characteristics and hence advantages and disadvantages depending on the developers decisions. For example, some models employ global hydrologic models (GHMs) as computational core, typically simulating the global water balance at coarse spatial resolution, whereas others use hydrodynamic models, computing surface flow at much fine spatial resolution than GHMs. The causes for inundations are a combination of several location-dependent factors, *inter alia* increased precipitation, soil saturation, channel and embankments dimensions, and backwater effects. Consequently, not every GFM is, depending on its computational core, equally suitable for certain applications as they may not be able to represent the relevant physical processes, boundary conditions, and model forcing for a given case and at a sufficient spatial resolution. Therefore, the need arises to further improve current global flood modelling practices by advancing their output accuracy as well as breadth of applicability.

One avenue to advance the accuracy of physically-based model output is model coupling. By combining processes from different models in a smart way, advantage can be taken from the individual model strengths whilst model weaknesses would be compensated for. Some of the issues associated with the GFM specific properties could therefore potentially be solved by coupling the model components of different GFMs which excel at their specific objective, creating a new flood model. For example, a coarse-resolution hydrologic model calculating surface runoff could be utilized to force a hydrodynamic model for improved flood wave simulations and inundations maps at a finer spatial resolution. By forcing hydrodynamic models with hydrologic output, it becomes easier to represent the impact of climate change on meteorology and in turn predict future flood hazard as well. Vice versa, subjecting floodplain inundation in hydrodynamic models to hydrologic processes such as groundwater infiltration and evaporation could improve simulations in areas where those processes are essential, such as the Niger Inland Delta (cf. Dadson et al. (2010), Mahe (2009), and Mahe et al. (2009)).

Reflecting the wide range of computer models and possible applications available, there are several technical ways to couple models and it is of course not limited to inundation modelling only. Amongst others, hydrologic, hydrodynamic, glacier, atmospheric, and groundwater models were coupled in different ways and different combinations (section 1.3). While certain model combinations may be advantageous, other combinations may not improve flood

simulations. With an ensemble of GFM's available now, it would first be required to make an inventory of why certain models outperform others and how underlying model process descriptions could be supplemented with other model routines in a smart and lean way. So far, such a systematic analysis of GFM's is still limited to few studies (Trigg et al., 2016; Bernhofen et al., 2018) and thus establishing a GFM validation platform may not only be a step towards more systematic GFM assessments, but would also facilitate analysing potential model coupling designs.

Better informed model coupling strategies are key as the processes to be integrated across may differ per application in mind. For instance, large-scale routing applications would not require the detailed representation of floodplain elevation and channels while locally applicable inundation modelling could not do without it. Once we obtain a clearer picture where which process dominate, smart and objective-oriented model coupling of GFM components may help producing a new generation of flood hazard estimates and, in turn, planning of mitigation and adaption measures.

1.2 General research objectives

Based on the preceding section, the research objectives of this thesis are formulated as:

1. *Establish a flexible end-user oriented scientific software framework to facilitate integration across physical processes, spatial resolutions, and flood hazard drivers.*
2. *Improve the understanding of how model discretization as well as the complexity of simulated physical processes influence simulated discharge and inundation extent.*

1.3 Model Coupling in Hydrologic Sciences

Model coupling itself is a proven technique – not only for flood modelling purposes, but also in earth system sciences in general and its subsidiary fields of climate science, carbon and vegetation modelling, and hydrology (Bierkens, 2015). For instance, hydrologic or land surface models have been coupled with atmospheric models (Butts et al., 2014; Zabel and Mauser, 2013). Other examples show coupling between a glacier and a land surface model (Naz et al., 2014; Zhao et al., 2013) or of a groundwater and hydrologic model (Sutanudjaja et al., 2014; Maxwell and Miller, 2005). It is also possible to combine different hydrodynamic models (Alarcon et al., 2014).

Yet, there are different ways how models can be coupled. Hereafter, a brief overview of different coupling strategies and their technical implementations is provided.

1.3.1 Coupling strategies and applications

Based on reviewed literature, three categories are here defined and outlined briefly. A major constraint thereby was that many studies couple models but do not specify which methods were used to establish links between models. To fit within the scope of this thesis, we decided to differentiate between external, internal, and flexible coupling. It must be noted though that there are multiple, not necessarily corresponding, definitions of model coupling strategies as well (for instance by Morita and Yen (2002) who discern simultaneous, alternating iterative, and external coupling).

External coupling: output files are input files

A simple model coupling cascade can be established by using the output files from model A as input files for lateral inflows into model B as done in various studies (cf. Biancamaria et al. (2009), Felder et al. (2018), Lian et al. (2007), and Schumann et al. (2013)). While such an approach should yield identical results as other more complex coupling strategies, output and input files are often aligned using auxiliary scripts which renders external coupling inflexible if the coupling design should change. An example of reading from model output files in a standardized framework is the Delft-FEWS system (Werner et al., 2013) where files are stored in XML format as proposed by Kokkonen et al. (2003). Also, computational efficiency is hampered due to increased efforts required for file input/output (I/O) operations. Since models have to be run consecutively, computational efficiency is further reduced, which is particularly disadvantageous for large-scale applications.

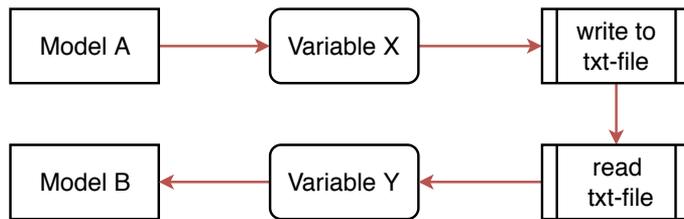


Figure 1.2: Simplified visualization of the external coupling approach where output from model A is first saved to file (here exemplified with a txt-file) and subsequently used as input file to model B.

Internal coupling: explicitly merging and hard-coding processes

In case the processes of two models shall be merged for one specific purpose, hard-coding the equations, processes, and dependencies of different models into one is an option. While internal (also named “tightly-coupled”) approaches allow for fully controlling the modelling process, they require consistent internal conventions (eg. for data structures) within the models (Jiang et al., 2017). As a consequence, internal coupling is in most cases likely to require more work than external coupling. However, once the coupling is established, no alignment and storage of output and input files is needed anymore. Internal coupling is favourable if the coupling is designed for one bespoke application but reaches its limits rather quickly if alternative objectives are pursued as flexibility is curbed.

Such coupling was applied by, for instance, Sutanudjaja et al. (2014) who established a link between the bucket-type groundwater storage of PCR-GLOBWB (Sutanudjaja et al., 2018; van Beek et al., 2011) and the groundwater model MODFLOW (Harbaugh et al., 2000). Similarly, Maxwell and Miller (2005) coupled the Common Land Model with the groundwater model Parflow (Ashby and Falgout, 1996). Other examples comprise the work of Davison et al. (2018) who dynamically coupled the HydroGeoSphere (Brunner and Simmons, 2012) model with the Weather Research and Forecasting model.

Flexible coupling: exchanging information via interfaces

This “loosely-coupled” approach provides most freedom to applicants as the coupling scheme between models can be defined depending on the study’s objective. Since all processes of the coupled models are executed near-simultaneously, it comes closest to full coupling (Morita and Yen, 2002). However, it also requires most preparatory work as interfaces and frameworks are needed to communicate with and between models. On the upside, the internal model structure can remain unchanged as only interfaces have to be established (Jiang et al.,

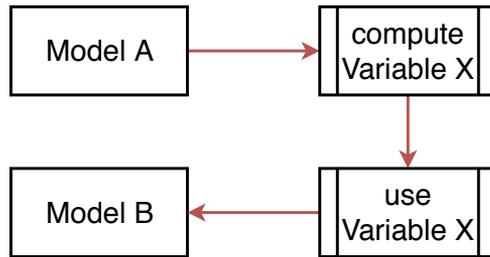


Figure 1.3: Simplified visualization of the internal coupling approach where code of models A and B are combined, and variable information exchanged within the newly created coupled model.

2017). By coupling models via interfaces, they can be executed (near-)simultaneously, which is beneficial from a run time perspective as well as for implementing dynamic feedback loops on a time step basis.

Typically, Environmental Modelling Frameworks (EMFs) such as the Earth System Modeling Framework (ESMF; Hill et al. (2004)), Object Modeling System (OSM; David et al. (2013)), OpenMI (OpenMI (2018); see section 1.3.2) or Basic Model Interface (BMI; Peckham et al. (2013); see section 1.3.2) follow this approach as it provides most freedom in design and execution. As such, the flexible coupling approach is widely used in efforts to couple different models and model components.

For instance, the HIRHAM regional climate model was coupled with the hydrologic model MIKE SHE by using OpenMI (Larsen et al., 2014). Another example is the Web Modeling Tool (<https://csdms.colorado.edu/wmt/>) and the PyMT toolkit (<https://csdms.github.io/pymt/>), both developed by the Community Surface Dynamics Modeling System (CSDMS). Currently being in the start-up phase, coupling various models via standardized interfaces is the goal the eWaterCycle2 project (Hut et al., 2018). The ESMF was used to establish the Earth System Prediction Suite (ESPS) integrating various U.S. weather and climate models and model components (Theurich et al., 2015). Since models can communicate via interfaces, they can also be executed via web services (Jiang et al., 2017; Castronova et al., 2013b).

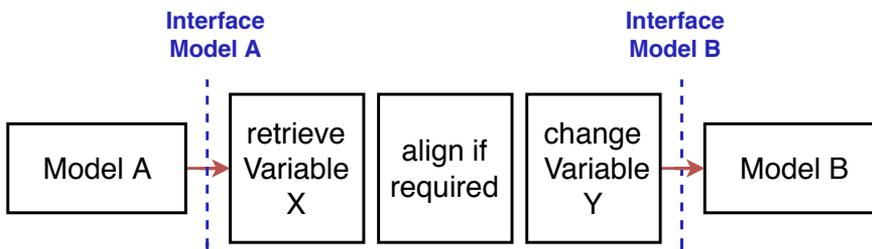


Figure 1.4: Simplified visualization of the flexible coupling approach where variable information of model A is retrieved, aligned to meet requirements of model B, and added to model B via model interfaces.

So far, the benefits of model coupling for inundation modelling have only been shown in bespoke showcase studies (eg. Zischg et al. (2018) and Felder et al. (2017)) or were designed as a proof-of-concept. Despite the great potential of model coupling to advance global flood modelling, no study yet exists investigating the possibilities and limitations of such an approach. Therefore, it was necessary to establish a standardized coupling framework to gain insight on how different flexible coupling set-ups compare with each other. Answering

research question 1 will show to what extent such frameworks have added value for large-scale inundation modelling. To develop ways how the gained insights can be translated into more physically-based and hence better-informed flood hazard estimates simulated by GFMs, research question 2 is formulated.

1.3.2 Technology for exchanging model information

There are different ways on how model information can be exchanged between models. Here, we want to focus on two common interfaces, namely the Open Modeling Interface (OpenMI) and the Basic Model Interface (BMI). Discussing other EMFs in detail is outside the scope of this thesis.

Open Modeling Interface (OpenMI)

The Open Modeling Interface (OpenMI) Standard is a software component interface definition (OpenMI, 2018). It is an OGC (Open Geospatial Consortium) standard and thus integrates well with other standards acknowledged by the OGC. OpenMI was originally designed to facilitate the simulation of interacting (environmental) processes by enabling independent models to exchange data per time step. Model components complying with the OpenMI standard can, without any programming, be coupled to OpenMI modelling systems.

To make a model OpenMI-compliant, at least two files (typically in XML-format) must be defined *a priori* to the model run: the omi-file describing the OpenMI-compliant models and component-related information, and the opr-file specifying the data exchange between the components (Becker and Burzel, 2016). With the files, it is distinguished which models are exchanging which variable and how.

As the OpenMI standard already exists for many years, the number of models being compliant is extensive. *Inter alia*, the SOBEK model of Deltares (<https://www.deltares.nl/en/software/sobek/>) and the MIKE models of DHI (<https://www.mikepoweredbydhi.com/>) can all be coupled using this standard. Besides, it is applied in several studies exploring possibilities of coupling models of different origin and purpose (cf. Becker and Burzel (2016), Becker and Talsma (2014), Buahin and Horsburgh (2018), Castronova et al. (2013a), and Larsen et al. (2014)).

Basic Model Interface (BMI)

The Basic Model Interface (BMI; Peckham et al. (2013)) is a set of functions facilitating a structured and uniform way to advance a models state. It was originally developed to enable the model coupling framework of the “Community Surface Dynamics Modelling System (CSDMS)”. Each model being coupled in this way requires a model-specific BMI to provide the necessary information about model grid, variable names, time, and so forth. Adding the BMI to (existing) model code is non-invasive and therefore the overall model functionality is not altered even if BMI functions are not used. Before the BMI functionality can be used, the needed information must be exposed via dedicated code. Once this is achieved, model coupling via BMI is framework-agnostic, increasing the models versatility. Also, the functions provided allow for a fine-grained control over model time stepping and variable states. By being based on functions instead of files, the BMI allows for interactive modelling and thus for more flexibility in altering the coupling structure compared to the OpenMI standard where the exchanges have to be defined unchangeably before the simulation commences. Even though the BMI is applicable in any programming language (eg. Fortran), it may be necessary to add a “wrapper”, translating it to another, central coupling language (eg. Python).

So far, the BMI was applied in several studies but is not (yet) as widespread as the OpenMI. For instance, it was already tested within a Experimental Modeling Environment for Linking

and Interoperability (EMELI) (Peckham, 2014) whose aim is to facilitate smart and automatic coupling of self-describing models. Similarly, the BMI concept was employed to couple different model at the time-step level, for example river morphology and ocean dynamics (Ratliff et al., 2018). Recently, EMELI-Web was developed allowing for model coupling via BMI-compliant web services (Jiang et al., 2017). Work is on-going to provide interoperability between BMI and OpenMI to further advance the integration models and model components (Goodall and Peckham, 2016).

1.4 Global Flood Modelling

1.4.1 Global Flood Modelling – Why and why not?

With flood hazard and risk becoming increasingly important for global stakeholders (for example re-insurance industry, financial institutions or aid organizations), the development and application of GFMs increased too. Since GFMs were and are developed with specific objectives in mind, the benefits and challenges of their application in contrast to other models (eg. local-scale model) should be discussed (see Table 1.1 for a SWO analysis).

For instance, GFMs can help identifying hotspots where current or future flood risk requires mitigation or adaption measures (Ward et al., 2015). Based on such assessments, more detailed analyses can be conducted to further delineate possible realizations of flood risk by employing local-scale models. Related to that are more operational versions of GFMs, allowing for detecting areas at risk of flooding based on meteorological forecasts such as GloFAS (Alfieri et al., 2013; Emerton et al., 2018). Such information is pivotal to duly initialize preparatory actions aimed at mitigating consequences, for example the distribution of protection and aid goods, or other measures such as forecast-based financing (Jjemba et al., 2018; Coughlan de Perez et al., 2015). GFMs can provide fast estimates of flood hazard and risk for countries, regions, and provinces which do not have a detailed local flood model in place. Such information is essential in data-sparse countries to stimulate stakeholder involvement and to facilitate decision making.

Another reason for applying GFMs is that most major flood events (such as the Elbe or Danube flood event in 2009 and 2012, respectively) are trans-national hazards and can affect several countries simultaneously. As fluvial flood events are mostly triggered by large-scale synoptic meteorological events, using global models allows for a seamless representation of both cause and effect of floods which otherwise would not be possible with limited local or regional modelling approaches (Ward et al., 2015; Alfieri et al., 2018). Pappenberger et al. (2012) points out that tiling (that is, the combination of flood maps from different model sources) should be avoided as it introduces uncertainties associated with resolution, input data, boundary conditions, and so forth.

Despite these advantages, there are still several challenges associated to the current versions of GFMs. One of the main challenges is that robust information on flood protection measures such as dikes or other small flood-controlling features are missing in most global models, although first attempts to catalogue protection standards were made (Scussolini et al., 2016). Depending on local and flood wave properties, this may paint an erroneous picture of flood risk as it a) may be overpredicted and b) spatial patterns may be shifted (Triet et al., 2017). Similarly, urban inundation cannot be resolved sufficiently well yet with the current generation of models due to the lack of information about man-made structures such as culverts, railway levees or ditches. In terms of boundary conditions, only few global models are capable of accounting for dynamic sea level boundaries, although this represents a major

| Strengths and opportunities | Weaknesses |
|---------------------------------------|--|
| Hotspot identification | Local flood risk assessments |
| Climate change assessments | No/little representation of hydraulic structures |
| Application in data-sparse areas | Urban inundation modelling |
| starting point for local developments | |

Table 1.1: Strengths, weaknesses, and opportunities (SWO) analysis of GFMs.

driver of flooding due to compound events, ie. the coincidence of different hazards such as high river discharge and high sea levels (Ikeuchi et al., 2017; Ward et al., 2018).

A further challenge is to convey the information GFMs produce, particularly if the user acts at a different, mostly smaller, spatial scale. Using output from GFMs at a regional or local scale may be difficult as many processes and data, which local decision-makers may find relevant, are not implemented or not represented at a locally relevant accuracy (Ward et al., 2015). Besides, many GFMs run at spatial resolutions which are not sufficient for stakeholder involvement (Beven et al., 2015). Recently, GFMs evolved towards even finer spatial resolutions, although being merely applied for the Continental United States (CONUS) (Wing et al., 2017; Wing et al., 2018; Quinn et al., 2019) and at high computational costs. Making GFM output actionable is therefore one of the main challenges the global flood risk community is facing (Global Flood Partnership, 2016, 2017).

1.4.2 Global Flood Models – an overview

Currently the most fully developed and openly accessible state-of-the-art GFMs are CaMa-Flood, GLOFRIS, JRC, CIMA-UNEP, as well as the Fathom and the ECMWF model. Even though all models compute inundation maps at various return periods, they can be subdivided into two main categories (see Figure 1.5). For detailed information of common and different aspects, please consult the supplementary material of Trigg et al. (2016).

Those belonging to the climate cascade models usually force a land surface model (LSM) or global hydrologic model (GHM) with meteorological data, often reanalysis products such as ERA-Interim or ERA5. Simulated runoff is then routed (either as part of the LSM/GHM or with an additional model) and resulting discharge is used to perform a stochastic flood frequency analysis. Besides, obtained inundation maps can in some cases downscaled to finer spatial resolutions in a post-processing step.

Models employing gauged flow data follow a different approach. They typically make use of observed streamflow data and create synthetic flood waves with certain return periods (magnitudes) based on a regional flow frequency analysis. The obtained flood waves are then used as upstream boundary conditions for river and floodplain flow simulations, resulting in flood extent and depth estimates corresponding to the return periods used for flood wave generation.

In a nutshell, the different operations at various model stages portrayed in Figure 1.5 result in a range of modelling approaches, each one using its own input data (for example meteorological forcing or flood wave with specific return period) and methods of calculating floodplain inundation.

This range of modelling approaches has consequences on how runoff is routed along a channel network and subsequently on resulting inundation maps. It is thus paramount to understand how differences in routing scheme may affect simulation results and whether the applied scheme is fit for the envisaged application. Moreover, the way GFMs are discretized

in terms of input data and spatial resolution differs per modelling approach. While input data is often defined by what is available, the chosen spatial resolution is a deliberate decision of the developer. For instance, GLOFRIS runs at a 30 arc-min spatial resolution (that is, around 50 km x 50 km at the Equator) before post-processing and downscaling to 1 km, whereas the Fathom model yields output directly at 90 m. What is therefore relevant to know is whether and how the currently followed model discretization approaches affect simulation results and if there are possible ways forward in case that is needed.

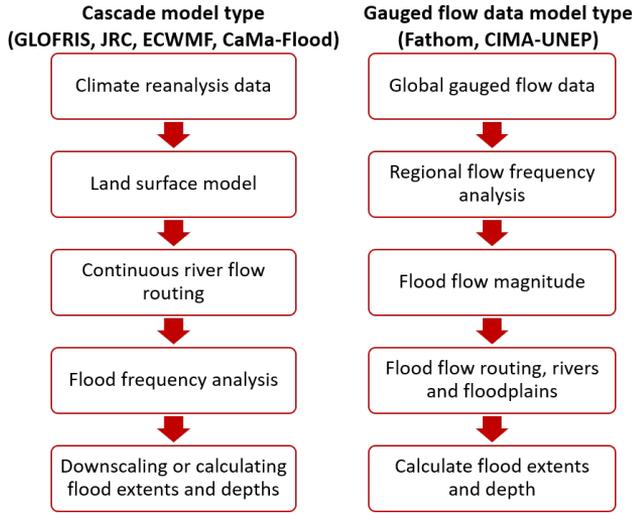


Figure 1.5: Modelling steps required for the two sub-categories of global flood models; from Hoch and Trigg (2019), based on Trigg et al. (2016)

The different flavours of routing

As mentioned above, different GFM employ different routing schemes. The coarse spatial resolution and large domain of LSMs or GHMs impedes employing the full shallow water equations or Saint-Venant equations. Instead, they often must limit themselves to simpler approximations, such as the kinematic wave equation. Hydrodynamic models, for instance CaMa-Flood or the LISFLOOD-FP model driving the Fathom GFM, can account for more physical processes by solving higher-order approximations such as the local inertia equations (Bates et al., 2010).

Both hydrologic and hydrodynamic models solve (parts of) the depth-averaged Saint-Venant equations (or “shallow water equations”) of continuity (Eq. 1.1) and momentum (Eq. 1.2) (Chow et al., 1988), either in longitudinal direction only (1D models) or additionally in latitudinal direction (2D models) (Dimitriadis et al., 2016).

In the continuity and momentum equations, Q denotes discharge [$\text{m}^3 \text{s}^{-1}$], A the wetted area [m^2], g the gravitational acceleration [m s^{-2}], h the water depth [m], z the bed elevation [m], R the hydraulic radius [m], and n the Mannings friction coefficient [$\text{m}^{-1/3} \text{s}$].

$$\frac{\delta A}{\delta t} + \frac{\delta Q}{\delta x} = 0 \quad (1.1)$$

$$\frac{\delta Q}{\delta t} + \frac{\delta(Q^2/A)}{\delta x} + gA \frac{\delta(h+z)}{\delta x} + \frac{gn^2|Q|Q}{R^{4/3}A} = 0 \quad (1.2)$$

The first, second, third, and fourth term of the momentum equation (Eq. 1.2) represent the local acceleration, advection (or convective acceleration), water slope (gravity and pressure term), and friction slope, respectively.

Compared to the full shallow water equation, the local inertia equation assumes that the advection term of Eq. 1.2 is negligible compared to the other three terms. The kinematic wave approximation further simplifies flow description by assuming that flow is uniform, and that the friction slope follows the water slope.

Because of these simplifications, LSMs and GHMs lack important discharge dynamics compared to hydrodynamic models, especially for flat areas where ignoring the water level and pressure slope is insufficient. This is shown by previous research illustrating that simulating full hydrodynamics is important to capture flood-triggering processes such as backwater effects (Meade et al., 1991; Moussa and Bocquillon, 1996) and peak discharge (Zhao et al., 2017). Running GHMs at a spatial resolution sufficiently fine to sensibly apply higher-order approximations is, however, currently still a “grand challenge” due to limits in computational power, data for constraining the model, and the risk of trading resolution as accuracy, as discussed in several contributions (cf. Beven et al. (2015), Beven and Cloke (2012), Bierkens et al. (2015), Bierkens (2015), and Wood et al. (2011)).

As outlined above, one possible avenue to overcome the issue of requiring large-scale hydrologic simulations and their seamless output across borders as boundary conditions on the one side and advanced routing schemes including flood-triggering processes on the other side is hydrologic-hydrodynamic model coupling. Research question 3 therefore aims at identifying the potential benefit of such hydrologic-hydrodynamic coupling for large-scale flood wave propagation.

Differences in model grids

As mentioned before, the discretization of models of GFMs can influence results and thus a few words should be spent on the role of spatial resolution and grid design for large-scale inundation modelling.

Except for CaMa-Flood, which makes use of unit catchments, the hydrologic and hydrodynamic models driving the above-listed GFMs typically employ uniform regular grids. Regular grids do have the advantages that input resolution can be aligned with model resolution and that the quality as well as uncertainty of model output does not differ spatially within a discretization (Savage et al., 2016a; Savage et al., 2016b).

The application of regular grids, however, is not necessarily the computationally most efficient approach as fluvial inundations often occur in a fraction of the model, typically close to the river channels. One viable approach to achieve a more computationally efficient discretization is the use of flexible meshes (also named unstructured grids). By means of flexible meshes, cells can have various geometrical shapes (triangles, rectangles, heptahedron, and so forth) and sizes. Employing flexible meshes hence has the advantage that only for areas where dynamics are important, for example floodplains, river bends or areas with a strong concentration of exposure, a fine spatial resolution is employed, while for other less important cells a larger grid size can be applied, yielding a reduction in run time (Figure 1.6; Castro Gama et al. (2013)).

Flexible meshes themselves are not a novelty as such. In fact, there are several hydrodynamic models employing them (for instance ParBreZo or TELEMAC). However, they are mostly applied to local studies (Schubert et al., 2008; Horritt and Bates, 2001b) making use of a sumptuously designed mesh, which requires detailed local data and, hence, renders them not fully transferrable to global studies where often only little data is available. Since

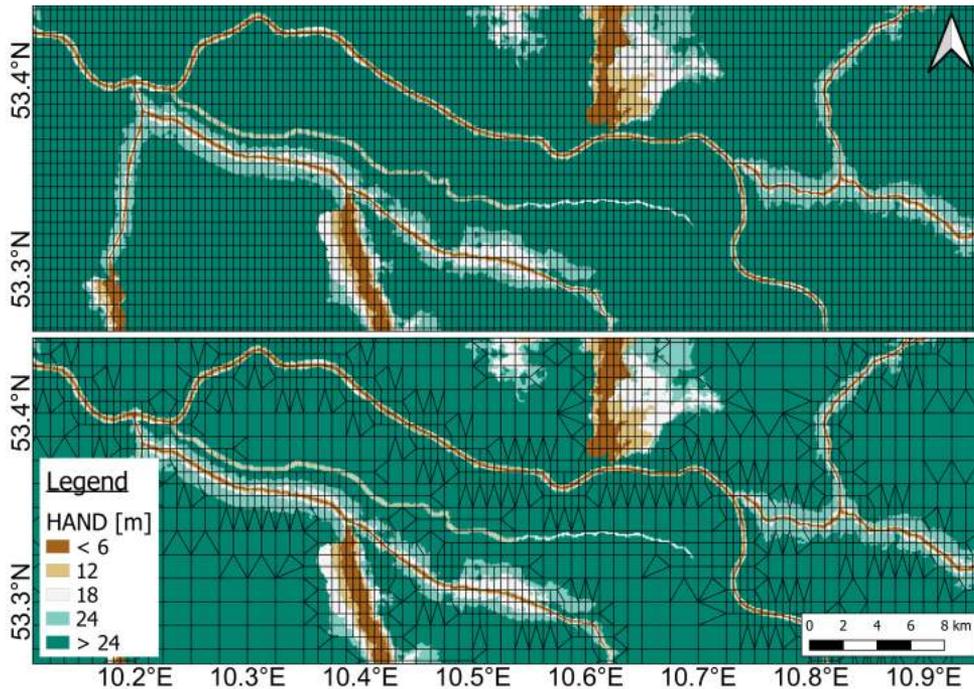


Figure 1.6: Visualization of differences in required grid cells between regular grid (top) and flexible mesh (bottom) for a part of the Elbe river (also see chapter 4); flood-prone areas are indicated using the Height Above Nearest Drainage (HAND; Rennó et al. (2008)) value with lower values being easier to be flooded

they were not applied for larger scales yet, there are no lessons learnt although the application of flexible meshes could potentially prove a time-saving and equally accurate alternative for large-scale fluvial inundation modelling. Research question 4 therefore tries to assess how applying flexible meshes discretized with merely global data can contribute to improved large-scale inundation estimates.

1.5 Research questions and Thesis outline

Following the review section above, this dissertation aims at answering the following research questions.

With respect to research objective 1, the subsequent research questions are formulated:

1. *How can a standardized model coupling framework help benchmarking the performance of different hydrodynamic models?*
2. *How can a model coupling framework be applied to advance the development of global flood models?*

Related to research objective 2, these research questions are raised:

3. *To what extent can model coupling improve the skill of large-scale discharge and inundation extent simulations?*

4. What are pros and cons of employing flexible meshes for fluvial inundation modelling compared to regular grids?

Hence, it is the overarching research theme of the present thesis to what extent and how model coupling can help gaining insight into model applicability and sensitivity as well as how it can improve large-scale inundation estimates.

Consequently, this thesis provides the methodology, applications, and results required for modular coupling of hydrologic and hydrodynamic models for large-scale applications. It begins with a first proof-of-concept of model coupling and extends then to the first development, testing, application as well as evolution of GLOFRIM, a **g**lobally applicable **f**ramework for integrated hydrologic-hydrodynamic **m**odelling. While the eventual objective of GLOFRIM is to enrich global-scale flood models, it can also be applied for specific basin-scale test cases as done in this thesis due to computational and time constraints.



Figure 1.7: The hand-made GLOFRIM logo accompanying the 2.0 release

The remainder of the thesis is structured as follows. In chapter 2, a first proof-of-concept of hydrologic-hydrodynamic model coupling is presented. Chapter 3 then presents GLOFRIM 1.0, a globally applicable framework for integrated hydrologic-hydrodynamic modelling, which is a formalized tool of the code used for the previous chapter. Since applying flexible meshes for large-scale riverine inundation modelling is promising yet still novel, we present an assessment of different mesh discretization approaches in chapter 4. Chapter 5 sketches a platform for standardized GFM validation and outlines ways how a model coupling framework could contribute in advancing the state of GFMs. On-going developments towards GLOFRIM 2.0 and applications of GLOFRIM are outlined and discussed briefly in chapter 6. The dissertation is concluded with a synthesis of the presented work, including final remarks and recommendations for further research.

Chapter 2

Assessing the impact of hydrodynamics on large-scale flood wave propagation – a case study for the Amazon Basin

Large-scale flood events often show spatial correlation in neighbouring basins, and thus can affect adjacent basins simultaneously, as well as result in superposition of different flood peaks. Such flood events therefore need to be addressed with large-scale modelling approaches to capture these processes. Many approaches currently in place are based on either a hydrologic or a hydrodynamic model. However, the resulting lack of interaction between hydrology and hydrodynamics processes, by for instance implementing groundwater infiltration on inundated floodplains, can hamper modelled inundation and discharge results where such interactions are important. In this study, the global hydrologic model PCR-GLOBWB (PCR) at 30 arcmin spatial resolution was one-directionally and spatially coupled with the hydrodynamic model Delft3D Flexible Mesh (DFM) for the Amazon River basin at a grid-by-grid basis and at daily time step. The use of a flexible unstructured mesh allows for fine-scale representation of channels and floodplains, while preserving a coarser spatial resolution for less flood-prone areas, thus not unnecessarily increasing computational costs. In addition, we assessed the difference between a 1D-channel/2D-floodplain and a 2D schematization in DFM. Validating modelled discharge results shows that coupling PCR to a hydrodynamic routing scheme generally increases model performance compared to using a hydrodynamic and hydrologic model only for all validation parameters applied. Closer examination shows that the 1D/2D schematization outperforms 2D for r^2 and RMSE whilst having a lower KGE. We also found that spatial coupling has the significant advantage of a better representation of inundation at smaller streams throughout the model domain. A validation of simulated inundation extent revealed that only those set-ups incorporating 1D channels can represent inundations for reaches below the spatial resolution of the 2D mesh. Implementing 1D channels is therefore particularly of advantage for large-scale inundation models as they are often built upon remotely sensed surface elevation data which often enclose a strong vertical bias, hampering downstream connectivity. Since only a one-directional coupling approach was tested, and therefore important feedback processes are not incorporated, simulated discharge and inundation extent for both coupled set-ups is generally overpredicted. Hence, it will be the subsequent step to extend it to a two-directional coupling scheme to obtain a closed feedback loop between hydrologic and hydrodynamic processes. The current findings demonstrating the potential of one-directionally and spatially coupled models to obtain improved discharge estimates form an important step towards a large-scale inundation model with a full dynamic coupling between hydrology and hydrodynamics.

Based on: Hoch, J. M., Haag, A. V., van Dam, A., Winsemius, H. C., van Beek, R., Bierkens, M.F. P. (2017), Assessing the impact of hydrodynamics on large-scale flood wave propagation – a case study for the Amazon Basin, *Hydrology & Earth System Sciences* 21 (1), 117–132.

2.1 Introduction

Global flood risk is increasing at an accelerating rate due to a combination of changed climatic conditions and intensified urbanization in proximity to rivers (cf. Ceola et al. (2014), Ward et al. (2015), Jongman et al. (2012), and Winsemius et al. (2016)). This is reflected by a significant increase in economic losses in the latter half of the 20th century associated with flooding. In 2012 alone, economic losses exceeded USD19 billion, comprising one-third of all losses due to natural hazards (MunichRe, 2010; UNISDR, 2015a; Visser et al., 2012). To better understand current and future hazards and risks, and to facilitate robust climate change adaptation and mitigation measures, this study aims to show the strengths, weaknesses, and opportunities of spatially coupled hydrologic-hydrodynamic models compared to mere hydrologic and hydrodynamic models, respectively. We believe that coupling models is a pivotal cornerstone for more realistic, robust, and integrated flood hazard and risk assessments.

Recently, modelling flood hazards and risks experienced a boost in attention as flood hazard maps are paramount for sound flood risk assessments (Hagen and Lu, 2011). In many cases, however, flood hazard maps are computed for geographically limited areas only. Because flood waves show strong spatial correlation in different but neighbouring basins, they can be considered to be large-scale phenomena, and, in turn, demand large-scale modelling approaches (Jongman et al., 2014), especially over data-scarce areas (Ward et al., 2015). The outcome of such large-scale models may be beneficial for global stakeholders as the United Nations Office for Disaster Risk Reduction (UNISDR) or the World Bank, for instance, to facilitate discussions with stakeholders risks, better allocate their funding, but also for re-insurance companies or governmental entities (Ward et al., 2015). Tiling small-scale maps from different small-scale studies to obtain the required large-scale estimates is not a viable alternative, as it introduces many sources of uncertainty and inconsistencies (Pappenberger et al., 2012; Pappenberger et al., 2006) and does not account for any spatio-temporal correlation. Recent studies aimed to model large-scale flood hazard by dividing the model domain into various catchments (Alfieri et al., 2014; Dottori et al., 2016; Sampson et al., 2015). Notwithstanding the promising results, such approaches still require upstream boundary forcing, additional efforts due to division and merging, and still cannot fully account for the aforementioned spatial correlation of flood events in neighbouring basins, as they use synthetic flood events.

Triggered by an increase in computational capacities and in availability of remotely sensed data for parameterization, calibration, and validation, research on large-scale inundation modelling was intensified in past years. For example, a range of global data sets is by now freely available such as, inter alia, digital elevation maps (DEMs) (eg. HydroSHEDS, Lehner et al. (2008); ASTER; GTOPO30), water body maps (eg. G3WBM Yamazaki et al. (2015)), global river width and depth (eg. GWD-LR, Yamazaki et al. (2014a); GRWL, Allen and Pavel-sky (2018); Andreadis et al. (2013)), or observed river discharge (Global River Discharge Centre (GRDC); Global River Discharge Project (RivDIS)). In addition, algorithms to quantitatively describe topography (height above nearest drainage (HAND), Rennó et al. (2008)), or to apply surface reconditioning (Yamazaki et al., 2012a) were presented.

With these data sets and algorithms being available, large-scale flood hazard modelling approaches are strongly facilitated. Most of the approaches can be categorized by (a) the processes represented and (b) the model schematization. While the latter category comprises possible schematizations such as 2D grids, 1D channels, or coupled 1D/2D models, the first contains the possibility to include or exclude several hydrologic or hydrodynamic models or their components in the computational backbone.

Global hydrologic models (GHMs), such as PCR-GLOBWB (PCR; van Beek et al. (2011) and Sutanudjaja et al. (2018)), WaterGAP (Döll et al., 2003), or the variable infiltration capacity

(VIC) model (Liang et al., 1994; Wood et al., 1992), are capable of modelling water balances, and hence available surface water volumes, at the global scale. Another advantage is that hydrologic models can easily be forced with ensembles of global climate models (GCMs), which is beneficial for predictions of future changes in flood hazard and risk (cf. Hirabayashi et al. (2013), Jongman et al. (2014), Winsemius et al. (2016), and Sperna Weiland et al. (2010)). However, large-scale hydrologic models strongly depend on the quality of their input data and robustness of their process descriptions, which may differ remarkably between individual catchments (Kling et al., 2015; Li et al., 2015). Besides, many GHMs are relatively coarse scale, with the finest spatial resolution for global models currently being 5 arcmin or 10 km x 10 km the Equator (Bierkens, 2015). Although sub-grid post-processing can be used to meliorate outcomes as done, for instance, in the Global Flood Risk with IMAGE Scenarios (GLOFRIS) framework (Winsemius et al., 2013), this may reduce model accuracy, since important floodplain properties and channel-floodplain dynamics can only be implemented in a simplistic manner.

Dedicated hydrodynamic models, on the other hand, put their emphasis on the correct simulation of surface water flow and levels, and hence consider important factors such as inertia terms of channel geometry, in more detail than most large-scale hydrologic models, as the latter often employ kinematic wave or Muskingum-Cunge approaches only. Thus, hydrodynamic models allow for simulating backwater effects which are pivotal flood-triggering processes (Moussa and Bocquillon, 1996; Paiva et al., 2013). Hydrodynamic models are usually forced with upstream boundary conditions based on regionalization of observation stations (Huang et al., 2014; Sampson et al., 2015; Wilson et al., 2007). Yet, using observed boundary conditions makes them highly dependent on the presence and spacing of the stations.

The aforementioned spatial correlation of flood waves can thus not realistically be modelled, as important spatially distributed flood-triggering processes such as precipitation events over large surface areas would not necessarily be captured by the stations (for instance, the El Niño–Southern Oscillation (ENSO) phenomenon in the Amazon River basin (Molinier et al., 2009)). Most hydrodynamic modelling approaches are implemented by employing 1D, 2D, or 1D/2D schematizations. Mere 1D models, however, have difficulties with modelling surface flow over larger areas and floodplains specifically, while regular 2D models inevitably lead to an increase in required computational power, especially if results need to be computed at a fine spatial resolution (Finaud-Guyot et al., 2011; Liu et al., 2015). In addition, 2D models experience problems in case the actual river width is smaller than the grid size and also in case there are multiple rivers within one cell, although it is possible to partly overcome that by applying sub-gridding routines (Neal et al., 2012a). Besides, flow resistance to surface roughness is overestimated in 2D set-ups.

In addition to the currently employed techniques, use of flexible meshes is emerging, which allows for both a fine spatial resolution in more relevant areas while at the same time not unnecessarily increasing computational costs where only limited dynamics and changes are expected. Such flexible gridding over the model domain may moreover be a viable avenue to meet the debated grand challenge of hyper-resolution modelling (Bierkens et al., 2015; Wood et al., 2011). Yet, the application of flexible meshes focussed so far mostly on oceanic and coastal computations (Chen et al., 2003; Muis et al., 2016; Martyr-Koller et al., 2017) and less on the representation of rivers and floodplains, although studies corroborate its high potential (Castro Gama et al., 2013).

Based on this, a call for a more holistic large-scale modelling approach can be formulated. Coupling existing models may provide a way forward as the strengths of individual models are maintained and weaknesses compensated. In fact, many studies already integrate various

disciplines by model coupling, for instance, hydrologic with atmospheric models (Senatore et al., 2015; Wagner et al., 2016), climate models (Zabel and Mauser, 2013; Butts et al., 2014) or glacier models (Naz et al., 2014; Zhao et al., 2013). To obtain information about inundation patterns, approaches to couple hydrology with hydrodynamics were already explored in previous studies, but either at the sub-catchment scale only (Paiva et al., 2013; Rudorff et al., 2014a; Rudorff et al., 2014b); by using a land surface model (LSM) to obtain input (Pappenberger et al., 2012); by employing VIC to compute boundary discharge for LISFLOOD-FP in the Lower Zambezi River (Schumann et al., 2013); by using output from a hydrologic model as lateral inflow for LISFLOOD-FP to model inundation dynamics in the Ob River (Biancamaria et al., 2009); or by using used output from the Global Flood Awareness System (Alfieri et al., 2013; GloFAS, 2018; Emerton et al., 2018) with hydrodynamics to obtain synthesized floods with different return periods (Dottori et al., 2016).

Notwithstanding the contributions of these studies to current flood risk understanding, they still lack the capability to produce hydrological forcing within the actual model domain and are thus not able to simulate the feedback between hydrology and inundation processes on floodplains.

In the present study, we present a one-directional and spatially explicit coupling approach between PCR and the hydrodynamic model DFM, allowing for the exchange change of information throughout the entire model domain. To our knowledge, this is a novelty in large-scale inundation modelling. Moreover, the exchange of variables between hydrology and hydrodynamics takes place on a grid-to-grid basis at the time-step or even sub-time-step level. This approach allows for online coupling, thus providing the potential to eventually perform two-directional exchange of information. The Amazon River basin was schematized with both a 2D flexible mesh and a 1D/2D set-up, allowing us to test potential (dis)advantages between both set-ups. Additionally, the hydrologic and hydrodynamic models were also run in a stand-alone mode to fully assess the added value of model coupling. The utilization of only global data sets and algorithms ensures transferability to other basins as well as a straightforward scalability of our approach to larger scales. It is moreover a part of the study's aim to detect the most suitable model set-ups to continue with future extensions and larger-scale applications of our coupling technique. With our approach, we are confident in our ability to close the gap between hydrology and hydrodynamics, and to make a step towards a global, fully fledged inundation model. Such a model set-up can provide information on spatial correlations and interrelations between flood events, ultimately facilitating current large-scale flood hazard and risk assessments. Eventually, this can be used for the formulation of more robust climate change adaptation and mitigation measures, and to further inform global flood risk policies.

2.2 Methodology

The two models used for this study are the global hydrologic model PCR-GLOBWB (PCR; van Beek et al. (2011), van Beek (2008), and Sutanudjaja et al. (2018)), and the hydrodynamic model Delft 3D Flexible Mesh (DFM; Kernkamp et al. (2011)). To test the added value of our coupling approach as well as the differences between 2D and 1D/2D schematization, the following experimental set-up was designed, consisting of five modelling runs: (i) PCR-GLOBWB with its DynRout extension to obtain purely hydrology-based results; (ii) a 2D and (iii) 1D/2D DFM schematization both forced with discharge observed at GRDC stations to obtain purely hydrodynamic-based results; (iv) and (v) the same two DFM schematizations forced with output from PCR. For all runs with DFM, a constant water level of 0.0 m is assumed at the river mouth as a downstream boundary. Even though the influence of ocean

tides is reported to be significant (Lima et al., 2003), tidal dynamics were not considered in the present study, as this exceeds the scope of the work.

Each set-up was applied for the Amazon Basin for the period from 1 January 1985 to 31 December 1990. This early period had to be chosen, as for some GRDC stations no more recent discharge data are available. Output of all cases was validated against observed GRDC discharge data at Óbidos (GRDC station no. 3629000), the most downstream GRDC station available (Figure 2.1). To this end, three functions were applied for validation: the coefficient of determination (r^2) to assess the reproduction of the shape of the hydrograph, the root mean squared error (RMSE) to assess the water balance, and the Kling–Gupta efficiency (KGE) (Gupta et al., 2009) to evaluate the models skill. In addition, we qualitatively inspected the inundation extent and water levels for the various model runs. We employed LandSat imagery taken on 1 July 1989 to validate simulated inundation extent, as it is one of the few cloud-free images at this time, and represents inundation patterns during peak season. To compare simulated water levels, four observation points on floodplains along the main river reach were defined (Figure 2.1): Loc1 close to the delta (1.62° S, 52.46° W); Loc2 downstream of Óbidos (2.15° S, 54.55° W); Loc3 just upstream of Óbidos (2.45° S, 56.81° W); Loc4 even further upstream (2.97° S, 58.35° W).

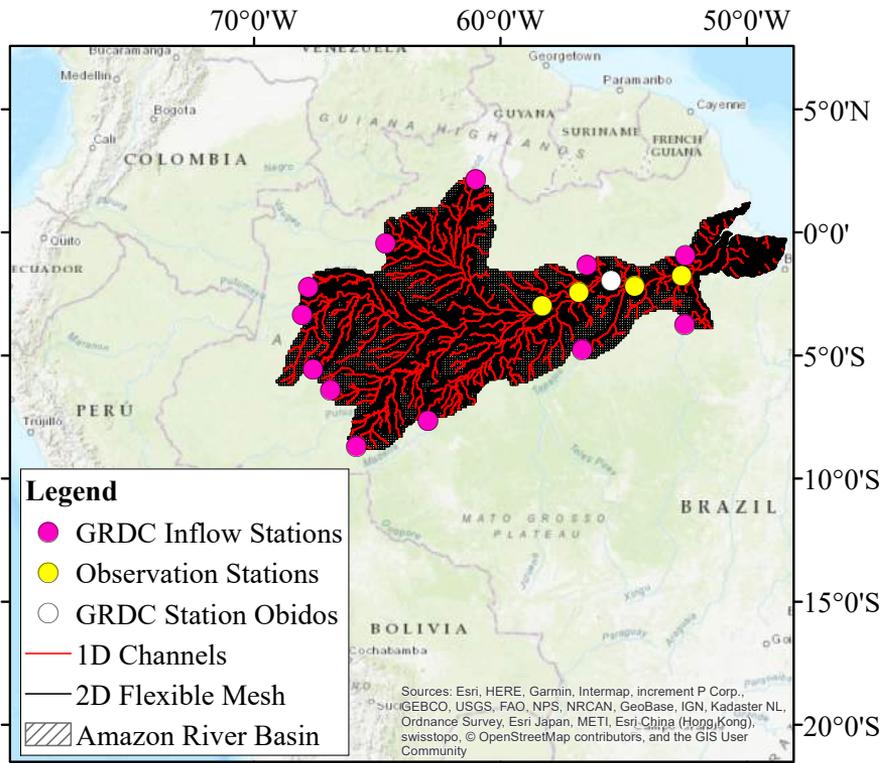


Figure 2.1: Map of the extent of 2D grid and 1D channels as part of entire Amazon River basin; additionally shown are the water level observation stations 1-4 counting from delta to upstream, as well as GRDC station Óbidos for discharge measurements and validation of simulations, and all GRDC input stations.

2.2.1 The hydrologic model: PCR-GLOBWB

To generate hydrologic input, PCR was applied at 30 arcmin resolution (approximately 55 km x 55 km at the Equator). It is entirely coded in PCRaster Python (Karssenberg et al., 2010). PCR distinguishes between two vertically stacked soil layers, an underlying groundwater layer, and a surface canopy layer. Water can be exchanged vertically, and excess surface water can be routed horizontally along a local drainage direction (LDD) network. In the present study, the kinematic wave approach was used for routing, and Manning's surface roughness coefficient was uniformly set to $0.03 \text{ m}^{-1/3} \text{ s}$. This value is in line with other studies in the Amazon Basin (cf. Paiva et al. (2013), Rudorff et al. (2014a), Rudorff et al. (2014b), Trigg et al. (2009), and Yamazaki et al. (2011)). A uniform value was chosen to eliminate this factor as a cause for differences when comparing the stand-alone runs as well as 1D/2D set-ups. The model was forced with Climate Research Unit (CRU) precipitation and temperature data (Harris et al., 2014), and evaporation was computed using the Penman-Monteith equation. Data sets were downscaled to daily fields for the period from 1957 to 2010 using ERA40/ERA-Interim (Kállberg et al., 2005; Uppala et al., 2005). For more information on PCR, we refer to van Beek et al. (2011), van Beek (2008), and Sutanudjaja et al. (2018).

PCR was already applied in various studies: Sperna Weiland et al. (2010) investigated how forcing from different global circulation models can reproduce global discharge variability; Candogan Yossef et al. (2012) concluded that PCR shows skill when used for flood forecasting; Wanders and Wada (2015) employed the model to assess the impact of humans and climate on drought in the 21st century; de Graaf et al. (2015) fully coupled PCR with a physically based groundwater model capable of simulating lateral flows.

From a priori runs, we were informed that PCR underestimates discharge in the Amazon Basin. To eventually obtain discharge values that are close to observed values and enhance the significance of the validation procedure, we decided to apply a simple regional model parameter optimization technique. To this end, we tested the models performance sensitivity to a range of multipliers for these parameters, using the log-scaled Nash-Sutcliffe coefficient of simulated discharge at Óbidos as a performance indicator. Based on performance, we then chose the combination of multipliers resulting in the highest log-scaled Nash-Sutcliffe coefficient. Consequently, the minimum soil depth fraction for which interflow is calculated, the log-scaled saturated hydraulic conductivity of groundwater flow (k_{sat}), and the log-scaled recession coefficient were multiplied by 0.5. The general parameterization of PCR, however, remained unaffected and no further local calibration was performed to preserve the global applicability of the model. PCR also has the option to include human water use from irrigation, households, and industry as an integral part of its model runs. In our application, however, we decided to simulate river discharge under natural flow conditions.

2.2.2 The hydrodynamic model: Delft 3D Flexible Mesh

For hydrodynamic calculations, the hydrodynamic model DFM was employed (Kernkamp et al., 2011). It allows the user to schematize the model domain with a flexible mesh in 1D/2D/3D, and therefore supports the computationally efficient schematization of topographically challenging areas such as river bends or irregular slopes. The model solves the full Saint-Venant equations, or shallow water equations (SWEs; see Eqs. 1.1, 1.2). Solving the SWEs is, as stated before, a major advantage compared to most large-scale hydrodynamic and hydrologic models because this is essential to account for important flood-triggering processes such as back-water effects (Moussa and Bocquillon, 1996; Paiva et al., 2013). In analogy to PCR, the surface roughness coefficient was set to $0.03 \text{ m}^{-1/3} \text{ s}$ to guarantee comparability. We expressly desisted from calibrating model parameters for any of our hydraulic model set-

ups due to two reasons. First, calibration may obscure the actual performance of the model set-ups with respect to real governing hydraulic processes as their quality may possibly be governed by calibration. Second, we aim to apply the presented coupling scheme at other basins. Locally calibrating FM for the Amazon Basin may introduce inconsistencies among the global data sets used for model set-up, and jeopardize their validity for ungauged basins. Due to its very recent publication, only a limited number of published studies using DFM are available (cf. Muis et al. (2016), Martyr-Koller et al. (2017), and Chen et al. (2003)). Besides, Castro Gama et al. (2013) applied DFM successfully to model flood hazard at the Yellow River and concluded that applying a flexible mesh reduces computation time by a factor of 10 compared to square grids with equal quality of model output.

2.2.3 Defining the 1D network

The course of the 1D river channels as well as effective river width w were derived based on the Global Width Database for Large Rivers (GWD-LR) algorithm by Yamazaki et al. (2014a), hence already accounting for river braiding and islands. Comparing both the course and the computed width of the obtained 1D network schematization with OpenStreetMaps (OSM) yielded an overall good fit with lower goodness of fit in meandering and delta regions (Figure 2.2). River depth d [m] was subsequently estimated from river width w [m] by combining the following equations from Paiva et al. (2011), with A_d in km^2 being the upstream area of one point along the river:

$$w = 0.81 * A_d^{0.53} \quad (2.1)$$

$$d = 1.44 * A_d^{0.19} \quad (2.2)$$

to the following width–depth relation:

$$d = 1.55 * w^{0.36} \quad (2.3)$$

Benchmarking the resulting river depths obtained with Eq. 2.3 with those found in a global river bankfull width and depth database by Andreadis et al. (2013) showed better results than those obtained with the widely used width–depth relation proposed by Leopold and Maddock (1953). By means of the aforementioned equations, a maximum depth of 54 m, a minimum depth of 5 m, and an average depth of 13 m were computed. Finally, width and depth information was stored in cross sections along the network with a spacing of around 20 km (Figure 2.2).

