Rebuttal Letter Referee #1, Anonymous Reviewer

We thank the anonymous referee #1 for the kind words on our manuscript and the points brought forward as they resulted in an improvement of the submitted manuscript. We have added the outcome of the sensitivity analysis of different LISFLOOD-FP parameters. Below, we repeat the reviewer's comments, and provide our response in italics. In the revised manuscript, the changes made to the manuscript are highlighted in yellow.

Page 2, Line 5: 'Sound inundation estimates'. Could the authors use a less colloquial term.

We adjusted the manuscript accordingly and replaced too colloquial terms in general.

Page 4, Line 36: Please provide the spatial resolution of the CRU data.

Thank you for making us aware of the missing reference which was added to the revised manuscript.

Section 2: A schematic of the models used would be beneficial to the reader.

As reaction to your useful comment we added the models currently available within GLOFRIM to Figure 1 to not only provide textual, but also graphical information to the reader.

Page 7, Line 37: Provide reference for the SRTM data.

Thank you for making us aware of the missing reference which was added to the revised manuscript.

Page 11, Line 7: Did you take into account uncertainties in the discharge at Obidos?

Thank you for addressing this aspect. We neglected the uncertainty of observed discharge at Obidos as we assume that it is insignificant compared to other possible uncertainties, for instance parameterization of PCR-GLOBWB or surface roughness of the hydrodynamic models, particularly in large-scale modelling studies. As Clarke et al. (2000) reported an uncertainty of around 16 % of year-to-year variability, we added this reference and a brief comment to facilitate the reader's comprehension of model validation and its limitations.

Page 11, Lines 14-20: Why are the results of the sensitivity analysis not included? The results are not surprising but you need to provide evidence.

Thanks to your useful comment, an explanatory figure (Figure 6a and 6b) was added to the revised version of the manuscript, although we initially decided to not provide an additional plot as we assumed that this may distract the reader from the core of the manuscript. In addition to the plot, we extended the results section 5.2 accordingly, addressing the results and their implications. It is now possible to obtain a better idea of why variations in surface roughness and meandering coefficient can be excluded as cause for deviating discharge simulations between LISFLOOD-FP and Delft3D Flexible Mesh.

Page 11, Line 25-26: What are the different gridding approaches applied? I'm not sure if this is stated elsewhere in the manuscript.

To avoid any confusion, we named the gridding approaches explicitly. Thank you for pointing out this lack of clarity!

With the improvements made to the manuscript based on the valuable and critical reviewer's remarks, we are convinced to have responsibly addressed all ambiguities and shortcomings of the initially submitted version.

Rebuttal Letter Referee #2, Dai Yamazaki

We thank referee #2, Dr. Dai Yamazaki, for the kind words on our manuscript and the points brought forward as they resulted in an improvement of the submitted manuscript. Below, we repeat the reviewer's comments, and provide our response in italics. In the revised manuscript, the changes made to the manuscript are highlighted in light blue.

P7. L32: "If RFS is activated, water volume is directly coupled to the 1-D channels of the hydrodynamic model while, when RFS is inactive, water is distributed over all grid cells of the 2-D domain": The description of RFS is not sufficient. Please explain the relationship between the hydrology model grid and hydrodynamic model pixels in a more detailed manner. I guess, "water volume of each coarse-resolution hydrology model grid" is distributed to the "corresponding high-resolution hydrodynamic model cells within the coarse-resolution grid". And the difference due to RFS is whether water volume is distributed only to river cells or both river and floodplain cells within each coarse-resolution grid box. Readers who are not familiar with this topic might misunderstood water volume is distributed uniformly all-over the calculation domain (not to the corresponding cells).

Thank you very much for pointing out the lack of clarity in our wording. Indeed, the River-Floodplain-Scheme functions as guessed by you. We thus re-wrote the explanation of the River-Floodplain-Scheme to avoid any ambiguity, and to improve the reader's understanding of the coupling scheme.

P7. L35: "the accuracy of the 2-D elevation data which is known to contain strong vertical bias, in particular when derived from remotely sensed global data": The elevation data is affected not only by vertical bias but also by various random/ systematic noises. I recommend to add reference to the latest research on this topic [Yamazaki et al., 2017].

Since at the time of writing the current manuscript the MERIT DEM was not yet published, we could not refer to it despite it's positive contribution to the current state of knowledge. However, we now added the reference to it in the revised manuscript.

P8. L4: "We found that updating fluxes reduces run times compared to states, and hence advise opting for for this option": How downstream boundary conditions are treated. The two test cases executed in the manuscript assumes the downstream boundary is river mouth (0m constant). Some potential users might be interested to simulate flooding in middle-stream, that requires a setting of downstream boundary conditions. Without a reasonable treatment of the downstream boundaries, it is difficult to state that the developed framework is "globally applicable". [I also note that "for" appears twice in the sentence.]