We are confident that applying a hydro-geomorphic relation between river depth and observed width, as also applied by Neal et al. (2012a), is valid in this case due to three reasons. First, the equations were constructed based on a large sample of cross-section information, and thus can be extrapolated over larger areas of application. Second, we consider the width information of GWD-LR to reflect local conditions better than the more one-dimensional approach of relating bathymetric information on only one predicting variable such as upstream area or discharge. Last, bathymetric information is internally interpolated in the model, preventing any extreme local variations in observed river width and depth, consequently also avoiding improbable local flow hindrance.

2.2.4 Defining the 2D flexible mesh

For surface elevation values, we used the HydroSHEDS data set, which was derived from the Shuttle Radar Topography Mission (SRTM) (Lehner et al., 2008). Because significant vertical

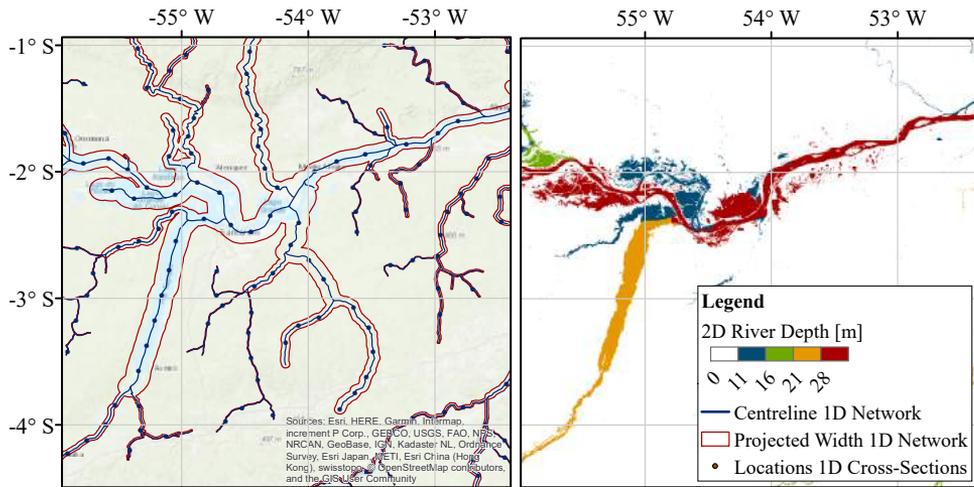


Figure 2.2: Comparison of the 1D network centre line as well as its computed (projected) width with river course and width according to OpenStreetMaps. Additionally shown are the locations of the cross sections in this area and the computed river depth for all cells defined as permanent water bodies in G3WBM.

measurement errors emanate from the C-band synthetic aperture radar (SAR) used by SRTM, extensive hydrologic conditioning was carried out in this study to remediate the most relevant errors in currently available data sets.¹

First, noise by vegetation cover was reduced. This is essential as the radar signal cannot fully penetrate dense canopy, leading to quality degradation especially in rainforests (Berry et al., 2007). As a result, absolute vertical errors of around 22 m were found in the Amazon Basin (Carabajal and Harding, 2006; Sanders, 2007). The approach used in the present study to account for vegetation cover is described in detail by Baugh et al. (2013). For the present study, 50% of canopy heights reported by Simard et al. (2011) were subtracted from original elevation values, as proposed by Baugh et al. (2013).

Even after vegetation was removed, flow connectivity can be hindered by grid cells surrounded by higher elevated cells which can stem from elevation irregularities such as islands, bridges, or other residues. Thus, these local depressions were removed in a second step to guarantee downstream flow connectivity along flow paths. Conventional procedures, such as lifting downstream cells or stream burning, fail, however, to adequately address this issue as the land surface is altered to be one-sided, and thus should not be applied to rivers in flat environments such as the Amazon River (Getirana et al., 2009). Hence, a more advanced algorithm based on the work of Yamazaki et al. (2012a) was applied. This algorithm either digs or fills along a flow path, as defined by the HydroSHEDS LDD, resulting in smoothed elevation values along downstream flow paths as demonstrated for two flow paths in Figure 2.3.

While, for 1D/2D applications, the 1D vector channel data are embedded into the smoothed 2D elevation, it was necessary to compute bathymetric information for the 2D schematizations. This is because the DEM used lacks reliable information about river bathymetry as the SRTM radar signal is not able to fully penetrate deeper water bodies. To derive bathymetry information, current research projects aim to exploit available remotely sensed data or aerial

¹Many of the issues related to the original SRTM elevation data were corrected since writing this article. Currently, the MERIT-DEM (Yamazaki et al., 2017) represents the most error-free global elevation data set.

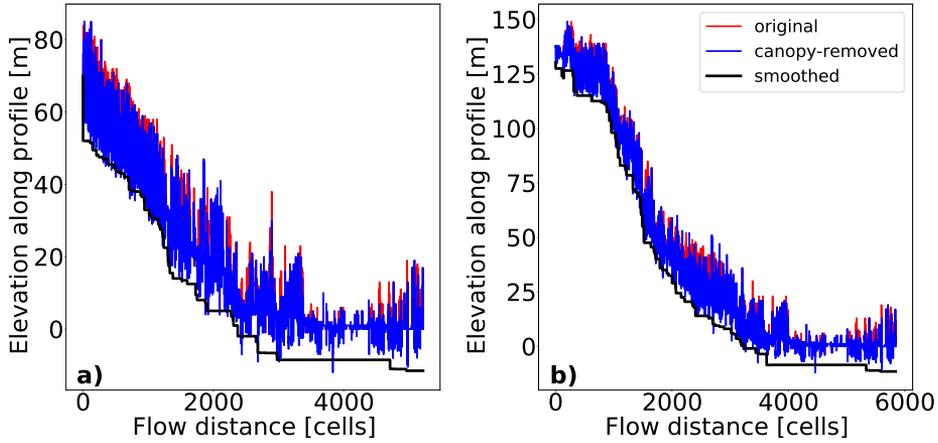


Figure 2.3: Impact of vegetation removal (“canopy-removed”) and surface reconditioning (“smoothed”) on surface elevation along two exemplary flow paths compared to original HydroSHEDS DEM data (“original”). Flow path **a)** starting at index [2500, 3000], flow path **b)** starting at index [4500, 3500]

photography (cf. Kinzel et al. (2013), Legleiter (2015), and Legleiter (2016)). Yet, obtaining satisfactory information for large-scale river bathymetry remains a major research challenge. For the present study, river depth d was computed as a function of upstream area A_d as follows: for all grid cells where $A_d > 10^4 \text{ km}^2$, Eq. 2.2 was applied to compute d on a grid-by-grid basis. The threshold of 10^4 km^2 was chosen after trial and error to filter many small and short reaches which were not represented by the 1D network. Due to the differences in the 1D vector network and LDD map used for the 2D raster data, it was, however, not possible to precisely apply the same equations. Despite these minor differences in methodology, manual inspection of computed river depths, computed for 1D channels and 2D bathymetry, revealed no major discrepancies in our model domain, and we therefore consider both ways to compute valid bathymetry, particularly in light of the limited availability of bathymetry data for large-scale applications. The computed depth of one specific pixel was then spread to all cells whose distance is shortest to the pixel under consideration.

Subsequently, the resulting bathymetry map was created by lowering elevation values of only those pixels defined as permanent water bodies in the Global 3-second Water Body Map (G3WBM) developed by Yamazaki et al. (2015) (see Figure 2.2). The computed elevation values were then interpolated over the flexible mesh, and elevation values per DFM cell are obtained by unweighted spatial averaging of the computed elevations at the cell vertices.

Since the hydrodynamic computations and model coupling still require significant computational power for multiyear simulations, the modelling domain of DFM was limited to flood-prone areas. To derive a suitable extent, the height above nearest drainage (HAND) algorithm was applied (Rennó et al., 2008), as it yields relative terrain elevation to the nearest hydrologically connected drainage. The flexible mesh was then obtained by automatic local grid refinement of a coarser regular grid based on the obtained HAND values and limiting it to grids where computed HAND values are less than or equal to 25 m; that is, until terrain reached an elevation of 25 m above the nearest water body. The final model domain is presented in Figure 2.1 and still encompasses an area of around $1.2 \cdot 10^6 \text{ km}^2$, which is nearly a

fifth of the entire Amazon River basin. The threshold was chosen arbitrarily but model results showed that it is sufficiently large.

By establishing the refinement on this algorithm, the flexible mesh has the finest spatial resolution (2.5 km x 2.5 km) for areas with lowest HAND values, such as water bodies and floodplains, while areas with higher HAND values, and hence areas more remote from water bodies, are modelled with coarser spatial resolution up to 10 km x 10 km per grid. In these latter regions, the number of grid cells is thus reduced by a factor of 16, benefitting the stability-limited computational time step and significantly reducing overall computation times.

2.2.5 Coupling the models

Coupling PCR with DFM was achieved by means of the basic model interface (BMI). Peckham et al. (2013) proposed the BMI as a tool within the Community Surface Dynamics Modeling System (CSDMS) project to exchange information between separate models at any given time step. By exposing certain internal state variables of the model by means of the BMI, interactive modelling is facilitated, as these variables can be modified during the model execution.

Generally, each BMI has several functions that can be called from external applications like, as in this case, a Python script. First, models need to be initialized. Second, the BMI enables the user to retrieve variables, and to manipulate them if required, for instance, to convert units or to add values. Third, the manipulated variables can be set back to the original model or can be used to overwrite variables in one or multiple other models, given that they agree to the internal data structure of those models. Fourth, models connected to a BMI can be updated at a user-specified time step. This way, it is possible to get, change, and set variables during the execution of the models in use. In a last step, models can be finalized to end the computations. It has to be noted that for each model involved, one specific BMI adapter has to be developed with respect to the specific internal model structure and programming language.

Whilst PCR is already in Python and its BMI implementation is hence straightforward, DFM offers a native C-compliant BMI implementation which can be called from within Python using the BMI Python package (see www.github.com/openearth/bmi-python). For further information on the BMI, refer to Peckham et al. (2013), the related website CSDMS (2019a), and chapter 1.3.2.

In order to be able to spatially couple both models, it is required to overlay the model extent of both DFM and PCR. To this end, the centroid of each 2D DFM cell was computed, and a DFM cell is then considered to be coupled to PCR if its centroid is located within the bounds of the PCR cell.

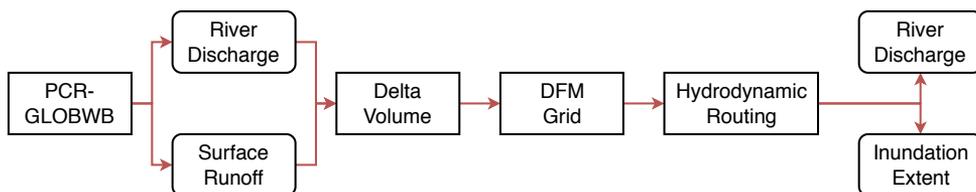


Figure 2.4: Flow diagram of the various steps required for model coupling. Retrieving the river discharge and surface runoff information as well as adding the delta volume onto the DFM grid was achieved by means of the BMI functions.

The coupling algorithm (Figure 2.4) was employed at a daily time step: first, PCR was run for one day; then, a daily delta volume (that is, the volume to be added to DFM with the day's time step) was computed for every coupled PCR cell as the sum of daily river discharge inflows at the boundary of DFM and local surface runoff throughout the model domain. The daily delta volume was subsequently divided over and added to all DFM cells within this specific PCR cell. Note that this explicit spatial forcing of DFM is fundamentally different from the GRDC-fed runs, where only upstream discharge boundary conditions are applied, and no spatially distributed forcing is active.

As only the most downstream part of the Amazon basin is schematized in DFM, no coupling was performed for the upper part of the basin. For these uncoupled areas, PCR is run in stand-alone mode, and water is routed towards the coupled domain using the kinematic wave approximation. Within the coupled area, the LDD of PCR was deactivated to prevent further routing in the hydrologic model. As a last step in the coupling algorithm, DFM was updated and integrated forward in time until it reached the same model time step as PCR to compute daily inundation and discharge values. Since only a one-way coupling approach is tested, water added to DFM can only be routed downstream, but cannot infiltrate or evaporate, most likely leading to overestimation of modelled discharge and inundation.

2.3 Results and discussion

2.3.1 Discharge simulation at Óbidos

Hydrology-only runs: PCR-DynRout

PCR-DynRout reproduces low flows well, but fails in reproducing the observed variation in discharge as shown by a low coefficient of determination (Table 2.1). This low value can be attributed to the rugged hydrograph obtained, as shown in Figure 2.5. The strong fluctuations cannot be fully explained, but we assume that they may be related to the simplistic routing scheme used, as discharge results for the coupled run do not show such behaviour, although they receive the same hydrologic input. In addition, peak discharge is generally modelled too early. This low performance is related to PCR-DynRout being a global hydrologic model, thus not specifically designed for simulating discharge at the basin scale despite the regional optimization technique applied for this study. The employed kinematic wave approximation as well as the coarse resolution of 30 arcmin can be identified as factors currently hampering a more accurate simulation of discharge.

Hydrodynamic-only runs: DFM forced with GRDC data

Forcing DFM only with discharge observed at GRDC stations, we found that the aggregated input discharge as obtained from upstream GRDC station observations (Figure 2.1) accounted for only 59% of the discharge generated in the basin as observed at Óbidos (Figure 2.5). This underrepresentation can be linked to the discrepancy between catchment area at Óbidos and summed catchment area of all input stations upstream of Óbidos. Comparing both, we found that only 63% of upstream catchment area at Óbidos is accounted for by input stations (Table 2.2).

The differences between accumulated and observed discharge can therefore be attributed to the additional discharge created in the intermediate area between Óbidos and the upstream inflow stations. To avoid the expectable discharge estimates that are too low and facilitate comparability with other model runs, we therefore decided to scale the input discharge values accordingly. The results then reveal that the strength of purely hydrodynamic runs is the correct reproduction of discharge variability, as shown by high coefficients of determination.

| | PCR-DynRout | 2D GRDC | 1D/2D GRDC | 2D 1way | 1D/2D 1way |
|-------|-------------|---------|------------|---------|------------|
| r^2 | 0.49 | 0.92 | 0.85 | 0.77 | 0.83 |
| RMSE | 34,100 | 16,229 | 18,735 | 21,451 | 19,548 |
| KGE | 0.64 | 0.80 | 0.86 | 0.84 | 0.79 |

Table 2.1: Performance of model runs in objective functions for both actual and scaled model input. The term 1way indicates one-way coupled runs.

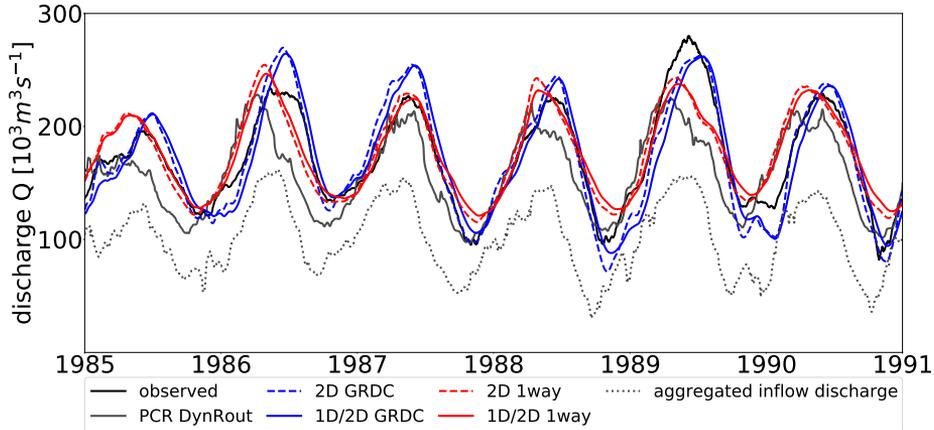


Figure 2.5: Plot of simulated discharge obtained from all model runs and observed discharge values at GRDC station Óbidos. Also plotted is the aggregated input discharge as computed from all GRDC stations upstream of Óbidos (see also table 2.2).

Still, model results obtained with only DFM resulted in lagged discharge, with the 1D/2D schematization having lower discharge results and a larger time lag. We suspect that the obtained attenuation and time lag for both 2D and 1D/2D schematization result from the absence of any internal forcing. By using only upstream discharge boundaries and neglecting internal sources, discharge will need longer to propagate until Óbidos due to the larger average travel distance. It should be noted that, from a computational point of view, the 1D/2D set-up has the advantage of a 25% lower wall clock time required to finish the simulation period compared to the 2D set-up.

Coupled hydrologic-hydrodynamic runs: DFM forced with PCR output

Assessing model results for the coupled runs, we see that the simulated discharge is higher than that of both the purely hydrology-based and purely hydrodynamic-based models. Deviations between coupled and GRDC runs can be ascribed to differences in forcing, which are not only different in terms of input volumes but also in terms of input locations. We also find that the coupled runs do not reach the same variability in discharge as the GRDC-forced runs, although they are employing the same model schematizations. This may be related to a higher proportion of overland flow resulting from distributing water volumes over the DFM cells, which would reduce discharge dynamics. The disparities in discharge of coupled runs compared to PCR-DynRout, however, have to be attributed to a combination of differences

| Type | Name | Number | Catchment area |
|---|-----------------------|---------|----------------|
| i | Caracaraí | 3618500 | 124,980 |
| i | Uaracu | 3618950 | 40,506 |
| i | Acanaui | 3621200 | 242,259 |
| i | São Paulo de Olivenca | 3623100 | 990,781 |
| i | Gavião | 3624120 | 162,000 |
| i | Arumã-Jusante | 3625310 | 359,853 |
| i | Porto Velho | 3627040 | 954,285 |
| i | Ji-Paraná (Rondônia) | 3627408 | 32,606 |
| i | Estirao Da Angelica | 3628500 | 26,040 |
| $\Sigma(\text{catchment area of input stations})$ | | | 2,933,310 |
| o | Óbidos | 3629000 | 4,640,300 |
| <i>Proportional representation catchment area</i> | | | 63% |

Table 2.2: List with catchment area per GRDC station located upstream of Óbidos (type i) compared to catchment area of observation station Óbidos (type o). All data are sourced from the official GRDC website (<http://www.grdc.br/unh.edu>).

between model schematizations and process representation as we have carefully examined the water balance throughout the entire coupling process, and therefore can exclude volume errors as sources of deviations. First, DFM and PCR-DynRout differ in their spatial resolution, with the latter having a much coarser spatial resolution.

Eventually, this difference can have an impact on modelled discharge accuracy, because the role of channel-floodplain interaction is pivotal for inundation and discharge estimates and so is schematization of connecting channels (Neal et al., 2012a; Savage et al., 2016a; Rudorff et al., 2014a) which both are facilitated by using finer spatial resolution. This is underlined by the smoothed daily discharge which results when replacing the simple kinematic wave routing at 30 arcmin spatial resolution with a hydrodynamic model at fine spatial resolution, even though both are subject to the same meteorological forcing as well as hydrologic processes. Second, differences in process description can lead to improved discharge estimates compared to PCR-DynRout. In particular, solving the SWE – as implemented DFM – instead of the kinematic wave approximation may have influenced results, as it accounts for back-water effects which play an important role in the Amazon Basin because of its low gradients (Meade et al., 1991; Moussa and Bocquillon, 1996; Paiva et al., 2013). Third, our coupled set-ups may yield higher discharge than PCR-DynRout due to the one-directional coupling scheme implemented. For peak flow conditions, the higher discharge can be attributed to the absence of important groundwater infiltration and evaporation processes on inundated areas, resulting in increased surface water volumes routed downstream. Note that in PCR-DynRout flooded areas are subject to evaporation which can partly explain the higher discharge resulting from the one-directionally coupled model. During low flow conditions, however, the excess water that remained on the floodplains, although it should have infiltrated or evaporated, can return into the channel, resulting in higher discharge too. Comparing our results to other studies, we find that both coupled runs have remarkably lower RMSE than those reported in Alfieri et al. (2013) for GloFAS. The obtained coefficients of determination come close to those by Yamazaki et al. (2011) and Yamazaki et al. (2012b), who connected runoff from a land surface model with a riverfloodplain routing scheme.

2.3.2 Water levels

Assessing modelled inundation water levels, we find that, because discharge results are almost identical, simulated water levels for the GRDC-fed runs differ only slightly between 2D and 1D/2D schematization, with the latter generally showing lower water levels (Figure 2.6). This is the result of the 1D channels providing better hydraulic connectivity throughout the study area since also smaller side channels below the spatial resolution of the 2D mesh are accounted for (Figure 2.7). Results furthermore show that for some observation locations, the GRDC runs yield higher water level values than one-way coupled runs and vice versa at other locations.

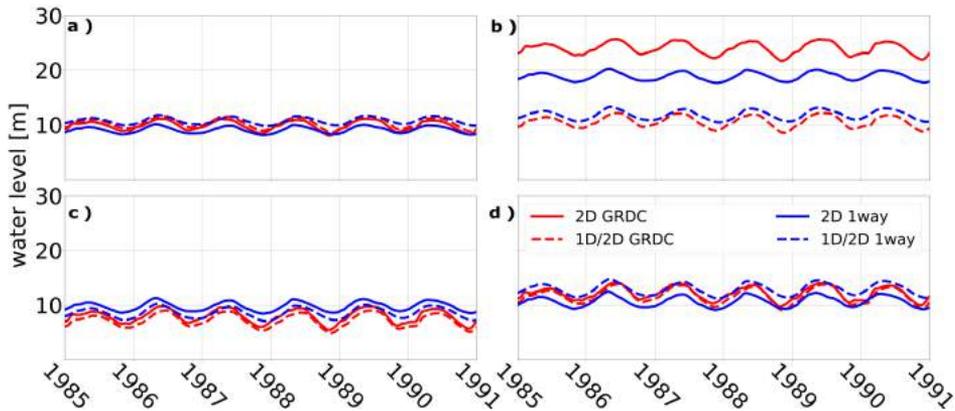


Figure 2.6: Plot of simulated water levels at four different observation locations throughout the study domain. Plots a) to d) correspond with the location numbers 1-4 as depicted in figure 2.1.

As the model schematizations are exactly the same, these local differences can be related to the difference in volume input into the DFM model domain (dividing over DFM cells with PCR output versus upstream boundaries with GRDC data), as well as local influence of precipitation events within the intermediate catchment area on water level dynamics. The discrepancy between simulated water level for 1D/2D and 2D set-ups at Loc2 exemplifies the impact vertical errors in input elevation data can have on 2D schematizations. While the area where the location was placed could be conveyed by the 1D network, this was not possible in the 2D set-up, thus resulting in local accumulation of water in a local depression. Results also indicate that locations closer to the delta (see Loc4 as an example) are less influenced by river dynamics or precipitation events, but more by the downstream water level boundary, for which smaller differences in simulated water level between model runs are revealed. From a holistic point of view, large-scale water level dynamics are correctly represented with only minor differences between model set-ups, despite the results at Loc2 as mentioned above.

2.3.3 Inundation extent

In terms of inundation extent, we performed a first-order and qualitative validation of simulated against observed water extent for all runs except the DynRout set-up. Our results indicate that the 1D/2D schematization with GRDC forcing performs particularly well (see Figure 2.7). This demonstrates that the advantage of implementing 1D channels as inundation extent is modelled more accurately, especially for smaller side branches of the stream where the 2D resolution does not allow for detailed simulation of channel-floodplain interaction. This finding is in line with the observations made by Neal et al. (2012a), who employed

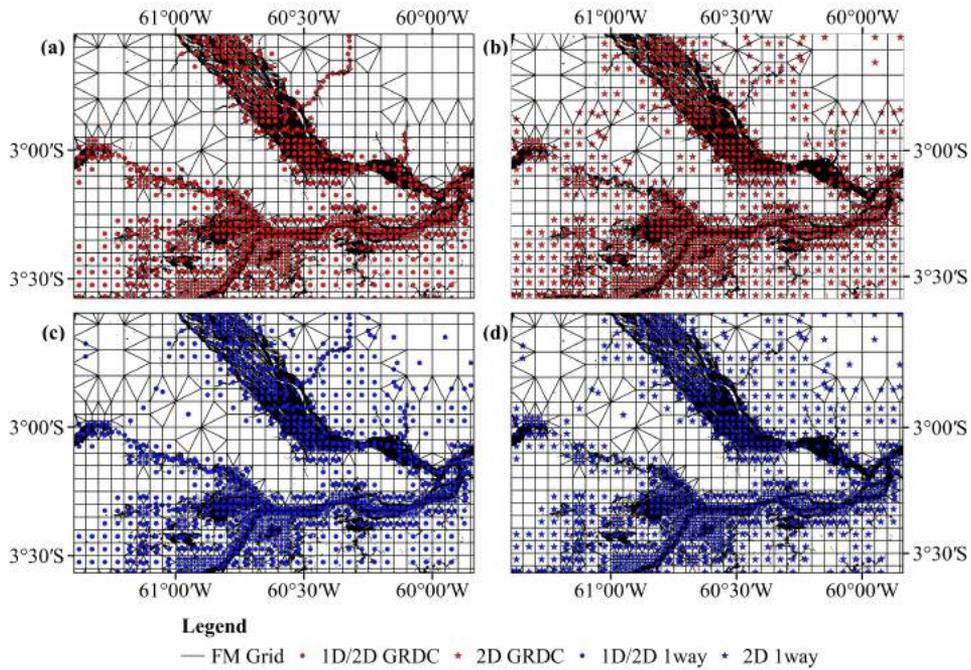


Figure 2.7: Plot of simulated inundation extent per model set-up compared to observed water body extent as observed by Landsat imagery on 1 July 1989; the validation is performed for (a) the 1D/2D GRDC run, (b) the 2D GRDC run, (c) the 1D/2D one-way (1way) run, and (d) the 2D one-way (1way) run.

a sub-gridding scheme. For the coupled set-ups, water extent is well modelled for the main reaches of the river, but overpredicted for floodplain areas. We attribute these deviations to both the quality of remotely sensed input elevation and the coarse spatial resolution of the flexible mesh which may overly facilitate flow over floodplains. Besides, distributing water volumes over the FM cells in the coupling process may also have led to stronger inundation on floodplain areas than point inflow from GRDC stations. Assessing simulated water extent over the entire study area, we again find that the use of 1D channels can highly improve the level of detail for river streams and bends for both the main branch as well as more remote areas, as shown in Figure 2.8. Similar to the local water level validation, we found that the areas where inundation is modelled differ strongly compared to the GRDC runs. While inundation for those runs is limited to streams that are connected to upstream discharge boundaries, spatially coupling hydrology with hydrodynamics additionally yields inundation information for smaller reaches throughout the entire model domain which otherwise would not be fed with water. In particular, for the 1D/2D run, this results in an overall good representation of inundation along rivers throughout the entire model domain.

This constitutes a major improvement, and is a strong hint that model coupling can indeed contribute to better inundation extent estimates. Notwithstanding this achievement, we again see that water can accumulate locally, which can partially be related to the presence of temporarily filled depressions during rainfall, and partially to the spatial resolution of the hydrodynamic model in combination with the quality of the elevation data used for model schematization. Also, in the big picture, the local accumulation of water is less severe in

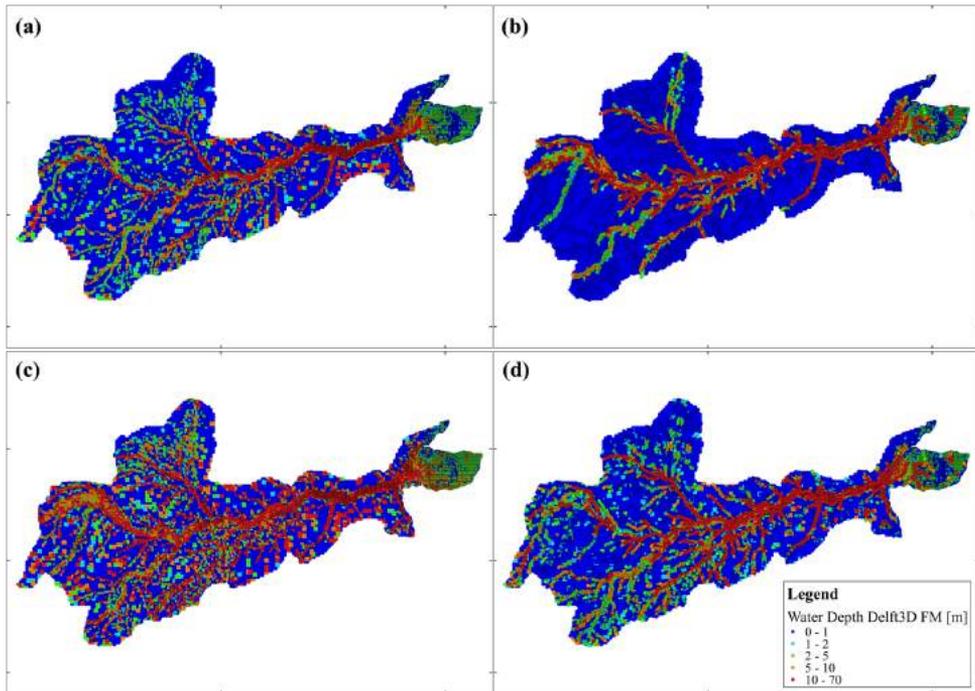


Figure 2.8: Plots of simulated water depth for 31 October 1990 for all runs with Delft 3D FM and the entire study domain: (a) 2D GRDC run; (b) 1D/2D GRDC run; (c) 2D one-way (1way) run; and (d) 1D/2D one-way (1way) run.

the 1D/2D than in the 2D set-up due to a facilitated hydrologic connectivity within the river basin.

2.4 Conclusion and recommendations

In the present study, we spatially coupled the global hydrologic model PCR-GLOBW with the hydrodynamic model Delft3D Flexible Mesh, and compared resulting discharge and inundation extent with estimates obtained from stand-alone runs as well as actual observations to investigate possible strengths, weaknesses, and opportunities of model coupling for large-scale inundation modelling.

Our results showed that hydrology-only runs conducted with PCR-DynRout have the least accurate discharge simulation of all runs. Particularly discharge variability could not be captured by a global hydrology model due to its coarse spatial resolution and its kinematic wave approximation of surface water flow in an area with limited topographic gradients. The question therefore remains: what is more important, the coarse resolution or the simple hydrodynamics? Therefore, once PCR at 5 arcmin spatial resolution is fully tested and available, the model runs should be repeated to better understand whether results can be improved by finer spatial resolution or are constrained by the employment of a kinematic wave approach. Besides, fine-tuning of sensitive parameters of PCR at a global scale seems to be required to obtain a better-timed peak flow, not only for those optimized so far but also others such as Manning's surface roughness coefficient. Comparison revealed that runs forced with ob-

served discharge from GRDC, once the underrepresentation of water volume in the systems was accounted for, outperform hydrology-based models in resembling discharge dynamics.

While validation of GRDC-forced runs against observed discharge showed good performance, the disadvantage of such set-ups is the limitation of discharge to river reaches fed by the discharge boundaries. As a result, inundations along reaches that start within in the domain or along reaches not being fed by upstream discharge boundaries cannot be simulated. A first qualitative validation of simulated inundation extent with Landsat imagery showed that, for those rivers connected to upstream discharge boundaries, the 1D/2D schematization with GRDC forcing showed the best performance of all runs.

Representation of 1D channels results in a better conveyance of surface water in the model domain and consequently less flood artefacts, in particular where 1D channel dimension is below the grid size of the 2D grid cells. We also found that GRDC-forced runs show stronger attenuation and lagged peak discharge due to the longer average travel time required to propagate from the boundaries through the model domain. Both 1D/2D and 2D coupled runs were able to capture the peak flow better than GRDC runs, and to follow the discharge dynamics better than the simple kinematic wave model. The fact that they overpredict peak discharge for some years can be attributed to the absence of a feedback loop to hydrological processes on floodplains, such as groundwater infiltration and evaporation. It will be the aim of a follow-up study to implement a fully dynamic coupling scheme, whereby information is exchanged between hydrology and hydrodynamics at each time step, and water on the floodplains is allowed to evaporate or recharge the groundwater store. We expect that this will lead to lower and more accurate discharge estimates. Replacing the simplistic routing scheme of PCR with a full hydrodynamic model remarkably improves the coefficient of determination as well as the models skill.

From our results we conclude that spatially coupling hydrology and hydrodynamics merges the best of two worlds, namely water volume accuracy and routing scheme. From a computational point of view, the use of a 1D/2D set-up is favourable, as it requires less computational time. At the same time, it yields a better spatial resolution of the river network than the 2D set-up because it decreases dependency on quality of space-borne DEM data sets which are known for introducing errors in large-scale inundation models. Especially for the coupled runs, these vertical errors are partly responsible for overestimated inundation extent and local water levels, in particular in floodplain regions. Another part of the overestimation may lie in the way water volumes are distributed over the 2D grid. It needs to be researched in more detail how the distribution of volumes impacts model results, and whether other techniques such as adding water directly into the 1D channels than onto the 2D grid may improve model performance. Besides, a future study should contain an assessment of the impact of varying spatial resolution of both the hydrologic and the hydrodynamic model as well as their interplay to obtain a better picture of the potential of model coupling at larger scales. In this study, we used only global data sets for both the hydrological and the newly developed hydrodynamic model DFM. Thus, the presented set-up can easily be applied in other river basins as well. In the long term, we are confident that the proposed spatially coupled model set-up can eventually contribute to a better assessment of both current and future flood hazard and risk.

Author contributions

AVH prepared the code for model coupling. AvD supported the application of DFM and RvB provided information for PCR and PCR-DynRout runs. RvB, HCW, and MFPB supervised the

research and provided important advice. JMH designed and executed the research, and also prepared the manuscript, with contribution from all co-authors.

Acknowledgements

The authors declare that they have no conflict of interest. The authors want to acknowledge the valuable contributions of Gennadii Donchyts, Herman Kernkamp, and Robert Leander from Deltares as well as Edwin Sutanudjaja from Utrecht University. This study was financed by the EIT Climate-KIC programme under project title “Global high-resolution database of current and future river flood hazard to support planning, adaption and re-insurance”. We want to acknowledge the constructive contributions of one anonymous reviewer and Dai Yamazaki who helped to strongly increase the quality of the manuscript.

Chapter 3

GLOFRIM v1.0 – A globally applicable computational framework for integrated hydrological-hydrodynamic modelling

We here present GLOFRIM, a globally applicable computational framework for integrated hydrologic–hydrodynamic modelling. GLOFRIM facilitates spatially explicit coupling of hydrodynamic and hydrologic models and caters for an ensemble of models to be coupled. It currently encompasses the global hydrologic model PCR-GLOBWB as well as the hydrodynamic models Delft3D Flexible Mesh (DFM; solving the full shallow-water equations and allowing for spatially flexible meshing) and LISFLOOD-FP (LFP; solving the local inertia equations and running on regular grids). The main advantages of the framework are its open and free access, its global applicability, its versatility, and its extensibility with other hydrologic or hydrodynamic models. Before applying GLOFRIM to an actual test case, we benchmarked both DFM and LFP for a synthetic test case. Results show that for sub-critical flow conditions, discharge response to the same input signal is near-identical for both models, which agrees with previous studies. We subsequently applied the framework to the Amazon River basin to not only test the framework thoroughly, but also to perform a first-ever benchmark of flexible and regular grids on a large-scale. Both DFM and LFP produce comparable results in terms of simulated discharge with LFP exhibiting slightly higher accuracy as expressed by a Kling–Gupta efficiency of 0.82 compared to 0.76 for DFM. However, benchmarking inundation extent between DFM and LFP over the entire study area, a critical success index of 0.46 was obtained, indicating that the models disagree as often as they agree. Differences between models in both simulated discharge and inundation extent are to a large extent attributable to the gridding techniques employed. In fact, the results show that both the numerical scheme of the inundation model and the gridding technique can contribute to deviations in simulated inundation extent as we control for model forcing and boundary conditions. This study shows that the presented computational framework is robust and widely applicable. GLOFRIM is designed as open access and easily extendable, and thus we hope that other large-scale hydrologic and hydrodynamic models will be added. Eventually, more locally relevant processes would be captured and more robust model inter-comparison, benchmarking, and ensemble simulations of flood hazard on a large scale would be allowed for.

Based on: Hoch, J. M., Neal, J. C., Baart, F., van Beek, R., Winsemius, H. C., Bates, P. D., Bierkens, M.F. P. (2017), GLOFRIM v1.0 – A globally applicable computational framework for integrated hydrological-hydrodynamic modelling, *Geoscientific Model Development* 10, 3913–3929.

3.1 Introduction

In the latter half of the last century, losses due to riverine floods increased greatly, leading to economic losses of more than USD 1 billion and 220,000 casualties since 1980 (MunichRe, 2019; Visser et al., 2012). Much of this increase is thought to be due to continued settlement along rivers and shifts in climate patterns, meaning that this tendency will most likely be exacerbated in the future (Ceola et al., 2014; Hirabayashi et al., 2013; Winsemius et al., 2016). Robust inundation estimates are therefore paramount to enhance our process understanding and to provide better flood hazard estimates for risk models.

Since recent research showed that flood inundation can easily affect large areas, in particular neighbouring river basins (Jongman et al., 2014), it is vital that flood hazard models can simulate the relevant processes over large domains. Applying such large-scale models has the additional advantage of facilitating the identification of risk hotspots and providing critical insight into data-scarce areas (Ward et al., 2015). In fact, there are already a number of global-scale inundation models available (cf. Dottori et al. (2016), Pappenberger et al. (2012), Sampson et al. (2014), Winsemius et al. (2013), and Yamazaki et al. (2011)), differing in their process descriptions and computational engine. While some approaches derive flood hazard from a coarse-scale hydrologic model and subsequent downscaling, others force fine-scale hydrodynamic models with globally regionalized discharge data. A first inter-comparison of global flood hazard models by Trigg et al. (2016) for the African continent, however, revealed that they agree for only 30-40% of aggregated flood extent, thus indicating that the representativeness of local flood risk estimates may depend strongly on the computational engine opted for as well as on the model forcing applied. Identifying the exact reasons for model disagreement was impossible due to the diversity of methods and lack of a systematic approach to the inter-comparison where individual aspects of the modelling frameworks could be isolated.

Employing a global hydrologic model (GHM) such as PCR-GLOBWB (van Beek et al., 2011; van Beek, 2008; Sutanudjaja et al., 2018), WaterGAP (Alcamo et al., 1997; Döll et al., 2003) or VIC (Liang et al., 1994; Wood et al., 1992) has the benefit of providing spatially distributed surface runoff and routed discharge simulations, thereby facilitating direct forcing for spatially distributed inundation models. In addition, these models are usually forced by global meteorological data, hence diminishing the dependency on observed data as well as allowing for easier implementation of future climate scenarios. However, the routing schemes currently implemented in large-scale hydrologic models can generally be described as simplistic as they are based on gridded drainage networks at a coarse spatial resolution, with the currently finest spatial resolution of global hydrologic models being 5 arcmin or around 10 km x 10 km at the Equator (Bierkens, 2015). Furthermore, discharge accuracy may be reduced in low-gradient catchments since topography at this scale is generally parameterized in distribution functions and river routing is often represented by a simple scheme, such as the kinematic wave approximation.

Hydrodynamic models, on the other hand, can be built in numerous ways for inundation modelling, typically in 1D, 2D or combined 1D/2D, and are mostly forced with gauged discharge data or synthesized flood waves. While such approaches do not require rainfall-runoff conversion, they are problematic for studies concerning large-scale climate change impacts or the seamless simulation of flood events and their spatial correlation (Jongman et al., 2014). Some models like CaMa-Flood (Yamazaki et al., 2011) route *a priori* computed hydrology-based surface runoff with 1D hydrodynamics and parameterized 2D floodplain storage. Applying such a 1D/2D approach, however, does not allow for explicit modelling of floodplain flow pathways as well as channel-floodplain interactions. Explicitly representing these pro-

cesses would be beneficial as they are known to greatly influence inundation dynamics and patterns (Neal et al., 2012a; Trigg et al., 2009). Compared to hydrologic models, hydrodynamic models solving the full shallow water equation (SWE) or at least a more advanced approximation such as the local inertia equations (LIEs) have the advantage of providing a better representation of backwater effects, which are important flood-triggering processes (chapter 1.4.2; Meade et al. (1991), Moussa and Bocquillon (1996), and Paiva et al. (2013)). Another difference to GHMs is that current applications of hydrodynamic models on the large to global scale can run at spatial resolutions of up to 1 km (Sampson et al., 2014), greatly facilitating the representation of both relevant channel-floodplain interactions (Rudorff et al., 2014a; Rudorff et al., 2014b) and flow pathways on floodplains (Rudorff et al., 2014a; Tayefi et al., 2007) as well as enhancing the usability for decision-making processes (Beven et al., 2015; Trigg et al., 2016). Notwithstanding these advantages, most hydrodynamic models applied for large-scale inundation modelling lack an advanced implementation of hydrologic processes and thus may overpredict both inundation extent and depth as, for instance, groundwater infiltration and evaporation from inundated floodplains are currently not (yet) fully accounted for.

Large-scale flood hazard estimates may thus benefit from increased integration of hydrology and hydrodynamics in inundation models to allow for physically more integrated assessments and to compensate for their respective shortcomings. In fact, hydrologic-hydrodynamic coupling was already applied in a number of studies (cf. Biancamaria et al. (2009), Lian et al. (2007), Schumann et al. (2013), and Kim et al. (2012)). For example, output from hydrologic or land-surface models was used as input to the 1D/2D hydrodynamic model LISFLOOD-FP (Bates and De Roo, 2000; Bates et al., 2010) at a number of locations. While such approaches reduce the dependency on gauged data or synthesized flood waves, they cannot fully account for important and spatially distributed hydrologic flood-triggering processes within the model domain. This would, however, be advantageous to support the assessment of spatial correlations of flood waves in adjacent river basins, which are shown to increase transnational flood risk (Jongman et al., 2014).

A further valuable contribution for promoting the coupling of models from different disciplines was realized by the Community Surface Dynamics Modelling Systems group (CSDMS) with their development of the Web Modelling Tool (WMT; CSDMS (2019b)). This tool enables the user to create a coupled model from a list of readily available models and run it on a server of the CSDMS. Whilst this is an important step towards integrated modelling between disciplines, applicability is hampered by the fact that model code is not openly accessible and that the number of available models is limited and predefined.

Hoch et al. (2017a) coupled PCR-GLOBWB (hereafter PCR) with the hydrodynamic model Delft3D Flexible Mesh (hereafter DFM; Kernkamp et al. (2011)) for the Amazon River basin to integrate the hydrologic and hydrodynamic processes occurring over the entire study area. Results indicate that spatially explicit coupling of hydrologic and hydrodynamic models can improve the representation of inundation for all river reaches, not only those that are connected to upstream boundary conditions. Findings also corroborate that spatially distributed forcing retrieved from a hydrologic model in combination with a sophisticated river routing scheme outperforms results obtained with both models run in stand-alone mode.

Even though these results are promising, it has to be acknowledged that the accuracy of a hydrologic and hydrodynamic model can vary strongly, depending on the chosen study area, model parameterization, model structure, numerical scheme or the use of different input data (Li et al., 2015; Trigg et al., 2016). It would hence be advantageous to base the choice of the

coupled models on their local performance, potentially outperforming predefined set-ups, or simply on the model schematization at hand.

To facilitate such model selection and to further promote the coupling of large-scale hydrologic and hydrodynamic models, we developed GLOFRIM, a GLOBally applicable computational FReamework for Integrated hydrologic–hydrodynamic Modelling. In addition to the work of Hoch et al. (2017a), it includes the widely used hydrodynamic model LISFLOOD-FP (hereafter LFP; Bates et al. (2010)) and an improved as well as extended coupling algorithm, thus catering for a wider range of model schematizations and applications. As we believe that relevant processes can be captured better by combining the locally best-performing hydrologic and hydrodynamic models, GLOFRIM is designed in an expandable way to eventually incorporate more models. Furthermore, the framework is openly available under the GNU GPL v3.0 license to stimulate collaboration and idea exchange within the scientific community. Key assets of the framework are its free and open accessibility, its global applicability, its versatility, and its potential to be further developed to a full two-dimensional coupling scheme between hydrology and hydrodynamics, which would play a particularly crucial role in basins in semi-arid climates such as the Niger (Dadson et al., 2010; Mahe et al., 2009).

In the remainder of this paper, we first describe the model components of the framework and thereafter the framework and its functionalities in detail. Subsequently, we compare the two hydrodynamic models in a simple synthetic test case to obtain a first understanding of possible differences, in particular in terms of their numerical schemes. As a means of benchmarking, we assess simulated discharge along the flow paths as well as run times for a 1D and 2D set-up individually. We then apply GLOFRIM to one-directionally couple PCR with both DFM and LFP and benchmark the set-ups for an actual test case in the Amazon River basin, hence also constituting a first comparison of flexible and regular grids for large-scale applications. For model benchmarking, we assess simulated discharge, water levels, run times, and inundation extent. Pearsons correlation r , the root mean square error (RMSE), and the Kling–Gupta efficiency (KGE; Gupta et al. (2009)) are determined by comparison to observed discharge data from the Global Runoff Data Centre (GRDC) at Óbidos, Brazil. We opt for GRDC data as the presented approach is merely based on input data sets with global coverage. Simulated water levels are compared at an upstream, midstream, and downstream station to assess (a) whether water level dynamics are correctly represented and (b) to what extent DFM and LFP differ or agree in their water level computations. Computational efficiency is assessed by comparing the run times of the coupled set-ups. To benchmark inundation extent from DFM with LFP, we determine the hit rate H , false alarm ratio F , and the critical success index C based on inundation maps of both models at the end of the simulation. No validation of simulated inundation extent was performed as Hoch et al. (2017a) already showed good agreement of results obtained with DFM for the same study domain.

This openly available computational framework makes a valuable contribution to current inundation modelling on the large scale by enhancing the integration of hydrologic and hydrodynamic model processes, which eventually may lead to improved decision-making and planning of adaption and mitigation measures.

3.2 Models

Currently, GLOFRIM includes the hydrologic model PCR as well as the hydrodynamic models DFM and LFP. Hereafter, an overview of the main features of the models is provided. For further details regarding model development and model set-up, we refer to the specific manuals or websites.

3.2.1 PCR-GLOBWB

To generate hydrologic input, the PCR is currently incorporated in the framework. It can be applied at 30 arcmin resolution (approximately 55 km x 55 km at the Equator) and at 5 arcmin resolution (approximately 10 km x 10 km at the Equator), which may increase accuracy but also runtime. PCR is entirely coded in PCRaster Python (Karssenberget al., 2010) and distinguishes between two vertically stacked soil layers, an underlying groundwater layer, and a surface canopy layer. Water can be exchanged vertically, and excess surface water can be routed horizontally along a local drainage direction network employing the kinematic wave approximation.

The model is forced with Climate Research Unit (CRU) precipitation and temperature data (Harris et al., 2014) at a 30 arcmin spatial resolution, and evaporation is computed using the Penman–Monteith equation. Data sets are downscaled to daily fields for the period from 1957 to 2010 using ERA40/ERA-Interim (Källberg et al., 2005; Uppala et al., 2005). Besides, PCR is able to account for domestic and industrial water consumption by accounting for water demand data (FAO STAT 2017). For more detailed information on CRU forcing, its processing, and PCR in general, we refer to van Beek (2008), van Beek and Bierkens (2008), van Beek et al. (2011), and Sutanudjaja et al. (2018). PCR was already applied for a wide range of studies such as flood and drought forecasting (Candogan Yossef et al., 2012), human impact on droughts (Wanders and Wada, 2015), global water stress (van Beek et al., 2011), and global groundwater simulations (de Graaf et al., 2015).

More relevant to this study, PCR constitutes the computational backbone of the Global Flood Risk with IMAGE Scenarios framework (GLOFRIS; Winsemius et al. (2013)), which is also used as the basis for the Aqueduct Global Flood Analyzer of the World Resources Institute (Aqueduct Global Flood Analyzer 2019).

3.2.2 Delft3D Flexible Mesh

DFM allows the user to schematize the model domain with a flexible mesh in 1D/2D/3D, and therefore supports the computationally efficient schematization of topographically challenging areas such as river bends or irregular slopes. The model solves the full Saint-Venant equations, or shallow water equations (SWEs). The main partial differential equations solved by DFM are

$$\frac{\delta h}{\delta t} + \nabla(h * \mathbf{u}) = 0 \quad (3.1)$$

$$\frac{\delta \mathbf{u}}{\delta t} + \frac{1}{h}(\nabla(h\mathbf{u}\mathbf{u}) - \mathbf{u}\nabla(h\mathbf{u})) = -g\nabla\zeta + \frac{1}{h}\nabla(vh(\nabla\mathbf{u} + \nabla\mathbf{u}^T)) + \frac{1}{h}\frac{\tau}{\rho} \quad (3.2)$$

with

$$\nabla = \left(\frac{\delta}{\delta x}, \frac{\delta}{\delta y}\right)^T \quad (3.3)$$

ζ being the water level, h the water depth, \mathbf{u} is the velocity vector, g the gravitational acceleration, v the viscosity, ρ the water mass density, and τ the bottom friction.

For 1D flow, the equations remain the same except that the viscosity v does not contain horizontal eddy viscosity. For further technical details and derivation, we refer to the technical manual (Delft3D Flexible Mesh Technical Manual 2019). DFM is an openly accessible model and can be obtained by contacting Deltares (<https://www.deltares.nl/en/software/delft3d-flexible-mesh-suite/>). Besides riverine flood hazard modelling, it also caters for a

wider range of applications, for instance groundwater flow, sediment transport, and water quality simulations in 1D, 2D, and 3D. For more information regarding the application of DFM, we refer to the user manual (*Delft3D Flexible Mesh User Manual* 2019).

Due to its quite recent publication and still on-going development, only a limited number of published studies using DFM are available. It was, for instance, applied in a global-scale reanalysis for extreme sea levels (Muis et al., 2016). In another study, Castro Gama et al. (2013) applied DFM to model flood hazard at the Yellow River, and concluded that applying a flexible mesh reduces computation time by a factor of 10 compared to square grids with equal quality of model output.

3.2.3 LISFLOOD-FP

LFP is a widely used, raster-based model to compute floodplain inundation. Since its first version (Bates and De Roo, 2000), it has regularly been adapted and improved (Bates et al., 2010), for instance by adding a sub-gridding scheme to account for channel flow within cells (Neal et al., 2012a). It is possible to run LFP with different set-ups: a 2D only, a 1D, a 1D/2D or a sub-grid model, with the latter being the most accurate for large-scale inundation modelling approaches as it greatly increases floodplain connectivity (Neal et al., 2012a). When using the sub-grid scheme, LFP solves the subsequent equation for channel flow that is based on a simplification of the SWE ignoring advection (Bates et al., 2010; Neal et al., 2012a).

For channel flow, the one-dimensional local inertia equations are solved (see Eqs. 1.1 and 1.2 for their derivation from the shallow water equations).

Once bankfull depth is exceeded, water flow on the overlying floodplain is similarly described in terms of continuity and momentum equations. Changes of water height of a single cell within a grid of square cells is then calculated as:

$$\frac{\delta h^{i,j}}{\delta t} = \frac{Q_x^{i-1,j} - Q_x^{i,j} + Q_y^{i,j-1} - Q_y^{i,j}}{\Delta x \Delta y} \quad (3.4)$$

Volumetric flow rates between cells are simulated as:

$$Q_x^{i,j} = \frac{h_{flow}^{5/3}}{n} \left(\frac{h^{i-1,j} - h^{i,j}}{\Delta x} \right)^{0.5} \Delta y \quad (3.5)$$

Here, $h^{i,j}$ is the water height at index (i, j) , Δx and Δy the cell dimensions, n Manning's surface roughness coefficient, Q_x and Q_y the flow rates between cells with Q_y defined analogously with Q_x , and h_{flow} the possible flow depth between two cells.

For further information about model development, derivation of numerical solutions, assumptions, and validations, we refer you to the above-mentioned papers and the LFP webpage (<http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/>). LFP is specifically developed to model floodplain inundation and has been used in a wide range of studies. Most notable in the context of large-scale flood hazard modelling is the work by Sampson et al. (2014), who applied LFP to compute global estimates of flood hazard and risk and by Schumann et al. (2013) and Biancamaria et al. (2009), who used LFP to simulate inundation in the Zambezi and Ob rivers, respectively, forced with lateral input from a land surface model.

The Basic Model Interface (BMI) adapter (see subsequent section) was implemented for LFP version 5.9, which provides all relevant features, in particular the sub-gridding scheme, to model large-scale inundation.

3.2.4 Basic Model Interface (BMI)

Generally, the BMI has several functions that can be called from external applications like, as in this case, a Python script. To make these functions available for a model, a BMI adapter needs to be developed for each model with respect to the specific internal model structure and programming language. Whilst PCR is already written in Python and its BMI implementation is hence straightforward, DFM offers a native C-compliant BMI implementation. For LFP, which is written in C++, the code and file structure had to be slightly adapted to agree with the requirements for the BMI.

Once a BMI adapter is developed, it is possible to execute a set of functions: first, the user can initialize the models by using the BMI adapter. Second, the BMI adapter allows for retrieving a set of variables from memory. The variables exposed through the BMI adapter can be defined during the development of the BMI adapter and is thus not limited to a preset range. Third, the manipulated variables can be set back to the original model or can be used to overwrite variables in one or multiple other models, given that they agree to the internal data structure of those models. Fourth, models connected to a BMI adapter can be updated at a user-specified time step, hence enabling online coupling of models. In this way it is possible to get, change, and set variables during the execution of the models in use on a time step basis. Last, models can be finalized to end the computations.

Both DFM and LFP, although not being coded in Python, can be called from within Python using a Python wrapper (see <https://github.com/openearth/bmi-python>). For further information regarding the BMI, we refer to Peckham et al. (2013), the related website CSDMS (2019a), and section 1.3.2.

3.3 The computational framework GLOFRIM

The computational framework presented here consists of two key elements, (a) the actual code and (b) a settings-file. Here, a brief overview is given of their main properties. More detailed information and an outline is provided in the files themselves.

The computational backbone of GLOFRIM is entirely written in Python 2.7 and was developed and tested on Ubuntu systems. By means of a python file (*couplingFramework_v1.py* in the downloadable data), the steps for model coupling are executed (see Figure 3.1 for a flow chart). The models are first initialized: the model configuration files of each model are read and the internal steps required to obtain an initial state of the models are prompted by the BMI adapter.

Thereafter, the BMI adapter is used to retrieve all required model variables, especially geometry information. This information is subsequently used to construct the grids of the models and to spatially couple them by overlay and grid-to-grid assignment. A many-to-one assignment based on raster indices is performed and the routing computations in PCR are turned off for all cells signalled as coupled. In case no 1D or 2D hydrodynamic cells are located within a PCR cell, this cell is therefore not considered to be coupled and the routing scheme as implemented in PCR prevails. Further information about the spatial coupling can be found in Hoch et al. (2017a).

Once the models are spatially coupled, the update loop commences. During execution of this loop, PCR will be updated at each time step – typically one day – and surface runoff and discharge output will be retrieved and adapted to agree with the data structure of the chosen hydrodynamic model. Then, either the water depth or a flux variable in the hydrodynamic model will be overwritten, and finally the hydrodynamic model will be updated until

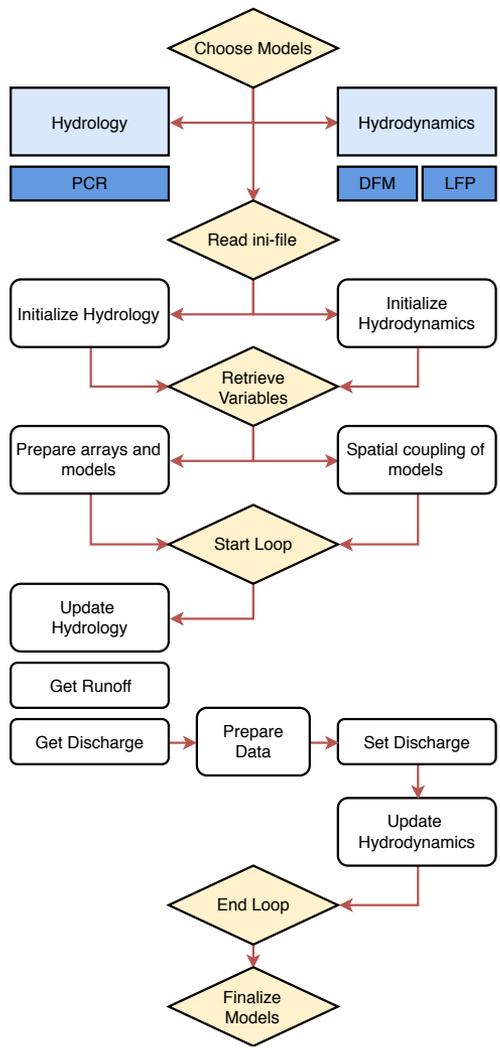


Figure 3.1: Flow diagram of steps executed in GLOFRIM as well as the models currently available within the framework.

it reaches the same simulation time as PCR. The loop is exited once a user-specified number of time steps is reached.

It should be noted that in the current version of the framework, only one-directional coupling from hydrology to hydrodynamics is supported, possibly leading to local overprediction of simulated discharge as there is, for instance, no re-infiltration of water going overbank. Future research will thus focus on extending this to a full two-directional coupling scheme with feedback loops from hydrodynamics to hydrology. Such two-way coupling would, for instance, contain explicit modelling of hydrologic processes over inundated areas in the hydrodynamic model.

To specify all relevant information about the coupling run to be performed, a configuration file is needed (*default.ini* in the downloadable data). Besides all critical paths to model data, other model settings can be defined in the configuration file, for example the number of model time steps. In general, settings defined in the ini-file overrule those specified for the individual models. In the current version of GLOFRIM, three options need to be specified to realize model coupling: by activating the so-called River-Floodplain-Scheme (RFS), by specifying the variables to be updated, and by choosing for hydrodynamic models in either spherical or projected coordinate systems.

First, the RFS defines where output from PCR is coupled to. If RFS is activated, water volume of one PCR cell is directly coupled to the 1D channels of the hydrodynamic model within the corresponding PCR cell. If RFS is inactive, water is distributed over all 2D grid cells within the corresponding PCR cell. Applying the RFS has two major advantages: first, it reduces run times as data exchange and computations need to be performed for a smaller number of cells; second, using RFS in large-scale applications with sufficient channel information reduces the dependency on the accuracy of remotely sensed 2D elevation data such as Shuttle Radar Topographic Mission (SRTM) data (Farr and Kobrick, 2000). Recent research showed that such global data sets contain strong vertical bias as well as systematic and random noise (Yamazaki et al., 2017). In particular, simulating flow over vertically irregular terrain resulting in supercritical regimes is contraindicated for LFP because of its use of the LIE. In case overland flow needs to be modelled by LFP, we advise to take measures accordingly, for instance by limiting flow velocities. For DFM we found that runs are more stable, yet slower, when deactivating the RFS.

Second, it is possible to force the hydrodynamic models by updating the water depth variable (in metres) or by updating fluxes, which are expressed in LFP as discharge (in $\text{m}^3 \text{d}^{-1}$) and in DFM as precipitation (in mm d^{-1}). For DFM, added daily water depth is divided into a number of user-specified time steps hence reducing the computational load, while fluxes are daily constants. We found that updating fluxes reduces run times compared to states, and hence advise opting for this option. While it is also possible to perform state-updating in LFP, test runs showed that this option should be used carefully as it easily increases run times. This is because it is currently not possible to update LFP at a user-specified time step due to the Courant–Friedrichs–Lewy condition. It may hence happen that gradients between added daily water depths are too steep, increasing the risk of model instability. We therefore recommend applying flux-updating in LFP instead.

Third, it is possible to use the hydrodynamic models with Cartesian coordinates, although PCR runs in non-Cartesian coordinates. By providing the projected coordinate system the model is based on, the computational framework can translate the grid into spherical coordinates and perform the grid overlay and cell assignment, thus guaranteeing the applicability of all already existing hydrodynamic schematizations. All other computations remain unaffected

by the coordinate system in use as the coordinate information is solely required for spatially coupling the grids.

As expressed before, GLOFRIM employs the BMIs functionalities to couple hydrologic to hydrodynamic processes. Even though the current version of GLOFRIM only supports one-directional coupling, basing it upon the BMI yields strong advantages for future two-directional coupling as coupled models do not get unnecessarily entangled, so-called “integronsters” (Voinov and Shugart, 2013). Such two-directional coupling is currently not yet available for GLOFRIM due to on-going testing as well as concept development and will be provided in a future version of the framework.

Besides being openly accessible and thus adaptable and extendable to the users preferences or individual modelling requirements, GLOFRIM contains a number of additional advantages: first, by having PCR, or any other global hydrologic model, as the hydrologic output creator, the framework can easily be applied anywhere on the globe given a hydrodynamic schematization; second, models to be coupled may be selected depending on their local performance, thus possibly capturing more relevant processes; third, the spatially explicit coupling scheme can be extended to a full feedback loop between hydrology and hydrodynamic steps, also incorporating important groundwater infiltration and evaporation processes; fourth, by guaranteeing identical hydrologic forcing, applying the computational framework facilitates benchmarking of hydrodynamic models by eliminating sources of difference, potentially supporting hydrodynamic ensemble modelling approaches.

3.4 The synthetic test case

3.4.1 Set-up

To gain insight into possible differences in model behaviour between LFP and DFM, we created two synthetic test cases, one being set-up as 1D only (STC 1D) and the other as 2D only (STC 2D). For the latter, both models were schematized such that they cover a domain of 11 cells by 500 cells, with the cell resolution being 1 km. For the 1D only design, the channel had a length of 500 cells with a 1 km resolution, a uniform channel width of 500 m, and a uniform channel depth of 3 m.² As default settings, we applied Mannings surface roughness coefficients of $0.04 \text{ m}^{-1/3} \text{ s}$ for the 1D only run and $0.07 \text{ m}^{-1/3} \text{ s}$ for the 2D only run.

Both synthetic test cases were forced with an artificial upstream discharge boundary spanning 1 year and consisting of 2 peak flow moments to introduce variability in model dynamics, thus not employing GLOFRIM for those test cases. As downstream boundary condition, a constant water level of 0 m was set. The entire simulation period was 3 years to ensure it exceeds the time of concentration. To assess model output, 7 cross sections were defined, hence capturing the downstream propagation of the artificial flood waves and facilitating the assessment of possible attenuation and dampening effects. For benchmarking the models, we then compared discharge along the cross sections and run times to obtain a first indication of how the different computational schemes might vary (Figure 3.2).

3.4.2 Results and discussion

Assessing the results for both 2D and 1D, we find that both models simulate the same responses to the input signal applied (Figure 3.3). Due to the higher friction coefficient and the wider flow area, it takes the 2D schematization almost the entire simulation period to entirely

²The dimensions were arbitrarily chosen. The only criterium to be fulfilled was that they should resemble flow over a larger area and sub-critical flow

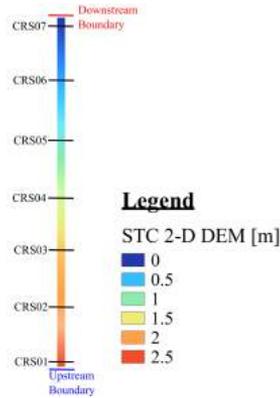


Figure 3.2: Digital elevation model (DEM) of the 2D synthetic test case (STC 2D) for LFP and DFM. Also shown are the locations of the cross sections along the flow path enumerated from 01 (downstream) to 07 (upstream).

| | 2D | 1D |
|-----|----------|---------|
| DFM | 19.5 min | 5.5 min |
| LFP | 2.1 min | 2.6 min |

Table 3.1: Run times of different set-ups in the synthetic test case.

convey the water volumes to the downstream boundary. In the 1D schematization, however, all water is already drained after around 30% of the entire simulation period.

The similarity of simulated discharge between LFP and DFM is, despite the models differences in complexity and design, in line with the findings made by Neal et al. (2012a) and de Almeida and Bates (2013). In the latter study, differences in governing equations were assessed analytically for various flow regimes ranging from sub- to supercritical flow. It was concluded that for applications with low Froude numbers ($Fr \approx 0.5$), such as the synthetic test case used here, no significant differences occur between models solving the LIEs and those solving the full dynamics of the SWEs.

Also, Neal et al. (2012b) showed that it seems unnecessary to employ models solving the SWEs for flow gradually varying in time and for subcritical flow regimes. In addition, the study showed that for those applications, run times of local inertia models are shorter than those of models solving the full SWEs. The run times measured for the various synthetic test cases used here underpin this finding as LFP exhibits shorter run times, especially for the 2D schematization (Table 3.1).

To facilitate comparability, we a priori set the maximum solver time step in DFM to the average of the time steps required by LFP. It is noteworthy that the differences in run times may not merely be attributable to varying solver complexity but partially also to the programming language and compiler used as well as to general model complexity and level of code optimization applied.

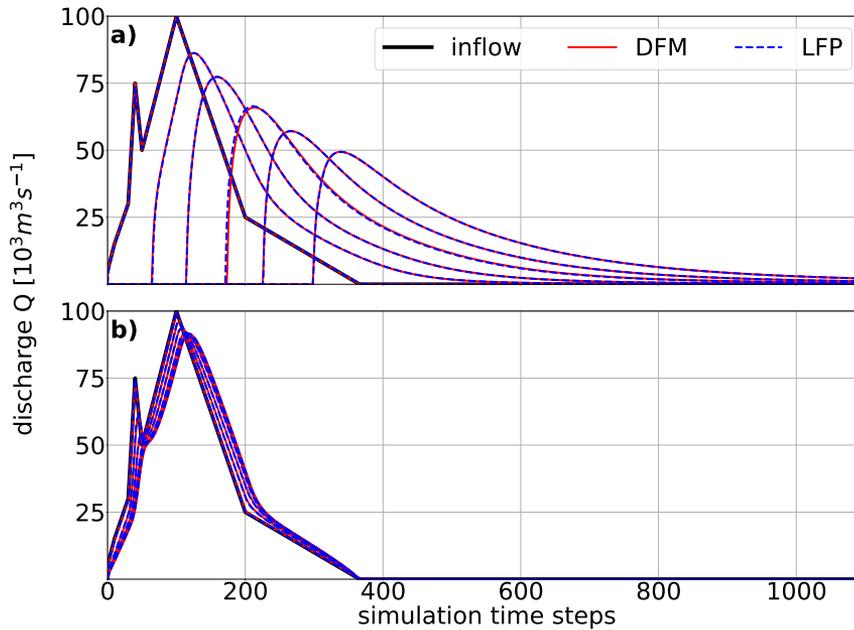


Figure 3.3: Inflow discharge imposed at the upstream boundary as well as simulated discharge of (a) 2D and (b) 1D synthetic test case for the seven cross sections depicted in Figure 3.2.

3.5 Test case: the Amazon River basin

3.5.1 Set-up

To test GLOFRIM in an actual test case and to benchmark the flexible and regular grids, the framework was applied in the Amazon River basin with DFM and LFP being schematized as a flexible mesh and regular grid, respectively. The methods applied to derive the hydrodynamic schematization of the Amazon River basin for DFM are explained in detail in Hoch et al. (2017a).

First, a regular 2D grid at 10 km x 10 km resolution refined until a grid size of 2 km x 2 km was locally obtained, based on the Height Above Nearest Drainage algorithm (HAND; Rennó et al. (2008)). Thereby areas with low HAND values were more strongly refined than those with higher values, resulting in a finer mesh along and next to river channels. This implies a major difference to the synthetic test case above, as we now employ a flexible mesh instead of a regular grid for DFM. As input elevation, canopy-free elevation data at 15 arcsec spatial resolution was applied (Baugh et al., 2013; O’Loughlin et al., 2016) and subsequently smoothed to eliminate local depressions and other residues due to vertical errors of SRTM data (Yamazaki et al., 2012a). Elevation data were then assigned to the flexible mesh by spatial averaging.

For the 1D channel network and bathymetry, river width data of the Global Width Database for Large Rivers (GWD-LR; Yamazaki et al. (2014a)) was employed which was combined with the equations from Paiva et al. (2011) to derive bathymetry information.

For further information, we refer to the relevant papers.

To obtain a LFP schematization equivalent to the DFM schematization, elevation data as well as both river width and river depth information were processed to agree with the requirements of LFP. For river channel properties, the depth and width information stored in the vector data used for DFM were rasterized, and for the elevation data the smoothed canopy-free elevation data were upscaled to a 2 km spatial resolution, employing the nearest neighbour technique, to match the finest spatial resolution of the DFM schematization (Figure 3.4).

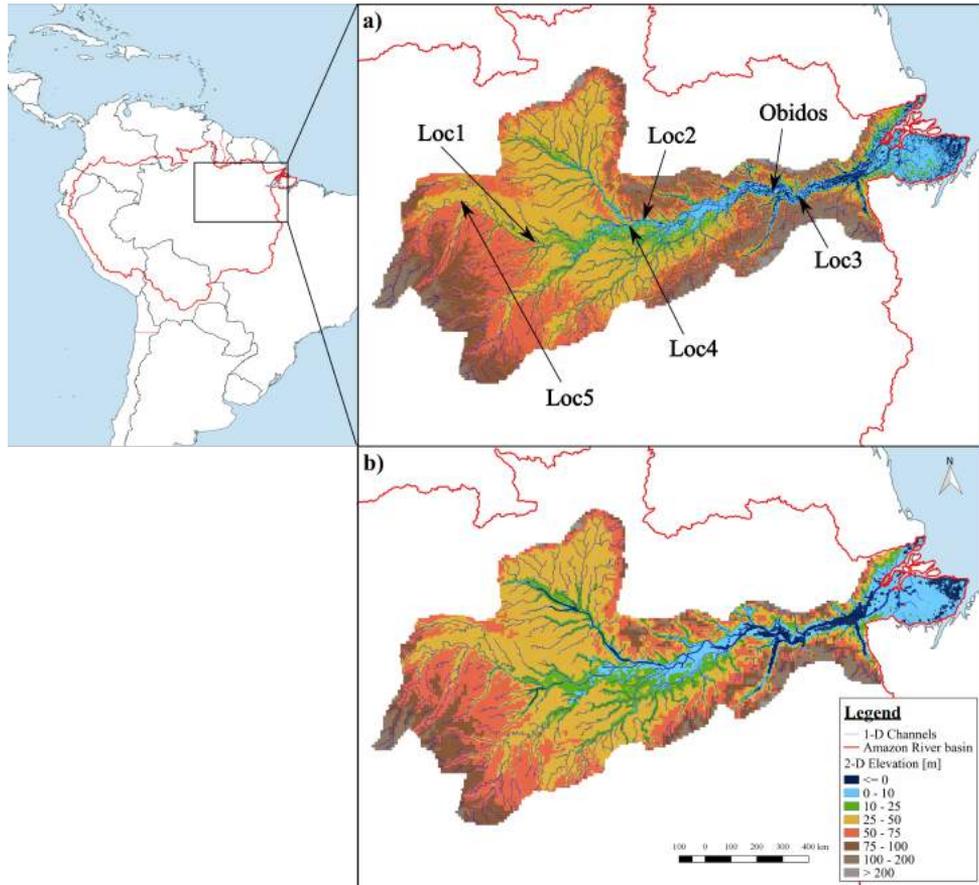


Figure 3.4: Digital elevation model (DEM) and 1D channel network as used in a) LFP and b) DFM; discharge was benchmarked and validated at Óbidos while water levels were compared at three locations throughout the domain; note the gradually increasing difference in spatial resolution for more upstream areas between the two DEMs.

From Figure 3.4 it is visible that LFP contains a greater level of detail in areas farther upstream due to the finer spatial resolution uniformly applied. Consequently, the total number of cells in LFP exceeds the number of 2D cells in DFM by a factor of 4 (Table 3.2). Furthermore, only around 10% of the entire schematization represents 1D channels in LFP, while the channel network of DFM was based on around 30% of all DFM cells. For both DFM and

| | 2D cells | 1D cells | smallest cell | largest cell |
|-----|----------|----------|-----------------------|-------------------------|
| DFM | 41,207 | 12,185 | 2 × 2 km ² | 10 × 10 km ² |
| LFP | 174,982 | 17,119 | 2 × 2 km ² | 2 × 2 km ² |

Table 3.2: Properties of the two hydrodynamic models with DMF employing a flexible mesh and LFP a regular grid.

LFP, Manning’s surface roughness coefficient was uniformly set to $0.03 \text{ m}^{-1/3} \text{ s}$ for channel and floodplains, which is consistent with other case studies in the Amazon (cf. Paiva et al. (2013), Rudorff et al. (2014a), Rudorff et al. (2014b), Trigg et al. (2009), and Yamazaki et al. (2011)).

As downstream boundary, we imposed a constant water level of 0m at the rivers delta. It is noteworthy that GLOFRIM supports the coupling of any hydrodynamic schematization, not only those bordering at a delta but also midstream applications, for instance, if the internal hydrodynamic model requirements are satisfied. Additionally, it should be mentioned that the 1D channels of both schematizations, even with the GWD-LR accounting for islands and thus providing an effective width, do not capture the impact of both braiding and river bifurcation, which may potentially impact model results, especially at the river mouth (Yamazaki et al., 2014b). This is, however, not due to the inability of the hydrodynamic models to account for them, but merely because the chosen algorithm to derive 1D network properties does not allow for it.

For the hydrologic model PCR, the kinematic wave approach was used for routing outside of the coupled domain. This is required as the hydrodynamic schematizations in this test case do not cover the entire extent of the Amazon River basin, even with the kinematic wave approximation potentially introducing an error to the upstream boundary inflow applied. Since simulated discharge from PCR for the Amazon substantially under-predicts observations, we decided to apply an optional regionalized optimization technique facilitating comparison between simulated and measured discharge values (Hoch et al., 2017a). As such an optimization technique is optional and only advisable for catchment studies, a global application is thereby not constrained. In analogy to the hydrodynamic models, the surface roughness coefficient of PCR was uniformly set to $0.03 \text{ m}^{-1/3} \text{ s}$.

Model output of both set-ups was validated against observed GRDC discharge at Óbidos, the most downstream station of the GRDC network in the Amazon River basin (Figure 3.4). To that end, Pearson’s r , the RMSE, and the KGE (Gupta et al., 2009) were computed. Possible uncertainties in observed discharge (Clarke et al., 2000) were thereby omitted. Besides, simulated discharge was qualitatively compared at two locations further upstream (Loc1 and Loc2).

The model time covers the period from January 1984 until December 1990 with the first year being used for spin-up of the coupled settings. This period had to be chosen due to the limitation of available GRDC data for model validation. As with the synthetic test case, run times were compared. To be able to understand water level dynamics as simulated by both models, we compared them at three locations throughout the basin (Figure 3.4). The locations were chosen such that they represent the downstream (Loc3), midstream (Loc4), and upstream dynamics in the basin (Loc5). Besides, inundation extent was benchmarked by applying three evaluation functions, using the LFP inundation results as the benchmark data set.

First, the hit rate H was computed based on the subsequent equation:

| | r | RMSE | KGE | runtime |
|-----|----------|-----------------------|------------|----------------|
| DFM | 0.92 | 25,289 m ³ | 0.76 | 7h |
| LFP | 0.89 | 22,291 m ³ | 0.82 | 6h |

Table 3.3: Results of Pearson's coefficient r , root mean square error (RMSE), Kling–Gupta efficiency (KGE), and runtime obtained to the two coupled runs.

$$H = \frac{N_{DFM} \cap N_{LFP}}{N_{LFP}} \quad (3.6)$$

N_{LFP} and N_{DFM} indicate thereby the number of inundated cells in LFP and DFM at the same moment in time, respectively. To perform consistent benchmarking, the flexible cells of DFM were resampled to the resolution of LFP. The hit rate can vary between 0, signalling that DFM and LFP have no inundated cells in common and 1, indicating that all cells in LFP are also inundated by DFM.

In addition, we determined the false alarm ratio F to also consider false positive alarms. The false alarm ratio can be obtained with

$$F = \frac{N_{DFM} \setminus N_{LFP}}{N_{DFM} \cap N_{LFP} + N_{DFM} \setminus N_{LFP}} \quad (3.7)$$

In the optimal situation, F would be 0 showing that no cells are incorrectly marked as flooded in DFM, whereas a value of 1 indicates that all cells are classified as false alarms.

Last, we assessed the critical success index C , which combines both hit rate and false alarm ratios into one parameter which can vary between 0 in the worst and 1 in the best scenario, indicating perfect match between both inundation maps:

$$C = \frac{N_{DFM} \cap N_{LFP}}{N_{DFM} \cup N_{LFP}} \quad (3.8)$$

For both set-ups, the River-Floodplain-Scheme was activated and flux-updating was opted for. All simulations were performed in a Linux environment with an Intel i7-4790 core at 3.90 GHz and 16 GB memory.

3.5.2 Results and discussion

Simulated discharge

Benchmarking discharge results against observation from GRDC at Óbidos shows that both models behave similarly. However, LFP tends to compute earlier peak flow and earlier and lower low flow (Figure 3.5). Therefore, obtained coefficients of correlation are lower for LFP, while the model's skill as expressed by KGE is higher for LFP and the RMSEs are comparable (Table 3.3).

The general deviation of simulated results to observations can be due to a range of factors, for example the lack of channel bifurcations in the schematization, the already less-accurate upstream inflow as simulated with the kinematic wave approximation or the general overprediction of discharge by PCR (Hoch et al., 2017a), but have not been further explored as this would exceed the scope of this study.

Even though the discrepancies in simulated discharge between the two models are not remarkable, they require further investigation as they cannot be exhaustively explained with our current process understanding. Based on the results obtained in the synthetic test case

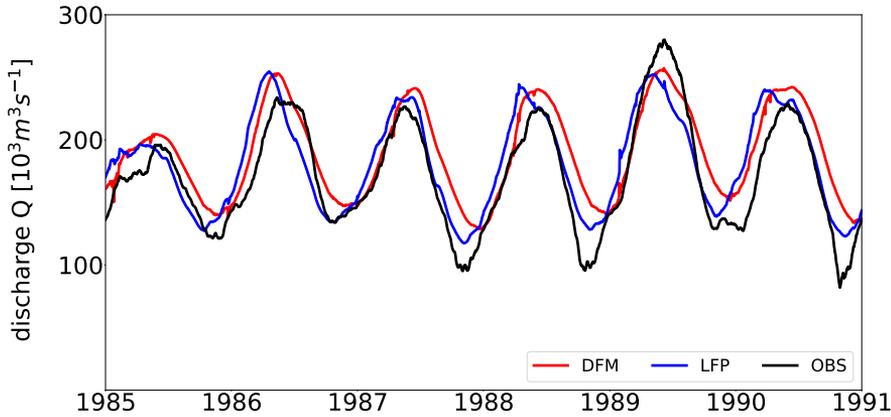


Figure 3.5: Comparison of simulated discharge from both DFM and LFP with observed discharge (OBS) from the Global Discharge Data Centre (GRDC) at Óbidos (see figure 3.4 for the location).

and since the hydrologic forcing of both models is equal in terms of water volumes, spatial distribution, and timing, we decided to evaluate the impact of the following parameters: the actual river length and dimension in LFP compared to DFM and the sensitivity of LFP to Manning's surface roughness coefficient over large areas.

Since the routing scheme of LFP is based on a D4 system where water can flow in a southerly, northerly, easterly or westerly direction, channel length and dimension in LFP tend to differ from other hydrodynamic models that are not based on such a system, for example DFM. Reducing or increasing the unitless meandering coefficient in LFP to river-scale dimensions, however, did not show any significant impact on simulated discharge (Figure 3.6a). After investigating how changes in surface roughness values in LFP may close the gap to DFM, we indeed found a more pronounced response, yet it cannot satisfactorily explain the difference in simulated discharge either (Figure 3.6b). Since in the synthetic example both models can produce near-identical results if using the same friction coefficient and, because the flow regime in the Amazon basin can be described as sub-critical, different sensitivities to surface roughness over large areas can thus also be disregarded as the cause for discharge discrepancies.

For the remaining gap in simulated discharge, we can at this point only make assumptions about the cause. Possible reasons include differences in internal processing of 1D channel bathymetry, channel-floodplain interaction, and input elevation assignment due to the different grid schematizations of a flexible mesh and regular grid.

For a further first-order assessment of a possible impact of spatial resolution, we compared simulated discharge at two stations further upstream, Loc1 and Loc2 (Figure 3.4). Results indeed suggest that the differences in upstream spatial resolution result in different flood wave propagations (Figure 3.6c); covered flow distance and peak discharge in LFP is increasingly delayed compared with DFM, presumably due to the larger floodplain cells in DFM. Besides, the timing of the rising and falling limb is affected. Higher simulated discharge by LFP than DFM at Loc1 does not only indicate that the impact of cell resolution is reduced with downstream distance and additional tributaries contributing to the flood wave but also

| | Loc3 | Loc4 | Loc5 |
|---------------------------------|---------------------|---------------------|----------------------|
| Input elevation (m) | 4.0 | 7.0 | 44.5 |
| Elevation LFP (m) | -0.2 | 2.4 | 37.4 |
| Elevation DFM (m) | 0.5 | 4.9 | 42.5 |
| Cell area LFP (m ²) | | 4*10 ⁶ | |
| Cell area DFM (m ²) | 7.7*10 ⁶ | 7.7*10 ⁶ | 30.9*10 ⁶ |

Table 3.4: Local properties of water level observation stations; input elevation refers to values obtained after hydraulic conditioning of canopy-free SRTM elevation data at 15 arc-sec spatial resolution but before resampling to or interpolating across the hydrodynamic grid.

that discharge computations in upstream areas can be easily affected as the discrepancy in cell size is largest there.

Simulated water levels

Assessing differences in simulated water level dynamics at the observation locations, we cannot find any particularly prevailing difference between the models response to hydrologic forcing (Figure 3.6d). In general, we observe that modelled water levels are comparable, yet with locally differing patterns. While at the most upstream station, Loc5, DFM simulates lower water levels than LFP, this is opposite at the most downstream station Loc3, and at Loc4 both models provide comparable results. Besides differences in actual water levels, both models show a comparable response to model input, yet LFP tends to yield earlier peak water levels than DFM, which concurs with the discharge dynamics observable.

The reason for differences in simulated water levels and their dynamics could not be fully attributed to one specific cause. For example, the more pronounced difference in water levels at Loc1 may be a local effect due to spatial feedback dynamics between neighbouring cells of an observation station (Hardy et al., 1999), may be related to slight differences in model schematization at the downstream boundary or to backwater effects in the delta regions as a result of different influences of the downstream water level boundary. Furthermore, discrepancies are likely to be related to differences in surface elevation simulated at the observation stations due to the differences in gridding between DFM and LFP.

Indeed, assessing the local properties of the observation stations revealed that the surface elevation in DFM is higher than in LFP (Table 3.4). Last, results indicate that differences in gridding and therefore cell size may thus have locally impacted the overall water levels too since the above-discussed discharge simulations in upstream areas exhibited clear deviations between both models (Figure 3.6c).

Run times

Regarding the run times of the two coupled set-ups, we find that it takes LFP around 6 h to simulate the entire simulation period of 7 years, that is model time plus spin-up, while performing the same simulation with DFM takes around 7 h (Table 3.3). The difference in run times is less pronounced than for the synthetic test case, which can be related to the lower number of cells in DFM compared to LFP due to use of a flexible mesh. In addition, a more computationally expensive interaction between the 1D and 2D domain in DFM could also affect run times. As DFM is in general a multi-purpose tool whose application is not limited to inundation modelling, it is not unexpected that it may be slightly slower than programmes specifically tailored for efficient large-scale inundation modelling such as LFP.

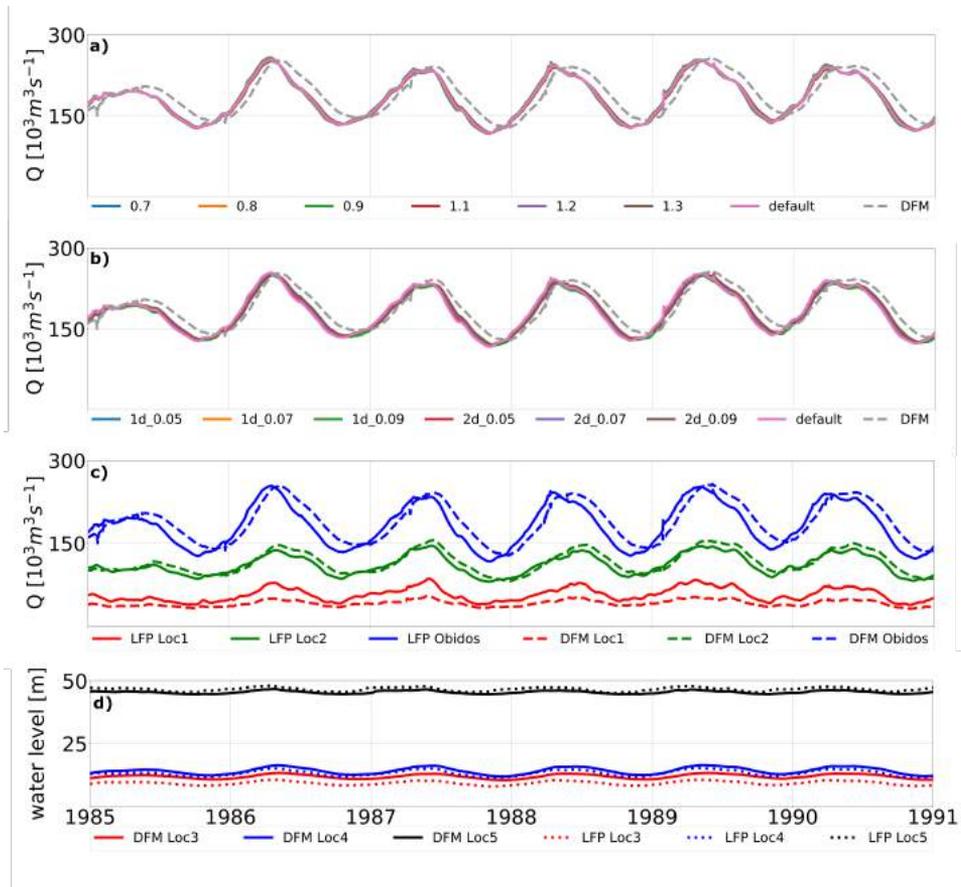


Figure 3.6: Results of the sensitivity analysis of **a)** the meandering coefficient and **b)** both 1D and 2D surface roughness coefficients in LFP. Since the D4 system in LFP can both decrease and increase effective river dimension, the dimensionless meandering coefficient was not only reduced from default (1.0) to 0.09, 0.08, and 0.07, but also increased to 1.1, 1.2, and 1.3. As default Mannings surface roughness is already low ($0.03 \text{ m}^{-1/3} \text{ s}$, coefficients were increased to 0.05, 0.07, and 0.09; **c)** compares simulated discharge across the basin to assess impact of spatial resolution on simulated discharge; to that end two additional observations upstream of Óbidos were introduced, Loc1 (most upstream) and Loc2 (intermediate upstream) – see figure 3.4 for their locations; **d)** depicts the comparison of simulated water depth at three different locations (Loc3, Loc4, and Loc5) randomly picked within the domain (see figure 3.4).

| | H | F | C |
|---------|------|------|------|
| LFP/DFM | 0.85 | 0.50 | 0.46 |

Table 3.5: Resulting hit rate H, false alarm ratio F, and critical success index C for benchmarking inundation extent between LFP and DFM.

Simulated inundation extent

We find that inundation extents obtained at the end of the simulation runs with DFM and LFP are comparable, yet far from identical (Figure 3.7). Due to the larger inundation extent of DFM, a hit rate of 0.85 is obtained, indicating that 85% of extent as simulated by LFP is also simulated by DFM. Especially differences in inundated extent in upstream areas and along small reaches can explain the obtained false alarm ratio of 0.50 (Table 3.5). These differences are also responsible for the critical success index of 0.46 corroborating that in bit less than half of the cells inundation extent is simulated by both models. A model agreement of 46% is slightly higher than the 30-40% found by Trigg et al. (2016) for a benchmarking study of global flood hazard models.

This, in fact, suggests that the choice of numerical scheme and model schematization alone can greatly impact upon inundation, confirming that differences in model forcing and boundary conditions do not act alone as a cause of modelled inundation difference, which could have been the case in the results obtained Trigg et al. (2016).

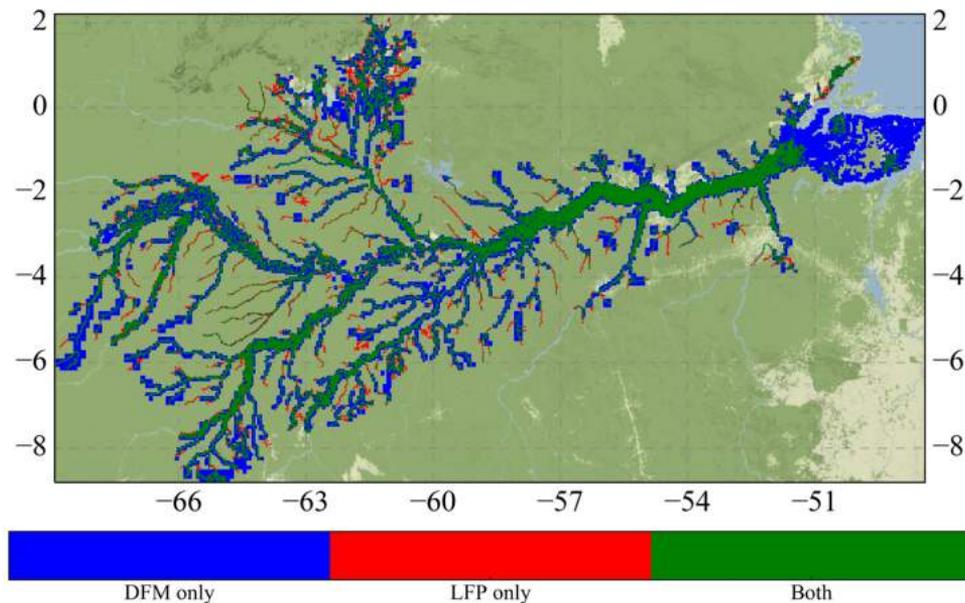


Figure 3.7: Visualization of the benchmarking of simulated inundation extent between DFM and LFP for the entire study area with some cells being flooded only by DFM, only by LFP or for both models.

A main cause for the differences observed for regions further upstream is that DFM tends to compute larger flood extent than LFP: with DFM having larger cells in upstream areas due to the flexible meshing, a larger 2D area is instantly marked as inundated for DFM once overbank flow occurs. This loss of level of detail in DFM is the concession to be made for a reduced number of grid cells and hence potentially faster computations in the 2D domain. For more downstream regions, differences in inundation extent are primarily present at small river channels while floodplain inundation is comparable. This, however, can to some extent be attributed to differences in how the 1D domain is implemented in the models, with DFM using grid-size independent vectors and LFP using grids at the overall spatial resolution of the schematization. Given the overall larger inundation extent simulated by DFM, the above-

discussed deviations in simulated discharge and in particular the more pronounced wave attenuation in DFM may be explained as return flows from the floodplain to the channel as they seem to be faster in LFP than in DFM.

3.6 Conclusion and recommendations

In this study, we presented GLOFRIM, a GLOBally applicable computational FRamework for Integrated hydrologic hydrodynamic Modelling. In its current version, it provides an environment to one-directionally couple the global hydrologic model PCR with two hydrodynamic models: DFM solving the full SWEs, and LFP solving the LIEs.

By linking hydrology to hydrodynamics, it is possible to take advantage of the strengths of both while at the same time compensating their weaknesses. We define five main assets of GLOFRIM: (i) it is openly accessible and hence can be directly applied and adapted to specific purposes, and extended with other models; (ii) by employing a global hydrologic model to obtain model forcing, the framework can easily be applied globally; (iii) models to be coupled may be selected depending on their local performance and thus more relevant processes can be captured; (iv) the spatially explicit coupling scheme can be extended to a full feedback loop between hydrology and hydrodynamics; (v) thorough benchmarking and ensemble modelling of hydrodynamic models is supported by providing identical hydrologic forcing for experiments.

GLOFRIM at present provides a range of options for model coupling. Users can choose between coupling PCR to either the 1D or 2D domain, can specify whether to update hydrodynamics through states or fluxes, and can run hydrodynamic models in both non-Cartesian spherical and projected coordinate systems. It is generically written and does not require any a priori knowledge of the code as all important settings are specified in a separate settings file.

Besides PCR as well as DFM and LFP, there are many other global hydrologic and hydrodynamic models available which have their individual advantages. As the framework is freely and openly available, its design can easily be extended and adapted to cater for the coupling of other hydrologic or hydrodynamic models, merely requiring the implementation of the BMI into each model to be added.

Eventually, adding a 1D continental hydrodynamic model such as CaMa-Flood (Yamazaki et al., 2011) would allow for replacing the kinematic wave approximation of PCR to provide more accurate upstream boundary inflow to the domain with explicit high-resolution 2D floodplain computations.

Employing a BMI does not change the model functionality but it does provide a range of added functions. Furthermore, not all model variables need to be exposed, only those to reproduce model geometry, distinguish between 1D and 2D cells, and to update model states. We therefore recommend considering this option for future model developments and will also aim to incorporate other models ourselves. To our knowledge, spatially explicit model coupling on a global scale by means of such a framework is unprecedented. Consequently, user experiences and lessons learnt are still sparse and any initiatives regarding framework extension are therefore kindly received by the authors, as well as feedback and experiences. We also recommend the testing and application of it in other study areas and under different boundary conditions to further evaluate the code, process flow, and applicability.

Before applying GLOFRIM in an actual test case, we performed a simple synthetic test case to obtain a first-order insight into how both models may differ regarding their computational complexity. Thereby both the 1D and 2D domain were forced by a synthetic inflow signal

and simulated discharge was evaluated along the flow path. Results show that both models produce the same response to the signal despite the difference in solver complexity.

The results obtained are in line with previous studies showing that for sub-critical flow regimes discharge results should be similar (de Almeida and Bates, 2013; Neal et al., 2012a). Both hydrodynamic models were then applied within GLOFRIM for the Amazon River basin and evaluated regarding simulated discharge, water levels, run time, and inundation extent, also constituting a first comparison of large-scale flexible mesh and regular grid applications.

Assessing simulated discharge shows that both models exhibit comparable results with LFP tending to compute earlier and slightly increased peak discharge estimates. As thorough testing of plausible causes did not show significant improvements, we speculate that differences in processing of 1D channel bathymetry, interaction between 1D channels and 2D floodplains or assignment of input surface elevation data to the different grids may impact discharge results. The latter is supported by discharge observations made in farther upstream areas where differences in grids are largest. A more in-depth analysis of these differences was, however, outside the scope of this study and thus needs to be performed in a follow-up study.

As the general overprediction of observed discharge at Óbidos can partly be attributed to the absence of hydrologic processes on inundated floodplains, it is envisaged to extend the current code such that it also caters for a full feedback loop between hydrodynamics and hydrology.

Water levels simulated by both models differ locally, yet only slightly. These discrepancies between both models are most likely due different grid schematizations in DFM and LFP, which results in locally differing elevation values and cell areas and thus influences simulated water levels. Due to differences in model structure and design, downstream boundary conditions had to be implemented slightly differently, possibly also impacting water level results in particular for more downstream stations. As it was the aim of this paper to introduce the computational framework applied, a more elaborated evaluation of causes for water level deviations is future work.

A key parameter for large-scale modelling is run time. In the current study, the schematization of LFP contains more than 4 times the number of 2D cells than DFM while the number of 1D cells is 40% higher in LFP than in DFM. Despite the greater number of cells, LFP has a slightly shorter run time. This is in line with the results obtained in the synthetic test case, yet the relative difference is reduced due to the application of flexible meshes for the 2D domain and the nature of the coupling algorithm applied: because water was coupled directly into the 1D channels, flow over the 2D domain was limited and, as a result, so was the impact of differences in computational efficiency of the models.

Differences in run times may also be related to more fundamental factors, such as the degree of code optimization applied. Additionally, DFM was, in contrast to LFP, not explicitly developed for efficient inundation modelling, but as a multi-purpose tool including several additional physical processes, such as the potential to simulate 3D flow, estuarine processes or hydrogeomorphologic dynamics, which could also result in longer run times. To better understand causes of run time discrepancies, further model development, testing, and evaluation is therefore recommended.

To benchmark LFP and DFM in terms of simulated inundation extent in the Amazon River basin, the hit rate H , the false alarm ratio F , and the critical success index C were determined. In general, both models agree about as often as they disagree, indicating that both DFM and LFP predict simulation extent for around half of all cells. This level of agreement is slightly higher than the one obtained by Trigg et al. (2016) and is a strong indication that the model

geometry and numerical scheme play a similarly strong role in influencing model accuracy as the boundary conditions and model forcing applied in global flood hazard models.

Moreover, a higher value could not be obtained due to the impact of the flexible mesh, especially for upstream areas where DFM runs at cells that are a factor of 25 larger than in LFP. While such large cells contribute strongly to shorter run times, they may also have implications for detailed flood hazard estimates which can be strongly hampered. In the case of employing a flexible mesh, it seems as if an a priori decision has to be made where and to which extent such models are supposed to provide fine-scale results or whether computational efficiency is the main aim both at the same time does not seem to be feasible from our results.

We hence recommend testing the application of flexible meshes for large-scale riverine inundation modelling in more detail to obtain a better understanding of the trade-off to be made between grid refinement and model accuracy.

With the presented computational framework GLOFRIM and the satisfactory results obtained, we trust to have contributed to the current development of model coupling and integration, and to have provided an openly accessible tool that facilitates more accurate large-scale flood hazard estimates.

We hope that, eventually, the integration of hydrologic and hydrodynamic models will lead to improved flood risk assessments and planning of climate change impact mitigation and adaption measures.

Code and data availability

The code of GLOFRIM 1.0 and the BMI versions of LFP and PCR are openly accessible and freely downloadable at <https://doi.org/10.5281/zenodo.597107> (Hoch et al., 2018b).³

Author contributions

FB and JMH developed the BMI adapter for LISFLOOD-FP. JMH, RvB, and JCN developed the code of the computational framework. JCN, RvB, HCW, PDB, and MFPB supervised the research and provided important advice. JMH designed and executed the research, and also prepared the manuscript as well as code, with contribution from all co-authors.

Acknowledgements

This study was financed by the EIT Climate-KIC programme under project title “Global high-resolution database of current and future river flood hazard to support planning, adaption and re-insurance”. We also want to acknowledge the contributions of Climate-KIC and University of Bristol to realize a research stay at the University of Bristol. Special thanks are reserved for Arthur van Dam and Herman Kernkamp from Deltares for their support in applying Delft3D Flexible Mesh and Edwin Sutanudjaja for PCR-GLOBWB advice. Last, we want to express our gratitude to an anonymous reviewer and Dai Yamazaki for evaluating a previous version of the manuscript and providing invaluable feedback.

³Note that computer code and its structure is under constant development. It may hence be that the most recent release deviated from the one used here. However, functionality should not have changed.

Chapter 4

Benchmarking flexible meshes and regular grids for large-scale fluvial inundation modelling

Damage resulting from flood events is increasing world-wide, requiring the implementation of mitigation and adaptation measures. To facilitate their implementation, it is essential to correctly model flood hazard at the large scale, yet fine spatial resolution. To reduce the computational load of models, flexible meshes are an efficient means compared to uniform regular grids. Yet, thus far they have been applied only for bespoke small-scale studies requiring a high level of a priori grid preparation. To better understand possible advantages as well as shortcomings of their application for large-scale riverine inundation simulations, three different flexible meshes were derived from Height Above Nearest Drainage (HAND) data and compared with regular grids under identical spatially explicit hydrologic forcing by using GLOFRIM, a framework for integrated hydrologic-hydrodynamic inundation modelling. By means of GLOFRIM, output from the global hydrologic model PCR-GLOBWB was passed to the hydrodynamic model Delft3D Flexible Mesh. Results show that applying flexible meshes can be beneficial depending on the envisaged purpose. For discharge simulations, similar model accuracy was obtained between flexible and regular grids, with the former generally having shorter run times. For inundation extent simulations, however, the coarser gridding of flexible meshes in upstream areas results in a poorer performance if assessed by contingency maps. Moreover, while the ratio between minimum and maximum spatial resolution of flexible meshes has limited impact on discharge simulations, water level estimates may be stronger influenced by the application of larger grid cells. As this study presents only a small set of possible realizations, additional research needs to unravel how the data and methods used as well as the choices for discretizations influence model performance. Generally, the application and particularly discretization process of flexible meshes involves more options, bringing more responsibilities for the user. Once an a priori decision is made on the model purpose, flexible meshes can be a valuable addition to modelling approaches where short run times are essential, facilitating large-scale flood simulations, ensemble modelling or operational flood forecasting.

Based on: Hoch, J. M., van Beek, R., Winsemius, H. C., Bierkens, M.F. P. (2018), Benchmarking flexible meshes and regular grids for large-scale fluvial inundation modelling, *Advances in Water Resources* 121, 350–360.

4.1 Introduction

In recent years, losses due to riverine inundations increased strongly: between 1980 and 2013, they exceeded USD 1 trillion of direct economic losses and more than 220,000 fatalities (MunichRe, 2010). This development can be attributed to the growth of both population and asset values in floodplains (Ceola et al., 2014; Winsemius et al., 2016) as well as changes in river regimes (cf. Jongman et al. (2012), MunichRe (2010), Visser et al. (2012), and Winsemius et al. (2016)). Despite inherent uncertainties, several studies indicate that flood risk will enhance in the future (Hirabayashi et al., 2013; Jongman et al., 2014; Winsemius et al., 2016).

To capture the driving climate-flood interactions and processes world-wide, it is beneficial to apply global hydrologic models (GHMs) to guarantee seamless large-scale inundation modelling across basins and borders. Besides modelling flood hazard at such scale, information should be provided at a spatial resolution sufficiently fine to be locally relevant (Bierkens et al., 2015), facilitating stakeholders involvement (Beven et al., 2015). However, the finest spatial resolution achieved for GHMs is currently 10km x 10 km at the Equator (Bierkens, 2015).

One way to improve the applicability of GHMs would be by simulating lateral floodplain flow and channel-floodplain interactions at a finer scale. Moving to a finer scale is, however, not straightforward as the current debate about hyper resolution shows (cf. Beven et al. (2015), Bierkens et al. (2015), Wood et al. (2011), and Beven and Cloke (2012)).