We thank you for mentioning the missing information. It is, in fact, possible to employ any hydrodynamic schematization within GLOFRIM if it complies with Delft3D Flexible Mesh and LISFLOOD-FP requirements, respectively. This means that also other downstream boundaries besides constant water levels at a river mouth are feasible, for instance time-varying water depths at a midstream observation point. For improved clarity, we extended and clarified the paragraph (section 5.1.) in the revised version of the manuscript.

P8. L9: "although PCR runs in spherical coordinates": Given that a spherical coordination can be organized by a Cartesian system, it's better to clarify that PCR-GLOBWB runs at "non-Cartesian spherical system" while it is possible to use "regular lon-lat Cartesian system" for hydrodynamic models.

Thank you very much for this comment. We updated section 3 accordingly.

P8. L30. "4 The Synthetic Test Cases": It's better to explicitly state that PCR-GLOBWB is not used in the synthetic test case. I think this test case is done only for comparing Delft3D and LISFLOOD-FP under an ideal situation. Thus, this test case is directly not related to the GLOFRIM framework, thus readers might be confused.

Based on your useful remark, we re-wrote the paragraph in the revised manuscript to improve clarity.

P8. L34: "0.04 s m-1/3 for the 1-D run and 0.07 s m-1/3 for the 2-D run": Does this mean the same roughness coefficient is used for river channel and floodplains in 2D run? Usually, river channels have smaller roughness compared to floodplains.

Thank you very much for your comment. It is worth mentioning that we performed a 1-D only and a 2-D only run, thus the latter did not contain any 1-D features. As this misunderstanding has also let to your remark regarding our choice of surface roughness values, we updated section 5.1 to increase clarity that the synthetic test case is really 2-D only in this case.

P9. L25: "5.1 Set-up": Please describe how the downstream boundary was treated as this is critical for simulations. Please also explain how the complex channel network of the delta, bifurcating sections, and braided streams were treated. If they are treated differently by Delft3D and LISFLOOD-FP, this difference could be a potential cause of the disagreement of the simulation results.

We thank you for mentioning this aspect. Since the schematization of LISFLOOD-FP is derived from Delft3D Flexible Mesh, both models employ the same 1-D network. In contrast to CaMa-Flood, the schematizations employed in the current manuscript do not account for bifurcations and thus the channel complexity of the delta had to be captured with only one channel. To acknowledge this shortcoming in our model schematizations as possible cause for deviations between simulated and observed discharge, we added information on not only the channel network in the delta, but also on how downstream boundaries were treated in section 5.1. Besides, we referred to it as possible source for the deviation of simulated and observed values in the discussion section 5.2.

P10. L6. "for the elevation data the smoothed 5 canopy-free elevation data was upscaled to a 2 km spatial resolution": Please explicitly explain how the DEM was upscaled, because this has large impact on flood inundation. Did the authors took the mean within a cell, or the minimum elevation?

As you rightly mention the influence of the chosen upscaling approach, we have extended the revised manuscript, noting that we used the nearest neighbour approach to avoid undesired smoothing effects in the elevation values.

P10. L11. "roughness coefficient was uniformly set to 0.03 s m-1/3 for channel and floodplains": Was the same roughness used for channel and floodplain? If so, please clearly state, because different values are usually used.

Thanks for your comment. Indeed, you are right that usually different surface roughness values are used for channel and floodplain, respectively. We, however, decided to use a uniform surface roughness value as this was also done by other studies as stated in the current manuscript. Hence, we desisted from further elaborating on this aspect in the revised manuscript.

P10. L14: "For the hydrological model PCR-GLOBWB, the kinematic wave approach was used for routing outside of the coupled domain": Please explain that the simple kinematic routing may result in poor upstream boundary inflow, as backwater effects or river floodplain interactions can be neglected. Also, this approach could be a limitation for generating a realistic downstream boundary condition. Probably, using continental-scale hydrodynamic model (such as MGB-IPH or CaMa-Flood) as an intermediate step between the hydrology model and high-resolution hydrodynamic model can be a solution.

We thankfully acknowledge your remark on the downsides of the kinematic wave approximation with respect to simulated upstream boundary flow computations. As you rightfully state, the upstream inflow signal can already deviate from observations due to the use of the kinematic wave approximation. Hence, we added this relevant aspect to the results section 5.2. Besides, we added your valuable proposition to employ as 1-D model such as MGB-IPH or CaMa-Flood to the concluding section 6.

P10. L16: "we decided to apply a regionalized optimization technique": This "regionalized optimization" could be a limitation for using the proposed framework for "global application". The "modelling of flood inundation" may be possible at a global scale, but "global application" can be restricted by the quality of input/boundary datasets. Please state this limitation in the conclusion section.

Thanks for pointing out that a global application can be locally restricted by data availability for model set-up. In the revised manuscript, we pointed out that the "regionalized optimization" is optional and PCR-GLOBWB can also be run as is in its default parameterization, thus not posing any constraint to a global application. We described both aspects in more detail in the description of the set-up (section 5.1) to avoid any ambiguity in this matter.