In contrast to GHMs, hydrodynamic models can run at a finer spatial resolution, for instance 90 m globally (Sampson et al., 2015) or 30m for the Continental United States (Wing et al., 2017; Wing et al., 2018). A downside of hydrodynamic models, however, is that they often use observed discharge as model forcing or employ synthesized flood waves, hence not accounting for all relevant hydrological processes. Consequently, the spatial correlation of large-scale flood events as well as the impact of climate change on flood hazard and risk can be simulated only with concessions.

One way to circumvent the problems associated with coarse spatial resolutions of GHMs and data dependency of hydrodynamic models is hydrologic-hydrodynamic model coupling. On smaller scales, this was already achieved (Felder et al., 2017; Kim et al., 2012; Viero et al., 2014), and in a more recent study Hoch et al. (2017a) coupled large-scale hydrologic and hydrodynamic models. Besides spatial extent, the latter approach distinguishes itself from others such that it employs the Basic Model Interface (BMI; Peckham et al. (2013)), providing a flexible coupling design avoiding changes to, and entanglement of model code. By means of this interface, output from PCR-GLOBWB (PCR; Sutanudjaja et al. (2018)) forced the hydrodynamic model Delft3D Flexible Mesh (DFM; Kernkamp et al. (2011)). While discharge simulations improved, the extent to which the chosen flexible mesh impacted results remained unclear, calling for additional research on the use of flexible meshes for large-scale inundation modelling.

Despite the large number of studies employing flexible meshes for fluvial flooding (cf. Castro Gama et al. (2013), Kim et al. (2014), Kumar et al. (2009), Sanders et al. (2010), and Schubert et al. (2008)), none explicitly assesses the role of different mesh configurations, let alone for large-scale applications. While for these bespoke studies an efficient mesh was usually created first, such fine-tuning is too time-consuming for large-scale inundation modelling which may encompass several larger catchments. What is rather needed are fast approaches to generate flexible meshes over large areas covering a grand variety of topological properties. To make maximum use of the potential of flexible meshes, a resolution sufficiently fine to provide “locally relevant” (Bierkens et al., 2015) results has to be determined a priori. How

various degrees of mesh refinement impact large-scale inundation modelling results is hardly researched until now and thus additional insight is needed.

In contrast to flexible meshes, there is a multitude of studies investigating various aspects of regular grid refinement. For instance, it was found that spatial resolution impacts the accuracy of inundation estimates (Savage et al., 2016b), water depth estimates and floodplain drainage flow (Savage et al., 2016a), and channel flow through near channel storage effects (Horritt and Bates, 2001a). Hardy et al. (1999) concluded that grid resolution impacts simulated discharge linearly and water depth in a less structured way due to the impact of the geographical surrounding of each observation location. Comparable results were obtained by Fewtrell et al. (2008) in a small urban environment.

In this study, we will add considerations for large-scale (potentially even global-scale) flood hazard models using a fast set-up of flexible meshes. We present a first benchmark and sensitivity analysis to advance our understanding how model accuracy scales with flexible mesh discretization in large-scale studies. Eventually, we want to better understand (a) how different configurations of flexible meshes influence model accuracy, (b) how results differ between flexible meshes and regular grids, and (c) what lessons can be learned for future applications.

The analysis was performed by employing GLOFRIM, a globally applicable framework for integrated hydrologic-hydrodynamic modelling (Hoch et al., 2017b). GLOFRIM is an openly accessible, modular, and extensible tool facilitating model coupling, currently allowing for spatially coupling PCR with DFM or LISFLOOD-FP (LFP; Bates et al. (2010)). What was decisive to apply GLOFRIM was the requirement to guarantee identical spatially varying model forcing for all discretizations as well as the need to include all river reaches and floodplains in the analysis.

Three different 1D/2D flexible meshes of the lower Elbe basin (Figure 4.1) were forced with identical output from PCR at 30 arcmin spatial resolution. All flexible meshes were created based on the Height Above Nearest Drainage (HAND) method (Rennó et al., 2008). We decided to use HAND as it requires only little input data and is fast in computing topographical gradients with respect to the channel network. To benchmark model results of flexibly gridded meshes, we also applied GLOFRIM to three regular grids. All model results were then validated against observed discharge values as well as benchmarked with respect to their simulated water levels, water volume, run time, and inundation extents.

4.2 Models and methods

4.2.1 PCR-GLOBWB

The global hydrologic model PCR-GLOBWB (PCR; Sutanudjaja et al. (2018)) distinguishes between two vertically stacked soil layers, an underlying groundwater layer, and a surface canopy layer. Water can be exchanged vertically, and excess surface water can be routed horizontally along a local drainage direction network, employing the kinematic wave approximation. The model was forced with Climate Research Unit (CRU) precipitation and temperature data (Harris et al., 2014), and potential evaporation was computed using the Penman–Monteith equation. Data sets were downscaled to daily fields for the period from 1957 to 2010 using ECMWF (European Centre for Medium Weather Forecasts) re-analysis products (ERA40/ERA-Interim; Kållberg et al. (2005) and Uppala et al. (2005)) as outlined in van Beek (2008). PCR furthermore takes into consideration irrigation water demand and industrial and domestic water abstraction based on reported water demand (FAOSTAT 2017). To use the best possible hydrologic forcing for the hydrodynamic model, we applied a regional

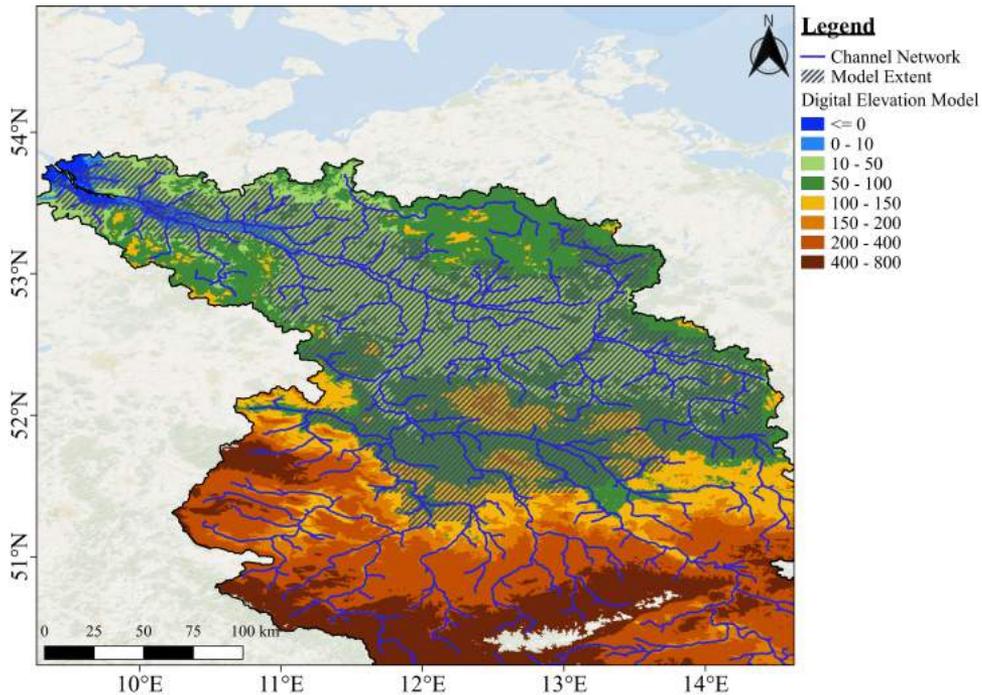


Figure 4.1: Plot of the digital elevation model (DEM) and channel network in the Elbe basin as well as the actual model extent for which hydrodynamic simulations were run.

optimization scheme to find the parameterization yielding the most accurate discharge estimations at Neu-Darchau (Figure 4.2). Further explanation regarding the optimization technique can be found in Hoch et al. (2017a). Manning's surface roughness coefficients of $0.04 \text{ m}^{-1/3} \text{ s}$ and $0.07 \text{ m}^{-1/3} \text{ s}$ were used for river channel and floodplain, respectively.

4.2.2 Delft3D Flexible Mesh

Delft3D Flexible Mesh (DFM; Kernkamp et al. (2011)) allows its user to discretize the 2D model domain with a flexible mesh, applying different geometrical shapes at various resolutions to discretize the study area, or regular grids, using the same spatial resolution over the entire domain. The application of a flexible mesh with DFM is both mass and momentum conservative as a) the continuity equation is formulated in a conservative way and b) requirements of orthogonality must be met for stable model runs. For instance, triangles must be acute, that is none of the internal angles must be larger than 90° . For further information on the use as well as technical descriptions of DFM, we refer to the user manual and technical reference manuals (*Delft3D Flexible Mesh User Manual 2019*; *Delft3D Flexible Mesh Technical Manual 2019*). The numerical solutions of DFM are provided by Eq. 3.1 and Eq. 3.2. In contrast to PCR, DFM solves the full shallow water equations and thus can capture important flood triggering processes such as backwater effects (Moussa and Bocquillon, 1996). To maintain comparability, DFM also employs Manning's surface roughness coefficients of $0.04 \text{ m}^{-1/3} \text{ s}$ and $0.07 \text{ m}^{-1/3} \text{ s}$ for 1D channels and 2D floodplain flow, respectively.

4.2.3 GLOFRIM

GLOFRIM is a globally applicable framework for integrated hydrologic-hydrodynamic modelling (Hoch et al., 2017b). With GLOFRIM, it is possible to perform spatially explicit coupling between hydrologic and hydrodynamic models at a time step basis. With the current version of GLOFRIM, PCR can be coupled to either the DFM model used here or LISFLOOD-FP (Bates et al., 2010). Applying GLOFRIM has two major advantages: first, identical model forcing is provided through PCR output, guaranteeing reproducibility and comparability; and second, setting up a coupled hydrologic-hydrodynamic model is greatly facilitated due to the pre-defined workflow. GLOFRIM is built upon the Basic Model Interface (BMI; Peckham et al. (2013)). In contrast to other model coupling studies employing other approaches to exchange information (chapter 1.3.1), using the non-invasive BMI allows continuing separate development of the models and avoids the entanglement of model code. By means of the BMI it is possible to retrieve, manipulate, and place model data during model execution. Hence, spatial coupling can be achieved by overlaying grids from two models, and assigning hydrodynamic to hydrologic cells on a grid-to-grid basis. Consequently, the hydrodynamic model is forced with output from PCR by exchanging runoff and discharge volumes between corresponding PCR and DFM cells. For further information regarding GLOFRIM, we refer to Hoch et al. (2017b) and chapter 3.

4.2.4 Hydrodynamic discretizations

Six different DFM discretizations of the lower Elbe basin were designed (Figure 4.2): three with spatially varying grid size and three with uniform grid size. For the flexible meshes, we designed these set-ups: F1 used a length of 1600 m for its coarsest resolution whereas F2 used 3200 m. Both employed 400 m for its smallest cells. Comparable to F1, F3 also used 1600 m for its coarsest resolution but used only 800 m for its finest cell, in order to evaluate the effects of both the largest and finest cell lengths. For the regular grids, we set up discretizations with 400 m (R1), 800 m (R2), and 1600 m spatial resolution (R3) to be compared with the flexible meshes. Additional descriptive statistics can be found in Table 4.1. None of the hydrodynamic discretizations were calibrated as the impact of model parameters scales with spatial resolution (Fewtrell et al., 2008).

To derive the model discretization for DFM, we used HydroSHEDS surface elevation and drainage network data at 15 arcsec (Lehner et al., 2008) to apply the Height Above Nearest Drainage algorithm (HAND; Rennó et al. (2008)).

We opted for HAND as it provides a tool for fast grid generation in terms of both data requirements and execution time and is thus well suited for large-scale applications. Besides, HAND was applied for other inundation modelling studies (Nobre et al., 2016; Speckhann et al., 2018) and has only a user-defined upstream area threshold as possible source of uncertainty. Various levels of grid refinement were achieved by using different initial grid sizes as well as varying values for both minimum grid cell size and maximum model time step.

As 2D floodplain elevation values we employed the canopy-removed bare earth SRTM data (O'Loughlin et al., 2016) and hydraulically smoothed it to account for the vertical measurement errors inherent in remotely sensed elevation data (Yamazaki et al., 2017; Yamazaki et al., 2012a) before assigning it to the 2D part of the grids.

We based both the network and river width information of the 1D channels on the Global Width Database for Large Rivers (GWD-LR; Yamazaki et al. (2014a)), while river depth information was derived by applying the equations of Leopold and Maddock (1953). Bathymetric information was stored at cross-sections with a spacing of around 10 km and subsequently interpolated between cross-sections along the river network. For both flexible meshes and

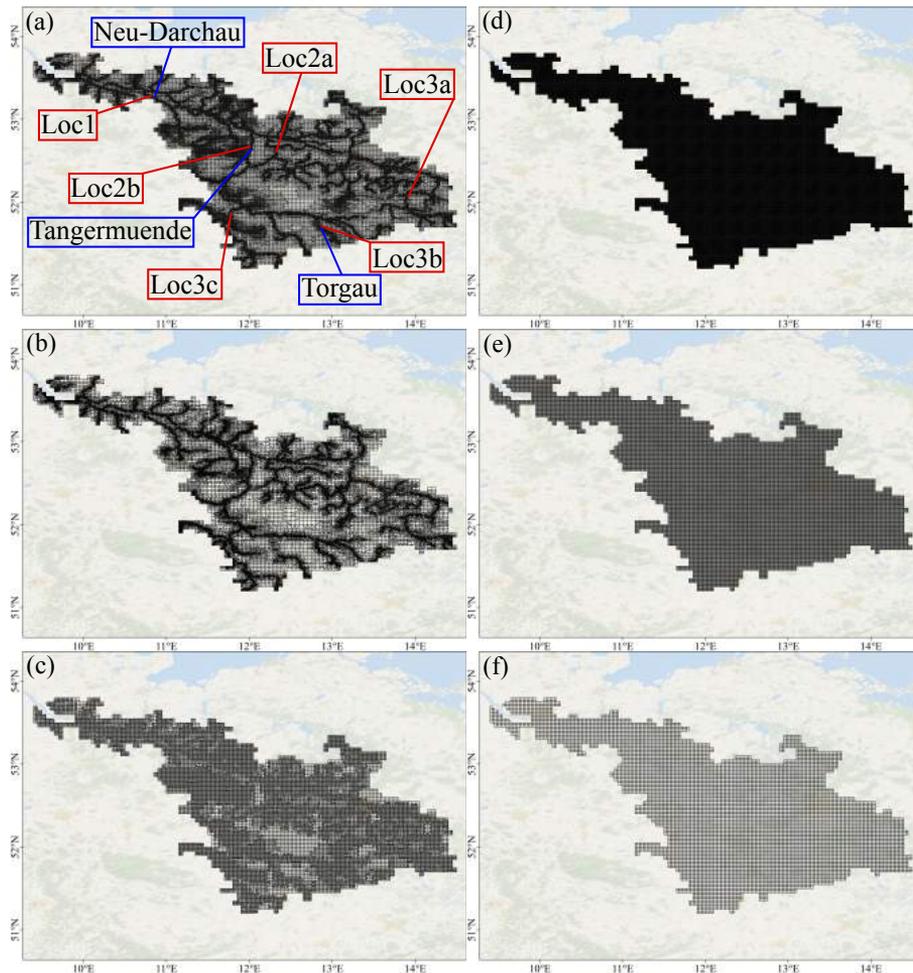


Figure 4.2: Locations of observation stations used for both discharge validation and water level analysis as well as different discretizations of the Elbe Basin in DFM: **a)** F1; **b)** F2; **c)** F3; **d)** R1; **e)** R2; and figure 3.4 R3. For more detailed properties of grids see Table 4.1.

regular grids, the 1D channel discretization remained unaltered to guarantee consistency between model runs.

We did not account for dikes and other man-made structures due to the lack of reliable global data for our large-scale applications and implementing them would otherwise introduce additional uncertainty to model results. Due to the same reason, we desisted from using sub-grid elevation data or spatially heterogeneous surface roughness values which would typically be done for catchment-scale studies.

4.2.5 Assessment of model results

All test cases were run for the period 01 January 2002 until 31 December 2010 after two years of spin-up. Simulated discharge was validated at Neu-Darchau (ND), Tangermuende (TM), and Torgau (TG) (Figure 4.2) by computing the coefficient of correlation (r), the root mean

| | F1 | F2 | F3 | R1 | R2 | R3 |
|----------------------------------|-----------|------------|-----------|---------|-----------|-----------|
| Min. cell length | 400 m | 400 m | 800 m | 400 m | 800 m | 1600 m |
| Max. cell length | 1600 m | 3200 m | 1600 m | 400 m | 800 m | 1600 m |
| Number of 2D cells | 130,532 | 44,213 | 27,806 | 138,710 | 34,684 | 8,669 |
| Min. cell area [m ²] | 142,982 | 143,153 | 572,761 | 285,906 | 1,143,690 | 4,575,274 |
| Max. cell area [m ²] | 4,811,793 | 19,081,223 | 4,846,141 | 303,290 | 1,213,058 | 4,851,411 |
| Ration max/min | 33 | 133 | 8 | 1 | 1 | 1 |
| Avg. cell area [m ²] | 626,893 | 925,359 | 1,471,248 | 294,897 | 1,179,564 | 4,718,120 |
| Med. cell area [m ²] | 299,654 | 576,330 | 1,184,197 | 295,197 | 1,180,852 | 4,723,911 |

Table 4.1: Measured and derived properties of different DFM discretizations; note that differences in minimum and maximum cell area for regular runs with uniform cell length are due to the use of a spherical coordinate systems which results in variations of area depending on latitude and longitude.

square error (RMSE), and the model's skill expressed as the Kling–Gupta efficiency (KGE; Gupta et al. (2009)). That way it is possible to assess the impact of different discretizations under different discharge regimes. The required discharge observations were kindly provided by the German Waterway and Shipping Administration (Wasser- und Schifffahrtsverwaltung des Bundes; WSV) via the Federal Institute of Hydrology (Bundesanstalt für Gewässerkunde; BfG).

To obtain an impression how simulated water levels differ throughout the basin, they were compared qualitatively at six observation stations covering the up-, mid-, and downstream part of the basin (Figure 4.2).

Inundation extent was benchmarked for all discretizations similar to the approach and reasoning of Fewtrell et al. (2008). Thereby, the hit rate H , the false alarm ratio F , and the critical success index C were determined for each inundation map with respect to the map with the highest spatial resolution $R1$. H , F , and C were computed with the subsequent equations where N_{R1} and N_{comp} indicate the number of inundated cells in of the benchmark map obtained with $R1$ and the map of the discretization to be compared, respectively.

$$H = \frac{N_{comp} \cap N_{R1}}{N_{R1}} \quad (4.1)$$

$$F = \frac{N_{comp} \setminus N_{R1}}{N_{comp} \cap N_{R1} + N_{comp} \setminus N_{R1}} \quad (4.2)$$

$$C = \frac{N_{comp} \cap N_{R1}}{N_{comp} \cup N_{R1}} \quad (4.3)$$

All parameters can vary between 0 and 1. While $H=1$ shows that all inundated cells in the benchmark data are also inundated in the comparison data, $F=1$ indicates that the inundated cells in the comparison are entirely false alarms with respect to the benchmark. The critical success rate C , in turn, should be 1 for perfect agreement, thereby penalizing for both under- and overprediction.

Unfortunately it was not achievable to validate simulated inundation extent against observations due to the lack of embankment height information and the resulting overestimation of simulated inundation extent.

Simulated discharge, water levels, and inundation extent were put into perspective by assessing simulated water volumes, which functions as a proxy for overbank water storage. In addition, run times are reported to evaluate the computational efficiency of the different grids.

4.3 Results and Discussion

4.3.1 Simulated discharge

Three observations can be made across all six discretizations regardless the gridding scheme. First, computed discharge exceeds observations for regular flow regimes, but underpredicts discharge for peak flow conditions (Figure 4.3). Further investigation revealed that this is mostly due to the discharge overpredicted by PCR which thus already determines the potential accuracy of the coupled output. Second, the magnitude of exceedance increases downstream, as expressed by the increase in RMSE (Table 4.2a). We postulate that this larger bias is caused, at least partly, by the absence of hydrological processes in the hydrodynamic model, such as groundwater infiltration or evaporation. Last, the absence of dikes influences the shape of all simulated hydrographs. Without dikes, simulated discharge is smoother due to less flow constriction, dampening and lagging particularly peak discharge. While the different aspects do affect model accuracy, all discretizations are, however, affected equally and hence further benchmarking is not hampered.

We assess the influence of the gridding technique applied first. Comparing the KGEs of the discretizations with 400m (F1, F2, and R1) and those with 800m finest spatial resolution (F3 and R2) reveals that the application of a regular grid improves model's skill insignificantly compared to a flexible mesh discretization if the same finer spatial resolution is applied (Table 4.2a). Additionally, results obtained for the regular grid runs indicate that further coarsening of the grid from 800m to 1600m impacts discharge results less drastically than from 400m to 800m, especially with respect to peak discharge computations.

Evaluating the impact of spatial resolution on discharge estimates, we find at ND that simulated discharge deviates only slightly between spatial resolutions (Figure 4.3a and Table 4.2a). For flexible meshes, F1 and F2 show near-identical discharge results while F3 yields lower estimates. Similarly, F1 and F2 yield comparable discharge results at TM and TG. The near-identical results of F1 and F2 at all three stations suggest that the choice of the finest spatial resolution within a flexible mesh strongly determines the accuracy of discharge simulations while the coarsest resolution is less influential. At these farther upstream stations, however, the deviation of F3 from F1 and F2 as well as of R3 from R1 and R2 is larger than at ND (Figure 4.3c-f).

Because discharge at ND differs hardly between discretization, the overall discharge volumes passing TM and RT should also be comparable to exclude any water balance errors. As this is not the case here, we re-run all discretization with cross-sections covering the entire floodplain width to exclude uncaptured floodplain flow as cause.

Comparing discharge obtained from channel flow (Figure 4.3) with the full floodplain discharge (Figure 4.4) suggest that with coarser cells a larger fraction of total downstream floodplain flow travels via the 2D floodplain cells, most likely due to the reduced number of 2D cells available to accommodate floodplain flow.

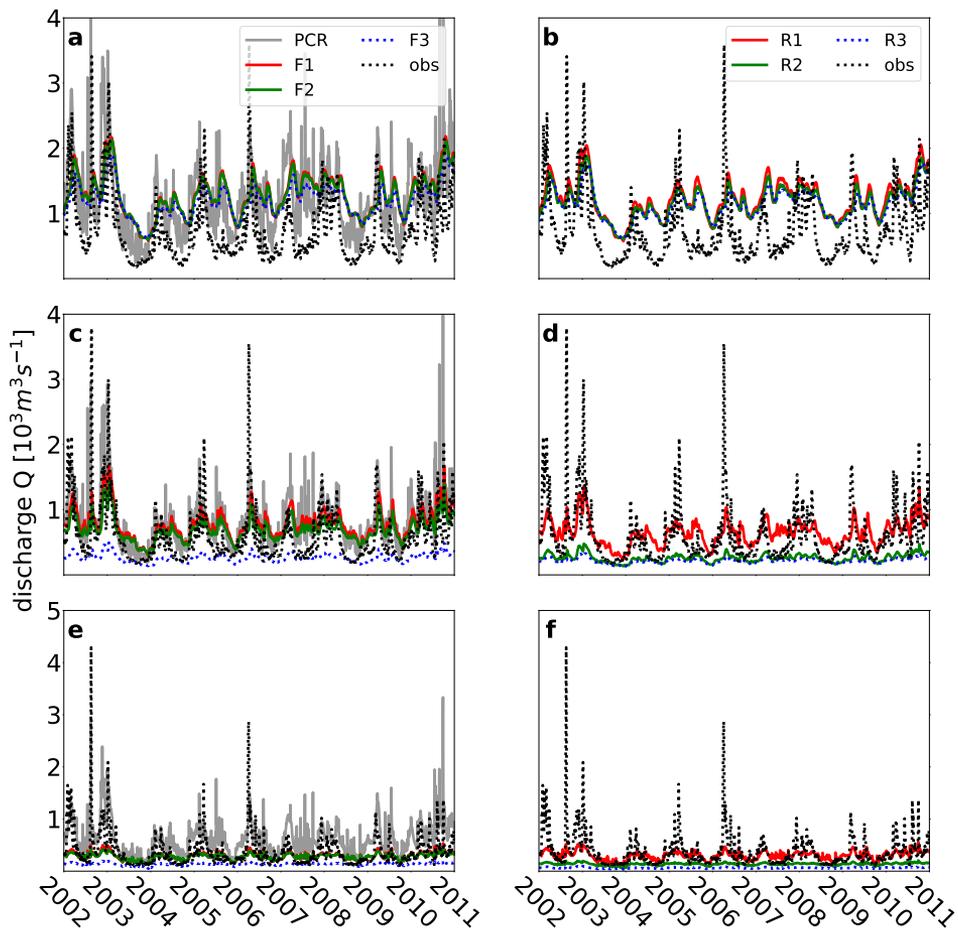


Figure 4.3: Simulated and observed discharge at three observation stations throughout the basin for cross-sections only capturing the flow in the 1D channels only, separately plotted for all flexible meshes (left) and regular grids (right): (a and b) Neu-Darchau, (c and d) Tangermuende, and (e and f) Torgau.

Since model skill is near-identical at each station across set-ups (Table 4.2a), the new results provide insight into flood wave propagation: at the most upstream station TR, PCR discharge and DFM discharge correlate very strongly ($r=0.94$), but with increasing downstream distance, the discrepancy between discharge simulated without and with GLOFRIM increases from $r=0.75$ at TM to $r=0.57$ at ND. This underpins the above made assumption that not accounting for open water evaporation and groundwater infiltration in hydrodynamic models can lead to a reduction of model accuracy. In the subsequent section, their potential influence is analysed in more depth.

As mentioned above, discharge peaks are not well simulated by the hydrodynamic model, and we thus performed a peak-above-threshold analysis to assess performance for peak flows separately (Table 4.2b). As threshold we used the long-term mean discharge per station as

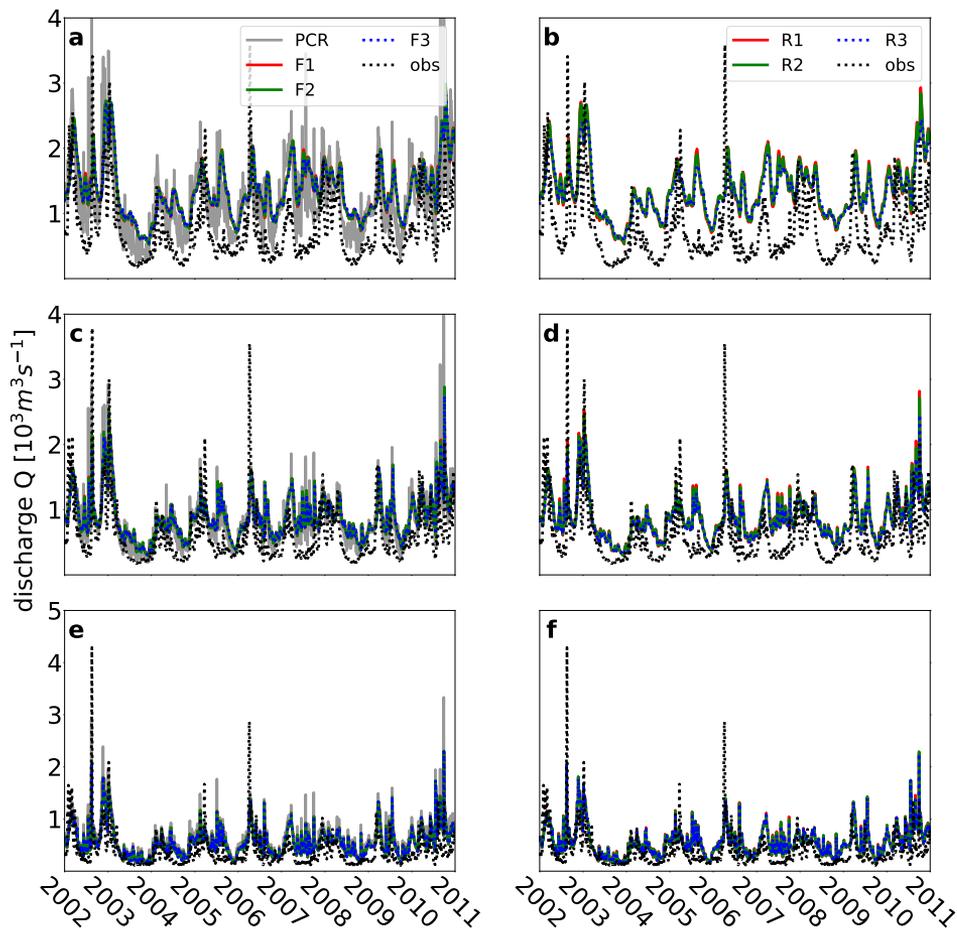


Figure 4.4: Simulated and observed discharge at three observation stations throughout the basin for cross-sections covering the 1D channels as well as the entire 2D floodplain, separately plotted for all flexible meshes (left) and regular grids (right): (a and b) Neu-Darchau, (c and d) Tangermuende, and (e and f) Torgau.

reported by the BfG⁴: for ND $705 \text{ m}^3 \text{ d}^{-1}$, for TM $562 \text{ m}^3 \text{ d}^{-1}$, and for TR $340 \text{ m}^3 \text{ d}^{-1}$. Results suggest that for peak flow conditions, the discretization approach opted for as well as the absence of dikes and other flood wave containing measures impacts model accuracy strongly and poses a limitation to using the here applied discretizations in an operational setting. Besides, results corroborate that capturing floodplain flow for coarser discretizations is even more important for peak flow conditions. These findings are, however, in line with expectations as we used only global data sets and thus applicability for local bespoke studies may be reduced (Ward et al., 2015).

⁴Hydrologic properties for each BfG station in the Elbe basin can be found on the Undine webpage: <http://undine.bafg.de/elbe/elbegebiet.html>.

| | F1 | F2 | F3 | R1 | R2 | R3 |
|---|--------------|--------------|--------------|--------------|--------------|--------------|
| <i>(a) Full time-series analysis</i> | | | | | | |
| Neu-Darchau | | | | | | |
| r | 0.33 | 0.31 | 0.28 | 0.33 | 0.29 | 0.25 |
| KGE | 0.04 / 0.07 | 0.03 / 0.06 | 0.05 / 0.05 | 0.08 / 0.07 | 0.07 / 0.06 | 0.06 / 0.05 |
| RMSE | 700 | 694 | 638 | 654 | 621 | 604 |
| Tangermuende | | | | | | |
| r | 0.39 | 0.37 | 0.31 | 0.40 | 0.32 | 0.18 |
| KGE | 0.16 / 0.39 | 0.22 / 0.39 | 0.05 / 0.38 | 0.23 / 0.38 | -0.02 / 0.38 | -0.11 / 0.39 |
| RMSE | 400 | 419 | 418 | 437 | 438 | 434 |
| Torgau | | | | | | |
| r | 0.65 | 0.60 | 0.60 | 0.69 | 0.59 | 0.63 |
| KGE | 0.12 / 0.39 | 0.06 / 0.39 | -0.19 / 0.39 | 0.47 / 0.39 | -0.19 / 0.39 | -0.32 / 0.39 |
| RMSE | 280 | 291 | 374 | 241 | 373 | 425 |
| <i>(b) Peak-over-threshold analysis</i> | | | | | | |
| Neu-Darchau | | | | | | |
| r | -0.01 | -0.02 | -0.04 | -0.01 | -0.04 | -0.05 |
| KGE | -0.03 / 0.15 | -0.04 / 0.15 | -0.06 / 0.10 | -0.03 / 0.14 | -0.06 / 0.12 | -0.08 / 0.05 |
| RMSE | 684 | 675 | 635 | 650 | 629 | 624 |
| Tangermuende | | | | | | |
| r | 0.04 | 0.01 | -0.05 | 0.05 | -0.03 | -0.13 |
| KGE | -0.15 / 0.35 | -0.08 / 0.36 | -0.29 / 0.33 | -0.02 / 0.34 | -0.34 / 0.32 | -0.42 / 0.30 |
| RMSE | 633 | 581 | 670 | 561 | 719 | 694 |
| Torgau | | | | | | |
| r | 0.47 | 0.46 | 0.46 | 0.48 | 0.47 | 0.43 |
| KGE | -0.11 / 0.37 | -0.17 / 0.37 | -0.34 / 0.37 | 0.27 / 0.37 | -0.34 / 0.37 | -0.44 / 0.36 |
| RMSE | 460 | 478 | 610 | 329 | 609 | 677 |

Table 4.2: Descriptive statistics of simulated discharge at stations Neu-Darchau, Tangermuende, and Torgau; results are separately assessed for (a) entire time series and (b) peaks above threshold. The threshold per station is reported BfG mean discharge. Also, the effect of including the entire floodplain flow is presented with descriptive statistics in italic.

4.3.2 Water volume

From Figure 4.5, three groups can be distinguished: the water volume of R3 which grossly exceeds all other simulated volumes; an intermediate group consisting of F1, R2, and F3; and the group of F2 and R1 containing the least water volume in the system. Overall, results show that the water volume stored in the system increases significantly when moving to a coarser discretization. The aggregation rate expressed as slope of the linear fit ranges between $11 \cdot 10^5 \text{ m}^3 \text{ d}^{-1}$ and $25 \cdot 10^5 \text{ m}^3 \text{ d}^{-1}$ for R1 and R3, respectively. Such accumulation may potentially lead to overestimation of simulated water levels and discharge (Table 4.4a).

One possible cause for the increase in water volume storage may be the absence of feedback loops between hydrodynamics and hydrology as discussed above. Another reason for the accumulation may be the absence of small 1D channels, hampering the drainage of floodplains, as well as the coarse 2D elevation information obstructing important floodplain-channel flows (Neal et al., 2012a).

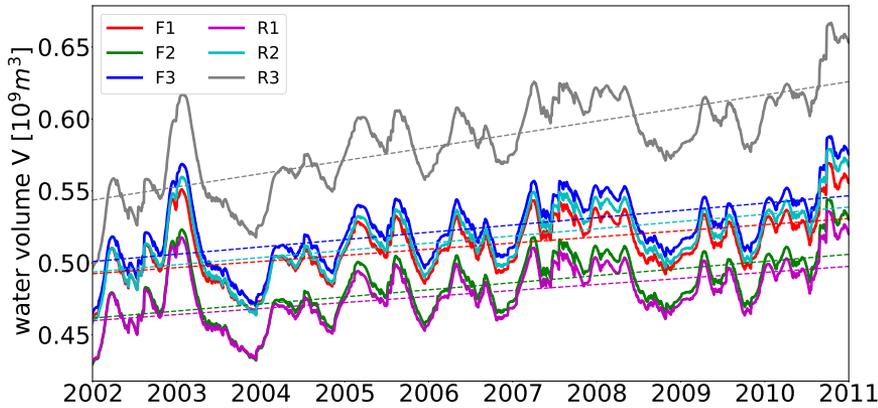


Figure 4.5: Water volume present on the 2D grid for all set-ups accumulated per time step over entire model domain including a linearly fitted trendline.

We conducted a first-order assessment whether estimates of spatial-temporal averages of both potential evaporation and groundwater infiltration rates could absorb the accumulated volumes (b). By multiplying the average of potential evaporation as used by PCR forcing (0.0019 m d^{-1}) with inundation area (Table 4.4b), a potential volumetric evaporation between $1.37 \cdot 10^{11} \text{ m}^3$ (for R1) and $2.06 \cdot 10^{11} \text{ m}^3$ (for R3) over the entire model period is obtained, both greatly exceeding the total accumulated water volume for any discretization. The average infiltration capacity expressed as the k_{sat} value is with 0.15 m d^{-1} even higher than potential evaporation. Since both values exceed the actually accumulated water volume, we cannot exclude the absence of hydrologic processes as cause for the aggregation.

A clear answer whether this or hindered dewatering of floodplains as reported by Neal et al. (2012a) was the main driver can, however, not be unambiguously be provided.

4.3.3 Water level

Results show (Figure 4.6) that the chosen spatial resolution impacts simulated water levels at all stations, regardless the application of flexible meshes or regular grids. Even though there are locally marked deviations, coarser spatial resolutions result in higher water levels at most of the stations. The main trend in higher water levels with coarser resolution is consistent with larger flood volumes during inundation.

The results can be explained by coarser spatial resolutions reducing connectivity as well as representation of both floodplain flow and floodplain-channel processes which may result in locally higher water levels (Altenau et al., 2017; Horritt et al., 2006; Neal et al., 2012a). Besides, coarser spatial resolutions reduce dynamics and, especially at upstream stations, do not capture all inundation events. Even though there is no linear relation between coarsening of grid size and change in surface elevation at all six measuring points, elevation values at observation stations tend to increase with spatial resolution, potentially limiting the magnitude of water level fluctuations (Table 4.3). This decrease of elevation with spatial coarsening is due to spatial averaging of input elevation values (Savage et al., 2016b).

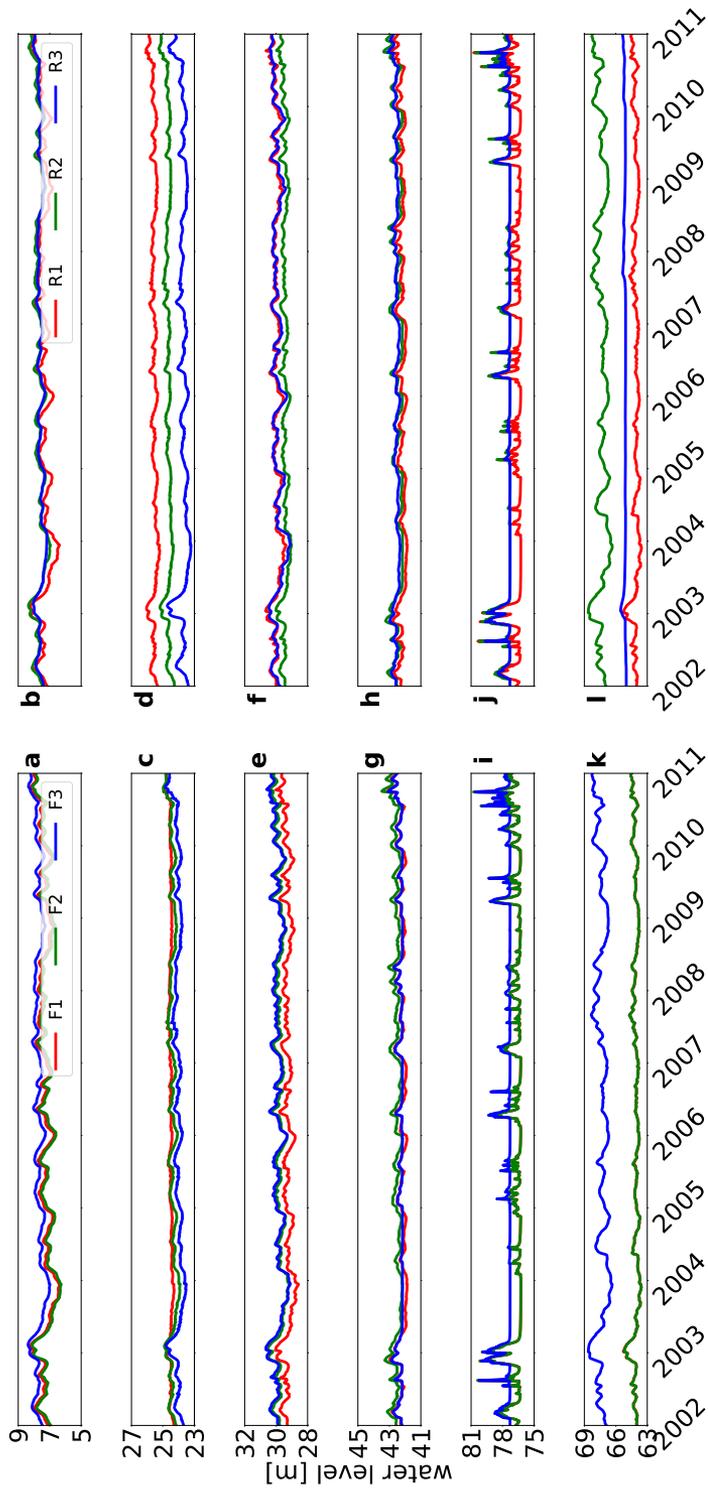


Figure 4.6: Simulated water levels at observation stations *Loc1*, *Loc2a*, *Loc2b*, *Loc3a*, *Loc3b*, and *Loc3c* (from top to bottom) throughout the basin, separately plotted for all flexible meshes (left) and regular grids (right). See figure 4.1 for the exact position of those locations.

| | Loc1 | Loc2a | Loc2b | Loc3a | Loc3b | Loc3c |
|----|------|-------|-------|-------|-------|-------|
| F1 | 3.57 | 19.65 | 25.28 | 36.86 | 76.21 | 62.85 |
| F2 | 3.57 | 19.65 | 24.16 | 36.86 | 76.21 | 62.85 |
| F3 | 1.07 | 18.48 | 21.16 | 42.16 | 77.28 | 60.49 |
| R1 | 3.57 | 19.65 | 25.28 | 36.86 | 76.21 | 62.85 |
| R2 | 1.07 | 18.48 | 24.16 | 42.16 | 77.28 | 60.49 |
| R3 | 7.17 | 18.48 | 18.18 | 38.64 | 77.28 | 65.00 |

Table 4.3: Surface elevation in meters above ordnance level at observation stations.

Also, studies report a non-linear connection between model results and bulk flow effects at coarser resolution as well as varying feedback loops at different resolution due to surrounding cells (Fewtrell et al., 2008; Hardy et al., 1999). An unambiguous answer which is the driving factor is unfortunately not possible due to the system's complexity.

Water levels of flexible meshes and regular grid per station generally compare well. The closest fit between flexible and regular grids could be found for the upstream stations Loc3b and Loc3c as well as the most downstream station Loc1 where the F1, F2, and R1 as well as F3 and R2, respectively, exhibit near-identical results.

4.3.4 Inundation extent

We benchmarked inundation extent at the end of the simulations of all discretizations with R1 as reference map and computed contingency maps for visualization of the hit rate H, false alarm ratio F, and critical success rate C (Figure 4.7).

It should be noted that, similar as for the discharge results, the absence of dikes and other man-made structures in our discretizations results in overestimations of inundation extent and thus we desisted from performing an actual validation against observed inundation extent. Besides, other factors potentially affecting inundation extent such as urban areas could not be included due to lacking data. Also, including sub-grid elevation data and spatially varying surface roughness values may have positively influenced the inundation extent obtained.

We find that not only the overall spatial resolution, but also the gridding approach greatly impacts the agreement of inundation extent at the finest level (Table 4.4b): although F1 has the same finest grid size as R1, they agree only to 74%. This suggests that accuracy of flexible meshes is reduced in those areas where a coarser spatial resolution is employed, which is mostly in upstream areas. This underlines the above-made suggestion that the coarsest grid size has a marked impact on simulated inundation extent. Besides, it seems that for a certain range of coarser discretizations it is inconsequential which cell size or gridding technique is opted for as H, F, and C are within close limits.

Coarser resolution models tend to predict larger inundation extent not only on floodplains, but also for areas farther away from the channels. This again can be related to a lack of 2D return flows with coarsened spatial resolution or missing hydrological feedback, as shown above and in previous studies. Besides, similar studies for regular grids also report an increase in inundation extent for coarser resolutions (Hardy et al., 1999) which, in turn, is linked to a reduction of contingency and representativeness (Altenau et al., 2017; Horritt and Bates, 2001a; Savage et al., 2016b). Altenau et al. (2017) also concluded that the critical success index drops for coarser spatial resolutions due to averaging of channel and floodplain properties.

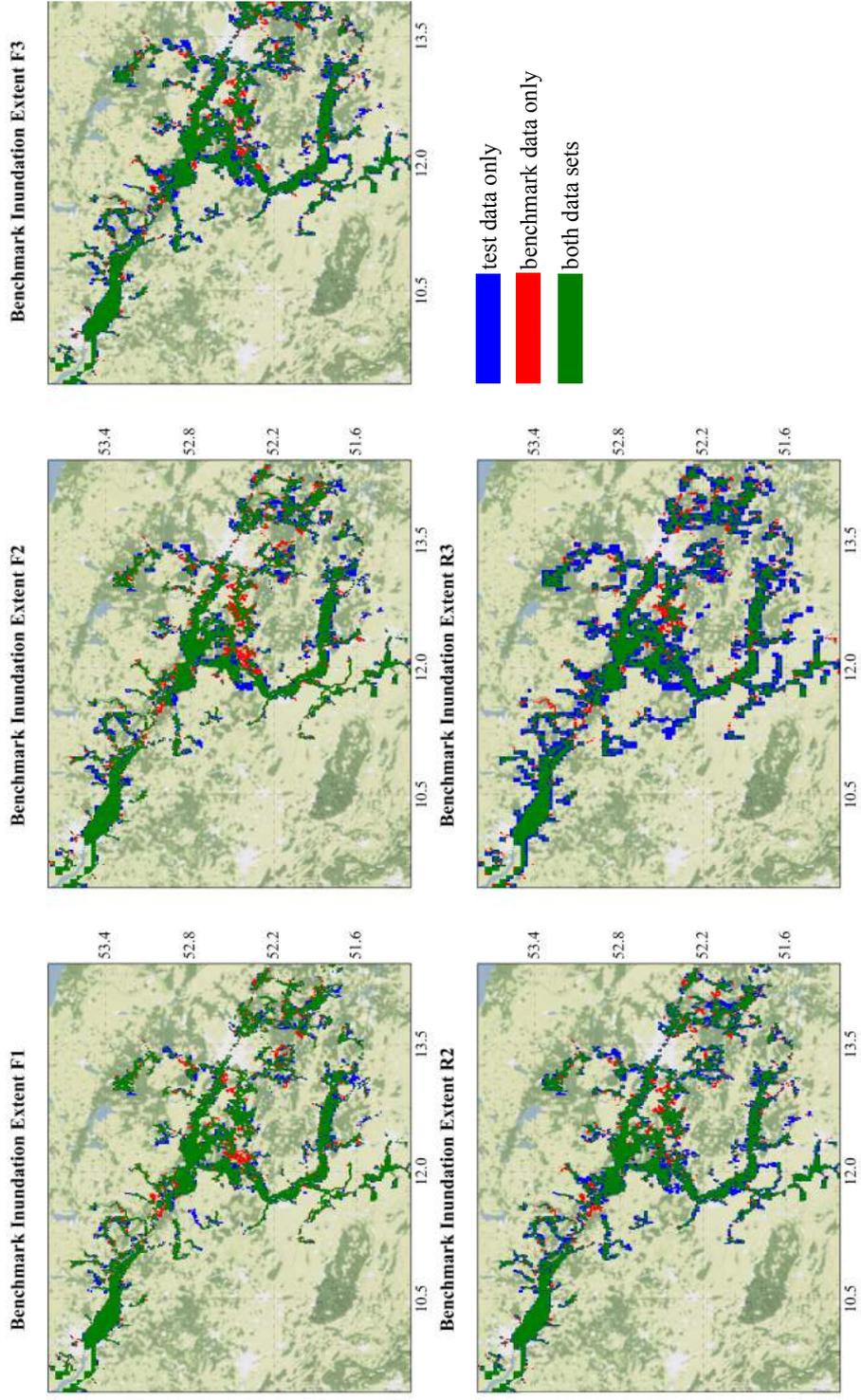


Figure 4.7: Contingency maps of inundation extent of all discretizations (test data) benchmarked against R1 (reference data)

| | F1 | F2 | F3 | R1 | R2 | R3 |
|--|------|------|------|------|------|-------|
| <i>(a) Water volume</i> | | | | | | |
| Offset | 8.04 | 9.39 | 9.65 | 7.92 | 9.53 | 17.73 |
| Slope [10^5 $m^3 d^{-1}$] | 12 | 13 | 14 | 11 | 14 | 25 |
| Increase [$10^{10} m^3$] | 0.95 | 1.00 | 1.10 | 0.85 | 1.10 | 1.60 |
| Increase [mm] | 475 | 500 | 500 | 472 | 500 | 593 |
| <i>(b) Inundation extent</i> | | | | | | |
| Inundated area [10^{10} m^2] | 2.0 | 2.0 | 2.2 | 1.8 | 2.2 | 2.7 |
| Hit rate | 0.87 | 0.84 | 0.87 | - | 0.87 | 0.86 |
| False alarm ratio | 0.17 | 0.24 | 0.28 | - | 0.29 | 0.44 |
| Critical success rate | 0.74 | 0.67 | 0.65 | - | 0.65 | 0.51 |
| <i>(c) Run time</i> | | | | | | |
| Duration [h] | 71 | 41 | 24 | 80 | 25 | 10 |
| Relative to R1 | 89% | 51% | 30% | - | 31% | 13% |

Table 4.4: Summary of benchmarking results: **a)** statistical analysis of the trend in water volume accumulation; **b)** metrical comparison of inundation extent; **c)** analysis of run time of different set-ups in relation to the fastest run R1.

We, nevertheless, should acknowledge that when other mesh generation approaches other than HAND would be applied, as for example GIS based approaches (Kumar et al., 2009), results may differ as mesh size properties are to some extent configuration dependent.

4.3.5 Run time

We find that especially for smaller resolutions the differences increase significantly which is in line with expectations and also found in comparable studies, although merely considering regular grids (Altenau et al., 2017; Savage et al., 2016b). Besides very similar performance in discharge computations, run times are almost identical for both R2 and F3. This is because only 2D cells adjacent to rivers will be inundated and thus run time does not depend on the overall number of 2D cells but is mainly governed by the number of 2D cells inundated. Even though F2 and F1 have the same finer spatial resolution, run times differ markedly with the latter having a factor 1.73 longer run time. The longest run time was, as expected, measured for R1 which is 12% longer than F1. While the difference between regular grids and flexible meshes is as expected, results show that major gains can be obtained if doubling the coarsest spatial resolution, for instance from 1600m (F1) to 3200m (F2) of the flexible mesh provided potential reduction of upstream model accuracy is acceptable.

4.4 Conclusion and Recommendations

To foster our understanding of differences between flexible meshes and regular grids as well as to better understand both advantages and shortcomings of using flexible meshes for large-

scale inundation modelling, we compared six hydrodynamic discretizations of the lower Elbe basin. To facilitate the fast generation of meshes, we used the Height Above Nearest Drainage (HAND) algorithm. Comparability between runs was ensured by employing the GLOFRIM framework, allowing for identical spatially varying and explicit forcing of hydrodynamic models with hydrologic output.

We conclude that the spatial resolution of the hydrodynamic model discretization influences model skill in simulating discharge, local water levels, and agreement between inundation maps, which complies with previous studies, although performed with different models and on other scales (cf. Altenau et al. (2017), Fewtrell et al. (2008), Hardy et al. (1999), Horritt et al. (2006), Savage et al. (2016a), and Savage et al. (2016b)). Even though the findings are configuration dependent and we test only a sample of all possible discretizations, those similarities across scales and applications are both confirmation of the robustness of our results and proof that these links are model independent and thus of more universal nature.

For discharge simulations, the finest spatial resolution in the grid controls output accuracy; that is, as long as the river and floodplains are schematized with sufficient detail, coarser cells farther away will be less influential. Furthermore, there are no significant differences between the application of a flexible or a regular mesh if comparable in resolution. This means that for discharge simulations less detailed discretizations are acceptable for areas farther away from the river system, if the finest spatial resolution is sufficiently fine to capture both channel and floodplain flow processes.

This is crucial since with coarser spatial resolution a higher fraction of overall flow is conveyed via the 2D part. To some extent adding more 1D channels could alleviate this, but on basis of our results we find the spatial resolution the larger impediment.

Since the here presented study solely employs large-scale data sets for a catchment-scale analysis, a peak-over-threshold assessment shows that such approaches should be critically examined for detailed flood hazard and risk assessments as not considering structures such as dikes and drivers like spatially varying roughness coefficients may reduce the accuracy of discharge simulations.

Unlike discharge, assessing inundation extent exhibited stronger deviations between the gridding techniques: generally, a uniform and fine spatial resolution outperforms any coarser or flexible grid. This is mostly due to the progressive coarsening of mesh size in upstream areas, leading to larger simulated inundation extent once bankfull discharge capacity is exceeded.

To better understand to which extent the application of HAND influenced the extent of simulated discharge and the models skill in simulating peak discharge situations, we recommend testing other mesh refinement approaches (see chapter 7.1.1 for recommendations).

Results suggest that applying a coarser spatial resolution enhances flow accumulation on the floodplains. Comparing the accumulated volumes with potential reduction due to evaporation or groundwater infiltration showed that these processes cannot be neglected. As most hydrodynamic models do not simulate evaporation or groundwater infiltration, accumulated water will remain there except for return flows. Consequently, future work should focus on establishing feedback processes between inundation floodplains and hydrologic processes. A first attempt to include two-directional coupling is shown in chapter 6.2.

What can be derived from these findings is that there is a threshold resolution defining the limits of meaningfulness of mesh refinement only if a certain minimum fineness of resolution is given, flow and inundation processes can be represented sufficiently well. While this was already found to be true for regular grids (Horritt and Bates, 2001a), this study illustrates similar patterns for flexible meshes.

As the relation between this resolution and model accuracy will most likely differ depending on basin and river dimensions as well as grid generation technique, we recommend further research to establish a relation between basin properties and mesh design. Such knowledge would be of invaluable use for any large-scale hydrodynamic study as essential time savings effects by grid size optimization could be achieved.

As a guideline for future applications of flexible meshes for large-scale inundation studies, we define three major aspects to consider if applying a flexible mesh for large-scale inundation studies:

- using HAND to generate large-scale flexible meshes is a fast and low-level approach for large-scale applications where more detailed topographical features can be neglected
- results for this test case underline the importance of including smaller topographic features for detailed catchment-scale flood hazard and risk assessments
- location-specific output, such as river discharge and floodplain water levels, is less sensitive to coarse-resolution flexible meshes in upstream areas
- A minimum fine spatial resolution must be met for floodplain areas to reduce volume conveyed via floodplains and facilitate return flows into channel
- domain-wide output, such as inundation extent, profits from the application of uniform fine-resolution regular grids

As this study is the first of its kind focussing on comparing flexible meshes with regular grids using methods and data for large-scale applications, the number of flexible meshes used was limited. To further increase our understanding of the confines of their applicability, we recommend a study merely focussing on the impact of mesh variations but then with a wider range of discretizations.

To conclude, we see potential in the application of flexible meshes for future “hyper resolution” large-scale inundation studies, but it also brings more responsibilities. Applicants of flexible mesh models need to put additional emphasis on the creation of the hydrodynamic discretization as it is the coarsest spatial resolution that may become the bottleneck of accuracy. While we used the HAND algorithm for fast and semi-automated mesh creations, we recommend testing other mesh generation approaches as well.

Generally, we think that mesh designs should be based on a number of considerations. For instance, for discharge simulation larger ratios between largest and smallest cell size are admissible, whereas for inundation extent computations this ratio should be minimized. Also, the context of the simulation needs to be considered: are detailed estimates required or are short run times essential? Does one need high accuracy output for the entire domain or just a small part of it? Once the user has a clear idea of the study objectives, the application of a flexible mesh can indeed serve as a computationally efficient alternative to regular grids, proving potentially useful for large-scale, operational, or ensemble modelling studies where results need to be computed in brief time.

Author contributions

JMH designed and executed the experiments as well as composed the manuscript. RvB, HCW, and MFPB supervised the work as well as provided invaluable feedback on the experimet design, results, and the manuscript.

Acknowledgements

This study was financed by the EIT Climate-KIC programme under project title “Global high-resolution database of current and future river flood hazard to support planning, adaption and re-insurance”. Furthermore, we kindly thank the German Waterway and Shipping Administration (Wasser- und Schifffahrtsverwaltung des Bundes; WSV) for providing the discharge measurements used for model validation. We also thank Edwin Sutanudjaja for support with PCR-GLOBWB as well as Herman Kernkamp and Arthur van Dam for advice on Delft3D Flexible Mesh. We also thank two anonymous authors for their critical remarks on a previous version of this article.

Chapter 5

Advancing global flood hazard simulations by improving comparability, benchmarking, and integration of global flood models

In recent years, a range of global flood models (GFM) were developed, each utilizing different process descriptions as well as validation data sets and methods. To quantify the magnitude of these differences, studies assessed the performance of GFMs on the continental and catchment level. Since the default models set-ups resulted in locally marked deviations, there is a clear need for further and especially more standardized research to not only maintain credibility, but also support the application of GFM products by end-users. Consequently, we here outline the basic requirements and challenges of a Global Flood Model Validation Framework for more standardized model validation and benchmarking in the hope of encouraging the much needed debate, research developments in this direction, and involvement of science with end-users. By means of the framework, it is possible to streamline the data sets used for input and validation as well as the validation approach itself. By subjecting GFMs to more thorough and standardized methods, we think their quality as well as acceptance will increase as a result, especially amongst end-users of their outputs. Otherwise GFMs may only serve a purely scientific purpose of continued model improvement but without practical use. Furthermore, we want to invite GFM developers to make their models more integratable which would allow for representation of more physical processes and even more detailed comparison on a model component basis. We think this is pivotal to not only improve the accuracy of model input data sets, but to focus on the core of each model, the process descriptions. Only if we know more about why GFMs deviate, are we able to improve them accordingly and develop a next generation of models, not only providing first-order estimates of flood extent but supporting the global disaster risk reduction community with more accurate and actionable information.

Based on: Hoch, J. M., Trigg, M. A. (2019), Advancing global flood hazard simulations by improving comparability, benchmarking, and integration of global flood models, *Environmental Research Letters* 14 (3), 034001.

5.1 Introduction

Economic damage and casualties due to flooding increased remarkably in recent decades. Due to a combination of factors, such as population growth, urbanisation, and a changing climate, flood risk will continue to rise world-wide (cf. Ceola et al. (2014), Winsemius et al. (2016), Ward et al. (2013), and Jongman et al. (2012)). Thus, the implementation of improved flood risk management, as well as efficient adaptation and mitigation measures are required and, more fundamentally, a better understanding of the processes driving flood events. With most riverine flood events simultaneously impacting multiple neighbouring countries and catchments (Jongman et al., 2014), declining availability of observed discharge data on the one hand, and increasing computational power on the other hand, the benefit of using global flood models (GFMs) was recognized as a key tool in tackling these challenges. Hence, hence their development and application increased rapidly in recent years (Bates et al., 2018; Ward et al., 2015).

All available GFMs are fit for the purpose of modelling global flood hazard and risk and validated to some extent during their development and dissemination. Yet, they all inherently have, depending on their governing processes and structure, distinct properties, strengths, and weaknesses. Since validation data, period, and location are usually not consistent between GFM description studies, model differences do not directly become visible while in fact they can result in locally remarkable deviations when compared with each other (Bernhofen et al., 2018; Trigg et al., 2016).

In contrast to GFMs, global hydrologic models (GHMs) are regularly compared, for instance their routing scheme (Zhao et al., 2017) or simulated runoff (Beck et al., 2017). Such model intercomparison projects are a great way to narrow the above-mentioned knowledge gap, let alone the stimulus for intensified scientific collaboration and exchange. Consequently, GFMs are behind in terms of collaborative testing as lack of more consistent and regular comparison, hampering a better understanding of the discrepancies in model outputs. This epistemic uncertainty could, we postulate, lead to problematic model equifinality as results may agree only coincidentally.

A better understanding of why and where each model may or should be used is, however, pivotal. Discerning not only a model's strengths, but also its uncertainties, limitations, and differences with respect to other models is a central pillar to put model outputs into perspective and increase their credibility. By virtue of a transparent comparison process with other models, individual model developers can see more clearly how to improve their own data sets and process representation where they may see these lacking.

Since the relative accuracy would become more tangible, the meaningfulness and applicability of each model for end-users would increase too. Various workshops aimed at bringing together researches and practitioners provide evidence that there is a growing demand for more transparency and better overview of GFMs as well as their characteristics and uncertainties (Watson, 2018; Partnership, 2016; Partnership, 2017). This is particularly important for non-expert users of model outputs who rely on a clear understanding of the appropriateness and limitations in order to use the data appropriately (Trigg et al., 2016; Ward et al., 2015).

So, what are some possible ways forward? First, to facilitate obtaining the required understanding, an easily accessible yet demanding validation and benchmarking framework could create a meaningful starting point. That this idea timely is shown by similar developments towards a framework for operational flood risk management (Alfieri et al., 2018) as well as from the above-mentioned need of end-users to get a better grasp of model properties. Second, models are in almost all cases closed systems where output is produced based on the

input provided and the subsequent model steps executed. While this works well for default model applications, it hampers the models extension and integration with new features, modules or even other models. In times of continued model integration, however, bridges with other models should be built if more (physical and non-physical) processes influencing flood hazard are required.

In the remainder of this article, we first present a range of GFMs and outline their specific properties. Second, we assess the different validation data sets, periods, and locations of these GFMs as published in peer-reviewed papers to supplement our call for streamlined validation approaches. Subsequently, we sketch a possible design of a Global Flood Model Validation Framework for model validation and benchmarking. Last, motivation and possible approaches to advance the openness and integration capability of GFMs is presented. The article is ended with conclusions, ideas on how to implement the presented ideas and recommendations for further improvement of comparability and applicability.

5.2 Current global flood models

Currently the most fully developed and openly accessible state-of-the-art GFMs are CaMa-Flood (Yamazaki et al., 2011), GLOFRIS (Winsemius et al., 2013), JRC (Dottori et al., 2016), CIMA-UNEP (Rudari et al., 2015), as well as the Fathom model (formerly SSBN; Sampson et al. (2015)) and the ECMWF model (Pappenberger et al., 2012). These models can be divided into two main categories of GFMs, depending on the flow derivation modelling steps taken (Figure 5.1).

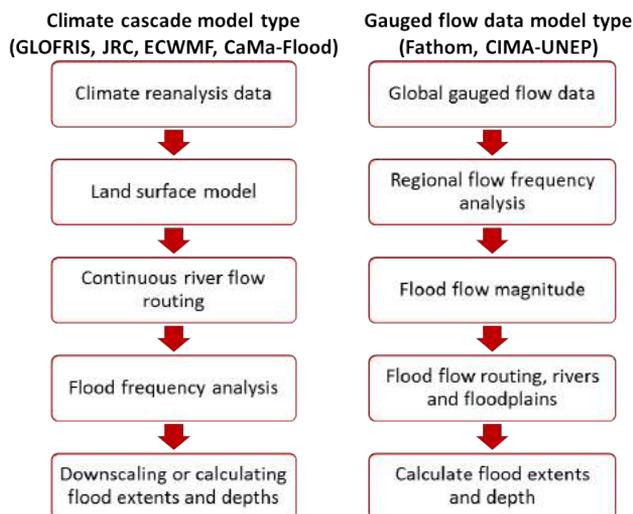


Figure 5.1: Modelling steps required for the two sub-categories of global flood models; modified from Trigg et al. (2016).

It must be noted that there is also a number of private or national CAT (“catastrophe models”) models that include global flood hazard, each also having its own properties, modelling cascades, and evaluation procedure and criteria. Obtaining information about these CAT models is, however, complicated due to the protection of Intellectual Property (IP) rights and competitive commercial advantage. The following comparison therefore represents only the

most open models and may need updating in the future if these commercial models become more transparent.

The differing operations at various model stages result in a range of modelling approaches, each one using its own input data, method of calculating floodplain inundation, and spatial resolution. For instance, GLOFRIS runs at a 30 arcmin spatial resolution before post-processing and downscaling to 1 km, whereas the Fathom model yields output directly at 90 m globally. Such discrepancy in spatial resolution is possible because the models simulate processes with different scaling potential (Bierkens, 2015). As a result, the models perform differently in these scale-dependent processes. For example, GFMs employing a land surface or hydrologic model excel in simulating processes such as open water evaporation and groundwater infiltration. Contrariwise, the routing schemes of land surface or hydrologic models (typically the kinematic wave approximation) are less sophisticated and therefore lack important discharge dynamics which can be obtained from models employing higher-order approximations of full shallow water equations.

Notwithstanding the differences, all models are applied regularly and used to inform policy-makers about flood hazard and risk. GLOFRIS, for example, is applied within the World Resources Institute (WRI) “Aqueduct Floods” tool, projecting current and future flood risk across the entire globe (*Aqueduct Global Flood Analyzer* 2019). The JRC model is applied as part of the Global Flood Awareness System GloFAS (Alfieri et al., 2013), and the Fathom model was recently used to compare flood risk with Federal Emergency Management Agency (FEMA) estimates across the continental United States (Wing et al., 2018). CIMA-UNEP was applied for estimating current and future flood risk for the Global Assessment Report (GAR) of the United Nations Office for Disaster Risk Reduction (Rudari et al., 2015; UNISDR, 2015a).

5.3 META study: validation of GFMs

Before employing a GFM for flood hazard and risk assessments, ideally it should undergo thorough testing and validation. Due to the wide range of available observation data sets and depending on the model period as well as study area opted for, all GFMs may obtain good validation results, yet without providing any insight into performance relative to other GFMs. To get a grasp of the differences in model validation, we here detail the various data sets, periods, and locations used for the above-mentioned models.

As table 5.1 shows, the spread in validation (or benchmarking) data sets used is tremendous. Partially, this can be explained by the particular moment of model publication and the availability of data sets at that time. It also shows that most GFMs are validated against inundation extent, but only a few compare simulated discharge and water surface elevation with observations, although these aspects are important for flood risk management as well. Besides, the river basins used for model validation differ widely between studies as does the number of scientific reports documenting the model development over time.

Trigg et al. (2016) showed that the GFMs listed in chapter 5.2 agree only for around a third of simulated flood extent in Africa. Since there can only be one actual realization of inundation at a time, this finding shows that a successful individual validation of models without comparison may be misleading with respect to the accuracy of the resulting inundation maps. Comparable findings were made by Bernhofen et al. (2018).

Together with the lack of congruency in model validation, the results from Trigg et al. (2016) bolster the above-made claim that, to really get an idea of why model results deviate and to eventually learn from each other, more standardized validation and benchmarking procedures could be useful.

| GFM | Study | River basin | Period | Discharge validation | Flood extent validation | WSL validation |
|------------------------|--|---|--|---|---|-------------------------------|
| CaMa-Flood | Yamazaki et al. (2011) Yamazaki et al. (2012b) Yamazaki et al. (2013) Yamazaki et al. (2014b) | a.o.: Amazon, Congo, Brahmaputra, Rhine, Ob Amazon Global maps used Mekong | <i>Varies per basin</i> 2003-2005 1991-2000 2001-2005 | GRDC ¹ ANEEL ² GRDC Inomata and Fukami (2008); MRC ³ | SAR (Hess, 2003); GIEMS (Prigent et al., 2007) GIEMS (Prigent et al., 2007) — — | — Envisat RA-2 — MRC |
| GLOFRIS | Winsemius et al. (2013) | Ganges, Brahmaputra a.o.: Severn, Thames, Elbe, | 1961-1990 | — | DFO ⁴ | — |
| JRC | Dottori et al. (2016) | Niger, Ganges, Mekong a.o.: Rhine, Columbia, Thames, Colorado, Yukon | 2000-2013 | Local observations, based on Hirpa et al. (2016) a.o.: GRDC, RivDis ⁷ | DFO, UNOSAT ⁵ | — |
| CIMA-UNEP ⁶ | Rudari et al. (2015) | Columbia, Germany and Thailand Bow River, North Saskatchewan, Red Deer; Severn, Thames | — | — | DFO | — |
| Fathom | Sampson et al. (2015) ⁸ | CONUS | <i>Comparing return periods</i> | — | Alberta State Government; JRC model | — |
| ECMWF | Wing et al. (2017) ⁸ Pappenberger et al. (2012) ⁸ | Various catchments <i>world-wide</i> | <i>Comparing return periods</i> <i>Comparing return periods</i> | — — | FEMA ⁹ , USGS ¹⁰ UNISDR flood maps | — — |

- ¹ GRDC, Global Runoff Data Center
² ANEEL, Agencia Nacional de Energia Electrica
³ MRC, Mekong River Commission
⁴ DFO, Dartmouth Flood Observatory
⁵ UNOSAT, United Nations Institute For Training And Research Operational Satellite Applications Programme
⁶ Development and validation of the CIMA-UNEP model was not published in peer-reviewed scientific journals
⁷ RivDis, Global River Discharge Database
⁸ these studies only performed benchmarks with inundation maps from other inundation models or databases for given return periods
⁹ FEMA, Federal Emergency Management Agency
¹⁰ USGS, United States Geological Survey

Table 5.1: Overview of validation data sets of discharge, inundation extent, and water surface level (WSL) as used in various scientific studies of GFM development.

5.4 Establishing a Global Flood Model Validation Framework

To facilitate standardized validation and benchmarking of models and their results, a framework facilitating these steps is needed. A first step towards more systematic benchmarking was set with the GLOFRIM framework (see Hoch et al. (2017b) and 3) which allows for forcing different hydrodynamic models with identical hydrologic output. Yet, it can only mark a first proof-of-concept since much more functionality would eventually be needed. Some key tasks of a Global Flood Model Validation Framework would be, amongst others, to provide a front-end where users can upload model results as well as a back-end to not only execute validation and benchmarking autonomously but also store validation and benchmarking results (Figure 5.2). Besides, the framework should provide input data sets to be used for each GFM run.

In its proposed form, the framework would be designed to only detect differences in simulated flood hazard. Since all GFMs employ different ways of how to determine risk by accounting for exposure and vulnerability, these aspects should be compared at a later stage as well. We here, however, focus on the physical modelling side of flood risk to keep the scope of the study and proposed framework manageable. Moreover, many end-users such as insurance companies do have their own exposure and vulnerability maps and rely on hazard estimates for risk assessments.

By means of the framework, it would not only be possible to provide standardized input and validation data, but also to clearly define model boundary conditions. Using identical data will improve the comparability of model validation as this is currently done independently, using different validation data products, time periods, and study areas as shown in chapter 5.3 and table 5.1.

5.4.1 Testing elements

We think it would be essential to test the models for the specific primary aspects listed below, yet this may be extended or altered if needed at any stage:

1. *Inundation extent*. A key output needed, for instance, by re-insurers to define flood-prone areas and determine premiums for portfolio exposure that intersects with the flood extent.
2. *Inundation depth*. Model output required by many risk assessments to assess potential damage via a depth-damage curve.
3. *Discharge hydrograph*. This is the fundamental driver of the out-of-bank flood processes.

By subjecting the GFMs to a thorough comparison and streamlining their input boundary conditions, the impact of the following secondary model aspects can also be tested:

1. *Forcing/Input data*. Assessing its impact is paramount to understand to which extent GFM accuracy is defined by model design or input/forcing data, something not covered by the study of Trigg et al. (2016).
2. *Regionality*. Here referred to as a models ability to perform in certain regions, differing in their meteorological, geographical, and other properties. Also, this could include an assessment whether GFMs perform well only for large rivers and where the threshold lies in the accurate representation of inundations along smaller reaches.

5.4.2 Testing challenges

Before establishing the framework, several decisions have to be made and existing challenges addressed, as Alfieri et al. (2018) also pointed out. These decisions require but are not limited to the following list:

1. *Test location.* First, it should be clear which river basins are ought to be used. As shown in table 5.1, most major river basins were already used for validation and thus it would be sensible to use one of them. Classic examples are the Amazon and Ganges-Brahmaputra basin as they both represent large low-lying floodplain areas where inundations occur regularly. The former is an indicator of performance in simulating large river flood extents while the latter is particularly relevant from a flood risk perspective due to a large population exposure. To be able to test for regionality, the chosen basins should also differ in their meteorological, geographical, and hydrological properties. Besides, only reasonably large catchments should be used to ensure that all models can sufficiently represent the processes despite differences in spatial resolution. As models and observations improve in resolution, the testing catchment testing scale can be adjusted appropriately.
2. *Forcing data.* Despite most model forcing data being openly accessible, a clear decision has to be made which dataset shall be used. For models based on meteorology (eg. GLOFRIS or the JRC and ECMWF models), recent global forcing data such as ERA-Interim or ERA5 should be provided (Berrisford et al., 2011; ECMWF, 2019). In case derivatives such as flood wave hydrographs are required (eg. Fathom, CIMA-UNEP), pre-processed data should be made available. In all cases, the data must be downloadable via the front-end of the framework.
3. *Downstream boundaries.* Even though not all GFMs can accommodate dynamic sea levels as downstream boundary condition, we recommend that in an initial approach they should be used if possible to facilitate comparison across default models settings. For advanced comparisons, the effect of changing downstream boundaries can be studied as well by de-activating them or, analogously, account for them once model development allows for it.

More challenging, validation data must be provided which meets the demands for state-of-the-art flood hazard modelling. State-of-the-art in this context also means that all validation data sets used must explicitly address possible uncertainties in observations. Hence, these additional aspects should be considered:

1. *Discharge data.* Required to validate the models skill to simulate discharge dynamics. Depending on the chosen test locations, different sources may be available, either global data sets or from local authorities. In case of the Amazon, for example, discharge data can be retrieved from ORE-HYBAM (ORE-HYBAM, 2018). One of the few global data bases of observed discharge is maintained by the Global Runoff Data Centre, currently containing data for more than 9,500 stations. To provide robust validation results, sufficiently long time series must be available. For those models simulating specific return periods only, corresponding discharge values should be derived from observations.
2. *Inundation maps.* Data ranging across various locations world-wide must be available, preferably open access remotely sensed satellite products to maintain global comparability. Since image quality may be hampered by cloud cover (Bernhofen et al., 2018), this step may require some pre-processing. Alternatively, maps from the Dartmouth

Flood Observatory (<http://floodobservatory.colorado.edu/>), the GIEMS data set (Papa et al., 2010; Prigent et al., 2007) or from Tellman et al. (2017) can be used. For those models simulating inundation extent for specific return periods, inundation maps for actual events with corresponding return periods should be used as much as possible for validation, possibly building upon recent methods (Giustarini et al., 2015; Huang et al., 2014).

3. *Water levels.* To guarantee a globally uniform approach, satellite products should be used, for example ICESAT, ICESAT 2, ENVISAT or SWOT once available. The locations used for validation and benchmarking should be chosen such that potential vertical inaccuracies are limited. Again, the data must be carefully selected and pre-processed, for example to remove measurements affected by land or vegetation signals, to streamline the entire validation and comparison process.

In a nutshell, the proposed frameworks objectives are threefold: (1) provide forcing data, (2) validate and benchmark model results, and (3) store reference model output per GFM (Figure 5.2). Once the user performed simulation runs with the provided forcing data, results can be uploaded via a front-end to the frameworks back-end. Here, both validation and benchmarking will be performed. For the validation, we suggest the following metrics: (i) for inundation extent, the hit ratio H , the false alarm ratio F , and the critical success index C ; (ii) for discharge, the Kling–Gupta efficiency KGE (Gupta et al., 2009), and its components; (iii) for surface water elevation, $RMSE$. Since these objective functions are only a recommendation, a definite choice should only be made after both developers and end-users agreed on common standards meeting their expectations and needs. This requires involvement of potential end-users in the development of the framework.

To perform the necessary operations, employing the increasing power of cloud computing could be a viable option. For benchmarking purposes, the model results will be stored in cloud-optimized format (for example cloud-optimized GeoTIFF; COG) and version-controlled according to the version number of tested GFM in a reference data repository, hence containing the most recent outputs of GFMs and allowing for tracking the impact of model developments on output. The reference observation data sets will then be used to apply the same objective functions. Once all steps are successfully executed, the resulting validation and benchmark statistics will be made available to the user via the front-end again.

The framework and data could be hosted by a neutral institution or other body, for instance within the Global Flood Partnership (GFP) which already collected first experiences with a common tool for operational flood risk management (Alfieri et al., 2018). Alternatively, such a framework could be hosted under the umbrella of the upcoming "Global Risk Assessment Framework (GRAF)" which aims at implementing a range of models and with a particularly end-user orientation (UNISDR, 2018).

We are aware that setting up such a framework requires both financial and time resources. Yet, we believe that once validation and benchmarking of GFMs is streamlined, they will benefit by reducing uncertainty associated with model output and its application. Using centrally provided data would also enhance the reproducibility of model output, as work flows and data use would become more transparent. We are confident the efforts made will eventually pay off as model output uncertainty will be reduced while scientific discourse will be improved, leading to better informed decisions and reduced economic damage and casualties.

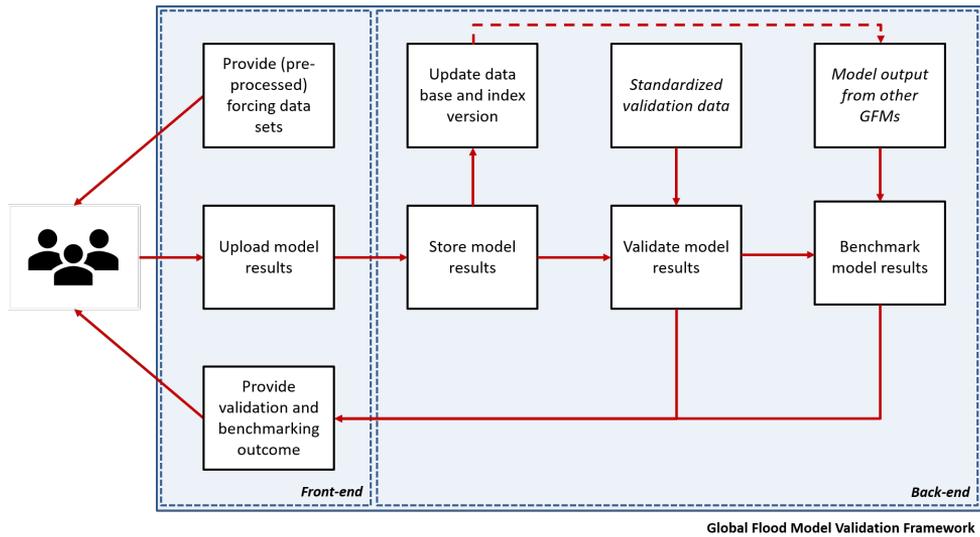


Figure 5.2: Conceptual design of a (cloud-based) Global Flood Model Validation Framework for model validation and benchmarking as well as maintaining a reference data base.

5.5 Opening the black box of model code

With the scientific funding bodies increasingly requiring research to be openly available, most (unfortunately not all) GFMs can be downloaded freely, advancing the usability and impact of the models. However, even with open code and model output availability, most models follow a black box modelling approach of reading input data, executing a prescribed and model dependant set of processes, and thereafter providing output data (Figure 5.1). Such approaches, nonetheless, pose a major limitation to making GFMs more integratable, intuitive, and interactive due to the lack of process accessibility. However, we consider process integration as key to better comparability as well as future improvements, and thus think that global flood hazard simulations can greatly benefit from opening the black box.

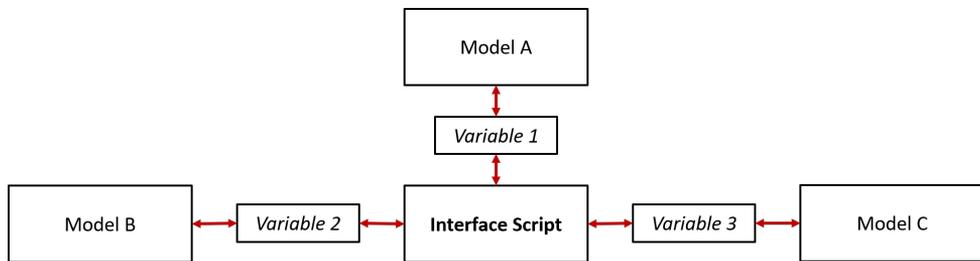


Figure 5.3: Schematic of coupling models as well as exchanging model information via BMI and a central interface script

Admittedly, the integration of different models is not rocket science and was already achieved as extensively discussed in chapter 1.3.

To facilitate interactive and intuitive model coupling as well as to avoid “integronster”, ie. models whose combined code is hard to disentangle and uncertainties are hard to trace

(Voinov and Shugart, 2013), the Basic Model Interface (BMI; Peckham et al. (2013)) provides a powerful and flexible tool to exchange model information via an (user-defined) interface script (Figure 5.3) without the need of integrating actual model code into one overarching model.

The different models can for example be hydrologic or hydrodynamic models and exchange variables such as runoff or inundation depth, respectively. Yet, also other models could be linked up such as coastal or crop growth models or even non-physical models such as agent-based models.

Within the context of model benchmarking, implementing the BMI functionalities into GFM may facilitate forcing them with identical data and, in turn, more standardized validation and benchmarking. The applicability of the BMI concept was shown by applying the GLOFRIM framework to benchmark different hydrodynamic models (Hoch et al., 2017b) as well as different schematizations of the same hydrodynamic model (Hoch et al., 2018a).

Unfortunately, none of the above mentioned GFM currently contains any BMI functionality (or anything like it, for example OpenMI (see chapter 1.3.2)). Since the implementation of a BMI is non-invasive, we think more efforts should be directed towards advancing the accessibility of model processes and variables. In the long term, this would, besides supporting model validation and benchmarking, allow for a plug-and-play design where applicants can create their own GFM depending on their study specific needs and would also facilitate more efficient modelling efforts.

Conveniently, the proposed framework can help in identifying which components of which models excel. For example, if benchmarking results indicate that Model A may profit from more physical groundwater modelling, such a module from another Model B could be added and forced with variables from Model A, for instance surface water depth. Besides, output from other non-GFM models could be employed such as sea levels from a tide and surge model (Model C) which would even further increase the number physical processes representable (Figure 5.3).

We are aware that this would not only require opening the black box, but possibly also developers minds. Besides, possible issues with IP rights may have to be solved first. Still, we are confident that such inter-active model functionalities can become a core element of advancing model validation and benchmarking across scales and processes, as they may result in new and promising research possibilities.

5.6 Conclusion and recommendations

Many GFM found application in policy tools or operational systems but are still not well compared and consequently differences are not well understood. However, we think that this is pivotal for increased acceptance of GFM by end-users and thus the existing different approaches to simulate inundation data require a more thorough and streamlined validation and benchmarking procedure.

GFM were validated with a wide range of data sets for various time periods in numerous river basins all over the World. While the data used for validation is to some extent related to the date of model publication and the data availability at that time, the fact that all models are validated successfully for non-identical settings may lead to the misleading conclusion that all model perform equally well. Additionally, it does not support a clear conclusion as to why results differ between GFM.

Due to the range in validation approaches, we see great potential for models to improve by comparing with and learning from others. Therefore, we sketch a Global Flood Model Validat-

tion Framework serving multiple purposes. First, it provides identical model forcing. Second, it validates simulated discharge, inundation extent, and surface water elevations. Third, it serves as a repository and version-control of GFM output and thus also allows for benchmarking output from different models and model versions. By establishing such a framework, we can ensure that, despite all independent model development trajectories, the same data and criteria are applied for assessing model output.

Since the framework can only streamline external factors, there will probably still be deviations in model results due to differences in internal model structure, processes, and parameterization. This is perfectly acceptable as the proposed framework is not meant to converge all GFMs, but rather as a testing and learning environment for researchers to improve usability and acceptance by end-users.

By means of the framework, insights could be provided into the upsides and downsides of each tested model design. If the framework is applied for more river basins and hydrologic conditions, it would furthermore be possible to identify where and under which circumstances certain models perform best. Such knowledge can, in turn, be beneficial when it comes to communicating model strengths and limitations to policy and decision makers and provides them with a tool to identify which GFM may be most appropriate for a project or application in a specific region. Besides, the insights gained may be used to better point towards model shortcomings that could benefit from adopting methods implemented in better performing models.

For models to profit further from such insights, it could be necessary to open up the default black box of model processes. While a standard comparison framework may be sufficient for default applications of the models, implementing functions to allow for accessing and exchanging model variables could facilitate integrating components from other GFMs to improve model performance. Moving away from a black box approach may stimulate the benchmarking and comparison of GFMs as assessments could be performed at an unprecedented level of detail and flexibility, allowing ranking of the importance of different elements of GFMs. For example, the same spatially varying hydrologic output could be applied to all models, reducing the number of factors influencing model deviations. Vice versa, it could be possible to provide a clearer picture on how the routines calculating hydrologic forcing may differ by applying one routing scheme to all models designs. Ultimately, the GFMs would move closer together without abandoning their specific properties, and uncertainties surrounding flood hazard outputs could be reduced greatly.

We are aware that the presented framework and the required openness about model performance may discourage contributions from private CAT models. Nevertheless, we are convinced that an independent validation and benchmarking framework can be beneficial for the private sector too, as (a) data providers could present their results from commercial CAT models in a broader context, and (b) data users could first analyse which products fits their needs best before purchasing a flood product. We hope that thorough benchmarking of inundation maps becomes the new normal, eventually requiring vendors to improve their services and consequently resulting in better risk estimates for end-users. From a technical point of view implementing the CAT models into the proposed framework would be relatively straightforward as they essentially employ the same technology and input data types as the open scientific models. A major requirement for those model developers would of course be that outcomes are not necessarily made publicly available.

While the here proposed Global Flood Model Validation Framework focusses on differences in model design and associated differences in model output, more steps should be taken to improve the comparability and consequential uptake of GFMs. First, the nomen-

clature of model variables and components differs greatly between models, hampering the traceability of model work flows. By using more standardized terminology, for example the standard names proposed by the Community Surface Dynamics Modelling System (https://csdms.colorado.edu/wiki/CSDMS_Standard_Names), comparing GFMs would become easier, particularly for non-expert users. And second, comparability, inter-operability, and usability of model outputs would be greatly supported by agreeing on clear standards for files, for instance based on the guidelines of the Open Geospatial Consortium (<http://www.opengeospatial.org/>). Third, it is necessary that all GFMs (as for models in general) provide easily comprehensible description of how they work and what their outputs represent.

To establish a full comparison between GFMs, exposure and vulnerability data should be compared as well. Since these data layers are not based on a modelling cascades, the proposed Global Flood Model validation and benchmarking framework may not be the right means. Nevertheless, we think that further investigation is needed to better understand to which extent differences in simulated risk assessment outputs are dependent on hazard, exposure or vulnerability. Eventually, the three pillars of risk could be compared altogether. Such an extensive intercomparison project would help greatly to advance the current state of GFMs and to identify new research possibilities.

More efforts should be taken to advance our understanding of GFMs and their differences. With our proposed validation and benchmarking framework together with greater model accessibility, we see great potential for future model developments as well as an increased number of GFM applications and hope that model comparison will play a more significant role in future flood hazard modelling studies.

Author contributions

JMH was responsible for the initiative writing this paper and set out the main narrative. MT contributed strongly by adding relevant data and shaping the design of the framework.

Acknowledgements

JMH was funded by EIT Climate-KIC under the project “Global high-resolution database of current and future river flood hazard to support planning, adaption and re-insurance”. We thank two anonymous reviewers for the discussion of an earlier version of the manuscript.

Chapter 6

On-going developments of GLOFRIM

GLOFRIM is a dynamic scientific tool, always subject to changes in functionalities and code, driven by new insights, possible applications as well as changes in technology. As such, the first version presented in chapter 3 only depicts a snapshot of what GLOFRIM consists of and can be applied to. While it was successfully applied for benchmarking hydrodynamic models (chapters 3, 4), many developments have happened since. Neglecting smaller changes to the code, we here present two major developments and applications of GLOFRIM which are still on-going.

The first major update to GLOFRIM 2.0 is presented in section 6.1 which was motivated by the need to include a computationally efficient large-scale hydrodynamic model as the explicit 1D/2D computations of Delft3D Flexible Mesh and LISFLOOD-FP as provided in GLOFRIM 1.0 are often not needed for the entire basin but only specific regions. Also, the increased range of coupling possibilities demanded improvements in user-friendliness and code structure. We thus included the global routing model CaMa-Flood and tested the added value for simulating discharge and inundation extent in two case studies in the Amazon and the Ganges-Brahmaputra, respectively.

In GLOFRIM 1.0, only one-directional coupling was possible. However, this does not allow for dynamic feedback loops between hydrology and hydrodynamics. Findings in previous studies (for instance in chapter 3, 4) indicate that including such loops could be beneficial as groundwater infiltration and open water evaporation over inundated floodplains could be modelled by the hydrologic model. Particularly in (semi-)arid basins accounting for these processes is essential. Therefore, section 6.2 presents a two-directional coupling with feedback loops between PCR-GLOBWB, the groundwater model MODFLOW, and the hydrodynamic model LISFLOOD-FP as well as preliminary results for two case studies.

6.1 GLOFRIM 2.0: extending the modelling cascade

Based on: Hoch, J. M., Eilander, D., Ikeuchi, H., Baart, F., Winsemius, H. C. (2019), Integrating large-scale hydrology and hydrodynamics for nested flood hazard modelling from the mountains to the coast, *Natural Hazards & Earth System Sciences Discussions*.

Depending on model structure and workflow, different hydrodynamic and hydrologic models do not only have specific advantages and shortcomings, but will also yield different inundation extents (Bernhofen et al., 2018; Trigg et al., 2016). For instance, some models may require rainfall and evaporation fields as meteorological forcing, while others only make use of synthesized flood waves with a certain return period. Some resolve hydrodynamics in two dimensions, others only in one. Also, there are remarkable differences between the spatial resolutions, affecting both the range of physical processes to be simulated and the applicability of model output maps (Beven et al., 2015; Bierkens et al., 2015).

Generally, flood models focus on simulating fluvial inundations, following a pre-defined set of model steps. However, this falls short in representing the complexity of a multifaceted hazard like floods which may be caused by several drivers at the same time. To be able to really understand, predict, and prepare for current and future flood hazard, models should integrate across physical processes which are currently simulated by separate models. In a recent study, Hoch et al. (2017a) showed that coupling hydrologic processes with more advanced hydrodynamic processes improves both representation of inundation along reaches as well as the simulation of flood wave propagation. Together with the findings made by Zhao et al. (2017) this shows that an advanced routing scheme is a central driver missing in most hydrologic models.

Additionally, different physical processes may be governing at different spatial scales, for instance channel routing for large-scale routing or explicit floodplain flow for local assessments. Modelling set-ups must thus be able to reflect the importance of various spatially-varying flood hazard by integrating across spatial scales. Answering the question how much complexity is needed can have benefits in avoiding over-fitting of the problem (Neal et al., 2012b). For instance, simulating fine-scale floodplain dynamics is disproportionate for regions where flood wave propagation is the main physical process to consider while it is very much needed if inundation patterns in delta areas are simulated.

Current models, however, are tailor-made for certain applications and excel at, for instance, their representation of hydrologic processes, computationally efficient routing, or output maps at fine spatial resolution. Depending on the envisaged application, modelling set-ups, however, must thus be able to reflect the importance of various flood triggers by integrating across both physical processes and spatial scales.

For simulating and investigating the various dependencies between physical processes and hazards across spatial scales without adding just another new model, flexible computational frameworks are viable means as they can be designed depending on envisaged application. By providing flexibility in setting up various model couplings, the models required for a certain application can be combined to create a tailor-made coupled model.

To further enhance process and scale integration, we developed GLOFRIM 2.0. Compared to previous versions, additional models are now available. First, the global river routing model CaMa-Flood (Yamazaki et al., 2011) was added to improve runoff routing over large areas. And second, we added the wflow hydrologic modelling platform (Schellekens et al., 2018) to provide more options for hydrologic simulations at various spatial resolutions. Be-

sides, we entirely re-structured the code and re-designed GLOFRIM execution for improved usability, modularity, and flexibility.

We envisage two applications as core of the new GLOFRIM 2.0 framework: A) fast routing of runoff over large domains and B) detailed local inundation modelling for smaller nested areas such as river deltas. Besides, GLOFRIM can also be applied for benchmarking hydrologic and hydrodynamic models.

To what extent applying different model coupling set-ups is beneficial for large-scale discharge and inundation simulations is hence the overarching research question of this article.

By linking hydrologic, routing, and hydrodynamic models, we can establish a model cascade which can simulate a range of flood-triggering processes from the mountains to the coast. As such, GLOFRIM 2.0 can be a key tool for future more holistic modelling studies researching the effect of the interplay of meteorology and hydrology, river routing, and flood-plain dynamics on flood hazard and risk. Since GLOFRIM 2.0 has a modular design, other models can easily be added to even extend the range of simulated flood hazard drivers, for example coastal and oceanic processes.

In the remainder of this section, we will first describe what is new in GLOFRIM 2.0. The benefit of applying GLOFRIM 2.0 is then tested in the Ganges-Brahmaputra basin where two different coupling designs are benchmarked. We conclude with recommendations and an outlook for future applications in integrated flood hazard modelling and assessment.

6.1.1 What is new in GLOFRIM 2.0?

GLOFRIM 2.0 contains two new models, allowing for more coupling realizations than the previous GLOFRIM version. Besides, applying GLOFRIM was simplified. Hereafter, these novelties are briefly outlined.

CaMa-Flood

CaMa-Flood (CMF; Yamazaki et al. (2011)) simulates the floodplain hydrodynamics of continental-scale rivers globally. Since it solves the 1D local inertial equation (Yamazaki et al., 2013; Bates et al., 2010) and only changes in water storage are prognosticated, simulations are computationally efficient. Another advantage is that CMF is a global model and therefore model data exists for the entire terrestrial surface, reducing the need to manually set up the model. It is noteworthy that CMF can also be run with other input data sets as long as they fulfill the model-specific requirements.

The river network of CMF is derived from 18 arcsec (approximately 500 m at the Equator) elevation data using the FLOW method (Flexible Location of Waterways; Yamazaki et al. (2009)). It also containing river bifurcation which improves inundation representation in delta areas (Ikeuchi et al., 2015; Mateo et al., 2017; Yamazaki et al., 2014b). River channel width is based on the Global Width Database for Large Rivers (GWD-LR; Yamazaki et al. (2014a)). The model domain is discretized by unit catchments corresponding with each model calculation grid. Model output is provided at 0.25 degree spatial resolution (approximately 25 km at the Equator), but inundation depth can be downscaled to a resolution of 18 arcsec in a post-processing step (see supplementary material of Ikeuchi et al. (2015)).

For additional information about CMF, we refer to the above-mentioned articles as well as the to the manual (http://hydro.iis.u-tokyo.ac.jp/~yamadai/cama-flood/Manual_CaMa-Flood_v362.pdf). Download of the default CMF version (v3.6.2) and further information is possible at <http://hydro.iis.u-tokyo.ac.jp/~yamadai/cama-flood/>. A GLOFRIM compatible version of CMF model can currently only be obtained by contacting the authors.

wflow

WFL is a distributed hydrological modelling platform and part of Deltares OpenStreams project. Within WFL, a range of hydrologic models are available, for example HBV (Lindström et al., 1997), SBM (based on the TOPOG SBM model; Silberstein et al. (2007)), WR3A (van Dijk et al., 2013), and since recently PCR-GLOBWB (PCR; Sutanudjaja et al. (2018)). We tested GLOFRIM so far, however, only for the HBV model.

WFL was applied for different study objectives, such as assimilation of streamflow and soil moisture data (Lopez et al., 2016) or model intercomparison with other hydrologic models (de Boer-Euser et al., 2017).

Similar as PCR-GLOBWB (PCR), the kinematic wave approximation is used to route runoff along a drainage network. What is different is that WFL can be discretized at any spatial resolution. Also, model discretizations can be automatically obtained by using the built-in modelbuilder tool which makes use of the Google Earth Engine and global input data sets. To make WFL integrative with other models, it has a BMI adapter natively implemented. For more information about WFL, we refer to the online documentation at <https://wflow.readthedocs.io/en/latest/>. The WFL model code is available at <https://github.com/openstreams/wflow/releases>.

Possible coupling realizations with GLOFRIM 2.0

We envision GLOFRIM as a plug-and-play tool where the user can design the coupled model depending on the modelling needs, previous experience, or simply models at hand. With the increased number of models contained by GLOFRIM 2.0, the number of possible coupling cascades increased too. We defined three categories of models: i) hydrologic models computing runoff from meteorological data (PCR and WFL); ii) routing models focussing on simulating water transport simulation along a 1D river network (CMF); and iii) hydrodynamic models determining discharge for both 1D river network and 2D floodplain areas (Delft3D Flexible Mesh (DFM) and LISFLOOD-FP (LFP)).

In GLOFRIM 2.0 the combinations hydrology-routing, hydrology-hydrodynamics, and hydrology-routing-hydrodynamics are so far supported and tested (see Figure 6.1).

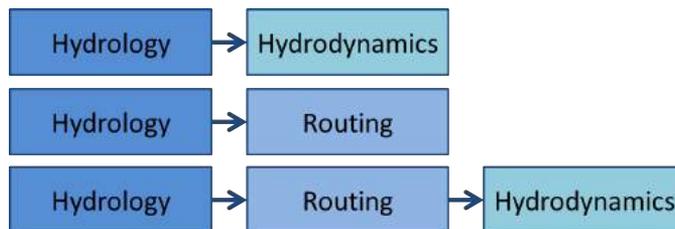


Figure 6.1: The coupling realizations we envision to be most sensible to define with the models contained in GLOFRIM.

With these combinations, various physical processes at different spatial resolutions can be simulated, starting from, for example, snow melt simulations in the hydrologic model via routing of excess runoff over large distances to the river mouth where detailed floodplain flow is simulated.

Applying GLOFRIM 2.0 as a “Human Model Interface (HMI)”

In contrast to GLOFRIM v1.0, we streamlined the execution of the models and limited the amount of information supplied by the user to the necessary minimum. Users only need to

reference the model's initialization files in an ini-file, which is sufficient for GLOFRIM 2.0 to interpret all model specifications including spatial domain, and to define the exchange of variables between models as well as the exchange interval. Note that the content of the ini-file changed fundamentally compared to previous versions of GLOFRIM (chapter 3). Also, we implemented a two-step model initialization in which all model-specific configuration files are first aligned to user-specified settings, such as the start and end times for simulation, and subsequently models are initialized based on the altered configuration.

GLOFRIM can be executed from command line specifying the path to the required settings file as well as start and end time:

```
python glofrim_runner.py run /path/to/glofrim.ini -s startdate -e enddate
```

By means of the environment file, where local paths to the model executables are provided, applying GLOFRIM in shared data environments is facilitated. To give users additional freedom, it is furthermore possible to not only execute coupled runs, but also run the models in stand-alone mode.

Initiating coupled model runs from command line is only possible after aligning each model and its BMI (after all, the BMI standards are implemented differently into each model) by inserting an additional layer of code into GLOFRIM. While it would also be possible to run models from there, on a very "basis" level, employing GLOFRIM is more convenient due to streamlining model settings, variable exchange, and model execution. As such, GLOFRIM essentially functions as a "Human Model Interface" facilitating the communication between user and code.

6.1.2 Applying GLOFRIM 2.0 in the Ganges-Brahmaputra basin: cascading model coupling and nested modelling

The experimental set-up

We here test the impact of model coupling cascades on flood wave propagation as well as simulated inundation extent by running the settings PCR, PCR→CMF, and PCR→CMF→LFP for the Ganges-Brahmaputra basin. To allow for a fair comparison of routed discharge, we ran PCR with its DynRout routing extension, adding simplistic volume distribution across floodplain areas and floodplain routing.

To make full use of the 1D hydrodynamic scheme, the extent of PCR and CMF are identical and thus runoff from each PCR cell is coupled to at least one CMF channel. Since applying LFP for the entire domain would be computationally very expensive and also not the purpose of the model, we nested a small LFP schematization of the river delta into the larger PCR and CMF domain. Within this nested setting, various boundary conditions apply for LFP: upstream discharge from CMF, local runoff from CMF, and downstream water level dynamics as prescribed within LFP itself (in this case 0.0 m).

To derive the local models, we used the global default model set-ups of PCR and CMF and clipped them to the extent of the Ganges-Brahmaputra basin. Subsequently, we performed a calibration of CMF channel depth and floodplain roughness coefficients due to initially insufficient accuracy of simulated discharge.

Based on calibration, we applied a Mannings coefficient of $0.03 \text{ m}^{-1/3} \text{ s}$ for PCR as well as for both river and floodplains in CMF. Channel depth was, also based on calibration, increased by changing the first factor from 0.14 to 0.20 in the following equation:

$$B = \max[0.14xR_{up}^{0.40}, 2.00] \quad (6.1)$$

where B is channel depth [m] and R_{up} is the annual maximum of 30-day moving average of upstream runoff [$\text{m}^3 \text{s}^{-1}$].

We consequently created an 18 arcsec (approximately 500 m at the Equator) LFP discretization of the delta region by utilizing the underlying input raster data of CMF at identical spatial resolution (Figure 6.2). To stay in line with the CMF discretization, we also applied a value of $0.03 \text{ m}^{-1/3} \text{ s}$ for both channel and floodplain friction.

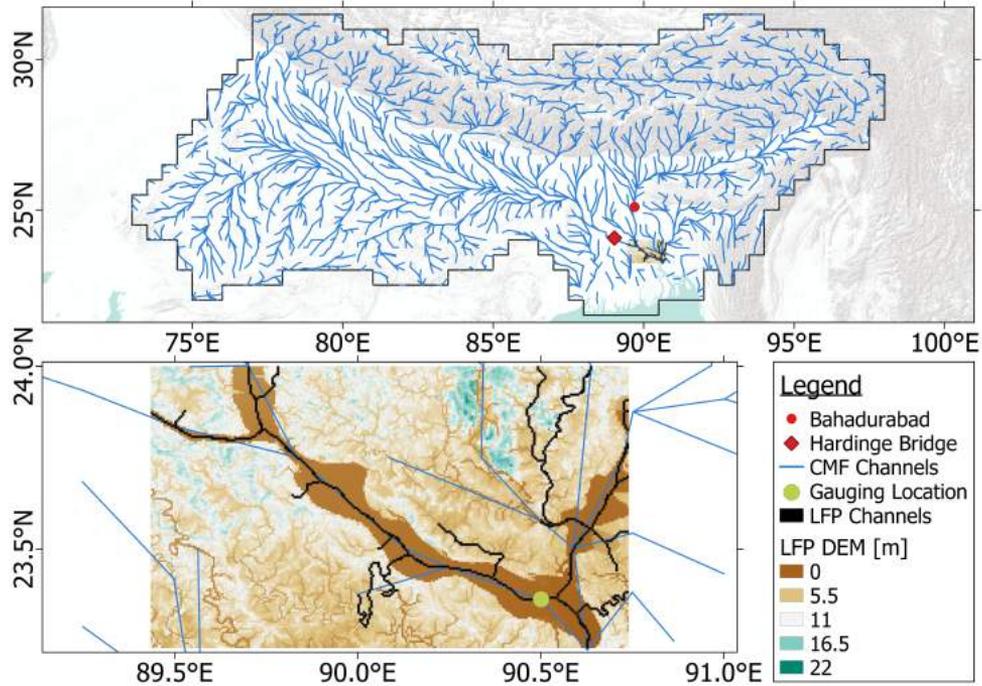


Figure 6.2: (top) CMF channel network in Ganges-Brahmaputra basin as well as locations of observation stations Hardinge Bridge and Bahadurabad for validating model output from both PCR and CMF; (bottom) zoom to LFP extent showing LFP DEM and channel network as well as gauging stations where output from PCR, CMF, and LFP is compared.