P11. L6: "5.2 Results and Discussion": More detailed analysis of the difference between Delft3D and LISFLOOD-FP is needed. As far as I guess from the figures, the flood peak of Delft3D is later than LISFLOOD-FP because it has larger inundation in upstream areas due to its coarser flexible cell resolution. The smaller water level amplitude in upstream must be also related to the larger inundation in upstream. Because floodplain inundation attenuate flood waves, suppress water level fluctuations, and delays flood peak, most of the disagreement between the two models can be explained consistently due to the inundation in upstream regions. The authors can analyze this effect easily, by comparing the simulated discharges by Delft3D and LISFLOOD-FP also at upstream locations other than the Obidos. By analyzing discharge, we can show where flood waves were attenuated. I suggest to include this discussion in the manuscript. And if the discussion above is true, the Delft3D simulation must be sensitive also to the spatial resolution in upstream regions, thus I recommend to include some sensitivity test on the spatial resolutions.

We are thankful for your extended remark on the model results. First, we agree with the assumption made that the spatial resolution applied by Delft3D Flexible Mesh in upstream areas may impact model results locally as well as further downstream. To shed light on this issue, we added a comparison of simulated discharge for two stations upstream of Obidos to obtain a first-order impression (Figure 4 and Figure 6c). Because we are currently working on a follow-up study concerning the relation of spatial resolution of flexible meshes and model results, we desisted from providing a too elaborate discussion in the current version of the manuscript and added a limited discourse only to section 5.2. Besides, we recommend further investigation of this linkage between cell size resolution and simulated discharge in section 6.

P11. L16: "Since the routing scheme of LFP is based on a D4 system, channel length and dimension in LPF tend to be longer than in other hydrodynamic models": This is not precise. The D4 river network can generate shorter channel length if the scale of channel meandering is smaller than the size of the cell. Please rewrite this sentence. Furthermore, the D4 system does not only change the flow length. It alters the connectivity of channels and floodplains. Some channels/floodplains which are connected in the D8 system (or a vector system) could be disconnected in the D4 system, because diagonal connectivity is not allowed. To avoid this problem, the DEM should be adjusted to ensure the D4 connectivity. Please make a discussion on this issue, as this could be one of the main reason of the difference between Delft3D and LISFLOOD-FP.

Thank you for the information on the D4 river network system and the related uncertainties. In the revised manuscript, we extended the description of the D4 system and its limitations. To show the impact of both increasing and decreasing channel dimension, we furthermore added a plot of the results from the conducted sensitivity analysis (Figure 6a) to better supplement our results and discussion section 5.2 which we updated accordingly, even with the results not indicating any notable change in discharge with varied meandering coefficients in LISFLOOD-FP.

P11. L29: "While at the most upstream station Loc3 DFM simulates lower water levels than LFP": This is probably due to larger flooding in Delft3D due to its coarser spatial resolution in upstream, as discussed above. Please clarify.

Thank you for this remark. We agree with your suggestion and have added the link between differences in simulated water level and simulated inundation extent to the revised manuscript. Besides, we qualitatively correlae it to the additional discharge simulations made in upstream areas.

P11. L34: "the more pronounced difference in water levels at Loc1 may simply be a local effect": What is the "local effect". Please explain in detail.

Thank you for pointing out the lack of clarity. A previous study already showed that the behaviour of simulated water level is not always predictable due to spatial feedback dynamics between neighbouring cells of an observation station (Hardy et al., 1999). Since we cannot really explain the more pronounced difference in water levels at this location with our current process understanding of the coupled set-up, we added this reference to the discussion of model results in the revised manuscript as a possible source of error.

With the additions made to the manuscript based on the valuable and critical reviewer's remarks, we are convinced to have responsibly addressed all uncertainties, ambiguities, and shortcomings of the initially submitted version.

Main Changes to Manuscript

Comments reviewer #1 (changes made highlighted yellow in revised manuscript):

1. Replaced too colloquial language in the manuscript.
2. Provided spatial resolution of the CRU-data as well as link to document describing the preparation of CRU-forcing in PCR-GLOBWB in section 2.1.
3. Added the models currently available within GLOFRIM to Figure 1.
4. Added reference to the original Shuttle Rader Topography Mission (SRTM) data to section 3.
5. Mentioned possible uncertainty in observed discharge data, and provided source quantifying this uncertainty in section 5.1.
6. Plotted results of sensitivity analysis of both meandering coefficient and surface roughness in LISFLOOD-FP in Figures 6a and 6b, respectively, and elaborated on it section 5.2.
7. Explicitly re-stated the different gridding techniques employed by the two hydrodynamic models in section 5.2.