Model validation

To assess the quality of discharge simulations, we calculated the Kling–Gupta efficiency (KGE; Gupta et al. (2009)) and its components as well as the Nash–Sutcliffe efficiency (NSE) based on simulated and observed values for two locations: Hardinge Bridge in the Ganges river and Bahadurabad in the Brahmaputra river (see 6.2). Observed values were kindly provided by the Institute of Water Modeling, Bangladesh, and the Bangladesh Water Development Board.

Unfortunately, both stations lie outside of the LFP domain and thus it was not possible to validate LFP discharge as well. Therefore, we performed a qualitative comparison between discharges from all coupling set-ups at a gauging location within the LFP domain (see Figure 6.2) to assess how output from PCR→CMF→LFP relates to the other two set-ups.

The simulated inundation maps were compared with observed imagery. Therefore, PCR and CMF maps were first downscaled to a resolution of 1 km and 500 m, respectively, making use of their model-specific downscaling routines. As validation data, 8-day composite

MODIS imagery of 2007 was used as this year was characterized by strong monsoon-induced inundations (Kotera et al., 2016). To guarantee comparability, all maps were clipped to the LFP model domain and resampled to 500 m spatial resolution applying the nearest neighbour approach.

Inundation extent was validated for all set-ups following the approach of Fewtrell et al. (2008). Thereby, the hit rate H , the false alarm ratio F , and the critical success index C were determined for each inundation map with respect to observed MODIS extent (see equations 3.6, 3.7, and 3.8). All parameters can vary between 0 and 1. While $H=1$ shows that all inundated cells in the benchmark data are also inundated in the comparison data, $F=1$ indicates that the inundated cells in the comparison are entirely false alarms with respect to the benchmark. The critical success rate C , in turn, should be 1 for perfect agreement, thereby penalizing for both under- and overprediction.

Results and discussions

Discharge validation When validating simulated discharge from PCR-DynRout and PCR→CMF, PCR-DynRout shows similar or more accurate results than the coupled run, with results generally being more accurate for the Ganges river (Figure 6.3 and Table 6.1).

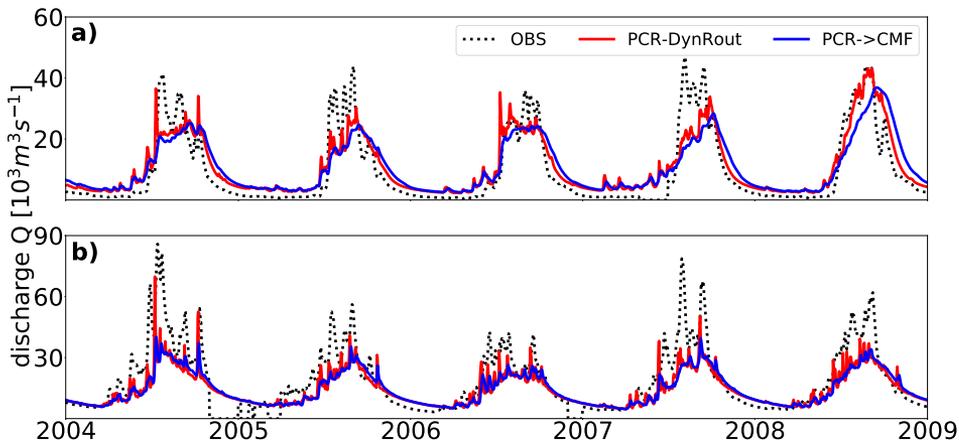


Figure 6.3: Simulated discharge by PCR-DynRout and PCR→CMF as well as observed discharge at **a)** Hardinge Bridge (Ganges) and **b)** Bahadurabad (Brahmaputra).

Compared with the coupled results shown in previous chapters (eg. in the Amazon, chapters 2 and 3), these results come as a surprise since the numerical scheme of CMF (the local inertia equations) accounts for more physical processes and therefore should yield better results than, as in this case, the kinematic wave approximation of PCR. Both equations are derived from the 1D depth-averaged Saint-Venant momentum equation (Eq. 1.2).

The local inertial equation is derived by neglecting only the advection term (second term in Eq. 1.2) of the shallow water equations, which is insignificant for many natural river and floodplain flow conditions (Yamazaki et al., 2013; Hunter et al., 2007). The kinematic wave approximation, however, only accounts for channel and friction slope (third and fourth term in Eq. 1.2). Please see 1.4.2 for an elaborated discussion.

| | KGE | KGE(r) | KGE(beta) | KGE(alpha) | NSE |
|-----------------|------|--------|-----------|------------|------|
| Hardinge Bridge | | | | | |
| PCR-DynRout | 0.71 | 0.89 | 1.15 | 0.78 | 0.77 |
| PCR→CMF | 0.63 | 0.83 | 1.15 | 0.70 | 0.66 |
| Bahadurabad | | | | | |
| PCR-DynRout | 0.46 | 0.84 | 0.79 | 0.52 | 0.55 |
| PCR→CMF | 0.44 | 0.86 | 0.79 | 0.50 | 0.54 |

Table 6.1: Assessment of performance of PCR, PCR→CMF, and PCR→CMF→LFP runs at both Hardinge Bridge (Ganges) and Bahadurabad (Brahmaputra); the coloured boxes indicate best performance compared to other set-ups.

Hence, we assume that the PCR results are “right for the wrong reasons” due to a fortunate combination of simulated runoff and routing. Given that CMF routing should produce better results when forced with accurate forcing, the results suggest that the runoff as simulated PCR is slightly off in timing. This points to one of the key structural challenges with such cascading model coupling as applied in GLOFRIM: while the most advanced hydrodynamic schemes can be added and even calibrated for best fit, the overall model accuracy still depends on the hydrologic forcing.

Benchmarking discharge from PCR→CMF→LFP From the previous section we know that discharge is reasonably well resembled by both PCR-DynRout and PCR→CMF. By sampling simulated discharge at a location which all three model set-ups have in common, we can assess how adding the 1D/2D hydrodynamic model LFP impacts discharge dynamics.

Figure 6.4 depicts that simulated discharge is generally less than of PCR→CMF, which more pronounced differences during high flow conditions. Despite the differences in flow magnitude, overall dynamics as well as timing is very comparable. We recall, however, that both CMF and LFP were discretized using the same, rather low, floodplain roughness coefficient. Consequently, the impact of simulating 2D flow on the floodplains in LFP is minimal. If higher values would have been used, however, we assume that differences in timing were more pronounced. Since already PCR→CMF underpredicted observed discharge in this case study (although at locations slightly more upstream), adding LFP does not necessarily improve discharge estimates.

Validating simulated flood extent If one were interested only in accurately simulated discharge dynamics, coupling PCR with CMF may be sufficient in many cases. However, a major indicator for flood hazard are inundation maps which may benefit greatly from adding actual floodplain flow computations. Validating the downscaled inundation maps from PCR and PCR→CMF and the modelled results of PCR→CMF→LFP with observed MODIS flood extent for the date 2007-08-18 shows significant deviations between model set-ups (Figure 6.5). In fact, results insinuate that acceptable representation of inundation patterns as expressed by the critical success index C can only be achieved by also accounting for floodplain flow and discharge through side channels (Table 6.2).

Table 6.2 furthermore shows that, despite having comparable false alarm ratios, the hit rate H is much higher for PCR→CMF→LFP and, in turn, so is the critical success index C. The differences in H are largely resulting from simulated inundations along smaller water

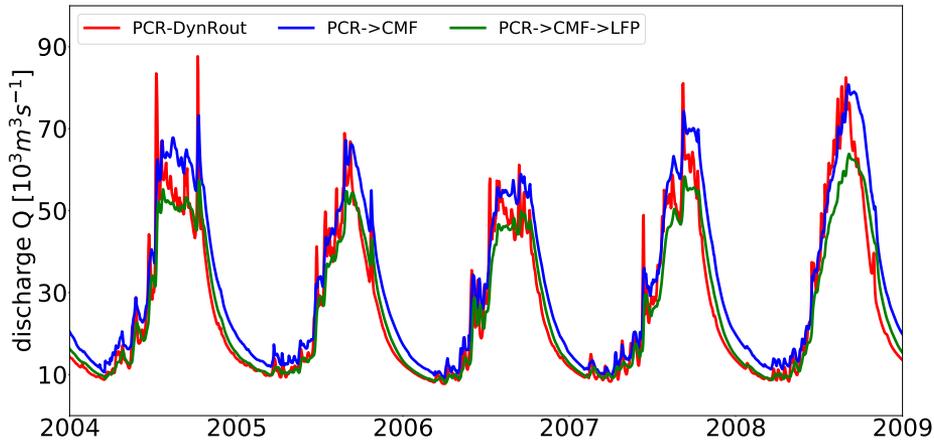


Figure 6.4: Simulated discharge for PCR-DynRout, PCR→CMF, and PCR→CMF→LFP at the common observation station (see figure 6.2).

| | PCR | PCR→CMF | PCR→CMF→LFP |
|------------------------|------|---------|-------------|
| Hit rate | 0.38 | 0.30 | 0.70 |
| False alarm ratio | 0.44 | 0.40 | 0.42 |
| Critical Success Index | 0.30 | 0.25 | 0.46 |

Table 6.2: Hit rate, false alarm ratio, and critical success index for the three model set-ups if compared to observed MODIS flood extent.

bodies, especially compared to PCR→CMF, and by simulating the extent across the entire river floodplain, which is particularly not the case for PCR (Figure 6.5). Despite all efforts to make the validation as fair as possible, there are still some limitations that must be kept in mind. For example, inundation patterns of the Ganges-Brahmaputra delta are largely affected by tide and surge dynamics (Ikeuchi et al., 2015).

Since we discretized all models with a steady 0 m water level boundary, it must be acknowledged that a perfect fit between observations and simulations would not be possible. Besides, the downscaling routines of PCR and PCR→CMF employ different approaches and data, resulting in locally marked differences in results. Aligning the routines is, however, outside of the scope of this work. The arising issues of different inundation extent due to different model routines and data was already discussed by other studies and remains subject to on-going debate how to minimise the gap between models (Bernhofen et al., 2018; Trigg et al., 2016; Hoch and Trigg, 2019).

Run time One aspect that can play a role when deciding between a more parsimonious or a more complex modelling approach is run time. We had to exclude PCR-DynRout from the comparison as the DynRout extension is not (yet) supported by GLOFRIM, and hence comparison is only possible for PCR→CMF and PCR→CMF→LFP. Results show that the addition of full 2D floodplain flow has great impact on run times, which increased from 38 minutes to 300 minutes on an Intel[®] Xeon[®] CPU E3-1276 processor with 3.60 GHz and 8 GB RAM.

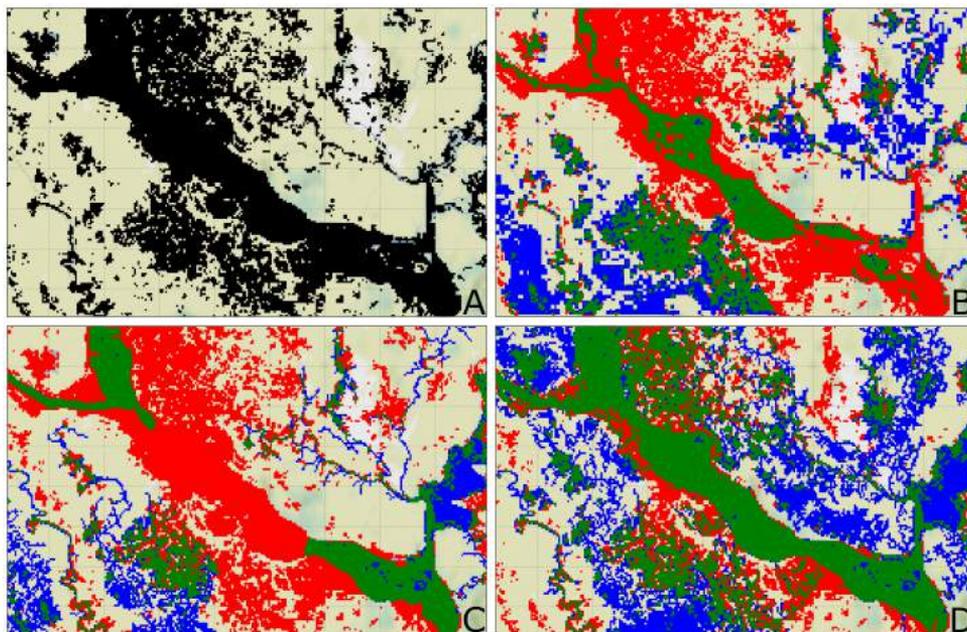


Figure 6.5: Inundation extent for the domain covered by LFP with **A** being observed flood extent according to MODIS. The benchmark of observed and modelled flood extent is shown for **B** downscaled PCR, **C** downscaled PCR→CMF, and **D** PCR→CMF→LFP. Thereby blue indicates model result only, red indicates observation only, and green indicates agreement between model result and MODIS observations.

Conclusion and recommendations

Flood hazard is driven by several physical processes which have varying influence depending on study area, study aim, scale, and so forth. GLOFRIM 2.0 is a modular computational framework, facilitating the simulation of by integrating across multiple physical processes and spatial scales. Besides, it can also be used to benchmark both hydrologic and hydrodynamic models. To that end, a range of hydrologic models (PCR-GLOBWB, wflow) can be spatially and online coupled to several hydrodynamic models (Delft3D Flexible Mesh, LISFLOOD-FP, CaMa-Flood).

By separately coupling PCR-GLOBWB to CaMa-Flood and to CaMa-Flood plus LISFLOOD-FP in the Ganges-Brahmaputra basin, we assessed how models with different complexity as well as focus perform in simulating flood wave propagation and inundation extent. While it is expected that employing the physically more complex routing scheme of CaMa-Flood would improve results, this is not the case in the here presented study. What we conclude from our results is that the benefit of replacing the routing scheme of a hydrologic model with a hydrodynamic model is limited by the accuracy in simulated runoff. For some case studies, for instance the Amazon, the timing and magnitude of runoff combines well with physically more complex routing scheme (Hoch et al., 2017a; Hoch et al., 2017b). For other basins, such as the Ganges, the additional flow attenuation introduced by CMF leads to reduced model accuracy, indicating that hydrology needs to be improved first.

Flood extent maps are a key indicator of hazard and regularly used for flood risk forecasting, definition of flood insurance premiums, and spatial planning. Our results show that applying a 1D/2D hydrodynamic model such as LFP, which includes floodplain flow and dis-

charge through secondary channels, is paramount for accurately simulated inundation maps. Notwithstanding the best performance of this set-up, a critical success index of 0.46 indicates that only around half of the actual extent is correctly captured by the model, leaving much room for improvement. For example, a thorough analysis of the used DEM may help to reveal whether water is trapped in local depressions, hampering return flows and therefore increasing the flood extent unnecessarily. Possible problems could be solved by hydraulic conditioning (Yamazaki et al., 2012a) or by updating model discretizations with more error-free elevation data sets such as the MERIT-DEM Yamazaki et al. (2017).

By linking hydrologic, routing, and hydrodynamic models, we can establish a model cascade which can simulate the inundation-driving processes from the mountains to the coast. As such, GLOFRIM 2.0 can be a key tool for future more holistic modelling studies researching the effect of the interplay of meteorology and hydrology, river routing, and floodplain dynamics on flood hazard and risk. Since GLOFRIM has a modular design, other models can easily be added to even extend the range of simulated flood hazard drivers, for example coastal and oceanic processes.

6.2 Implementing feedback loops with groundwater

Based on: Hoch, J. M., Sutanudjaja, E. H., van Beek, L. P. H., Winsemius, H. C., Bierkens, M. F. P. (in preparation), The benefit of integrated hydrologic-hydrodynamic-groundwater modelling for large-scale inundation modelling.

The development of GLOFRIM was and is aimed to improve the physical representation of surface water flow processes to better simulate inundation events. To that end, GLOFRIM allows to couple a variety of hydrologic and hydrodynamic models, *inter alia* the hydrologic model PCR-GLOBWB (PCR). So far, the default settings of PCR were employed to generate runoff fields which were then used to force hydrodynamic models in a spatially-explicit and online manner.

In its default design, PCR simulates for each grid cell and each time step moisture storage in two vertically stacked upper soil layers, as well as the water exchange among the soil, the atmosphere, and a underlying groundwater reservoir (Sutanudjaja et al., 2018). Such a rather simplistic module is computationally efficient but does not allow for simulating horizontal groundwater flow between cells, hampering the inclusion of important processes such as converging groundwater flow supporting baseflow during low flow conditions, and in turn also influencing discharge dynamics.

To improve physicality of simulated groundwater processes, a tight coupling was established between PCR and the groundwater model MODFLOW (chapter 1.3.1; Sutanudjaja et al. (2014) and de Graaf et al. (2015)), replacing the default groundwater module with the process descriptions MODFLOW (Harbaugh et al., 2000). So far, the coupling was used to assess how changes in water demand and withdrawal have on water resources and groundwater levels (de Graaf et al., 2017; Wada et al., 2016). However, no structural assessment has been made yet how coupling MODFLOW with PCR impacts the accuracy of simulated flood wave propagation.

Assessing the interplay between hydrology and groundwater is important to not only promote the evaluation of the coupled PCR-GLOBWB-MODFLOW model (PCR-MOD), but also because it could provide an additional physical process to be integrated within GLOFRIM. It would then be possible to integrate across even more physical processes, for instance by subjecting inundated floodplain areas simulated by hydrodynamics to groundwater infiltration computed with PCR-MOD. It is shown in previous work (see 4, Hoch et al. (2018a)) that not accounting for these processes may lead to an overestimation of flood volume, and thus establishing a two-way coupling scheme between hydrodynamics and PCR-MOD could advance the accuracy of large-scale inundation estimates.

Hereafter, we investigate the potential benefit of including the groundwater flow processes simulated by MODFLOW into GLOFRIM in a two-step procedure. First, we apply both PCR and PCR-MOD for the Mekong River basin to assess how the addition of horizontal and vertical groundwater flow impacts simulated discharge. We opted for this basin as here PCR already has very good performance in stand-alone mode which is needed for a unbiased comparison. Since evaporation and groundwater infiltration from floodplains are less pronounced in the Mekong, we then coupled PCR-MOD with LISFLOOD-FP (LFP) for a small test case in the Niger Inland Delta utilizing the model of Neal et al. (2012a) and qualitatively benchmarked simulated surface water elevation and inundation extent to obtain a first estimate of how including an online two-way coupling scheme may improve results.

6.2.1 Methodology

The PCR-GLOBWB-MODFLOW model

The coupling between PCR and MODFLOW is established by replacing the simple groundwater reservoir with a two-layer groundwater component, allowing not only for vertical, but also horizontal exchange of water. Originally, the MODFLOW model is forced with outputs from PCR, specifically net recharge and river levels. The net recharge is the input for MODFLOWs recharge package, while MODFLOWs river and drain packages are used to incorporate interactions between the groundwater and surface water. For more detailed information about the coupled PCR-MOD model, please see the relevant literature (de Graaf et al., 2017; de Graaf et al., 2015; Sutanudjaja et al., 2014).

For the here presented analysis of the interplay between hydrology and groundwater on simulated discharge, we made several additional changes to the original PCR-MOD code. While exchanges between models were originally on a monthly-basis due to the slow dynamics of groundwater processes, we changed that to daily exchange to minimize the possibility of water balance issues.

Besides, MODFLOW is not anymore forced by the head difference between groundwater and river water level only, but also between groundwater and floodplain water level. Infiltration from the floodplain is determined using the Arno scheme (Hagemann and Gates, 2003; Todini, 1996) which is implemented as default in PCR. The remaining water volume on the floodplain is then subject to open water evaporation calculated either following the Penman–Monteith equation (Allen et al., 1998) or using Hamon (1963) in case only daily mean temperature is available.

By exposing the PCR floodplain inundation volume via BMI, it is possible interactively update this model state based on model-external information, for instance from a hydrodynamic model. To that end, inundation volume as simulated by LFP was upscaled to PCR resolution and coupled to the floodplain inundation volume variable of PCR. As a result, hydrodynamics were coupled to hydrology and, due to the link between floodplain water level head and groundwater head, also driving groundwater processes. The coupling from PCR to LFP remained unaltered and follows the approach outlined in chapter 3, coupling runoff and/or discharge in a spatially explicit and online manner. By coupling PCR, MODFLOW, and LFP, a wide range of physical processes can eventually be simulated (Figure 6.6).

Mekong River basin: the impact of groundwater processes on discharge

To assess the impact of added groundwater process on simulated discharge, we opted for the Mekong River basin as a test study area. The underlying motivation to choose this basin was that we required good model performance already with default model settings to avoid unnecessary complications in the analysis of the coupled run. Both PCR and MODFLOW were discretised at 5 arcmin spatial resolution (or around 10 km x 10 km at the Equator). Note that this is the first time PCR was applied at this finer spatial resolution compared to all previous chapters. For parameterizations of both PCR and MODFLOW, we refer to the model description articles. As meteorological data, ERA-Interim re-analysis products were used (Berrisford et al., 2011).

To quantify the impact of the added feedback loop, we separately ran PCR stand-alone (PCR-only) and coupled with MODFLOW (PCR-MOD). The resulting discharge was validated against observation at Kratie, located at 12.480° E and 106.017° N (also see figure 6.7). Observed discharge was kindly provided by the Southern Institute for Water Resource Planning (SWIRP) in Vietnam. To assess model accuracy, we determined both the Kling–Gupta efficiency (KGE; Gupta et al. (2009)) and the Nash–Sutcliffe efficiency (NSE). Besides, we

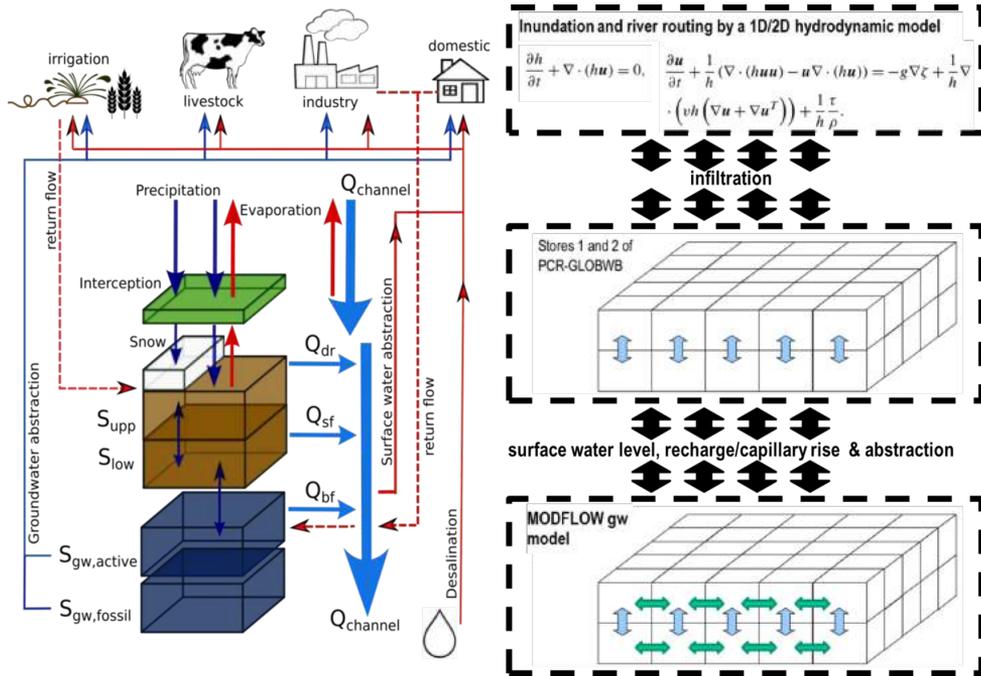


Figure 6.6: Schematic of model set-up PCR and MODFLOW coupling showing the various stocks and flows as well as interactions between models and model components.

analysed the contributions of groundwater discharge to total discharge to obtain a first estimate how groundwater processes can shape the hydrograph of resulting discharge.

Niger Inland Delta: the impact of adding hydrodynamics to PCR-MODFLOW

While the Mekong case merely assesses the impact of adding groundwater, we here investigate how a dynamic feedback coupling set-up between hydrology, hydrodynamics, and groundwater affects simulated inundation extent and water level in the hydrodynamic model.

To that end, the coupling between floodplain inundation depth and PCR-MOD as described in the previous section is added to the GLOFRIM framework (chapter 3). We tested the coupling design in the Niger Inland Delta (Figure 6.8) as this area is known for the strong impact of groundwater infiltration and evaporation on the water budget, and in turn inundation extent, during dry season (cf. Dadson et al. (2010), Mahe (2009), and Mahe et al. (2009)). In addition, the smaller size of the study area allowed for shorter run times as computational load was expected to rise due to additional exchanges, processes simulated, and grid cell combinations. The spatial resolution of LFP was 0.009 degree (translating to roughly 900 m cell length) allowing for fine-resolution output maps. The only change made to the original model set-up as described by Neal et al. (2012a) was to clip the model domain to the 30 m Height Above Nearest Drainage (HAND) extent (Rennó et al., 2008).

Adding a 1D/2D hydrodynamic model such as LFP is not *per se* required for improved discharge estimates (employing a large-scale routing model instead may be more efficient, see section 6.1) but can be beneficial to improve the representation of inundation extent and depth due to the explicit simulation of floodplain flow processes (chapters 2, 6.1). Hence, we qualitatively compared output for both from three runs: A) hydrodynamics only forced with

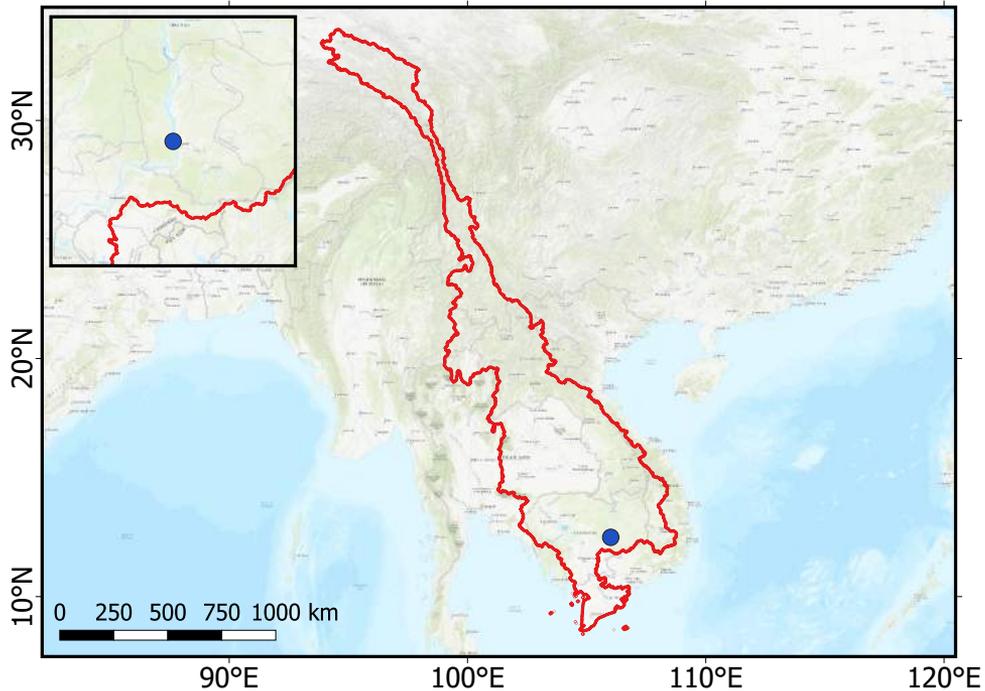


Figure 6.7: Depiction of the basin extent of the Mekong River basin (red shape) as well of Kratie where simulated discharge is validated (blue dot) in their wider topographical setting.

observed discharge data, named “LFP-only” hereafter; B) *idem*, but then additionally with one-way coupling of PCR runoff as described in chapter 3 named “PCR-LFP” hereafter; C) *idem*, but adding the feedback loop between PCR and MODFLOW as described above, named “PCR-MOD-LFP” hereafter.

Due to the early stage of the study, we limited ourselves to qualitatively benchmarking results from PCR-LFP and PCR-MOD-LFP, respectively, with the LFP-only run which we assume to be the most accurate due to the forcing with observed discharge data only. Whilst inundation extent was compared for the entire domain, simulated stage was recorded at several stations (Figure 6.8).

6.2.2 Results and discussion

The impact of adding MODFLOW on simulated discharge

Results indicate that adding a full groundwater model improves simulated discharge in the Mekong River basin. While the performance of PCR in reproducing discharge is already good without additional groundwater processes, both KGE and NSE increase further after adding MODFLOW (Figures 6.9 and 6.9, Table 6.3).

The increased model accuracy is mostly related to the higher peak discharges with PCR-MOD which compensate for the underprediction of peak discharge with PCR only. Further analysis reveals that the higher discharge is the result of the additional influx of groundwater discharge.

Despite the improvement, results also show that low flow conditions are still not simulated accurately. While low flow discharge is too high without considering MODFLOW, the reduc-

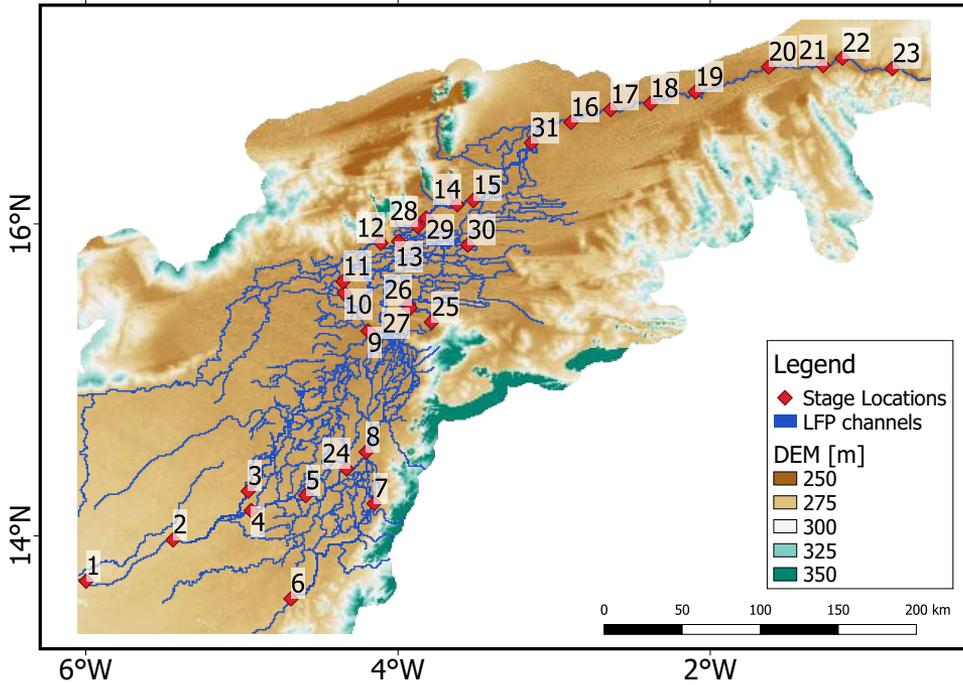


Figure 6.8: Study area of the Niger Inland Delta showing the Digital Elevation Map (DEM), the LFP channel network, and the locations including corresponding number (“station IDs”) where simulated water stage is recorded.

| | KGE | NSE |
|----------|------|------|
| PCR only | 0.72 | 0.82 |
| PCR-MOD | 0.88 | 0.90 |

Table 6.3: KGE and NSE of simulated discharge at Kratie obtained with uncoupled PCR (‘PCR only’) and coupled PCR-MODFLOW runs.

tion is too strong when adding the feedback between surface water and groundwater. What is striking is that total discharge is lowered to zero for parts of the low flow period, meaning the Mekong River is running dry. Similar as for peak flows, we see that the contribution of groundwater discharge to total discharge decreased when adding MODFLOW. Besides, the moments of zero discharge correspond well with zero groundwater discharge, indicating that the dynamics of groundwater discharge are a major control of surface water discharge under the used setting and parameterization.

This leads us to the question how the seasonal fluctuations of the groundwater storage (S_{gw} in Figure 6.6) influence groundwater discharge and in turn surface water discharge.

PCR conceptualizes net groundwater discharge volume Q_{gw} [$\text{m}^3 \text{ day}^{-1}$] to the surface water bodies as the difference between linear reservoir release baseflow and river bed infiltration Inf , both in m d^{-1} :

$$Q_{gw} = \text{baseflow}A + InfxA_{wat} \quad (6.2)$$

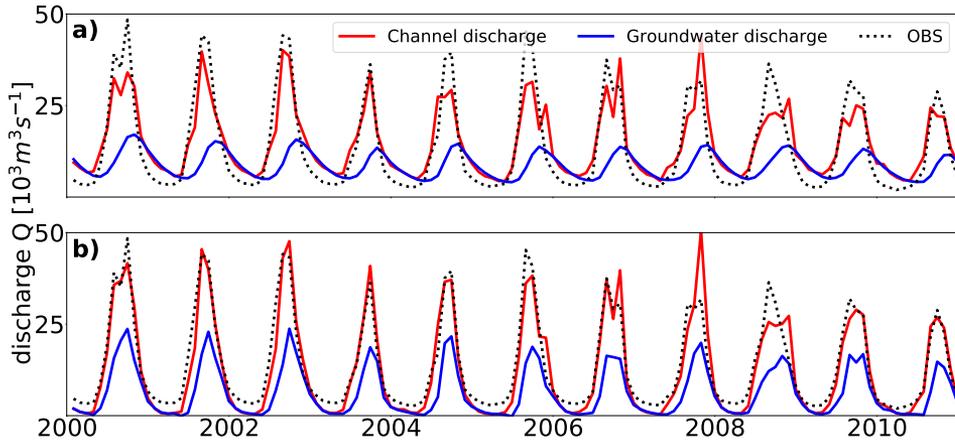


Figure 6.9: Simulated discharge of PCR-only as well as observed discharge at Kratie, also showing the contribution of groundwater discharge for both runs: **a)** without MODFLOW coupled and **b)** with daily coupling between PCR-GLOBWB and MODFLOW.

with the positive sign indicating water flowing from groundwater to surface water bodies, A and A_{wat} indicating total cell areas and surface water areas, both in m^2 .

The baseflow is the product of recession coefficient J [day^{-1}] and groundwater storage S_{gw} (m). For runs without MODFLOW, Inf can only have a negative value and occurs in areas where groundwater abstraction exceeds baseflow. It is limited by the assigned river bed conductivity. For runs with MODFLOW, Inf is conceptualized as in the RIVER package of MODFLOW (see Sutanudjaja et al. (2011)) and can have a positive value if groundwater head h [m] is higher than surface water elevation $HRIV$ [m]:

$$Inf = (1/res)(h - HRIV), \text{ when } h > RBOT \quad (6.3)$$

$$Inf = -(1/res)(HRIV - RBOT), \text{ when } h \leq RBOT \quad (6.4)$$

where $RBOT$ and res are surface water bed elevation [m] and resistance [day].

Consequently, during wet seasons when $h > HRIV$, the term in Eq. 6.3 is positive and suggests that the net groundwater discharge is faster in the run with MODFLOW (while in the run without MODFLOW this term is zero). In dry seasons when $h < HRIV$, Inf is limited in the run with MODFLOW by only available surface water, but it is not limited with river bed conductivity. Consequently, Inf can be higher (“more negative”) in the run with MODFLOW.

The impact of feedback loops on simulated inundation extent and depth

We see that adding a dynamic feedback loop between hydrology, hydrodynamics, and groundwater processes, does not necessarily improve simulated estimates of inundation extent and depth.

While the simulated stage (Figure 6.10) is only slightly increased if PCR runoff is added. This response is expected due to adding a new source of water but no sink.

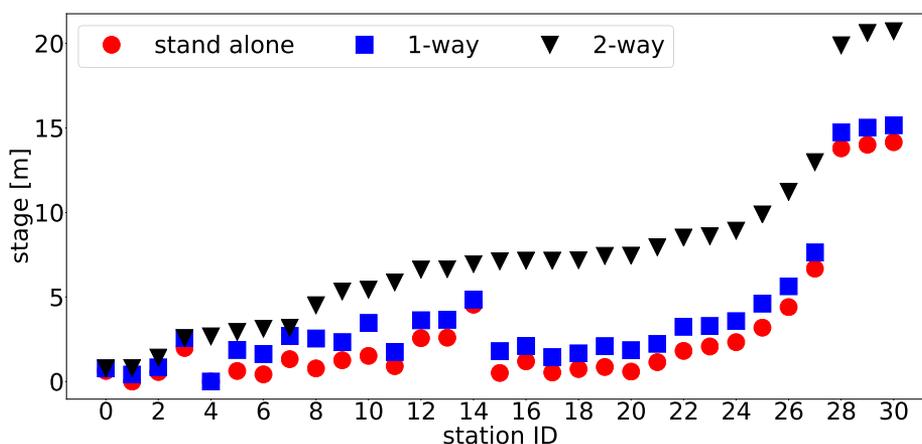


Figure 6.10: simulated stage at 30 observation locations (station IDs, see figure 6.8) throughout the Niger Inland Delta at the end of the modelling run for LFP stand-alone, one-way coupled with PCR, and two-way coupled with PCR-MODFLOW.

What is not expected, however, is that the offset between LFP-only and the two-way coupled PCR-MOD-LFP increases for all stations even though inundations should be subject to groundwater infiltration and evaporation processes. In addition to the higher simulated stage by PCR-MOD-LFP, simulated inundation extent is much larger compared to extents simulated by both LFP-only and PCR-LFP (Figure 6.11).

The reason for this behaviour could not yet been entirely disentangled due to the large number of processes influencing simulated inundation patterns. Nevertheless, we strongly suspect that the higher water stage as well as inundation extent are connected to increased peak flow which in turn may result from the increased contribution of groundwater discharge to river discharge, as shown in section 6.2.2. Additionally, higher peak flow may also lead to inundation of areas which cannot be drained during low flows, eventually yielding the greatly increased inundation extent.

6.2.3 Conclusion and recommendations

From the Mekong case, it can be hypothesized that the accuracy of the simulated groundwater discharge has a major influence of the accuracy of overall discharge estimates. It is encouraging to see that adding the MODFLOW model to PCR can help alleviate underpredictions of peak flow discharge, which is particularly important for flood risk management practices where capturing peaks is essential. Notwithstanding these improvements, the reduction of discharge to zero as shown in the Mekong River basin hints at several issues in the model set-up. Besides, the shown behaviour of the coupled model will most likely lead to a deterioration of results in cases where PCR already overpredicts observations by default.

The issues with the PCR-MOD model for inundation modelling are further illustrated in the Niger Inland Delta. We here see that only adding runoff to observed discharge as model forcing yields slightly higher river stages and inundation extent, while using a dynamic feedback loop with MODFLOW resulted in much larger inundation estimates. Since additional runoff may indeed increase available flood volume and in turn stages and extent, accounting for groundwater infiltration should in theory result in lower values.

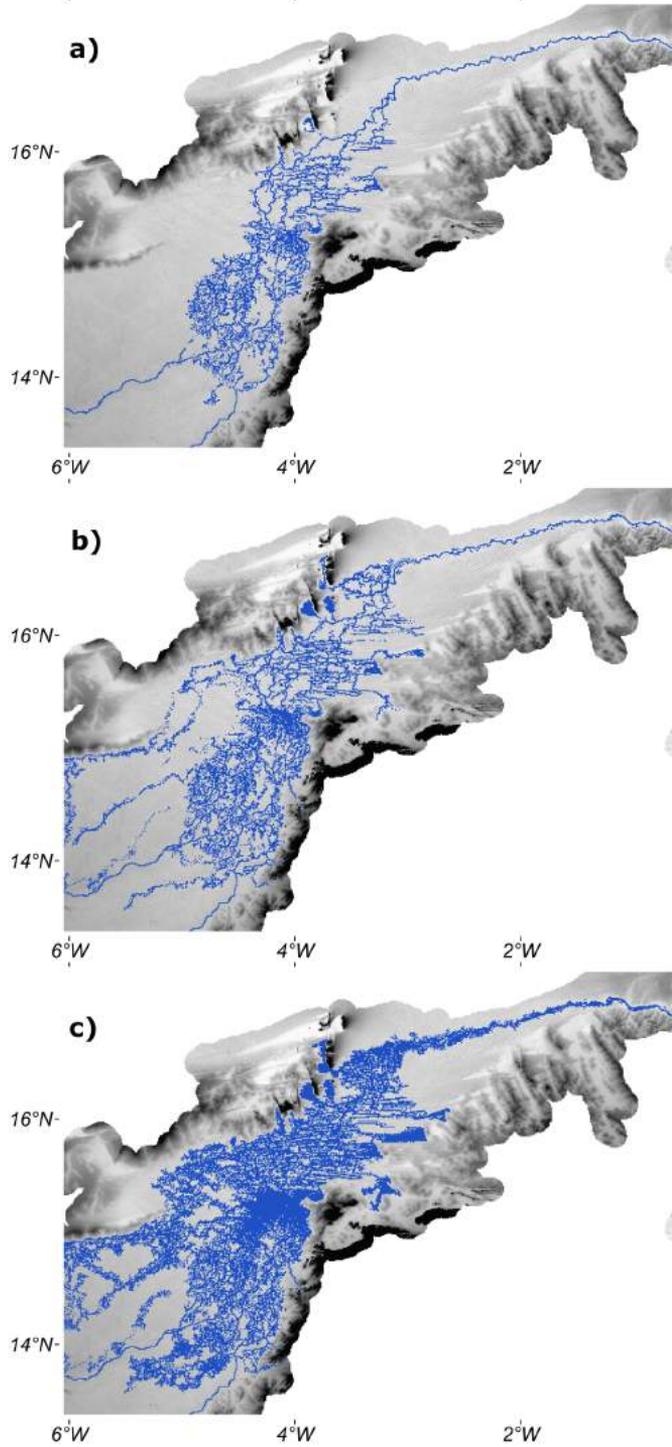


Figure 6.11: simulated inundation extent indicated in blue throughout the Niger Inland Delta at the end of the modelling run for a) LFP stand-alone, b) one-way coupled with PCR, and c) two-way coupled with PCR-MODFLOW.

As this is not the case, it can be concluded that the feedback with groundwater processes introduces an additional error in the computations. In fact, the PCR-MOD-LFP coupling mixes two approaches of coupling, namely internal coupling between PCR and MODFLOW, and flexible coupling between PCR and LFP (chapter 1.3.1). Unravelling the interdependencies in such a context between the hydrologic model, the groundwater model, and the inundation model becomes drastically more complex. However, as issues with PCR-MOD are already signalled in the Mekong case, it can be doubted whether the addition of a hydrodynamic model is the main reason for model behaviour in the Niger case.

Even though adding the coupling MODFLOW with PCR enabled several groundwater related studies, this cannot be directly translated to surface water studies yet. Before further development of the coupling between inundations and groundwater processes, it is recommended to further assess the parameterization and performance of the MODFLOW model, both in stand-alone and coupled mode. One of the first parameters to evaluate should be the river bed infiltration capacity as it has great control over both magnitude and timing of groundwater discharge. Only if the inter-dependencies are unravelled and individual model performance is optimized, the promising approach of fully integrating groundwater, hydrologic, and hydrodynamic models in one dynamically coupled framework can display its full potential.

Chapter 7

Synthesis

This thesis aimed to investigate possible technical implementations, opportunities, and challenges of integrated hydrologic-hydrodynamic modelling for improved large-scale inundation estimates, potentially also being applied in the context of global flood modelling. The research objectives of this thesis were:

1. *Establish a flexible end-user oriented scientific software framework to facilitate integration across physical processes, spatial resolutions, and flood hazard drivers.*
2. *Improve the understanding of how model discretization as well as the complexity of simulated physical processes influence simulated discharge and inundation extent.*

The objectives were achieved by developing and testing a coupled hydrologic-hydrodynamic model, which subsequently evolved to a flexible coupling framework named GLOFRIM, a **globally applicable framework for integrated hydrologic-hydrodynamic modelling**.

Model coupling can overcome major shortcomings of hydrologic and hydrodynamic models: while simplistic routing schemes and limitations to coarse-resolution drainage networks hinder hydrologic models to accurately simulate flood wave propagation for low-gradient areas and areas where backwater effects are predominant, hydrodynamic models typically depend on gauged data as model forcing which is problematic when it comes to the integration of spatially-varying hydrologic processes and future climate changes into models.

What has been made possible with the development of GLOFRIM is to combine the strengths of (global) hydrologic models in processing meteorological and land-use data to compute runoff with the more sophisticated routing schemes and finer spatial resolution of hydrodynamic models.

Hereafter, I first list and then elaborate the main conclusions of my thesis. This is followed by a discussion of the general and GLOFRIM-specific challenges of model coupling. I conclude this synthesis with recommendations and outlook for future research.

7.1 Main conclusions

Answering the research questions of this thesis, the main findings are:

1. Applying computational frameworks can facilitate model benchmarking (chapters 3, 4, 6.1). By employing identical hydrologic forcing to two or more different hydrodynamic models or to various discretizations of one model, the sensitivity of results to different model concepts or discretizations can be assessed efficiently in a standardized manner.
2. Global flood models have developed greatly in past years but must continue to do so. Since differences in model performance are not sufficiently well understood yet, a GFM

validation framework could help disentangling the causes for performance differences. A model coupling framework could then support the combination of the most accurate model components (chapter 5).

3. Replacing the default routing scheme of global hydrologic models with more advanced schemes enhances accuracy of simulated discharge. Additionally, by coupling models with varying spatial resolutions, the accuracy of simulated inundation patterns can be improved (chapters 2, 3, 6.1). Which models to couple should be chosen appropriate to the study objective.
4. Applying flexible meshes for large-scale inundation modelling has its opportunities, but also its challenges (chapter 4). While they allow for reduced run times, additional efforts have to be taken to develop techniques for accurate and efficient mesh generation based on global data only.

7.1.1 Applying frameworks for model benchmarking

Developing and applying computational frameworks with standardized workflows brings several advantages compared to bespoke model coupling efforts, for instance easier reproducibility, transferability to other basins, and increased user-friendliness. Additionally, standardizing workflows and model settings facilitates benchmarking of coupled models and their properties.

The role of model conceptualization

GLOFRIM is a globally applicable framework for integrated hydrologic-hydrodynamic modelling. While employing a computational framework with standardized workflow facilitates reproducible research, fundamental differences between both hydrologic and hydrodynamic model conceptualizations remain. As a result, coupling output from one hydrologic model to different hydrodynamic models, although using the identical framework, does not necessarily yield identical results.

Before applying GLOFRIM, the modeller must decide which models to couple. Depending on the modelers decision, output of the coupled models can differ not only due to subjective choices made (Melsen et al., 2019), but simply because model conceptualizations deviate too. In a coupled hydrologic-hydrodynamic model, differences in hydrodynamic model characteristics can have a marked influence on inundation patterns (Dimitriadis et al., 2016; Hunter et al., 2008), even if identical forcing (for instance, runoff from a hydrologic model) and boundary conditions are applied (chapter 3). While the coupled results therefore depend on the user's model choice and an educated decision must be made *a priori*, the standardized procedures of GLOFRIM can also be exploited to unravel which models and model parameterization work better than others and, more importantly, why.

By aligning model forcing and downstream boundary conditions as well as all internal parameters such as Manning's surface friction coefficients, the impact of these settings and parameters can be assessed (chapters 2, 3, 4, and 6.1). Unfortunately, some of the discrepancies in model conceptualization are so intrinsic that they cannot be aligned beforehand, if at all, and hence are intangible for structured model benchmarking studies. Examples encompass but are not limited to: differences in actual channel length, channel-floodplain interaction, and catchment delineation approach.

Consequently, it is not possible yet to entirely disentangle the reasons for the occurring differences between model output. Aligning model properties and discretizations could help achieving this, yet it is currently left with the user. To automate this, GLOFRIM or any comparable framework would have to be capable of aligning the above-mentioned discrepancies

on the fly as much as possible by not only replacing entire components such as the routing scheme, but also much more basic model information, for example on topography and channel geometry.

In this thesis, we show that these disparities can result in different simulated discharge and inundation extent (chapters 2, 3, 4, and 6). Even though it may seem trivial that different models produce different results, it is paramount to keep the discrepancies and their impact on results in mind to obtain best possible results as not each model is necessarily fit for any purpose at any location.

The standardized workflow of frameworks such as GLOFRIM can be employed to advance global flood models (GFMs). By coupling identical hydrologic forcing to hydrodynamic models, for instance, it could potentially be unravelled why simulated inundation extent for the same flood event deviates between GFMs using hydrodynamic models as their computational core (see chapter 4 and section 7.1.3).

In addition to benchmarking hydrodynamic models, understanding the uncertainties and sensitivities of hydrologic models in a coupled setting is crucial as well. With the currently developed GLOFRIM 2.0 (see chapter 6.1), the wflow modelling platform is added. Since this increases the number of hydrologic models available, GLOFRIM 2.0 could not only be employed to benchmark hydrodynamic models, but also hydrologic model conceptualizations. As above-shown results indicate that the hydrologic forcing in a coupled setting demarcates the possible improvements of model coupling (chapter 6.1), a benchmark of hydrologic models could bring clarity which hydrologic model to use for which application and/or under which circumstances. Additionally, employing ensemble runs could be a viable option to get an idea of the uncertainty inherent in hydrologic runoff and, in turn, discharge estimates from coupled hydrodynamic models.

Opportunities and challenges of alternative gridding techniques

If designing a coupled model set-up, it is not only the conceptualization of the selected models that will impact results, but also the way each model is parameterized and discretized. That modelers decisions on, *inter alia*, calibration data and period, spatial resolution, and topographic representation within one model can influence model results is well documented (chapters 2, 3; Melsen et al. (2019), Merwade et al. (2008), Savage et al. (2016a), and Savage et al. (2016b)).

GLOFRIM or similarly designed computational frameworks can be applied to investigate the impact of some of these factors, for example how the model is discretized for a specific study area in terms of spatial resolution or gridding approach.

Currently, most large-scale models and GFMs employ uniform regular grids. Notwithstanding the advantages of regular grids for data processing and model design due to their uniformity, their use has the disadvantage that the model may run at a finer spatial resolution than locally required (Figure 1.6), increasing run time unnecessarily (see chapters 1 and 4).

Even though they are computationally more efficient for large-scale problems, alternatives such as flexible meshes are not yet tested extensively enough for large-scale fluvial inundation modelling but limited to the use for oceanic or coastal modelling studies mostly (cf. Kernkamp et al. (2011), Martyr-Koller et al. (2017), and Muis et al. (2016)).

In addition to flexible meshes, the run times emerging from using regular grids may also be shortened by alternative ways of model discretization, for example by accounting for sub-grid terrain variability. Such approaches are already implemented in some hydrodynamic routines (such as HEC-RAS or 3Di) but are yet to be realized for GFMs.

By forcing all discretizations with identical output from the coupled hydrologic models, it was confirmed that the application of flexible meshes is a time-saving alternative compared to regular grids (chapter 4). However, it was also shown that flexible meshes need to be designed more carefully and depending on the envisaged study aim as the greater freedom in discretization can easily lead to a deterioration of model results, particularly for basin-wide output such as inundation extent maps (chapters 3, 4).

One of the major challenges is that, compared to applications simulating wave propagation across oceans and coastal zones, the bathymetry of the terrestrial surface is characterized by much more irregularities and shallower water depth. So far, mesh generation techniques have been presented as well as discussed for bespoke and small-scale studies (cf. Bahrainian and Mehrdoost (2012), Kim et al. (2014), and Schubert et al. (2008)). How to generate flexible meshes in ways appropriate to capture these irregularities for large-scale studies is still a major research challenge, particularly because approaches should require little data due to computational limitations and local data scarcity. Improvements could possibly be achieved by supplementing the here used Height Above Nearest Drainage (HAND) data with the “building-hole” approach (Schubert et al., 2008) to also resolve man-made structured within the domain. Besides, the recently published “Mesher” approach (Marsh et al., 2018) specifically designed for hydrologic and land surface models should be tested for large-scale applications too.

Hence, novel modelling approaches such as flexible meshes for large-scale fluvial inundation modelling have the potential to supplement or even replace the traditional regular grids. Before that is possible, however, additional research efforts, testing, and evaluation of mesh generation techniques have to be undertaken, particularly for large-scale applications.

7.1.2 GFMs must advance

In recent years, the development and application of GFMs has increased (Ward et al., 2015). Despite their application in policy-making and decision-making, there is still insufficient understanding why model results differ when applied for the same test cases (Bernhofen et al., 2018; Trigg et al., 2016). To increase the uptake of GFMs and make their results more actionable, additional light must be shed on the causes for these discrepancies and GFMs must advance consequently.

This confronts us with a two-step issue. First, it is necessary to identify internal differences between GFMs. Currently, assessing internal differences is hampered by using different model forcing, boundary conditions, and validation procedures (Bernhofen et al., 2018; Trigg et al., 2016; Hoch and Trigg, 2019). To obtain more insight, it is thus necessary to run GFMs with equal forcing and boundary conditions. Also, a more streamlined validation and benchmarking procedure of GFMs is required as it already happened in, for example, climate science (eg. the “Coupled Model Intercomparison Project (CMIP)” and the “Inter-Sectoral Impact Model Intercomparison Project (ISIMIP)”) and hydrology (eg. the “Structure for Unifying Multiple Modeling Alternatives (SUMMA)” project (Clark et al., 2015a; Clark et al., 2015b), the “Modular Assessment of Rainfall–Runoff Models Toolbox (MARRMoT)” (Knoben et al., 2019) or the recently initiated “eWaterCycle2” project (Hut et al., 2018; Dzigan et al., 2019; Pelupessy et al., 2019)).

To guarantee model comparability, the same objective functions, observations, date, and so forth must be used. All this could be achieved by means of a GFM validation framework (chapter 5), providing well-tested forcing data including their error structure and subjecting all GFM output to clear-cut and thorough validation and testing procedures. By hosting this framework online, model developers and users could obtain data from a central location.

Also, resulting benchmarks could be stored as well as version-controlled in the cloud, creating a central place for data and knowledge exchange similar to other approaches aiming at increasing model comparability (Ceola et al., 2015).

In a second step, the underlying causes of these differences should be tackled. Smart coupling of models and their components can help improving the accuracy of inundation extent (chapter 2), simulated discharge (chapters 2, 3), and potentially reduce the required run time (chapters 4, 6.1). Besides, the use of flexible meshes could help achieving fast and locally relevant inundation estimates (chapter 4). As most of the tested models in this thesis are parts of one of the available GFMs (section 1.4.2) already, model coupling could thus be applied to advance the state of GFMs.

Bespoke component combinations could then be created depending on envisaged purpose. To that end, flexible coupling frameworks such as EMELI (Jiang et al., 2017; Peckham, 2014) or GLOFRIM could be applied, creating the modelling environment necessary for demand-driven coupling approaches. That would require, of course, that models and their components are available within an (openly) accessible toolbox which is needed for this sort of “plug-and-play” approaches.

Alternatively, standardized “off-the-shelf” model couplings could be designed with *a priori* designed workflows, variable exchanges, and grid as well as numeric properties. Such pre-defined couplings may be easier to use, but due to the missing flexibility it would be necessary to create one per application, inflating the number of coupling set-ups which may over time reduce oversight and user-friendliness.

However, the success of such a framework depends on the willingness of GFM developers to not only contribute and subject their model to critical assessments, but also make their models more integrative by implementing interfaces like the Open Modelling Interface or the Basic Model Interface (chapters 1.3.2, 5). Even though this would not require opening up the source code *per se*, I personally would welcome moving towards increased open source software in global flood modelling. The current still rather siloed development and validation of GFMs should become more community-driven, eventually maybe even culminating in a “Global Flood Model Intercomparison Project (GFMIP)”. Making this happen is, besides addressing other limitations of GFMs such as the representation of flood defences, thus one of the central challenges for the next generation of GFMs. With flood risk increasing in the future, advanced GFMs are key for defining appropriate risk mitigation and adaption measures.

7.1.3 Model coupling for improved inundation simulations

Research shows that model choice has a marked effect on flood wave propagation and thus on the timing and magnitude of peak flow. Particularly for flat areas and during floods, the kinematic wave approximation employed by many GHMs and land surface models (LSMs) limits the accuracy of river discharge as important physical processes are neglected such as acceleration and pressure terms (chapters 2, 1.4.2; Bates et al. (2010), Yamazaki et al. (2013), and de Almeida and Bates (2013)), hampering the simulation of flood-triggering factors such as backwater effects (cf. Marengo and Espinoza (2016), Meade et al. (1991), and Rudorff et al. (2014a)).

It is important to correctly capture peak flow magnitude and timing as they do not only strongly determine the resulting inundation extent but are also used in operational flood forecasting systems such as the “Global Flood Awareness System (GloFAS)” (Alfieri et al., 2013; Emerton et al., 2018) and forecast-based financing approaches (Jjemba et al., 2018; Coughlan de Perez et al., 2015).

Replacing the default kinematic wave approximation of GHMs, for instance PCR-GLOBWB (PCR), with more sophisticated routing schemes such as the local inertia equation or shallow water equation improves the simulated flood wave propagation for many yet not all river basins, particularly concerning peak discharge and peak arrival (Zhao et al., 2017).

The important role routing plays when it comes to inundation modelling is confirmed by our studies replacing the kinematic wave approximation of PCR with the local inertia equations (LIEs) of LISFLOOD-FP (LFP) and CaMa-Flood (CMF) or the shallow water equations (SWE) of Delft3D Flexible Mesh (DFM) (chapters 2, 3, 6.1). Results also show that, if no supercritical flow conditions apply and flow changes gradually, employing the LIEs is sufficiently complex as the additional advection term in the SWE exerts no influence on simulated discharge (chapter 3; Neal et al. (2012b)). This implies, although not of immediate relevance for this study but maybe for long-term developments, that some models included in GLOFRIM should not be applied for applications where supercritical flow can occur, for instance urban flood modelling (Hunter et al., 2007; Schubert et al., 2008).

Coupling large-scale hydrology with fine-resolution hydrodynamics does not only improve discharge, but also representation of inundations. Chapter 2 shows that the spatial representation of river channels can be improved by routing runoff with a detailed hydrodynamic model instead with a global hydrologic model. It is not only that the general output contains higher resolution, but also inundations along reaches at the sub-grid scale of large-scale hydrology are better represented, especially if the hydrodynamic model contains a 1D network and does not only run in 2D. Yet, the chosen model discretization can also have a marked impact on resulting inundation extent (chapters 3, 4, 6.1) which underlines the need for an educated model choice before simulations should commence. In particular when locally relevant and accurate inundation maps are required (hence, applications that are typically outside the comfort zone of GFMs; see chapter 1.4.1), setting up detailed hydrodynamic models and nesting them into large-scale models can prove beneficial as issues with *a posteriori* downscaling can be avoided (chapter 6.1). Contrariwise, if downscaled inundation maps suffice, parsimonious approaches without nesting will be the less time-consuming way to go.

These results of course do not imply that the routing of a GHM must always be replaced by more advanced routing models. It is clearly the application in mind that must determine which model set-up is applied. The kinematic wave approximation performs well under certain flow conditions (Vieira, 1983; Philipp et al., 2012), for example high-gradient areas (such as mountainous areas) with clearly confined floodplains. As GHMs can also be used to calculate changes in other parts of the hydrologic cycle, for example the impact of water demand and abstractions (cf. De Graaf et al. (2014), Haddeland et al. (2014), Wada et al. (2016), and Wada et al. (2013)), the routing module does not necessarily have to have highest priority as its often only required to close the water cycle (Yamazaki et al., 2013).

Nevertheless, in context of global flood modelling, GHMs should be further developed towards containing more sophisticated routing schemes at finer spatial resolutions. Continuing this “grand challenge” may not only help obtaining “locally relevant” results, but hopefully yield more accurate runoff fields as well which would be key to fully exploit the potential of model coupling (chapters 6.1 and 7.3.1). With available computer power increasing constantly, the bottleneck of heavy computational load should be of the past soon too, increasing the skill of both simulated discharge and inundation extent.

7.2 Challenges of model coupling

The envisaged aim of the PhD project was to improve the simulation of inundation events by model coupling. Important conclusions about the potential of this approach could be drawn as outlined in the previous section. However, there are still quite some challenges for current and future model coupling efforts that have to be acknowledged in any expansion of GLOFRIM or development of other comparable coupling frameworks.

7.2.1 Avoiding "integronsters"

With GLOFRIM 1.0, the routing module of one model can be replaced with the routing scheme of a hydrodynamic model, following a coupled model cascade. Model or model component coupling is thus process-specific and one-directional and does not introduce any additional processes and interdependencies between coupled models or components.

Adding even more models or their components as building blocks for more complex systems representations may be tempting. If the coupling remains process-specific and one-directional, no issues should arise if the coupling follows the BMI and GLOFRIM conventions and, preferably, more general "good practices" of code contribution (Peckham, 2010). If not, however, it may become increasingly challenging to properly relate responses of the coupled model to the trigger in one of the models as (unintentionally) illustrated in chapter 6.2. This may be particularly true if models from "foreign" disciplines are added (Voinov and Shugart, 2013). Careful design of the coupling set-ups is thus required which, if properly implemented and sufficient data is available to validate the correct execution of the coupling procedures, can make coupling frameworks a powerful tool to investigate the interdependencies between different processes in different models. Otherwise, coupled models can be prone to equifinality if wrongly coupled processes still produce accurate results.

While the here presented framework GLOFRIM is less susceptible to intertwine code from different models due to its loose coupling via interfaces (section 1.3.1), it may still happen that coupled processes become too entangled and coupling is performed mechanically, creating so-called "integronsters – constructs that are perfectly valid as software products but ugly or even useless as models" (Voinov and Shugart, 2013). Or, in other words, you should not couple models simply "because you can". Such integronsters may work well in practice, but their scientific usability and meaningfulness are limited. Avoiding them must hence be paramount for any future extension of GLOFRIM or during development of other coupling frameworks. That said, it must always be kept in mind that many problems do not require advanced model coupling set-ups but can be solved with a more parsimonious approach.

For a more elaborated reasoning and (slightly philosophical) discussion on the beauty of model coupling, please consider reading the work by Voinov and Shugart (2013) who provide an in-depth discussion of both opportunities and challenges.

7.2.2 The choice of model parameter values in a coupled setting

The question which model parameter values to choose in a coupled setting is a rather philosophical question as well. As I see it, there are two possibilities: either the parameter sets of each coupled model are calibrated to yield the best possible performance, or a parameter set representative for the study area is chosen and all coupled models are aligned accordingly.

Whilst the first option results in the most accurate results per model and, in turn, of the overall model coupling chain, as exemplified in section 6.1, it would also require that the same model domain is parameterized differently per model although one might argue that there can only be one "true" parameterization representing reality.

Choosing for one parameter set across all coupled models, as done in chapters 2 and 3, would avoid this situation. However, changing the corresponding parameter of different models can have greatly deviating impacts on model results due to differences in model complexity, physics, and therefore sensitivity. For instance, changing the Manning’s floodplain roughness coefficient has less impact in CMF than in LFP (see Figure 7.1) since floodplain storage is only a prognostic variable in CMF whereas floodplain flow is simulated explicitly in LFP.

Hence, the modeller must choose between these two paths. By opting for the first option, the modeller acknowledges that using different sets of parameter values is justified by the fact that different models are different abstractions of reality and thus their parameterizations may very well deviate too. By opting for the second option, however, the modeller consistently reflects the characteristics of the study area in all models but must accept that this may result in less accurate estimates given the different abstractions of reality in computer models.

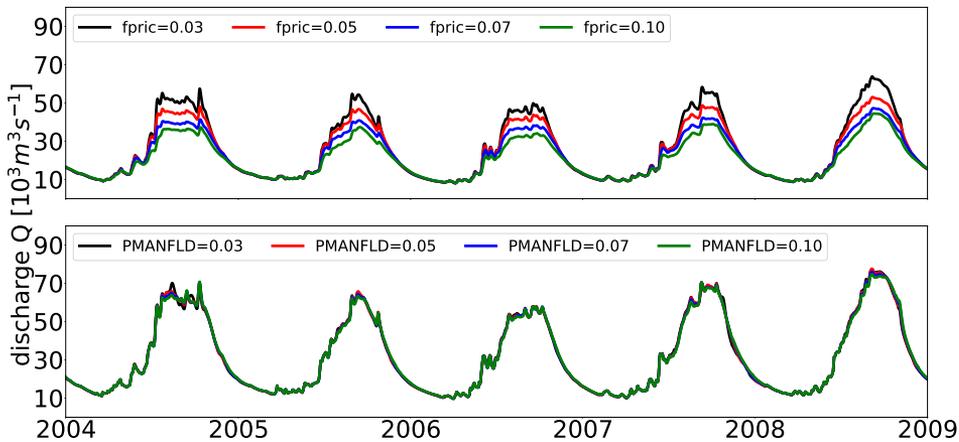


Figure 7.1: Simulated discharge at common observation point (see 6.2 for location) for LFP (top) and CMF (bottom), illustrating the impact variations in floodplain friction (named *fpric* in LFP and *PMANFLD* in CMF) has on model results.

7.3 Recommendations and outlook

7.3.1 The hydrologic model defines what’s possible

GLOFRIM was developed with the aim to facilitate integrating across models and scales, especially with respect to channel and floodplain flow. By employing the more sophisticated routing schemes of hydrodynamic models, the problems emanating from (kinematic wave) routing within GHMs can be overcome and simulated discharge can be improved.

These benefits apply only within certain boundaries since the overall model accuracy is greatly defined by the most upstream model in the coupled model cascade, in the here presented cases this is a hydrologic model. The accuracy of hydrologic models, unfortunately, depends per catchment (Kling et al., 2015; Li et al., 2015; Hattermann et al., 2017).

To illustrate which impact the chosen hydrologic model can have, we applied GLOFRIM 2.0 in the Amazon River basin for the same model domain as used in chapters 2 and 3.

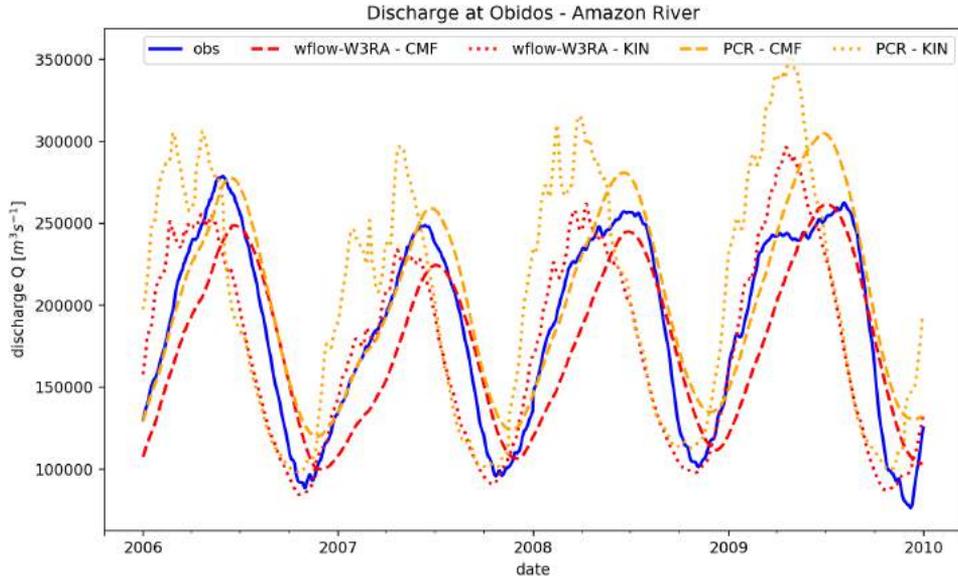


Figure 7.2: Simulated discharge at Obidos (see Figure 2.1) by wflow-W3RA and PCR with both their native kinematic wave routing and coupled to CMF.

| | KGE | KGE(r) | KGE(beta) | KGE(alpha) |
|---------|------|--------|-----------|------------|
| WFL-KIN | 0.78 | 0.80 | 0.94 | 0.83 |
| WFL-CMF | 0.78 | 0.87 | 0.94 | 1.03 |
| PCR-KIN | 0.53 | 0.62 | 1.08 | 1.28 |
| PCR-CMF | 0.88 | 0.93 | 1.08 | 0.93 |

Table 7.1: KGE and its components for runs with different WFL and PCR settings in the Amazon basin.

In addition to running PCR stand-alone and PCR coupled with CMF, we tested the W3RA model contained within wflow (WFL) both with its own kinematic wave routing and coupled with CMF. By using identical meteorological forcing as well as spatial resolution for PCR and W3RA and by coupling simulated runoff to the identical CMF discretization, the impact of using different hydrologic models could be estimated. Results (Table 7.1, Figure 7.2) show that the choice of hydrologic model strongly defines model accuracy, both for stand-alone and coupled runs. Indeed, previous research shows that there can be marked deviations between volume and timing of runoff when compared for various GHMs (Beck et al., 2017). Also, results insinuate that the possible improvement of results by using more complex routing schemes depends on the stand-alone hydrologic performance in the first place.

This implies that coupling hydrology with hydrodynamics can only alleviate situations where magnitude and timing of simulated runoff is (sufficiently) accurate. Since this is not the case for all basins (Zhao et al., 2017), improving runoff estimates in hydrologic models should hence be an important item on the research agenda. Based on remote sensing products (for example using surface water extent (Revilla-Romero et al., 2015) or soil moisture (Massari et al., 2014; Brocca et al., 2010)), it is possible to calibrate model parameter val-

ues for enhanced runoff estimates. Additionally, more (field) data is needed to validate the simulated runoff from hydrologic models and improve it where necessary.

Furthermore, it may be problematic if parameter values of hydrologic models were calibrated using simulated discharge from their own default routing scheme as dependent variable. This default scheme (for instance the kinematic wave approximation) may then produce the right results for the wrong reasons, but if replaced with other routing schemes (for instance the local inertia equation from CMF) it can happen that hydrologic model parameter values and coupled routing scheme do not match anymore, making a new calibration of the hydrologic model necessary if tuning hydrodynamic parameters does not suffice to reduce the bias.

These limitations of model coupling must be kept in mind if applying GLOFRIM. The sole fact that higher-order routing schemes are applied does not, although sounding auspicious, result in better inundation estimates *per se*. What is rather needed is a thorough sensitivity analysis of GHM input data and parameterization to ensure that the best possible runoff fields are simulated. The comparison by Beck et al. (2017) is a promising starting point for future research. It is only then that hydrologic-hydrodynamic model coupling can have a significant positive effect on model results.

7.3.2 Extending the model coupling cascade: adding groundwater, socio-economic impact, and meteorological forcing

Within this thesis, the development and testing of GLOFRIM is described. In its current version 1.0 and transition toward 2.0, it is still strongly focussed on integrating different hydrologic and hydrodynamic models. While this focus on hydrologic-hydrodynamic modelling makes it easier to track exchanged volumes between models, it limits itself to simulating drivers of flood hazard.

In chapter 6.2, first steps towards a coupled groundwater-hydrology-hydrodynamic model are described. Future developments should build upon the progress made to facilitate the integration of both horizontal and vertical groundwater flow on flood hazard. Besides, the importance of groundwater infiltration in many basins makes including a fully-fledged groundwater model a prerequisite for accurate hazard estimates. While we currently merely focus on surface water processes per catchment, including horizontal groundwater flow could allow for additionally considering the influence of inter-basin groundwater flow.

To move on from the hazard estimations and allow for more impact-based assessments, the interface-based coupling approach of GLOFRIM could be employed to establish links with non-hydrologic models focussing on other aspects of the risk calculation cascade.

In principle, each model could be linked if it has a BMI and is aligned with the GLOFRIM standards. Amongst many feasible examples, an interesting one would be to include a crop growth model (such as open source model AquaCrop-OS (Foster et al., 2017)) to better determine time-varying agricultural flood risk. By employing a hydrologic model to simultaneously force the hydrodynamic as well as the crop model, spatio-temporal changes in crop growth could be consistently linked with inundation patterns, yielding estimates of how (seasonal) agricultural flood risk may change over time. The fine-resolution output of hydrodynamic models furthermore allows for more detailed inundation information, potentially allowing for stakeholder assessments (Beven et al., 2015). What is more, additional inundation information can be gathered by simulating hydrodynamics, for instance the duration of inundation, the velocity of floodplain flow, and a more accurate depth compared to coarser hydrologic models.

Another option would be to incorporate the human dimension of flood risk by employing agent-based models. As such assessments are currently done in rather conceptual ways, using a bmi'ed model would allow for actual spatially-explicit assessments, possibly advancing the state of socio-hydrology (Di Baldassarre et al., 2018; Di Baldassarre et al., 2013b; Di Baldassarre et al., 2013a) and providing important insight in achievable (that is, socially accepted) flood mitigation practices.

But there are also other models to be added, with lesser focus on adding an impact layer to the simulations. As the previous chapters show, hydrologic output forces the model coupling cascade. Yet, global hydrology itself is forced with meteorological data. In addition to other factors such as parameterization, meteorological input data can have a marked impact on the accuracy of hydrologic models (cf. Yossef et al. (2013), Pan et al. (2010), and Biemans et al. (2009)).

The coupling cascade could hence be extended with an additional interface linking the hydrologic model with meteorological models. Unfortunately, meteorological models are computationally heavy and thus could become a severe bottleneck for the run times of GLOFRIM (although from the opposite point of view, the flood hazard simulations only have an marginal impact on overall run times of meteorological models). An alternative would therefore be to make use of web services (see eg. Jiang et al. (2017) for an example of models hosted at different servers being coupled via BMI and web services) to run meteorological models on external platforms. For convenience, other interfaces could also be implemented, for example by enabling the download of pre-processed meteorological data sets, for instance ECMWF data via the ECMWF WebAPI.

Either way, coupling via interfaces brings great possibilities and options. However, it must be ensured that the added models fit semantically and logically to GLOFRIM to avoid overfitting the problem. If model coupling is carefully implemented, the here presented GLOFRIM framework can be a starting point for more holistic modelling approaches, covering the entire flood risk chain: meteorological conditions, hydrologic response, rainfall-runoff conversion, routing along channels and floodplains, and socio-economic consequences. Including both hazard and risk into GLOFRIM would form the basis for a transition towards applying model coupling for policy and decision-making.

Appendix A

Data and code availability

GLOFRIM code

All code of GLOFRIM can be found online and is freely accessible. It is licensed under the GPL GNU v3.0 license – please consider the license-specific restrictions (see <https://choosealicense.com/licenses/gpl-3.0/>).

The code used for GLOFRIM 1.0 (as used in chapters 3 and 4) can be downloaded from <https://doi.org/10.5281/zenodo.1037990>⁵.

The code used for GLOFRIM 2.0 is not yet officially released as the associated publication is still under review. Chapter 6.1 contains a first introduction into GLOFRIM 2.0 and changes compared to earlier versions. Until GLOFRIM 2.0 is officially released it can be downloaded or pulled from the GLOFRIM openearth-repository at <https://github.com/openearth/glofrim> or as beta-version from <https://doi.org/10.5281/zenodo.2589891>.

GLOFRIM 1.0 and also 2.0 are documented using ReadTheDocs. The webpage, which is under constant development as code is still changing, can be found here: <https://glofrim.readthedocs.io/en/latest/>. It contains the most relevant information concerning the framework as well as its installation and execution, the models including their documentation, a brief installation guide, and information how to collaborate with and contribute to the GLOFRIM project.

PCR-GLOBWB

PCR-GLOBWB is a global hydrologic model developed at the Department of Physical Geography, Utrecht University. For information regarding model development, design, forcing, and applications, we refer to the key literature (Sutanudjaja et al., 2018; van Beek, 2008; van Beek and Bierkens, 2008; van Beek et al., 2011). By default, PCR-GLOBWB does not contain any BMI functionality.

The latest bmi'ed version of PCR-GLOBWB is hosted under <https://doi.org/10.5281/zenodo.1472347>. For installing PCR-GLOBWB and the required packages, please read the GLOFRIM online documentation. Before applying the model, please consider reading the following information points as well:

1. The bmi'ed PCR-GLOBWB version was built upon an up-to-date version at the begin of the framework development. Since development of the bmi'ed version, the default version proceeded in its development and functionality while the bmi'ed did not. Some

⁵Besides GLOFRIM code, the downloadable zip-file also contains the bmi'd versions of PCR-GLOBWB and LISFLOOD-FP. Note that after GLOFRIM 1.0 was released, various changes were made to the model code and therefore the downloadable PCR-GLOBWB and LISFLOOD-FP are not up-to-date. Please see the relevant chapters below for further information.

of the functions or processes you may be used to from the higher version numbers may thus not be available in the bmi'ed model.

2. The model was developed, tested, and applied on Linux system. Applying it on Windows platforms is thus not (yet) possible. Contributions in this matter are very welcome!
3. Additionally, there are also other bmi'ed versions of PCR-GLOBWB available (eg. for the eWaterCycle project; see <https://www.ewatercycle.org/>) which do not correspond with the model code used here and therefore may not be compatible with GLOFRIM. It's worth a test though!

WFLOW

wflow is a hydrologic modelling platform developed at Deltares and caters for a wide range of applications and set-ups due to the various models contained. Amongst others, it contains the hydrologic model routines of HBV, SBM, and W3RA. Since recently, it also contains PCR-GLOBWB. The wflow platform can be downloaded from GitHub (<https://github.com/openstreams/wflow/tree/python2>) and its documentation can be found at <https://wflow.readthedocs.io/en/latest/>.

Due to the on-going overall migration towards Python 3, please read the following important note:

- The latest release of wflow is developed using Python 3. Since GLOFRIM was not yet migrated to Python 3, the Python 2 version of wflow has to be used. Please ensure the right version is installed on your system to ensure compatibility.

CaMa-Flood

CaMa-Flood is a global hydrodynamic model developed at the University of Tokyo. Without any BMI functionality by default, a non-bmi'ed version of the most recent version of CaMa-Flood can be downloaded from <http://hydro.iis.u-tokyo.ac.jp/~yamadai/cama-flood/> while the manual can be found here http://hydro.iis.u-tokyo.ac.jp/~yamadai/cama-flood/Manual_CaMa-Flood_v362.pdf.

Unfortunately, the bmi'ed version is currently not yet openly available but can be obtained by contacting the GLOFRIM development team only. We hope to make the model openly accessible as soon as possible. If you want to use the bmi'ed version, please contact the CaMa-Flood or the GLOFRIM development team.

Please note that the bmi'ed version used for GLOFRIM 2.0 corresponds with the currently most up-to-date version of CaMa-Flood which is 3.6.2. In case the default version is updated in the future, we hope to also update the bmi'ed counterpart but cannot promise this at the moment. Please be hence aware of possible deviations and carefully check that you use the right model version.

LISFLOOD-FP

The LISFLOOD-FP model is a raster-based hydrodynamic model developed at the School of Geographical Sciences, University of Bristol. Similar to most models above, it does not contain a BMI by default. Besides, the default model code is not (yet) openly available. The

most up-to-date release of the bmi'ed LISFLOOD-FP model can be downloaded from <https://doi.org/10.5281/zenodo.1479836>. A description of the (not bmi'ed) model is provided at <http://www.bristol.ac.uk/geography/research/hydrology/models/lisflood/>. The model development of LISDFLOOD-FP is documented with several scientific publications, most notably Bates and De Roo (2000), Bates et al. (2010), and Neal et al. (2012a).

When applying the bmi'ed LISFLOOD-FP model, please note the following:

1. It was built upon version 5.9 of the original model. Since the development of the bmi'ed version, the default version proceeded in its development and functionality while the bmi'ed version did not. Some of the functions and processes you may be used to from the higher version numbers than 5.9 may thus not be available in the bmi'ed model.
2. The model was developed for Linux systems only. Applying it on Windows platforms is thus not (yet) possible. Feel free to add a compiled executable, it is very much welcomed.

Delft3D Flexible Mesh

Delft3D Flexible Mesh is a hydrodynamic model developed at Deltares allowing for the use of various grid types such as regular, irregular, and curvilinear grids. Its numerical solver is described by Kernkamp et al. (2011). It contains a wide range of possible applications which are documented in the technical manual to be found at https://content.oss.deltares.nl/delft3d/manuals/D-Flow_FM_Technical_Reference_Manual.pdf. A user manual is available at https://content.oss.deltares.nl/delft3d/manuals/D-Flow_FM_User_Manual.pdf.

Besides, the following is important to know if you are intending to apply Delft3D Flexible Mesh:

1. The Delft3D Flexible Mesh model is not yet openly available but it is estimated that this will happen in 2019. Until then, the model is unfortunately only available upon request.
2. The model code and thus model functions are still under development. It may be worthwhile checking for additional or new features which were not yet available at the time of writing.

References

- Alarcon, V. J., D. Johnson, W. H. McAnally, J. van der Zwaag, D. Irby, and J. Cartwright (2014). Nested Hydrodynamic Modeling of a Coastal River Applying Dynamic-Coupling. *Water Resources Management* 28.10, pp. 3227–3240. doi:10.1007/s11269-014-0671-6.
- Alcamo, J., P. Döll, F. Kaspar, and S. Siebert (1997). Global change and global scenarios of water use and availability: an application of WaterGAP 1.0. *Center for Environmental Systems Research (CESR), University of Kassel, Germany* 1720.
- Alfieri, L., P. Burek, E. Dutra, B. Krzeminski, D. Muraro, J. Thielen, and F. Pappenberger (2013). GloFAS – global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences* 17.3, pp. 1161–1175. doi:10.5194/HESS-17-1161-2013.
- Alfieri, L., P. Salamon, A. Bianchi, J. C. Neal, P. D. Bates, and L. Feyen (2014). Advances in pan-European flood hazard mapping. *Hydrological Processes* 28.13, pp. 4067–4077. doi:10.1002/HYP.9947.
- Alfieri, L. et al. (2018). A global network for operational flood risk reduction. *Environmental Science & Policy* 84, pp. 149–158. doi:10.1016/J.ENVSCI.2018.03.014.
- Allen, G. H. and T. M. Pavelsky (2018). Global extent of rivers and streams. *Science* 361.6402, pp. 585–588. doi:10.1126/SCIENCE.AAT0636.
- Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998). *Crop evapotranspiration*.
- Altenau, E. H., T. M. Pavelsky, P. D. Bates, and J. C. Neal (2017). The effects of spatial resolution and dimensionality on modeling regional-scale hydraulics in a multichannel river. *Water Resources Research* 53.2, pp. 1683–1701. doi:10.1002/2016WR019396.
- Andreadis, K. M., G. J.-P. Schumann, and T. M. Pavelsky (2013). A simple global river bankfull width and depth database: Data and Analysis Note. *Water Resources Research* 49.10, pp. 7164–7168. doi:10.1002/WRCR.20440.
- Aqueduct Global Flood Analyzer* (2019). URL: <http://floods.wri.org> (visited on 01/28/2019).
- Ashby, S. F. and R. D. Falgout (1996). A Parallel Multigrid Preconditioned Conjugate Gradient Algorithm for Groundwater Flow Simulations. *Nuclear Science and Engineering* 124.1, pp. 145–159. doi:10.13182/NSE96-A24230.
- Bahrainian, S. S. and Z. Mehrdoost (2012). An automatic unstructured grid generation method for viscous flow simulations. *Mathematics and Computers in Simulation* 83, pp. 23–43. doi:10.1016/J.MATCOM.2012.05.021.
- Baker, V. R. (1987). Paleoflood hydrology and extraordinary flood events. *Journal of Hydrology* 96.1, pp. 79–99. doi:10.1016/0022-1694(87)90145-4.
- Baker, V. R. (2006). Palaeoflood hydrology in a global context. *CATENA* 66.1, pp. 161–168. doi:10.1016/J.CATENA.2005.11.016.
- Ballesteros-Cánovas, J. A., M. Stoffel, S. St George, and K. Hirschboeck (2015). A review of flood records from tree rings. *Progress in Physical Geography* 39.6, pp. 794–816. doi:10.1177/0309133315608758.
- Bates, P. D. and A. P. J. De Roo (2000). A simple raster-based model for flood inundation simulation. *Journal of Hydrology* 236.1, pp. 54–77. doi:10.1016/S0022-1694(00)00278-X.
- Bates, P. D., M. S. Horritt, and T. J. Fewtrell (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology* 387.1-2, pp. 33–45. doi:10.1016/J.JHYDROL.2010.03.027.
- Bates, P. D., J. C. Neal, C. C. Sampson, A. Smith, and M. A. Trigg (2018). Chapter 9 - Progress Toward Hyperresolution Models of Global Flood Hazard. In: *Risk Modeling for Hazards and Disasters*. Ed. by G. Michel. Elsevier, pp. 211–232. doi:10.1016/B978-0-12-804071-3.00009-4.
- Baugh, C. A., P. D. Bates, G. J.-P. Schumann, and M. A. Trigg (2013). SRTM vegetation removal and hydrodynamic modeling accuracy. *Water Resources Research* 49.9, pp. 5276–5289. doi:10.1002/WRCR.20412.
- Beck, H. E., A. I. J. M. van Dijk, A. de Roo, E. Dutra, G. Fink, R. Orth, and J. Schellekens (2017). Global evaluation of runoff from 10 state-of-the-art hydrological models. *Hydrology and Earth System Sciences* 21.6, pp. 2881–2903. doi:10.5194/HESS-21-2881-2017.

- Becker, B. and A. Burzel (2016). Model Coupling with OpenMI - Introduction of Basic Concepts. In: *Managing the Complexity of Critical Infrastructures: A Modelling and Simulation Approach*. Ed. by R. Setola, V. Rosato, E. Kyriakides, and E. Rome. Studies in Systems, Decision and Control. Cham: Springer International Publishing, pp. 279–299. DOI:10.1007/978-3-319-51043-9_11.
- Becker, B. and J. Talsma (2014). On The External And Iterative Coupling Of Multiple Open Channel Flow Models With OpenMI. *International Conference on Hydroinformatics*.
- Bell, B. (1970). The Oldest Records of the Nile Floods. *The Geographical Journal* 136.4, p. 569. DOI:10.2307/1796184.
- Bernhofen, M. V. et al. (2018). A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique. *Environmental Research Letters* 13.10, p. 104007. DOI:10.1088/1748-9326/AAE014.
- Berrisford, P. et al. (2011). *The ERA-Interim archive Version 2.0*. 1. Shinfield Park, Reading: ECMWF, p. 23.
- Berry, P. A. M., J. D. Garlick, and R. G. Smith (2007). Near-global validation of the SRTM DEM using satellite radar altimetry. *Remote Sensing of Environment* 106.1, pp. 17–27. DOI:10.1016/J.RSE.2006.07.011.
- Beven, K. J. and H. L. Cloke (2012). Comment on Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water by Eric F. Wood et al. *Water Resources Research* 48.1. DOI:10.1029/2011WR010982.
- Beven, K. J., H. L. Cloke, F. Pappenberger, R. Lamb, and N. Hunter (2015). Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface. *Science China Earth Sciences* 58.1, pp. 25–35. DOI:10.1007/s11430-014-5003-4.
- Biancamaria, S., P. D. Bates, A. Boone, and N. M. Mognard (2009). Large-scale coupled hydrologic and hydraulic modelling of the Ob river in Siberia. *Journal of Hydrology* 379.1, pp. 136–150. DOI:10.1016/J.JHYDROL.2009.09.054.
- Biemans, H., R. W. A. Hutjes, P. Kabat, B. J. Strengers, D. Gerten, and S. Rost (2009). Effects of Precipitation Uncertainty on Discharge Calculations for Main River Basins. *Journal of Hydrometeorology* 10.4, pp. 1011–1025. DOI:10.1175/2008JHM1067.1.
- Bierkens, M. F. P. (2015). Global hydrology 2015: State, trends, and directions: Global Hydrology 2015. *Water Resources Research* 51.7, pp. 4923–4947. DOI:10.1002/2015WR017173.
- Bierkens, M. F. P. et al. (2015). Hyper-resolution global hydrological modelling: what is next?: Everywhere and locally relevant. *Hydrological Processes* 29.2, pp. 310–320. DOI:10.1002/HYP.10391.
- Brocca, L., F. Melone, T. Moramarco, W. Wagner, V. Naeimi, Z. Bartalis, and S. Hasenauer (2010). Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrology and Earth System Sciences* 14.10, pp. 1881–1893. DOI:HTTPS://DOI.ORG/10.5194/HESS-14-1881-2010.
- Brunner, P. and C. T. Simmons (2012). HydroGeoSphere: a fully integrated, physically based hydrological model. *Groundwater* 50.2, pp. 170–176. DOI:10.1111/J.1745-6584.2011.00882.x.
- Buahin, C. A. and J. S. Horsburgh (2018). Advancing the Open Modeling Interface (OpenMI) for integrated water resources modeling. *Environmental Modelling & Software* 108, pp. 133–153. DOI:10.1016/J.ENVSOFT.2018.07.015.
- Butts, M. et al. (2014). Embedding complex hydrology in the regional climate system Dynamic coupling across different modelling domains. *Advances in Water Resources* 74, pp. 166–184. DOI:10.1016/J.ADVWATRES.2014.09.004.
- Candogan Yossef, N., L. P. H. van Beek, J. C. J. Kwadijk, and M. F. P. Bierkens (2012). Assessment of the potential forecasting skill of a global hydrological model in reproducing the occurrence of monthly flow extremes. *Hydrology and Earth System Sciences* 16.11, pp. 4233–4246. DOI:10.5194/HESS-16-4233-2012.
- Carabajal, C. C. and D. J. Harding (2006). SRTM C-band and ICESat laser altimetry elevation comparisons as a function of tree cover and relief. *Photogrammetric Engineering & Remote Sensing* 72.3, pp. 287–298. DOI:10.14358/PERS.72.3.287.
- Castro Gama, M., I. Popescu, A. Mynett, L. Shengyang, and A. van Dam (2013). Modelling extreme flood hazard events on the middle Yellow River using DFLOW-flexible mesh approach. *Natural Hazards and Earth System Sciences Discussions* 1.6, pp. 6061–6092. DOI:10.5194/NHESSD-1-6061-2013.
- Castronova, A. M., J. L. Goodall, and M. B. Ercan (2013a). Integrated modeling within a Hydrologic Information System: An OpenMI based approach. *Environmental Modelling & Software*. Thematic Issue on the Future of Integrated Modeling Science and Technology 39, pp. 263–273. DOI:10.1016/J.ENVSOFT.2012.02.011.
- Castronova, A. M., J. L. Goodall, and M. M. Elag (2013b). Models as web services using the Open Geospatial Consortium (OGC) Web Processing Service (WPS) standard. *Environmental Modelling & Software* 41, pp. 72–83. DOI:10.1016/J.ENVSOFT.2012.11.010.

- Ceola, S. et al. (2015). Virtual laboratories: new opportunities for collaborative water science. *Hydrology and Earth System Sciences* 19.4, pp. 2101–2117. DOI:10.5194/HESS-19-2101-2015.
- Ceola, S., F. Laio, and A. Montanari (2014). Satellite nighttime lights reveal increasing human exposure to floods worldwide. *Geophysical Research Letters* 41.20, pp. 7184–7190. DOI:10.1002/2014GL061859.
- Chen, C., H. Liu, and R. C. Beardsley (2003). An Unstructured Grid, Finite-Volume, Three-Dimensional, Primitive Equations Ocean Model: Application to Coastal Ocean and Estuaries. *JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY* 20, p. 28.
- Chow, V., D. Maidment, and L. Mays (1988). Applied hydrology, 572 pp. Editions McGraw-Hill, New York.
- Clark, M. P. et al. (2015a). A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research* 51.4, pp. 2498–2514. DOI:10.1002/2015WR017198.
- Clark, M. P. et al. (2015b). A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resources Research* 51.4, pp. 2515–2542. DOI:10.1002/2015WR017200.
- Collenteur, R. A., H. de Moel, B. Jongman, and G. Di Baldassarre (2015). The failed-levee effect: Do societies learn from flood disasters? *Natural Hazards* 76.1, pp. 373–388. DOI:10.1007/s11069-014-1496-6.
- Coughlan de Perez, E., B. van den Hurk, M. K. van Aalst, B. Jongman, T. Klose, and P. Suarez (2015). Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts. *Natural Hazards and Earth System Sciences* 15.4, pp. 895–904. DOI:10.5194/NHESS-15-895-2015.
- CSDMS (2019a). *CSDMS Basic Model Interface (version 1.0)*. URL: https://csdms.colorado.edu/wiki/BMI_Description (visited on 01/11/2019).
- CSDMS (2019b). *CSDMS Web Modelling Tool*. URL: <https://csdms.colorado.edu/wmt/> (visited on 01/11/2019).
- Dadson, S. J., I. Ashpole, P. Harris, H. N. Davies, D. B. Clark, E. Blyth, and C. M. Taylor (2010). Wetland inundation dynamics in a model of land surface climate: Evaluation in the Niger inland delta region. *Journal of Geophysical Research: Atmospheres* 115 (D23). DOI:10.1029/2010JD014474.
- David, O., J. C. Ascough, W. Lloyd, T. R. Green, K. W. Rojas, G. H. Leavesley, and L. R. Ahuja (2013). A software engineering perspective on environmental modeling framework design: The Object Modeling System. *Environmental Modelling & Software*. Thematic Issue on the Future of Integrated Modeling Science and Technology 39, pp. 201–213. DOI:10.1016/J.ENVSOF.2012.03.006.
- Davison, J. H., H.-T. Hwang, E. A. Sudicky, D. V. Mallia, and J. C. Lin (2018). Full Coupling Between the Atmosphere, Surface, and Subsurface for Integrated Hydrologic Simulation. *Journal of Advances in Modeling Earth Systems* 10.1, pp. 43–53. DOI:10.1002/2017MS001052.
- De Almeida, G. A. M. and P. D. Bates (2013). Applicability of the local inertial approximation of the shallow water equations to flood modeling. *Water Resources Research* 49.8, pp. 4833–4844. DOI:10.1002/WRCR.20366.
- De Boer-Euser, T. et al. (2017). Looking beyond general metrics for model comparison – lessons from an international model intercomparison study. *Hydrology and Earth System Sciences* 21.1, pp. 423–440. DOI:10.5194/HESS-21-423-2017.
- De Graaf, I. E. M., E. H. Sutanudjaja, L. P. H. van Beek, and M. F. P. Bierkens (2015). A high-resolution global-scale groundwater model. *Hydrology and Earth System Sciences* 19.2, pp. 823–837. DOI:HTTPTS://DOI.ORG/10.5194/HESS-19-823-2015.
- De Graaf, I. E. M., L. P. H. van Beek, T. Gleeson, N. Moosdorf, O. Schmitz, E. H. Sutanudjaja, and M. F. P. Bierkens (2017). A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. *Advances in Water Resources* 102, pp. 53–67. DOI:10.1016/J.ADVWATRES.2017.01.011.
- De Graaf, I. E. M., L. P. H. van Beek, Y. Wada, and M. F. P. Bierkens (2014). Dynamic attribution of global water demand to surface water and groundwater resources: Effects of abstractions and return flows on river discharges. *Advances in Water resources* 64, pp. 21–33. DOI:10.1016/J.ADVWATRES.2013.12.002.
- Delft3D Flexible Mesh Technical Manual* (2019). URL: https://content.oss.deltares.nl/delft3d/manuals/D-Flow_FM_Technical_Reference_Manual.pdf (visited on 01/30/2019).
- Delft3D Flexible Mesh User Manual* (2019). URL: https://content.oss.deltares.nl/delft3d/manuals/D-Flow_FM_User_Manual.pdf (visited on 01/30/2019).
- Di Baldassarre, G., A. Castellarin, and A. Brath (2009). Analysis of the effects of levee heightening on flood propagation: example of the River Po, Italy. *Hydrological Sciences Journal* 54.6, pp. 1007–1017. DOI:10.1623/HYSJ.54.6.1007.

- Di Baldassarre, G., A. Viglione, G. Carr, L. Kuil, J. L. Salinas, and G. Blöschl (2013a). Socio-hydrology: conceptualising human-flood interactions. *Hydrology and Earth System Sciences* 17.8, pp. 3295–3303. DOI:10.5194/HESS-17-3295-2013.
- Di Baldassarre, G., M. Kooy, J. S. Kemerink, and L. Brandimarte (2013b). Towards understanding the dynamic behaviour of floodplains as human-water systems. *Hydrology and Earth System Sciences* 17.8, pp. 3235–3244. DOI:10.5194/HESS-17-3235-2013.
- Di Baldassarre, G. et al. (2018). Hess Opinions: An interdisciplinary research agenda to explore the unintended consequences of structural flood protection. *Hydrology and Earth System Sciences Discussions*, pp. 1–11. DOI:10.5194/HESS-2018-333.
- Dimitriadis, P., A. Tegos, A. Oikonomou, V. Pagana, A. Koukouvinos, N. Mamassis, D. Koutsoyiannis, and A. Efstratiadis (2016). Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. *Journal of Hydrology* 534, pp. 478–492. DOI:10.1016/J.JHYDROL.2016.01.020.
- Döll, P., F. Kaspar, and B. Lehner (2003). A global hydrological model for deriving water availability indicators: model tuning and validation. *Journal of Hydrology* 270.1-2, pp. 105–134. DOI:10.1016/S0022-1694(02)00283-4.
- Dottori, F., P. Salamon, A. Bianchi, L. Alfieri, F. A. Hirpa, and L. Feyen (2016). Development and evaluation of a framework for global flood hazard mapping. *Advances in Water Resources* 94, pp. 87–102. DOI:10.1016/J.ADVWATRES.2016.05.002.
- Dzigan, Y et al. (2019). eWaterCycle II: an online environment for explorative computational Hydrology. In: EGU General Assembly Conference Abstracts. Vol. 21, p. 12392.
- ECMWF (2019). *ERA5 data documentation*. URL: <https://confluence.ecmwf.int//display/CKB/ERA5+data+documentation> (visited on 01/28/2019).
- Emerton, R., E. Zsoter, L. Arnal, H. L. Cloke, D. Muraro, C. Prudhomme, E. M. Stephens, P. Salamon, and F. Pappenberger (2018). Developing a global operational seasonal hydro-meteorological forecasting system: GloFAS-Seasonal v1.0. *Geoscientific Model Development* 11, pp. 3327–3346. DOI:10.5194/GMD-2018-118.
- FAOSTAT (2017). URL: <http://www.fao.org/faostat/en/#home> (visited on 01/13/2017).
- Farr, T. G. and M. Kobrick (2000). Shuttle Radar Topography Mission produces a wealth of data. *Eos, Transactions, American Geophysical Union* 81.48, pp. 583–585. DOI:10.1029/EO0811048P00583.
- Felder, G., A. P. Zischg, and R. Weingartner (2017). The effect of coupling hydrologic and hydrodynamic models on probable maximum flood estimation. *Journal of Hydrology* 550, pp. 157–165. DOI:10.1016/J.JHYDROL.2017.04.052.
- Felder, G., J. J. Gómez-Navarro, A. P. Zischg, C. C. Raible, V. Röthlisberger, D. Bozhinova, O. Martius, and R. Weingartner (2018). From global circulation to local flood loss: Coupling models across the scales. *Science of The Total Environment* 635, pp. 1225–1239. DOI:10.1016/J.SCITOTENV.2018.04.170.
- Fewtrell, T. J., P. D. Bates, M. S. Horritt, and N. M. Hunter (2008). Evaluating the effect of scale in flood inundation modelling in urban environments. *Hydrological Processes* 22.26, pp. 5107–5118. DOI:10.1002/HYP.7148.
- Finaud-Guyot, P., C. Delenne, V. Guinot, and C. Llovel (2011). 1D2D coupling for river flow modeling. *Comptes Rendus Mécanique* 339.4, pp. 226–234. DOI:10.1016/J.CRME.2011.02.001.
- Foster, T., N. Brozovi, A. P. Butler, C. M. U. Neale, D. Raes, P. Steduto, E. Fereres, and T. C. Hsiao (2017). AquaCrop-OS: An open source version of FAO's crop water productivity model. *Agricultural Water Management* 181, pp. 18–22. DOI:10.1016/J.AGWAT.2016.11.015.
- Getirana, A. C. V., M.-P. Bonnet, and J.-M. Martinez (2009). Evaluating parameter effects in a DEM burning process based on land cover data. *Hydrological Processes* 23.16, pp. 2316–2325. DOI:10.1002/HYP.7303.
- Giustarini, L., M. Chini, R. Hostache, F. Pappenberger, and P. Matgen (2015). Flood Hazard Mapping Combining Hydrodynamic Modeling and Multi Annual Remote Sensing data. *Remote Sensing* 7.10, pp. 14200–14226. DOI:10.3390/RS71014200.
- GloFAS (2018). *Home*. Global Flood Awareness System global ensemble streamflow forecasting and flood early warning. URL: <http://www.globalfloods.eu/> (visited on 10/30/2018).
- Goodall, J. L. and S. D. Peckham (2016). Interoperability between the Basic Modeling Interface (BMI) and the Open Modeling Interface (OpenMI): A Step Toward Building the Earth System Bridge for Modeling Framework Interoperability. In: International Congress on Environmental Modelling and Software. Toulouse, p. 2.
- Greenbaum, N., U. Schwartz, G. Benito, N. Porat, G. C. Cloete, and Y. Enzel (2014). Paleohydrology of extraordinary floods along the Swakop River at the margin of the Namib Desert and their paleoclimate

- implications. *Quaternary Science Reviews* 103, pp. 153–169. DOI:10.1016/J.QUASCIREV.2014.08.021.
- Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology* 377.1, pp. 80–91. DOI:10.1016/J.JHYDROL.2009.08.003.
- Haddeland, I. et al. (2014). Global water resources affected by human interventions and climate change. *Proceedings of the National Academy of Sciences* 111.9, pp. 3251–3256. DOI:10.1073/PNAS.1222475110.
- Hagemann, S. and L. D. Gates (2003). Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations. *Climate Dynamics* 21.3, pp. 349–359. DOI:10.1007/s00382-003-0349-x.
- Hagen, E. and X. Lu (2011). Let us create flood hazard maps for developing countries. *Natural hazards* 58.3, pp. 841–843. DOI:10.1007/s11069-011-9750-7.
- Hamon, W. R. (1963). Computation of direct runoff amounts from storm rainfall. *International Association of Scientific Hydrology Publication* 63, pp. 52–62.
- Harbaugh, A. W., E. R. Banta, M. C. Hill, and M. G. McDonald (2000). MODFLOW-2000, The U. S. Geological Survey Modular Ground-Water Model-User Guide to Modularization Concepts and the Ground-Water Flow Process. *Open-file Report. U. S. Geological Survey* 92, p. 134.
- Hardy, R., P. Bates, and M. Anderson (1999). The importance of spatial resolution in hydraulic models for floodplain environments. *Journal of Hydrology* 216.1, pp. 124–136. DOI:10.1016/S0022-1694(99)00002-5.
- Harris, I. P. D. J., P. D. Jones, T. J. Osborn, and D. H. Lister (2014). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology* 34.3, pp. 623–642. DOI:10.1002/JOC.3711.
- Hattermann, F. F. et al. (2017). Crossscale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins. *Climatic Change* 141.3, pp. 561–576. DOI:10.1007/s10584-016-1829-4.
- Hess, L. (2003). Dual-season mapping of wetland inundation and vegetation for the central Amazon basin. *Remote Sensing of Environment* 87.4, pp. 404–428. DOI:10.1016/J.RSE.2003.04.001.
- Hill, C., C. DeLuca, Balaji, M. Suarez, and A. Da Silva (2004). The architecture of the earth system modeling framework. *Computing in Science & Engineering* 6.1, pp. 18–28. DOI:10.1109/MCISE.2004.1255817.
- Hirabayashi, Y., R. Mahendran, S. Koirala, L. Konoshima, D. Yamazaki, S. Watanabe, H. Kim, and S. Kanae (2013). Global flood risk under climate change. *Nature Climate Change* 3.9, pp. 816–821. DOI:10.1038/NCLIMATE1911.
- Hirpa, F. A., P. Salamon, L. Alfieri, J. Thielen-del Pozo, E. Zsoter, and F. Pappenberger (2016). The Effect of Reference Climatology on Global Flood Forecasting. *Journal of Hydrometeorology* 17.4, pp. 1131–1145. DOI:10.1175/JHM-D-15-0044.1.
- Hoch, J. and M. A. Trigg (2019). Advancing global flood hazard simulations by improving comparability, benchmarking, and integration of global flood models. *Environmental Research Letters* 14.3, p. 034001. DOI:10.1088/1748-9326/AAF3D3.
- Hoch, J. M., A. V. Haag, A. van Dam, H. C. Winsemius, L. P. H. van Beek, and M. F. P. Bierkens (2017a). Assessing the impact of hydrodynamics on large-scale flood wave propagation – a case study for the Amazon Basin. *Hydrology and Earth System Sciences* 21.1, pp. 117–132. DOI:10.5194/HESS-21-117-2017.
- Hoch, J. M., J. C. Neal, F. Baart, L. P. H. van Beek, H. C. Winsemius, P. D. Bates, and M. F. P. Bierkens (2017b). GLOFRIM v1.0 – A globally applicable computational framework for integrated hydrological-hydrodynamic modelling. *Geoscientific Model Development* 10.10, pp. 3913–3929. DOI:10.5194/GMD-10-3913-2017.
- Hoch, J. M., L. P. H. van Beek, H. C. Winsemius, and M. F. P. Bierkens (2018a). Benchmarking flexible meshes and regular grids for large-scale fluvial inundation modelling. *Advances in Water Resources* 121, pp. 350–360. DOI:10.1016/J.ADVWATRES.2018.09.003.
- Hoch, J. M., D. Eilander, and H. Ikeuchi (2018b). openearth/glofrim: Alpha dev towards GLOFRIM 2.0. *Zenodo*. DOI:10.5281/ZENODO.1452600.
- Horritt, M. S. and P. D. Bates (2001a). Effects of spatial resolution on a raster based model of flood flow. *Journal of Hydrology*, p. 11. DOI:10.1016/S0022-1694(01)00490-5.
- Horritt, M. S. and P. D. Bates (2001b). Predicting floodplain inundation: raster-based modelling versus the finite-element approach. *Hydrological Processes* 15.5, pp. 825–842. DOI:10.1002/HYP.188.