Comments reviewer #2 (changes made highlighted light-blue in revised manuscript):

1. Re-wrote the description of the River-Floodplain-Scheme in section 3.
2. Added reference to the MERIT DEM in section 3.
3. Updated and re-wrote section to clearer describe functionality of GLOFRIM with respect to the use of different types of downstream boundaries in section 5.1.
4. Explicitly stated that PCR-GLOBWB is not made use of in the synthetic test case in section 4.1
5. Clearly stated how downstream boundaries, river braiding, and bifurcations are treated in the hydrodynamic schematizations used in the study in section 5.1, and discussed possible shortcomings in the schematization with respect to model results in section 5.2.
6. Added the upscaling technique applied and gave reasoning in section 5.1.
7. Did not made changes to manuscript since reasoning and references for using a uniform surface roughness coefficient are already provided in manuscript
8. Mentioned use of kinematic wave approximation as potential source of error in section 5.1, and elaborated on it accordingly in discussion section 5.2; also, recommended use of large-scale 1-D models for upstream/midstream section in section 6.
9. Stated in section 5.1 more clearly that regional optimization is only optional and advised for catchment studies, hence not contradicting any global application.
10. Added two stations upstream of Obidos (see updated Figure 4) and assessed simulated discharge; added discussion of results to section 5.2 and recommended further investigation in section 6.
11. Improved description of D4 system and assessed impact of accounting for both over- and underestimation of channel dimension in Figure 6a; briefly elaborated on it in section 5.2.
12. Established stronger relation between spatial resolution and simulated water levels in section 5.2.; besides, used additional upstream discharge simulation to underpin the statement.
13. Added explanation and reference of the local effects that can be observed for water level simulations in section 5.2.

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

Jannis M. Hoch[1,2], Jeffrey C. Neal[3], Fedor Baart[2], Rens van Beek[1], Hessel C. Winsemius[2,4], Paul D. Bates[3], Marc F.P. Bierkens[1,2]

[1] Department of Physical Geography, Utrecht University, P.O. Box 80115, 3508 TC Utrecht, the Netherlands
[2] Deltares, P.O. Box 177, 2600 MH Delft, the Netherlands
[3] School of Geographical Sciences, University of Bristol, University Road, Bristol, BS8 1SS, UK
[4] Institute for Environmental Studies, VU University, De Boelelaan 1087, 1081 HV, Amsterdam, the Netherlands

Correspondence to: Jannis M. Hoch (j.m.hoch@uu.nl)

Abstract. ~~To increase the representation of physical processes in inundation modelling, current research approaches aim to integrate both hydrological and hydrodynamic models. A previous study by Hoch et al. (2017) showed that spatially explicit coupling approaches can outperform stand-alone runs by single-purpose models as they combine spatially distributed model forcing by hydrological models with more sophisticated routing schemes in hydrodynamic models. ~~We here present GLOFRIM, a globally applicable computational framework for integrated hydrological-hydrodynamic modelling~~,~~ GLOFRIM ~~to ~~facilitates ~~such ~~spatially explicit coupling ~~approaches ~~of hydrodynamic and hydrologic models and ~~to ~~caters for an ensemble of models to be coupled. It currently ~~allows for coupling~~encompasses the global hydrological model PCR-GLOBWB ~~with either~~as well as the hydrodynamic models Delft3D Flexible Mesh (DFM), solving the full shallow-water equations and allowing for spatially flexible meshing, ~~or ~~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 hydrological 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 ~~and, in addition,~~ but also to perform a first-ever benchmark of flexible and regular grids at the 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 ~~is ~~are to a large extent attributable to the gridding techniques employed. In fact, the result show that both the numerical scheme of the inundation model and the gridding technique can contribute ~~as strongly ~~to deviations in simulated inundation extent as~~, unlike the global flood model inter-comparison by Trigg et al. (2016), ~~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 ~~to be ~~easily extendable, and thus we hope that other large-scale hydrological and hydrodynamic models will be added.~~, eventually~~ Eventually, ~~capturing ~~more locally relevant processes would be captured ~~as well as~~and ~~allowing for ~~more robust model inter-comparison, benchmarking, and ensemble simulations of flood hazard at the large scale would be allowed for.

1 Introduction

In the latter half of the last century, losses due to riverine floods increased greatly, leading to economic losses of more than \$1 billion and 220,000 casualties since 1980 (Munich Re, 2013; 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). ~~Sound~~ 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 (Dottori et al., 2016; Pappenberger et al., 2012; Sampson et al., 2015; Winsemius et al., 2013; Yamazaki et al., 2011), differing in their process descriptions and computational engine. While some approaches derive flood hazard from a coarse-scale hydrological 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 hydrological model (GHM) such as PCR-GLOBWB (van Beek et al., 2011; van Beek and Bierkens, 2008), 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 hydrological models can generally be described as simplistic as they are based on gridded drainage networks at coarse spatial resolution, with the currently finest spatial resolution of GHMs 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 1-D, 2-D or combined 1-D/2-D, 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 1-D hydrodynamics and parameterized 2-D floodplain storage. Applying such a 1-D/2-D approach, however, does not allow for explicit modelling of floodplain flow pathways as well as channel-floodplain interactions. Explicitly representing these processes would be beneficial as they are known to greatly influence inundation dynamics and patterns (Neal et al., 2012a; Trigg et al., 2009). Compared to hydrological models, hydrodynamic models solving the full SWE or at least a more advanced approximation such as the local inertia equations (LIE) have the advantage of providing a better representation of backwater effects, which are important flood-triggering processes (Meade et al., 1991; Moussa and Bocquillon, 1996; Paiva et al., 2013). Another difference to GHMs is that current applications of hydrodynamic models at the large to global scale can run at spatial resolutions of up to 1 km (Sampson et al., 2015), greatly facilitating the representation of both relevant channel-floodplain interactions (Rudorff et al., 2014a, 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 hydrological processes and thus may overpredict both inundation extent and depth as, for instance, groundwater infiltration and evaporation from inundated floodplains are currently not 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, hydrological-hydrodynamic coupling was already applied in a number of studies (Biancamaria et al., 2009; Kim et al., 2012; Lian et al., 2007; Schumann et al., 2013)~~, but none of these studies coupled hydrology and hydrodynamics in a spatially explicit manner, that is on a grid-by-grid basis~~. ~~Instead~~For example, ~~they employed~~ output from hydrological or land-surface models was used as input to the 1-D/2-D hydrodynamic model LISFLOOD-FP ~~at a number of locations~~ (Bates et al., 2010; Bates and de Roo, 2000) 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 hydrological 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 trans-national 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 (2017)). This tool enables the user to create a coupled model from a list of readily available models and run it on a server of 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.

Recently, Hoch et al. (2017) 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 hydrological and hydrodynamic processes occurring over the entire study area. Results indicate that spatially explicit coupling of hydrological 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 hydrological 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 hydrological 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 hydrological and hydrodynamic models, we developed GLOFRIM, a GLObally applicable computational FRamework for Integrated hydrological-hydrodynamic Modelling. In addition to the work of Hoch et al. (2017), it includes the widely used hydrodynamic model LISFLOOD-FP (hereafter LFP; Bates and de Roo (2000)) and an improved as well as extended coupling algorithm, thus catering a wider range of model schematizations and applications. As we believe that by combining the locally best-performing hydrological and hydrodynamic models ~~can better capture all~~ relevant processes can be captured better, GLOFRIM is designed in an expandable way to eventually incorporate more models. Furthermore, the framework is openly available under GNU 3.0 license[1] 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 as for instance the Niger (Dadson et al., 2010; Mahe et al., 2009).

In the remainder of the 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 means for benchmarking, we assess simulated discharge along the flow paths as well as run times for a 1-D and 2-D 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. Pearson's 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. 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. (2017) 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 at the large scale by enhancing the integration of hydrological and hydrodynamic model processes, which eventually may lead to improved decision making as well as planning of adaption and mitigation measures.

2 Models

Currently, GLOFRIM includes the hydrological model PCR-GLOBWB as well as the hydrodynamic models Delft3D Flexible Mesh and LISFLOOD-FP. 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.

2.1 PCR-GLOBWB

To generate hydrological input, the global hydrological model PCR-GLOBWB (PCR) is currently incorporated in the framework. It can be applied at 30 arcmin resolution (approximately 55 km x 55 km at the Equator) as well as 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 (Karssenberg et 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 (LDD) network employing the kinematic wave approximation. The model is forced with Climate Research Unit (CRU) precipitation and temperature data ~~ ~~(Harris et al., 2014) at 30 arcmin

---

[1] The code and user manual of GLOFRIM is downloadable at doi.org/10.5281/zenodo.597107

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/ERAI (Kållberg 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, 2017). For more detailed information on CRU-fording, its processing, and PCR in general, we refer to the relevant literature (van Beek, 2008; van Beek et al., 2011; van Beek and Bierkens, 2008). PCR was already applied for a wide range of studies such as flood and drought forecasting (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 basis for the Aqueduct Global Flood Analyzer of the World Resources Institute (World Resources Institute, 2017).

2.2 Delft3D Flexible Mesh

Delft3D Flexible Mesh (DFM) allows the user to schematize the model domain with a flexible mesh in 1-D/2-D/3-D, 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 (SWE). The main partial differential equations solved by DFM are

$$\frac{\partial h}{\partial t}+\nabla\cdot\left(h\mathbf{u}\right)=0 \qquad (1)$$

$$\frac{\partial \mathbf{u}}{\partial t}+\frac{1}{h}\left(\nabla\cdot(\mathrm{h}\,\mathbf{uu})-\mathbf{u}\nabla\cdot\left(h\mathbf{u}\right)\right)=-g\nabla\zeta+\frac{1}{h}\nabla\cdot\left(vh\left(\nabla\mathbf{u}+\nabla\mathbf{u}^{T}\right)\right)+\frac{1}{h}\frac{\tau}{\rho} \qquad (2)$$

With

$$\nabla=\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y}\right)^{T} \qquad (3)$$

ζ being the water level, h the water depth, $\boldsymbol{u}$ is the velocity vector, g the gravitational acceleration, v the viscosity, ρ the water mass density, and τ the bottom friction. For 1-D 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 (Deltares, 2017a). 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 a wider range of applications, for instance groundwater flow, sediment transport, and water quality simulations in 1-D, 2-D, and 3-D. For more information regarding the application of DFM, we refer to the User Manual (Deltares, 2017b). Due to its very recent publication, 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 10 compared to square grids with equal quality of model output.