- Horritt, M. S., P. D. Bates, and M. Mattinson (2006). Effects of mesh resolution and topographic representation in 2D finite volume models of shallow water fluvial flow. *Journal of Hydrology* 329.1, pp. 306–314. DOI:10.1016/J.JHYDROL.2006.02.016.
- Huang, C., Y. Chen, and J. Wu (2014). Mapping spatio-temporal flood inundation dynamics at large river basin scale using time-series flow data and MODIS imagery. *International Journal of Applied Earth Observation and Geoinformation* 26, pp. 350–362. DOI:10.1016/J.JAG.2013.09.002.
- Huang, C. C., J. Pang, X. Zha, Y. Zhou, H. Su, and Y. Li (2010). Extraordinary Floods of 4100-4000a BP recorded at the Late Neolithic Ruins in the Jinghe River Gorges, Middle Reach of the Yellow River, China. *Palaeogeography, Palaeoclimatology, Palaeoecology* 289.1, pp. 1–9. DOI:10.1016/J.PALAEO.2010.02.003.
- Hunter, N. M. et al. (2008). Benchmarking 2D hydraulic models for urban flooding. *Proceedings of the Institution of Civil Engineers - Water Management* 161.1, pp. 13–30. DOI:10.1680/WAMA.2008.161.1.13.
- Hunter, N. M., P. D. Bates, M. S. Horritt, and M. D. Wilson (2007). Simple spatially-distributed models for predicting flood inundation: A review. *Geomorphology* 90.3, pp. 208–225. DOI:10.1016/J.GEOMORPH.2006.10.021.
- Hut, R., N. van de Giessen, and N. Drost (2018). The future of global is local. eWaterCycle II: bridging the gap between catchment hydrologists and global hydrologists. In: EGU General Assembly Conference Abstracts. Vol. 20, p. 10614.
- Ikeuchi, H., Y. Hirabayashi, D. Yamazaki, M. Kiguchi, S. Koirala, T. Nagano, A. Kotera, and S. Kanae (2015). Modeling complex flow dynamics of fluvial floods exacerbated by sea level rise in the Ganges-BrahmaputraMeghna Delta. *Environmental Research Letters* 10.12, p. 124011. DOI:10.1088/1748-9326/10/12/124011.
- Ikeuchi, H., Y. Hirabayashi, D. Yamazaki, S. Muis, P. J. Ward, H. C. Winsemius, M. Verlaan, and S. Kanae (2017). Compound simulation of fluvial floods and storm surges in a global coupled river-coast flood model: Model development and its application to 2007 Cyclone Sidr in Bangladesh. *Journal of Advances in Modeling Earth Systems* 9.4, pp. 1847–1862. DOI:10.1002/2017MS000943.
- Inomata, H. and K. Fukami (2008). Restoration of historical hydrological data of Tonle Sap Lake and its surrounding areas. *Hydrological Processes* 22.9, pp. 1337–1350. DOI:10.1002/HYP.6943.
- IPCC (2014). Terrestrial and Inland Water Systems. In: *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, pp. 271–359.
- Jiang, P., M. Elag, P. Kumar, S. D. Peckham, L. Marini, and L. Rui (2017). A service-oriented architecture for coupling web service models using the Basic Model Interface (BMI). *Environmental Modelling & Software* 92, pp. 107–118. DOI:10.1016/J.ENVSOFT.2017.01.021.
- Jjemba, E. W., B. K. Mwebaze, J. Arrighi, E. Coughlan de Perez, and M. Bailey (2018). Forecast-Based Financing and Climate Change Adaptation: Uganda Makes History Using Science to Prepare for Floods. In: *Resilience*. Elsevier, pp. 237–242. DOI:10.1016/B978-0-12-811891-7.00019-0.
- Jongman, B., P. J. Ward, and J. C. Aerts (2012). Global exposure to river and coastal flooding: Long term trends and changes. *Global Environmental Change* 22.4, pp. 823–835. DOI:10.1016/J.GLOENVCHA.2012.07.004.
- Jongman, B. et al. (2014). Increasing stress on disaster-risk finance due to large floods. *Nature Climate Change* 4.4, pp. 264–268. DOI:10.1038/NCLIMATE2124.
- Jongman, B., H. C. Winsemius, J. C. J. H. Aerts, E. Coughlan de Perez, M. K. van Aalst, W. Kron, and P. J. Ward (2015). Declining vulnerability to river floods and the global benefits of adaptation. *Proceedings of the National Academy of Sciences* 112.18, E2271–E2280. DOI:10.1073/PNAS.1414439112.
- Kållberg, P., P. Berrisford, B. Hoskins, A. Simmons, S. Uppala, S. Lamy-Thépaut, and R. Hine (2005). ERA-40 atlas, ERA-40 Proj. Rep. Ser. 19, Eur. Cent. for Medium Range Weather Forecasts. Reading, UK.
- Karssen, D., O. Schmitz, P. Salamon, K. de Jong, and M. F. P. Bierkens (2010). A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environmental Modelling & Software* 25.4, pp. 489–502. DOI:10.1016/J.ENVSOFT.2009.10.004.
- Kernkamp, H. W. J., A. Van Dam, G. S. Stelling, and E. D. de Goede (2011). Efficient scheme for the shallow water equations on unstructured grids with application to the Continental Shelf. *Ocean Dynamics* 61.8, pp. 1175–1188. DOI:10.1007/s10236-011-0423-6.
- Kim, B., B. F. Sanders, J. E. Schubert, and J. S. Famiglietti (2014). Mesh type tradeoffs in 2D hydrodynamic modeling of flooding with a Godunov-based flow solver. *Advances in Water Resources* 68, pp. 42–61. DOI:10.1016/J.ADVWATRES.2014.02.013.

- Kim, J., A. Warnock, V. Y. Ivanov, and N. D. Katopodes (2012). Coupled modeling of hydrologic and hydrodynamic processes including overland and channel flow. *Advances in Water Resources* 37, pp. 104–126. DOI:10.1016/J.ADVWATRES.2011.11.009.
- Kindel, P. J., C. J. Legleiter, and J. M. Nelson (2013). Mapping River Bathymetry With a Small Footprint Green LiDAR: Applications and Challenges. *JAWRA Journal of the American Water Resources Association* 49.1, pp. 183–204. DOI:10.1111/JAWR.12008.
- Kling, H., P. Stanzel, M. Fuchs, and H.-P. Nachtnebel (2015). Performance of the COSERO precipitation–runoff model under non-stationary conditions in basins with different climates. *Hydrological sciences journal* 60.7-8, pp. 1374–1393.
- Knoben, W. J. M., J. E. Freer, K. J. A. Fowler, M. C. Peel, and R. A. Woods (2019). Modular Assessment of Rainfall-Runoff Models Toolbox (MARRMoT) v1.0: an open- source, extendable framework providing implementations of 46 conceptual hydrologic models as continuous space-state formulations. *Geoscientific Model Development Discussions* 2019, pp. 1–26. DOI:10.5194/GMD-2018-332.
- Kokkonen, T., A. Jolma, and H. Koivusalo (2003). Interfacing environmental simulation models and databases using XML. *Environmental Modelling & Software* 18.5, pp. 463–471. DOI:10.1016/S1364-8152(03)00020-3.
- Kotera, A., T. Nagano, P. Hanittinan, and S. Koontanakulvong (2016). Assessing the degree of flood damage to rice crops in the Chao Phraya delta, Thailand, using MODIS satellite imaging. *Paddy and water environment* 14.1, pp. 271–280.
- Kumar, M., G. Bhatt, and C. J. Duffy (2009). An efficient domain decomposition framework for accurate representation of geodata in distributed hydrologic models. *International Journal of Geographical Information Science* 23.12, pp. 1569–1596. DOI:10.1080/13658810802344143.
- Larsen, M. A. D., J. C. Refsgaard, M. Drews, M. B. Butts, K. H. Jensen, J. H. Christensen, and O. B. Christensen (2014). Results from a full coupling of the HIRHAM regional climate model and the MIKE SHE hydrological model for a Danish catchment. *Hydrology and Earth System Sciences* 18.11, pp. 4733–4749. DOI:10.5194/HESS-18-4733-2014.
- Legleiter, C. J. (2015). Calibrating remotely sensed river bathymetry in the absence of field measurements: Flow REsistance Equation-Based Imaging of River Depths (FREEBIRD): Calibrating remotely sensed river bathymetry: freebird. *Water Resources Research* 51.4, pp. 2865–2884. DOI:10.1002/2014WR016624.
- Legleiter, C. J. (2016). Inferring river bathymetry via image-to-depth quantile transformation (IDQT). *Water Resources Research* 52.5, pp. 3722–3741. DOI:10.1002/2016WR018730.
- Lehner, B., K. Verdin, and A. Jarvis (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions, American Geophysical Union* 89.10, pp. 93–94.
- Leopold, L. B. and T. Maddock (1953). *The hydraulic geometry of stream channels and some physiographic implications*. Vol. 252. US Government Printing Office.
- Li, H., S. Beldring, and C.-Y. Xu (2015). Stability of model performance and parameter values on two catchments facing changes in climatic conditions. *Hydrological sciences journal* 60.7-8, pp. 1317–1330.
- Lian, Y., I.-C. Chan, J. Singh, M. Demissie, V. Knapp, and H. Xie (2007). Coupling of hydrologic and hydraulic models for the Illinois River Basin. *Journal of Hydrology* 344.3, pp. 210–222. DOI:10.1016/J.JHYDROL.2007.08.004.
- Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres* 99.D7, pp. 14415–14428. DOI:10.1029/94JD00483.
- Lima, I. B. T., R. R. Rosa, F. M. Ramos, and E. M. L. de Moraes Novo (2003). Water level dynamics in the Amazon floodplain. *Advances in Water Resources* 26.7, pp. 725–732. DOI:10.1016/S0309-1708(03)00052-6.
- Lindström, G., B. Johansson, M. Persson, M. Gardelin, and S. Bergström (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology* 201.1-4, pp. 272–288. DOI:10.1016/S0022-1694(97)00041-3.
- Liu, Q., Y. Qin, Y. Zhang, and Z. Li (2015). A coupled 1D2D hydrodynamic model for flood simulation in flood detention basin. *Natural Hazards* 75.2, pp. 1303–1325. DOI:10.1007/s11069-014-1373-3.
- Lopez, P. L., N. Wanders, J. Schellekens, L. J. Renzullo, E. H. Sutanudjaja, and M. F. P. Bierkens (2016). Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. *Hydrology and Earth System Sciences* 20.7, pp. 3059–3076. DOI:10.5194/HESS-20-3059-2016.
- Mahe, G. (2009). Surface/groundwater interactions in the Bani and Nakambe rivers, tributaries of the Niger and Volta basins, West Africa. *Hydrological Sciences Journal* 54.4, pp. 704–712.

- Mahe, G., F Bamba, A Soumaguel, D Orange, and J. Olivry (2009). Water losses in the inner delta of the River Niger: water balance and flooded area. *Hydrological Processes* 23.22, pp. 3157–3160. DOI:10.1002/HYP.7389.
- Marengo, J. A. and J. C. Espinoza (2016). Extreme seasonal droughts and floods in Amazonia: causes, trends and impacts. *International Journal of Climatology* 36.3, pp. 1033–1050. DOI:10.1002/JOC.4420.
- Marsh, C. B., R. J. Spiteri, J. W. Pomeroy, and H. S. Wheatler (2018). Multi-objective unstructured triangular mesh generation for use in hydrological and land surface models. *Computers & Geosciences* 119, pp. 49–67. DOI:10.1016/J.CAGEO.2018.06.009.
- Martyr-Koller, R., H. Kernkamp, A. van Dam, M. van der Wegen, L. Lucas, N. Knowles, B. Jaffe, and T. Fregoso (2017). Application of an unstructured 3D finite volume numerical model to flows and salinity dynamics in the San Francisco Bay-Delta. *Estuarine, Coastal and Shelf Science* 192, pp. 86–107. DOI:10.1016/J.ECSS.2017.04.024.
- Massari, C., L. Brocca, T. Moramarco, Y. Trambly, and J.-F. Didon Lescot (2014). Potential of soil moisture observations in flood modelling: Estimating initial conditions and correcting rainfall. *Advances in Water Resources* 74, pp. 44–53. DOI:10.1016/J.ADVWATRES.2014.08.004.
- Mateo, C. M. R., D. Yamazaki, H. Kim, A. Champathong, J. Vaze, and T. Oki (2017). Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods. *Hydrology and Earth System Sciences* 21.10, pp. 5143–5163. DOI:10.5194/HESS-21-5143-2017.
- Maxwell, R. M. and N. L. Miller (2005). Development of a Coupled Land Surface and Groundwater Model. *Journal of Hydrometeorology* 6.3, pp. 233–247. DOI:10.1175/JHM422.1.
- Meade, R. H., J. M. Rayol, S. C. Conceição, and J. R. G. Natividade (1991). Backwater effects in the Amazon River basin of Brazil. *Environmental Geology and Water Sciences* 18.2, pp. 105–114. DOI:10.1007/BF01704664.
- Melsen, L. A., A. J. Teuling, P. J. J. F. Torfs, M. Zappa, N. Mizukami, P. A. Mendoza, M. P. Clark, and R. Uijlenhoet (2019). Subjective modeling decisions can significantly impact the simulation of flood and drought events. *Journal of Hydrology* 568, pp. 1093–1104. DOI:10.1016/J.JHYDROL.2018.11.046.
- Merwade, V., A. Cook, and J. Coonrod (2008). GIS techniques for creating river terrain models for hydrodynamic modeling and flood inundation mapping. *Environmental Modelling & Software* 23.10, pp. 1300–1311. DOI:10.1016/J.ENVSOFT.2008.03.005.
- Molinier, M., J. Ronchail, J. L. Guyot, G. Cochonneau, V. Guimaraes, and E. De Oliveira (2009). Hydrological variability in the Amazon drainage basin and African tropical basins. *Hydrological Processes* 23.22, pp. 3245–3252. DOI:10.1002/HYP.7400.
- Morita, M. and B. C. Yen (2002). Modeling of Conjunctive Two-Dimensional Surface-Three-Dimensional Subsurface Flows. *Journal of Hydraulic Engineering* 128.2, pp. 184–200. DOI:10.1061/(ASCE)0733-9429(2002)128:2(184).
- Moussa, R. and C. Bocquillon (1996). Criteria for the choice of flood-routing methods in natural channels. *Journal of Hydrology* 186.1-4, pp. 1–30. DOI:10.1016/S0022-1694(96)03045-4.
- Mård, J., G. di Baldassarre, and M. Mazzoleni (2018). Nighttime light data reveal how flood protection shapes human proximity to rivers. *Science Advances*, p. 8.
- Muis, S., M. Verlaan, H. C. Winsemius, J. C. J. H. Aerts, and P. J. Ward (2016). A global reanalysis of storm surges and extreme sea levels. *Nature Communications* 7, p. 11969. DOI:10.1038/NCOMMS11969.
- MunichRe (2010). Topics Geo natural catastrophes 2009: Analyses, assessments, positions. *Munich: Munich Reinsurance Company*.
- MunichRe (2019). *NatCatSERVICE*. URL: <https://www.munichre.com/en/reinsurance/business/non-life/natcatservice/index.html> (visited on 01/11/2019).
- Naz, B. S., C. D. Frans, G. K. C. Clarke, P. Burns, and D. P. Lettenmaier (2014). Modeling the effect of glacier recession on streamflow response using a coupled glacio-hydrological model. *Hydrology and Earth System Sciences* 18.2, pp. 787–802. DOI:10.5194/HESS-18-787-2014.
- Neal, J. C., G. J.-P. Schumann, and P. D. Bates (2012a). A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. *Water Resources Research* 48.11. DOI:10.1029/2012WR012514.
- Neal, J. C., I. Villanueva, N. Wright, T. Willis, T. Fewtrell, and P. D. Bates (2012b). How much physical complexity is needed to model flood inundation? *Hydrological Processes* 26.15, pp. 2264–2282. DOI:10.1002/HYP.8339.
- Nobre, A. D., L. A. Cuartas, M. R. Momo, D. L. Severo, A. Pinheiro, and C. A. Nobre (2016). HAND contour: a new proxy predictor of inundation extent: Mapping Flood Hazard Potential Using Topography. *Hydrological Processes* 30.2, pp. 320–333. DOI:10.1002/HYP.10581.

- O'Loughlin, F. E., R. C. D. Paiva, M. Durand, D. E. Alsdorf, and P. D. Bates (2016). A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote Sensing of Environment* 182, pp. 49–59. DOI:10.1016/J.RSE.2016.04.018.
- OpenMI (2018). *New to OpenMI?* URL: <https://sites.google.com/a/openmi.org/home/new-to-openmi> (visited on 12/03/2018).
- ORE-HYBAM (2018). *HYBAM ORE-South America*. URL: <http://www.ore-hybam.org/index.php/eng/Data/Station-Access-Maps/HYBAM-ORE-South-America> (visited on 07/19/2018).
- Paiva, R. C. D., W. Collischonn, and C. E. M. Tucci (2011). Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach. *Journal of Hydrology* 406.3, pp. 170–181. DOI:10.1016/J.JHYDROL.2011.06.007.
- Paiva, R. C. D., W. Collischonn, and D. C. Buarque (2013). Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon. *Hydrological Processes* 27.3, pp. 333–346. DOI:10.1002/HYP.8425.
- Pan, M., H. Li, and E. Wood (2010). Assessing the skill of satellite-based precipitation estimates in hydrologic applications. *Water Resources Research* 46.9. DOI:10.1029/2009WR008290.
- Papa, F., C. Prigent, F. Aires, C. Jimenez, W. B. Rossow, and E. Matthews (2010). Interannual variability of surface water extent at the global scale, 19932004. *Journal of Geophysical Research* 115 (D12). DOI:10.1029/2009JD012674.
- Pappenberger, F., E. Dutra, F. Wetterhall, and H. L. Cloke (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade. *Hydrology and Earth System Sciences* 16.11, pp. 4143–4156. DOI:10.5194/HESS-16-4143-2012.
- Pappenberger, F., P. Matgen, K. J. Beven, J.-B. Henry, L. Pfister, and P. Fraipont (2006). Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources* 29.10, pp. 1430–1449. DOI:10.1016/J.ADVWATRES.2005.11.012.
- Partnership, G. F. (2016). *The Global Flood Partnership Conference 2016 - Linking global flood information with local needs*. Ispra, Italy: Joint Research Centre.
- Partnership, G. F. (2017). *The Global Flood Partnership Conference 2017 - From hazards to impacts*. Ispra, Italy.
- Peckham, S. D. (2010). *CSDMS Handbook of Concepts and Protocols: a Guide for Code Contributors*. URL: https://csdms.colorado.edu/wiki/Help:Tools_CSDMS_Handbook (visited on 12/14/2018).
- Peckham, S. D. (2014). EMELI 1.0: An Experimental Smart Modeling Framework For Automatic Coupling Of Self-Describing Models. In: International Conference on Hydroinformatics. CUNY Academic Works, p. 9.
- Peckham, S. D., E. H. W. Hutton, and B. Norris (2013). A component-based approach to integrated modeling in the geosciences: The design of CSDMS. *Computers & Geosciences* 53, pp. 3–12. DOI:10.1016/J.CAGEO.2012.04.002.
- Pelupessy, I. et al. (2019). HyMUSE: a multi-model framework for Hydrology. In: EGU General Assembly Conference Abstracts. Vol. 21, pp. 14271–1.
- Philipp, A., R. Liedl, and T. Wöhling (2012). Analytical Model of Surface Flow on Hillslopes Based on the Zero Inertia Equations. *Journal of Hydraulic Engineering* 138.5, pp. 391–399. DOI:10.1061/(ASCE)HY.1943-7900.0000519.
- Prigent, C., F. Papa, F. Aires, W. B. Rossow, and E. Matthews (2007). Global inundation dynamics inferred from multiple satellite observations, 19932000. *Journal of Geophysical Research* 112 (D12). DOI:10.1029/2006JD007847.
- Quinn, N., P. D. Bates, J. C. Neal, A. Smith, O. E. J. Wing, C. C. Sampson, J. Smith, and J. Heffernan (2019). The spatial dependence of flood hazard and risk in the United States. *Water Resources Research* 55.3, pp. 1890–1911. DOI:10.1029/2018WR024205.
- Ratliff, K. M., E. H. W. Hutton, and A. B. Murray (2018). Exploring Wave and Sea-Level Rise Effects on Delta Morphodynamics with a Coupled River–Ocean Model. *Journal of Geophysical Research: Earth Surface*. DOI:10.1029/2018JF004757.
- Rennó, C. D., A. D. Nobre, L. A. Cuartas, J. V. Soares, M. G. Hodnett, J. Tomasella, and M. J. Waterloo (2008). HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia. *Remote Sensing of Environment* 112.9, pp. 3469–3481. DOI:10.1016/J.RSE.2008.03.018.
- Revilla-Romero, B., H. E. Beck, P. Burek, P. Salamon, A. P. J. de Roo, and J. Thielen (2015). Filling the gaps: Calibrating a rainfall-runoff model using satellite-derived surface water extent. *Remote Sensing of Environment* 171, pp. 118–131. DOI:10.1016/J.RSE.2015.10.022.
- Rudari, R., F. Silvestro, L. Campo, N. Reborá, G. Boni, and C. Herold (2015). *Improvement of the global flood model for the GAR 2015*, p. 69.

- Rudorff, C. M., J. M. Melack, and P. D. Bates (2014a). Flooding dynamics on the lower Amazon floodplain: 1. Hydraulic controls on water elevation, inundation extent, and river-floodplain discharge. *Water Resources Research* 50.1, pp. 619–634. DOI:10.1002/2013WR014091.
- Rudorff, C. M., J. M. Melack, and P. D. Bates (2014b). Flooding dynamics on the lower Amazon floodplain: 2. Seasonal and interannual hydrological variability. *Water Resources Research* 50.1, pp. 635–649. DOI:10.1002/2013WR014714.
- Sampson, C. C., T. J. Fewtrell, F. O'Loughlin, F. Pappenberger, P. D. Bates, J. E. Freer, and H. L. Cloke (2014). The impact of uncertain precipitation data on insurance loss estimates using a flood catastrophe model. *Hydrology and Earth System Sciences* 18.6, pp. 2305–2324. DOI:10.5194/HESS-18-2305-2014.
- Sampson, C. C., A. M. Smith, P. D. Bates, J. C. Neal, L. Alfieri, and J. E. Freer (2015). A high-resolution global flood hazard model. *Water Resources Research* 51.9, pp. 7358–7381. DOI:10.1002/2015WR016954.
- Sanders, B. F. (2007). Evaluation of on-line DEMs for flood inundation modeling. *Advances in Water Resources* 30.8, pp. 1831–1843. DOI:10.1016/J.ADVWATRES.2007.02.005.
- Sanders, B. F., J. E. Schubert, and R. L. Detwiler (2010). ParBreZo: A parallel, unstructured grid, Godunov-type, shallow-water code for high-resolution flood inundation modeling at the regional scale. *Advances in Water Resources* 33.12, pp. 1456–1467. DOI:10.1016/J.ADVWATRES.2010.07.007.
- Savage, J. T. S., F. Pianosi, P. D. Bates, J. E. Freer, and T. Wagener (2016a). Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model. *Water Resources Research* 52.11, pp. 9146–9163. DOI:10.1002/2015WR018198.
- Savage, J. T. S., P. D. Bates, J. E. Freer, J. C. Neal, and G. Aronica (2016b). When does spatial resolution become spurious in probabilistic flood inundation predictions? *Hydrological Processes* 30.13, pp. 2014–2032. DOI:10.1002/HYP.10749.
- Schellekens, J., W. Van Verseveld, T. Euser, H. C. Winsemius, C. Thiange, L. Bouaziz, D. Tollenaar, and S. De Vries (2018). *openstreams/wflow: master-unstable*. Version master-unstable.
- Schubert, J. E., B. F. Sanders, M. J. Smith, and N. G. Wright (2008). Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding. *Advances in Water Resources* 31.12, pp. 1603–1621. DOI:10.1016/J.ADVWATRES.2008.07.012.
- Schumann, G. J.-P., J. C. Neal, N. Voisin, K. M. Andreadis, F. Pappenberger, N. Phanhuwongpakdee, A. C. Hall, and P. D. Bates (2013). A first large-scale flood inundation forecasting model: Large-Scale Flood Inundation Forecasting. *Water Resources Research* 49.10, pp. 6248–6257. DOI:10.1002/WRCR.20521.
- Scussolini, P., J. C. J. H. Aerts, B. Jongman, L. M. Bouwer, H. C. Winsemius, H. de Moel, and P. J. Ward (2016). FLOPROS: an evolving global database of flood protection standards. *Natural Hazards and Earth System Sciences* 16.5, pp. 1049–1061. DOI:10.5194/NHESS-16-1049-2016.
- Senatore, A., G. Mendicino, D. J. Gochis, W. Yu, D. N. Yates, and H. Kunstmann (2015). Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. *Journal of Advances in Modeling Earth Systems* 7.4, pp. 1693–1715. DOI:10.1002/2015MS000510.
- Silberstein, R., R. Vertessy, and V. Arora (2007). TOPOG. TOPOG online. URL: <http://www-data.wron.csiro.au/topog/> (visited on 10/29/2018).
- Simard, M., N. Pinto, J. B. Fisher, and A. Baccini (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research* 116 (G4). DOI:10.1029/2011JG001708.
- Speckhann, G. A., P. L. Borges Chaffe, R. Fabris Goerl, J. J. d. Abreu, and J. A. Altamirano Flores (2018). Flood hazard mapping in Southern Brazil: a combination of flow frequency analysis and the HAND model. *Hydrological Sciences Journal* 63.1, pp. 87–100. DOI:10.1080/02626667.2017.1409896.
- Sperna Weiland, F. C., L. P. H. van Beek, J. C. J. Kwadijk, and M. F. P. Bierkens (2010). The ability of a GCM-forced hydrological model to reproduce global discharge variability. *Hydrology and Earth System Sciences* 14.8, pp. 1595–1621. DOI:10.5194/HESS-14-1595-2010.
- Sutanudjaja, E. H., L. P. H. v. Beek, S. M. d. Jong, F. C. v. Geer, and M. F. P. Bierkens (2011). Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin. *Hydrology and Earth System Sciences* 15.9, pp. 2913–2935. DOI:HTTPS://DOI.ORG/10.5194/HESS-15-2913-2011.
- Sutanudjaja, E. H., L. P. H. van Beek, S. M. de Jong, F. C. van Geer, and M. F. P. Bierkens (2014). Calibrating a largeextent highresolution coupled groundwaterland surface model using soil moisture and discharge data. *Water Resources Research* 50.1, pp. 687–705. DOI:10.1002/2013WR013807.
- Sutanudjaja, E. H. et al. (2018). PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. *Geoscientific Model Development* 11.6, pp. 2429–2453. DOI:10.5194/GMD-11-2429-2018.
- Tayefi, V., S. Lane, R. Hardy, and D Yu (2007). A comparison of one- and two-dimensional approaches to modelling flood inundation over complex upland floodplains. *Hydrological Processes* 21.23, pp. 3190–3202. DOI:10.1002/HYP.6523.

- Tellman, B., J Sullivan, C Doyle, A Kettner, G. Brakenridge, T Erickson, and D. Slayback (2017). A Global Geospatial Database of 5000+ Historic Flood Event Extents. In: *AGU Fall Meeting Abstracts*.
- Theurich, G. et al. (2015). The Earth System Prediction Suite: Toward a Coordinated U.S. Modeling Capability. *Bulletin of the American Meteorological Society* 97.7, pp. 1229–1247. DOI:10.1175/BAMS-D-14-00164.1.
- Todini, E. (1996). The ARNO rainfallrunoff model. *Journal of Hydrology* 175.1, pp. 339–382. DOI:10.1016/S0022-1694(96)80016-3.
- Triet, N. V. K., N. V. Dung, H. Fujii, M. Kummu, B. Merz, and H. Apel (2017). Has dyke development in the Vietnamese Mekong Delta shifted flood hazard downstream? *Hydrology and Earth System Sciences* 21.8, pp. 3991–4010. DOI:10.5194/HESS-21-3991-2017.
- Trigg, M. A. et al. (2016). The credibility challenge for global fluvial flood risk analysis. *Environmental Research Letters* 11.9, p. 094014. DOI:10.1088/1748-9326/11/9/094014.
- Trigg, M. A., M. D. Wilson, P. D. Bates, M. S. Horritt, D. E. Alsdorf, B. R. Forsberg, and M. C. Vega (2009). Amazon flood wave hydraulics. *Journal of Hydrology* 374.1, pp. 92–105. DOI:10.1016/J.JHYDROL.2009.06.004.
- UNISDR (2015a). *Global Assessment Report*. URL: <https://www.unisdr.org/we/inform/gar> (visited on 10/30/2018).
- UNISDR (2015b). *Sendai Framework for Disaster Risk Reduction*. URL: <https://www.unisdr.org/we/coordinate/sendai-framework> (visited on 10/30/2018).
- UNISDR (2018). *Putting science to work for resilience*. URL: <https://www.unisdr.org/archive/58772> (visited on 07/23/2018).
- Uppala, S. M. et al. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* 131.612, pp. 2961–3012. DOI:10.1256/QJ.04.176.
- Van Beek, L. P. H. (2008). *Forcing PCR-GLOBWB with CRU data*.
- Van Beek, L. P. H. and M. F. P. Bierkens (2008). *The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification*.
- Van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources Research* 47.7. DOI:10.1029/2010WR009791.
- Van Dijk, A. I. J. M., J. L. Peña-Arancibia, E. F. Wood, J. Sheffield, and H. E. Beck (2013). Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide. *Water Resources Research* 49.5, pp. 2729–2746. DOI:10.1002/WRCR.20251.
- Vieira, J. H. D. (1983). Conditions governing the use of approximations for the Saint-Venant equations for shallow surface water flow. *Journal of Hydrology* 60.1, pp. 43–58. DOI:10.1016/0022-1694(83)90013-6.
- Viero, D. P., P. Peruzzo, L. Carniello, and A. Defina (2014). Integrated mathematical modeling of hydrological and hydrodynamic response to rainfall events in rural lowland catchments. *Water Resources Research* 50.7, pp. 5941–5957. DOI:10.1002/2013WR014293.
- Visser, H., A. Bouwman, W. Ligtoet, and A. C. Petersen (2012). A statistical study of weather-related disasters: past, present and future. *PBL Netherlands Environmental Assessment Agency, The Hague/Bilthoven, The Netherlands*.
- Voinov, A. and H. H. Shugart (2013). Integronsters, integral and integrated modeling. *Environmental Modelling & Software* 39, pp. 149–158. DOI:10.1016/J.ENVSOFT.2012.05.014.
- Wada, Y. et al. (2013). Multimodel projections and uncertainties of irrigation water demand under climate change. *Geophysical Research Letters* 40.17, pp. 4626–4632. DOI:10.1002/GRL.50686.
- Wada, Y., I. E. M. d. Graaf, and L. P. H. van Beek (2016). High-resolution modeling of human and climate impacts on global water resources. *Journal of Advances in Modeling Earth Systems* 8.2, pp. 735–763. DOI:10.1002/2015MS000618.
- Wagner, S., B. Fersch, F. Yuan, Z. Yu, and H. Kunstmann (2016). Fully coupled atmospheric-hydrological modeling at regional and long-term scales: Development, application, and analysis of WRF-HMS. *Water Resources Research* 52.4, pp. 3187–3211. DOI:10.1002/2015WR018185.
- Wanders, N. and Y. Wada (2015). Human and climate impacts on the 21st century hydrological drought. *Journal of Hydrology* 526, pp. 208–220. DOI:10.1016/J.JHYDROL.2014.10.047.
- Ward, P. J., B. Jongman, F. S. Weiland, A. Bouwman, L. P. H. van Beek, M. F. P. Bierkens, W. Ligtoet, and H. C. Winsemius (2013). Assessing flood risk at the global scale: model setup, results, and sensitivity. *Environmental Research Letters* 8.4, p. 044019. DOI:10.1088/1748-9326/8/4/044019.
- Ward, P. J. et al. (2015). Usefulness and limitations of global flood risk models. *Nature Climate Change* 5.8, pp. 712–715. DOI:10.1038/NCLIMATE2742.

- Ward, P. J., A. Couasnon, D. Eilander, I. D. Haigh, A. Hendry, S. Muis, T. I. E. Veldkamp, H. C. Winsemius, and T. Wahl (2018). Dependence between high sea-level and high river discharge increases flood hazard in global deltas and estuaries. *Environmental Research Letters* 13.8, p. 084012. DOI:10.1088/1748-9326/AAD400.
- Watson, W. T. (2018). *Insights from the Willis Re Flood Club: the weaknesses and strengths of flood modelling*. Willis Towers Watson Wire. URL: <https://blog.willis.com/2018/06/insights-from-the-willis-re-flood-club-the-weaknesses-and-strengths-of-flood-modelling/> (visited on 07/19/2018).
- Werner, M., J. Schellekens, P. Gijsbers, M. van Dijk, O. van den Akker, and K. Heynert (2013). The Delft-FEWS flow forecasting system. *Environmental Modelling & Software* 40, pp. 65–77. DOI:10.1016/J.ENVSOF.2012.07.010.
- Wilson, M., P. D. Bates, D. Alsdorf, B. Forsberg, M. S. Horritt, J. Melack, F. Frappart, and J. Famiglietti (2007). Modeling large-scale inundation of Amazonian seasonally flooded wetlands. *Geophysical Research Letters* 34.15. DOI:10.1029/2007GL030156.
- Wing, O. E. J., P. D. Bates, C. C. Sampson, A. M. Smith, K. A. Johnson, and T. A. Erickson (2017). Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resources Research* 53.9, pp. 7968–7986. DOI:10.1002/2017WR020917.
- Wing, O. E. J., P. D. Bates, A. M. Smith, C. C. Sampson, K. A. Johnson, J. Fargione, and P. Morefield (2018). Estimates of present and future flood risk in the conterminous United States. *Environmental Research Letters* 13.3, p. 034023. DOI:10.1088/1748-9326/AAAC65.
- Winsemius, H. C., L. P. H. van Beek, B. Jongman, P. J. Ward, and A. Bouwman (2013). A framework for global river flood risk assessments. *Hydrology and Earth System Sciences* 17.5, pp. 1871–1892. DOI:10.5194/HESS-17-1871-2013.
- Winsemius, H. C. et al. (2016). Global drivers of future river flood risk. *Nature Climate Change* 6.4, pp. 381–385. DOI:10.1038/NCLIMATE2893.
- Wood, E. F., D. P. Lettenmaier, and V. G. Zartarian (1992). A land-surface hydrology parameterization with subgrid variability for general circulation models. *Journal of Geophysical Research* 97 (D3), p. 2717. DOI:10.1029/91JD01786.
- Wood, E. F. et al. (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research* 47.5. DOI:10.1029/2010WR010090.
- Yamazaki, D., T. Oki, and S. Kanae (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* 13.11, pp. 2241–2251. DOI:10.5194/HESS-13-2241-2009.
- Yamazaki, D., S. Kanae, H. Kim, and T. Oki (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research* 47.4. DOI:10.1029/2010WR009726.
- Yamazaki, D., C. A. Baugh, P. D. Bates, S. Kanae, D. E. Alsdorf, and T. Oki (2012a). Adjustment of a spaceborne DEM for use in floodplain hydrodynamic modeling. *Journal of Hydrology* 436-437, pp. 81–91. DOI:10.1016/J.JHYDROL.2012.02.045.
- Yamazaki, D., H. Lee, D. E. Alsdorf, E. Dutra, H. Kim, S. Kanae, and T. Oki (2012b). Analysis of the water level dynamics simulated by a global river model: A case study in the Amazon River. *Water Resources Research* 48.9. DOI:10.1029/2012WR011869.
- Yamazaki, D., G. A. M. de Almeida, and P. D. Bates (2013). Improving computational efficiency in global river models by implementing the local inertial flow equation and a vector-based river network map: Speeding Up Global River Model Simulations. *Water Resources Research* 49.11, pp. 7221–7235. DOI:10.1002/WRCR.20552.
- Yamazaki, D., F. O'Loughlin, M. A. Trigg, Z. F. Miller, T. M. Pavelsky, and P. D. Bates (2014a). Development of the Global Width Database for Large Rivers. *Water Resources Research* 50.4, pp. 3467–3480. DOI:10.1002/2013WR014664.
- Yamazaki, D., T. Sato, S. Kanae, Y. Hirabayashi, and P. D. Bates (2014b). Regional flood dynamics in a bifurcating mega delta simulated in a global river model. *Geophysical Research Letters* 41.9, pp. 3127–3135. DOI:10.1002/2014GL059744.
- Yamazaki, D., M. A. Trigg, and D. Ikeshima (2015). Development of a global ~90m water body map using multi-temporal Landsat images. *Remote Sensing of Environment* 171, pp. 337–351. DOI:10.1016/J.RSE.2015.10.014.
- Yamazaki, D., D. Ikeshima, R. Tawatari, T. Yamaguchi, F. O'Loughlin, J. C. Neal, C. C. Sampson, S. Kanae, and P. D. Bates (2017). A high-accuracy map of global terrain elevations: Accurate Global Terrain Elevation map. *Geophysical Research Letters* 44.11, pp. 5844–5853. DOI:10.1002/2017GL072874.

- Yossef, N. C., H. C. Winsemius, A. Weerts, L. P. H. van Beek, and M. F. P. Bierkens (2013). Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. *Water Resources Research* 49.8, pp. 4687–4699. DOI:10.1002/WRCR.20350.
- Zabel, F. and W. Mauser (2013). 2-way coupling the hydrological land surface model PROMET with the regional climate model MM5. *Hydrology and Earth System Sciences* 17.5, pp. 1705–1714. DOI:10.5194/HESS-17-1705-2013.
- Zhang, Q., C. Zhu, C. Liu, and T. Jiang (2005). Environmental change and its impacts on human settlement in the Yangtze Delta, P.R. China. *CATENA* 60.3, pp. 267–277. DOI:10.1016/J.CATENA.2004.12.001.
- Zhao, F. et al. (2017). The critical role of the routing scheme in simulating peak river discharge in global hydrological models. *Environmental Research Letters* 12.7, p. 075003. DOI:10.1088/1748-9326/AA7250.
- Zhao, Q., B. Ye, Y. Ding, S. Zhang, S. Yi, J. Wang, D. Shangguan, C. Zhao, and H. Han (2013). Coupling a glacier melt model to the Variable Infiltration Capacity (VIC) model for hydrological modeling in north-western China. *Environmental Earth Sciences* 68.1, pp. 87–101. DOI:10.1007/s12665-012-1718-8.
- Zischg, A. P., G. Felder, M. Mosimann, V. Röthlisberger, and R. Weingartner (2018). Extending coupled hydrological-hydraulic model chains with a surrogate model for the estimation of flood losses. *Environmental Modelling & Software* 108, pp. 174–185. DOI:10.1016/J.ENVSOF.2018.08.009.

Acknowledgements

Successfully finishing a PhD project is truly not an easy task and requires the support of many people. I am very thankful to have been supported by such great family, friends, and colleagues throughout the entire project – both work-wise and, maybe even more importantly, outside of the office. Many many thanks y'all!

I want to start by thanking my housemates in Huize Scherpenburgh (Daan, Wim, AP, Merel and Merel, Marnix, Esther, Liesbeth) for giving me the opportunity to take my integration efforts one step further and letting "the German" living with them. Must have been a tough decision but guess the innumerable jokes about Germans and bicycles have made up for it! Feeling well integrated and "home" was indeed always important to me and a crucial role was played by my team mates from VV Voorwaarts Utrecht who also introduced me to the complicated politics of Dutch football... Fortunately I could take on a neutral role there!

At work, having colleagues to hang out with, annoy with complaints, share important milestones, but also to simply go for drinks is essential. A very big shout out to the founding fathers (literally!) of the "Man Cave" (Sepehr, Daniel, Jude, Yoeri) and those who joined them later – I really enjoyed the witty discussions over lunch with you as well as the highly scientific exchanges during the rest of the day... Not to mention all the memes I received in those few months! And certainly I want to thank my colleagues Yvonne, Winnie and Jakob who made things bearable if the academic FOMO or showing-off was on again. And everyone else who should be on this list of course!

I must of course not forget my friends back in Munich who always reminded me of my roots by regular Biergarten and Wirtshaus meetings. Thanks for still staying in touch with me after so many years abroad Consti, Anna, Pete, Mats, Pete, Toni as well as my Stammtisch Buam Valentin, Lars, Max und Moritz, and Andi (and all other new and old members)! Besides, there are quite some friends in the Netherlands as well and I want to mention here a few of them explicitly: Gert Jan, Mathias, and Moritz, many thanks and looking forward to the next round of beers with you.

During my PhD, I had the honour to collaborate with many great colleagues from other universities and research institutes. Without a particular order, I want to thank: Paul Bates and Jeff Neal as well as all post-grads and staff from University of Bristol for the great time there (I loved it!); Herman Kernkamp and Arthur van Dam from Deltares for their help with Delft3D Flexible Mesh; Fedor Baart for the help with (a lot of) computer code and BMI development; Edwin Sutanudjaja for sharing the office and for his support with PCR-GLOBWB; and Dirk Eilander from the VU Amsterdam, Hiroaki Ikeuchi from the University of Tokyo as well as Mark Trigg from the University of Leeds for the research collaborations. And of course anyone else who enabled my PhD behind the scenes!

My studies in both Munich and Enschede would not have been possible without the continuous support of my parents who always supported me during my career choices and never doubted that I would make the right decision (and of course also thanks for the financial support too...). This trust in me really means a lot to me and gave me all the freedom I needed to become the person I am now. Thanks Mama and Papa! An important factor was my grandpa

Walter who, as a Dr. rer. nat. (for non-Germans: a great natural scientist) but also a person with an incredibly rich history, was always (and still is) an immense source of inspiration for me. I also want to mention the rest of my family in Munich, most and foremost my two sisters Miriam and Mona, and all who came and went with the years.

It is just appropriate to also say a big "thanks!" to all my supervisors. Rens, thanks for all the detailed work and comments you provided me with on all sorts of manuscripts or other written pieces. Not to forget the interesting scientific discussions and for always being helpful when I had a question. Hessel, thanks for thinking along with me how to further develop my research project and for getting me in touch with so many people who eventually contributed to my thesis. I really like that you share the passion for model integration and believed in it even when I didn't. And Marc, thanks for giving me the space and time to develop my research into the direction I found most interesting and relevant while also having the big research lines and questions in mind. With your sheer all-encompassing knowledge you always knew a way forward when I had the feeling of being stuck. Without the help, support, and advice of you three I would not have been able to produce this thesis!

The two most important people still have to be mentioned: Carina, thanks for being with me and making my life so much more beautiful and better. It is beyond words to express all these intangible things I wanted to say, after all I am not good with words, but I know that without you and your love and support and help I would not have been able to finish my PhD at all. I also want to thank you for Casper, the greatest gift and wonder imaginable, and our new small family. I enjoy every minute I can spend with you two and am forever grateful for the support during the time of double responsibilities! And thanks Casper for setting my priorities right again. It is unbelievable how much joy you bring and the last months of my PhD were so much more bearable knowing I could take you in my arms again once back from office! Carina and Casper, the happiness of finishing this book is only marginal compared to the moments we spend together – looking very much forward to every happy moment with you that yet has to come!

List of Publications

Journal articles, as first author

- Hoch, J. M., Haag, A. V., van Dam, A., Winsemius, H. C., van Beek, R., Bierkens, M.F. P. (2017), Assessing the impact of hydrodynamics on large-scale flood wave propagation—a case study for the Amazon Basin. *Hydrology & Earth System Sciences*, DOI:10.5194/HESS-21-117-2017.
- Hoch, J. M., Neal, J. C., Baart, F., van Beek, R., Winsemius, H. C., Bates, P. D., Bierkens, M.F. P. (2017), GLOFRIM v1.0 – A globally applicable computational framework for integrated hydrological-hydrodynamic modelling. *Geoscientific Model Development*, DOI:10.5194/GMD-10-3913-2017.
- Hoch, J. M., van Beek, R., Winsemius, H. C., Bierkens, M.F. P. (2018), Benchmarking flexible meshes and regular grids for large-scale fluvial inundation modelling. *Advances in Water Resources*, DOI:10.1016/J.ADVWATRES.2018.09.003.
- Hoch, J. M., Trigg, M. A. (2019), Advancing global flood hazard simulations by improving comparability, benchmarking, and integration of global flood models. *Environmental Research Letters*, DOI:10.1088/1748-9326/AAF3D3.
- Hoch, J. M., Eilander, D., Ikeuchi, H., Baart, F., Winsemius, H. C. (2019), Integrating large-scale hydrology and hydrodynamics for nested flood hazard modelling from the mountains to the coast. *Natural Hazards & Earth System Sciences Discussions*, DOI:10.5194/NHESS-2019-75.

Journal articles, as co-author

- Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., van der Ent, R. J., de Graaf, I. E. M., Hoch, J. M., de Jong, K., Karssenbergh, D., Lopez Lopez, P., Pessensteiner, S., Schmitz, O., Straatsma, M. W., Vannamete, E., Wissler, D., Bierkens, M. F. P. (2018), PCR-GLOBWB 2: a 5arcmin global hydrological and water resources model. *Geoscientific Model Development*, DOI:10.5194/GMD-11-2429-2018.
- Towner, J., Cloke, H. L., Zsoter, E., Flamig, Z., Hoch, J. M., Bazo, J., Coughlan de Perez, E., Stephens, E. M. (2019), Assessing the performance of global hydrological models for capturing peak river flows in the Amazon Basin. *Hydrology & Earth System Sciences Discussions*, DOI:10.5194/HESS-2019-44.

Selected conference abstracts

- Hoch, J. M. (2018), GLOFRIM – A globally applicable framework for integrated hydrologic-hydrodynamic modeling, CSDMS conference 2018 "Geoprocesses, Geohazards" (Boulder, CO, United States of America).
- Hoch, J. M. (2018), GLOFRIM as an open-source modelling framework, NHV meeting "Open source software in de hydrologie" (RoyalHaskoningDHV, Amersfoort, the Netherlands)

About the author

Jannis Hoch was born on September 19th, 1988 in Munich (Germany) where he was raised as well. After obtaining his diploma, *nota bene* without any ambitions in natural sciences (particularly physics and maths), at the Wilhelmsgymnasium, a secondary school with focus on the classics, he decided – for whatever reason – to leave his humanistic education behind and started studying Environmental Engineering at the Technical University of Munich.

During his Bachelor studies, his interests were divided between water-related topics and urban infrastructure. Hence, his B.Sc. thesis focussed on the successful re-development of rural train stations with an example of Landsberg am Lech (probably better known as the city where Adolf Hitler was imprisoned). Jannis aimed at not only learning theory, but also gaining insight in engineering practices which is exemplified by his numerous work engagements at companies dealing with noise prevention, building physics, hydraulic engineering, soil remediation, and geology. Even though the thesis was awarded an excellent grade, Jannis decided to continue his studies focussing on “Water Engineering and Management” and hence applied for the homonymous Master of Science track at the University of Twente in Enschede (the Netherlands).

During his M.Sc., he gained knowledge in various subjects such as river geomorphology, water footprint assessment, hydraulic engineering, and sediment transport. Jannis decided to write his M.Sc. thesis on the last subject and developed as well as validated a novel global sediment supply model named PCR-GLOBST which was later shortlisted as the best thesis within Civil Engineering and Management.

In 2014, Jannis started his Ph.D. at the University of Utrecht. His research focussed on coupling hydrologic and hydrodynamic models for improved large-scale inundation modelling. During his Ph.D., Jannis collaborated with the University of Bristol, the Vrije Universiteit Amsterdam, the University of Tokyo, the University of Leeds, and Deltares. The work culminated in the development of GLOFRIM, a globally applicable framework for integrated hydrologic-hydrodynamic modelling, and his contribution to integrated modelling was honoured with the second runner-up of the Jay Syvitsky Student Modeller Award of the Community Surface Dynamics Modelling System (CSDMS) in Boulder, Colorado. Besides model integration, Jannis is interested in how to further advance the current state of global flood modelling. He furthermore contributed to teaching and education and supervised several M.Sc. and B.Sc. theses in cooperation with research partners such as Deltares and University of Twente.