2.3 LISFLOOD-FP

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 2-D only, a 1-D, a 1-D/2-D 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). Here *q* denotes the flow per unit width, *g* the gravitational acceleration, ζ the water level, *R* the hydraulic radius, *n* Manning's surface roughness, and ∇ the gradients in x- and y-direction as described in Eq. 3:

$$\frac{\partial q}{\partial t}+\nabla gh\zeta+\frac{gn^2q^2}{R^{4/3}h}=0 \tag{4}$$

Mass conservation is implemented as

$$\nabla(h+q)=0 \tag{5}$$

Whereby *Δt* denotes the time step, *Δx* the cell size and *i,j* the cell indices. For further information about model development, derivation of numerical solutions, assumptions, and validations, we refer to the above-mentioned papers.

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. (2015) who applied LFP to compute global estimates of flood hazard and risk as well as by Schumann et al. (2013) and Biancamaria et al. (2009) who used LFP to simulate inundation in the Zambezi River and Ob River, respectively, forced with lateral input from a land surface model.

The 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.

2.4 Basic Model Interface

Generally, the Basic Model Interface (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 ~~number~~ set of variables from memory. ~~This number~~The variables exposed through the BMI adapter can be defined during the development of the BMI adapter and is thus not limited to a pre-set 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. It is noteworthy that implementing the BMI functions does not alter any functionality or routines in the models. Both DFM and LFP, although not being coded in Python, can be called from within Python using the BMI-python package (see https://github.com/openearth/bmi-python). For further information regarding the BMI, we refer to Peckham et al. (2013) and the related website (CSDMS, 2016).

3 The computational framework GLOFRIM

The computational framework presented here consists of two key elements, a) the actual code and b) a settings-file. Hereafter, a brief overview is given of their main properties. More detailed information and 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 1 for a flow chart). The models are first initialized, that is, 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 1-D or 2-D 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. (2017). 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 as well as ~~externally~~ adapted to agree with the data structure of the chosen hydrodynamic model. Subsequently, either the water depth or a flux variable in the hydrodynamic model will be overwritten, and finally the hydrodynamic model will be updated until 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 hydrological 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”, by specifying the variables to be updated, and by choosing for hydrodynamic models in either spherical or projected coordinate systems.

The ~~so-called~~ “River-Floodplain-Scheme” (RFS) defines where output from PCR is coupled to. If RFS is activated, water volume of one PCR cell is directly coupled to the 1-D channels of the hydrodynamic model within the corresponding PCR cell, while, when RFS is inactive, water is distributed over all 2-D grid cells ~~of the 2-D domain~~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 ~~the~~ remotely sensed 2-D elevation data such as Shuttle Rader Topographic Mission (SRTM) data (Farr et al., 2007). Recent research showed that such global data sets~~which is known to~~ contain strong vertical bias as well as systematic and random noise (Yamazaki et al., 2017). ~~, in particular when derived from remotely sensed global data sets such as Shuttle Rader Topographic Mission (SRTM) data.~~ In particular, ~~simulation of~~simulating flow over vertically irregular terrain resulting in super-critical regimes is contra-indicated 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 *m* or by updating fluxes, which are expressed as discharge in LFP in m^3/s and as precipitation in *mm/d* in DFM. For DFM, added daily water depth is divided over 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 ~~for~~ this option. While it is also possible to perform state-updating in LFP, ~~we found~~test runs showed that this option ~~has to~~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 ~~spherical~~ 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 BMI's functionalities to couple hydrological 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). ~~Eventually, only certain arrays of, for example, inundation depth obtained in the hydrodynamic model needs to be linked with actual evaporation rates as well as groundwater infiltration.~~ 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 as well as extendable to the user's preferences or individual modelling requirements, GLOFRIM contains a number of additional advantages: first, by having PCR-GLOBWB, or any other GHM, as the hydrological 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, also incorporating important groundwater infiltration and evaporation processes; fourth, by guaranteeing identical hydrological forcing, applying the computational framework facilitates benchmarking of hydrodynamic models by eliminating a sources of difference, potentially supporting hydrodynamic ensemble modelling approaches.

4 The Synthetic Test Cases

4.1 Set-up

To gain insight in possible differences in model behaviour between LFP and DFM, we created two synthetic test cases, one being set-up as 1-D only (STC 1-D) and the other as 2-D only (STC 2-D). 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 1-D 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. As default settings, we applied Manning's surface roughness coefficients of 0.04 s $m^{-1/3}$ for the 1-D only run and 0.07 s $m^{-1/3}$ for the 2-D only run. Both synthetic test cases were forced with an artificial upstream discharge boundary spanning one year and consisting of two peak flow moments to introduce variability in model dynamics, thus not employing GLOFRIM for those test cases. As ~~a~~ downstream boundary condition a constant water level of 0 m was set. The entire simulation period was three years to ensure ~~that all water has drained before the end of the run~~it exceeds the time of concentration. To assess model output, seven 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~~models, we then compared discharge along the cross-sections as well as run times to obtain a first indication how the different computational schemes might vary (Figure 2).

4.2 Results and Discussion

Assessing the results for both 2-D and 1-D, we find that both models simulate the same responses to the input signal applied (Figure 3). Due to the higher friction coefficient and the wider flow area, it takes the 2-D schematization almost the entire simulation period to entirely convey the water volumes to the downstream boundary. In the 1-D schematization, however, all water is already drained after around 30 per cent 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. (2012b) 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* << 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~~Also, Neal et al. (2012b) showed that it appears 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~~in particular~~ for the 2-D schematization (Table 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.

5 Test case: the Amazon River basin

5.1 Set-up

To test GLOFRIM in an actual test case as well as to benchmark the flexible and regular grid, 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. (2017). First, a regular 2-D 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 stronger 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., 2012). Elevation data was then assigned to the flexible mesh by spatial averaging. For the 1-

D channel network and bathymetry, ~~global ~~river width data of the Global Width Database for Large Rivers (GWD-LR; ~~(~~Yamazaki et al.~~,~~ (2014)) 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 LPF 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 was upscaled to a 2 km spatial resolution, employing the nearest neighbour technique, ~~which equals~~to match the finest spatial resolution of the DFM schematization (Figure 4). From Figure 4 it is visible that LFP contains a greater level of detail in areas farther upstream due to the finer spatial resolution uniformly applied. ~~As a consequence~~Consequently, the total number of cells in LFP exceeds the number of 2-D cells in DFM by a factor 4 (Table 2). Furthermore, only around 10 per cent of the entire schematization represents 1-D channels in LFP, while the channel network of DFM was based on around 30 per cent of all DFM cells. For both DFM and LFP, Manning's surface roughness coefficient was uniformly set to 0.03 s m$^{-1/3}$ for channel and floodplains which is consistent with other case studies in the Amazon (Paiva et al., 2013; Rudorff et al., 2014a, 2014b; Trigg et al., 2009; Yamazaki et al., 2011). As downstream boundary, we imposed a constant water level of 0 m at the river's 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, is should be mentioned that the 1-D 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. This is, however, not due to the inability of the hydrodynamic models to account for them, but merely because the chosen algorithm to derive 1-D network properties does not allow for it.

For the hydrological model PCR-GLOBWB, 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 value (Hoch et al. (2017). As such optimization technique is optional and only advisable for catchment studies, a global application is not constrained thereby. In analogy to the hydrodynamic models, the surface roughness coefficient of PCR was uniformly set to 0.03 s m$^{-1/3}$.

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 4). To that end, Pearson's r, the relative mean square error RMSE, and the Kling-Gupta Efficiency 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 01/1984 until 12/1990 with the first year being used for spin-up of the coupled setting. 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 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 dataset. First, the hit rate H was computed based on the subsequent equation:

$$H = \frac{N_{DFM} \cap N_{LFP}}{N_{LFP}} \qquad (6)$$

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 resembled 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~~addition, we determined the false alarm ratio F to also ~~take into account~~consider false positive alarms. The false alarm ratio can be obtained with

$$F = \frac{N_{DFM} \setminus N_{LFP}}{N_{DFM} \cap N_{LFP} + N_{DFM} \setminus N_{LFP}} \qquad (7)$$

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 ratio into one parameter which can vary between 0 in the worst and 1 in the best scenario, indicating perfect match between both inundation maps:

$$C = \frac{N_{DFM} \cap N_{LFP}}{N_{DFM} \cup N_{LFP}} \qquad (8)$$

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

5.2 Results and Discussion

Benchmarking discharge results against observation from GRDC at Óbidos shows that both models behave similarly. However, LFP tends to compute earlier peak flow as well as earlier and lower low flow (Figure 5). ~~As a consequence~~Therefore, obtained coefficients of correlation are lower for LFP, while the model's skill as expressed by KGE are higher for LFP and the RMSEs are comparable (Table 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-GLOBWB (Hoch et al., 2017), 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 and since the hydrological 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 southerly, northerly, easterly or westerly direction, channel length and dimension in LPF tend to ~~be longer than in~~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 scale river ~~length~~dimensions, however, did not show any significant impact on simulated discharge (**Error! Reference source not found.**a). After investigating how changes in surface roughness values in LFP may ~~affect discharge estimates from LFP~~close the gap to DFM, we indeed found ~~different~~ a more pronounced response~~s to variations in surface roughness than DFM~~, yet it cannot satisfactorily explain the difference in simulated discharge either (**Error! Reference source not found.**b). ~~Yet, we know from~~Since in the synthetic example ~~that~~ both models can produce ~~similar~~ near-identical results ~~when~~ if using the same friction coefficient and, ~~since~~ because the flow regime in the Amazon basin can be described as sub-critical, different sensitivity to surface roughness over large areas can thus also be disregarded as 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 1-D channel bathymetry, channel-floodplain interaction, and input elevation assignment due to the different ~~gridding approaches applied~~grid schematizations of a flexible mesh and regular grid, respectively.

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 4). Results indeed suggest that the differences in upstream spatial resolution result in different flood wave propagation. (Figure 6c): with covered flow distance, peak discharge of 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, respectively, 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 that especially discharge computations in upstream areas can be easily affected as there the discrepancy in cell size is largest there.

Assessing differences in simulated water level dynamics at the observation locations, we cannot find any particularly prevailing difference between the models' response to hydrological forcing (**Error! Reference source not found.**~~Figure 6~~). In ~~general~~general, we observe that modelled water levels are comparable, yet with locally differing patterns. While at the most upstream station ~~Loc3~~ Loc5 DFM simulates lower water levels than LFP, this is opposite at the most downstream station ~~Loc1~~Loc3, and at ~~Loc2~~ 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 as well as their dynamics could not be fully attributed to one specific cause. For example, the more pronounced difference in water levels at Loc1 may ~~simply~~ be a local effect due to spatial feedback dynamics between neighbouring cells of an observation station (Hardy et al., 1999)~~and may,~~ may be related to slight differences in model schematization at the downstream boundary or to backwater effects in the delta regions ~~affecting results differently~~as a result of different influence 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, ~~A~~assessing the local properties of the observation stations revealed that the surface elevation in DFM is higher than in LFP (Table 4)~~, and due to the flexible meshing, cell size can vary greatly too (Figure 4).~~ Last, results indicate that d~~D~~ifferences in ~~cell size~~gridding and therefore cell size~~gridding~~ may thus ~~also~~ have locally impacted the overall water levels ~~as well~~too since the above-discussed discharge simulations in upstream areas exhibited clear deviations between both models (Figure 6c).

Regarding the run times of the two coupled set-ups, we find that it takes LFP around six hours to simulate the entire simulation period of seven years, that is model time plus spin-up, while performing the same simulation with DFM takes around seven hours (Table 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 1-D and 2-D 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.

We find that inundation extents obtained at the end of the simulation runs with DFM and LFP are comparable, yet far from identical (Figure 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 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 by Trigg et al. (2016) .

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 2-D 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 2-D 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 1-D 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 seem to be faster in LFP than in DFM.

6 Conclusion and recommendations

In this study, we presented GLOFRIM, a GLObally applicable computational FRamework for Integrated hydrological-hydrodynamic Modelling. In its current version, it provides an environment to one-directionally couple the global hydrological model PCR-GLOBWB (PCR) with two hydrodynamic models: Delft3D Flexible Mesh (DFM) solving the full shallow-water equations, and LISLFOOD-FP (LFP) solving the local inertia equations. 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, adapted to specific purposes, and extended with other models; (ii) by employing a global hydrological 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 hydrological forcing for experiments.

GLOFRIM at present provides a range of ~~possible~~ options for model coupling. Users can choose between coupling PCR to either the 1-D or 2-D 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 ~~a number of~~many other global hydrological 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 the coupling of other hydrological or hydrodynamic models, merely requiring the implementation of the BMI into each model to be added. Eventually, adding a 1-D continental hydrodynamic 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 2-D floodplain computations. ~~The~~ Empoying a Basic Model Interface (BMI) does not change the model functionality while at the same time providing a range of added functions.

Furthermore, not all model variables need to be exposed, only those to reproduce model geometry, distinguish between 1-D and 2-D cells, and ~~a variable~~ to~~ be~~ update~~d~~ 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 at 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 made. 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 in how both models may differ regarding their computational complexity. Thereby both the 1-D and 2-D 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., 2012b).

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 ~~for the test case in the Amazon River basin~~ shows that both models exhibit comparable results with LFP tending to compute earlier and slightly increased peak discharge estimates. As thorough testing of ~~possible causes~~plausible causes did not show significant improvements, we speculate that differences in processing of 1-D channel bathymetry, interaction between 1-D channels and 2-D 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 hydrological 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 four times the number of 2-D cells than DFM while the number of 1-D cells is 40 per cent higher in LFP as 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 2-D domain and the nature of the coupling algorithm applied: because water was coupled directly into the 1-D channels, flow over the 2-D 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 ~~a number of~~several additional physical processes, such as the potential to simulate 3-D 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, ~~C=0.46~~ 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 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 case of employing a flexible ~~mesh~~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 ~~related run time~~model accuracy. ~~Besides, further benchmarking of the impact of flexible meshes on model accuracy with respect to regular grids is recommended.~~

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 hydrological and hydrodynamic models will lead to improved flood risk assessments and planning of climate change impact mitigation and adaption measures.

Code and/or data availability. The code of GLOFRIM as well as the BMI-versions of LISFLOOD-FP and PCR-GLOBWB are openly accessible and freely downloadable at doi.org/10.5281/zenodo.597107

Author contributions. Fedor Baart and Jannis M. Hoch developed the BMI adapter for LISFLOOD-FP. Jannis M. Hoch, Rens van Beek, and Jeffrey C. Neal developed the code of the computational framework. Jeffrey C. Neal, Rens van Beek, Hessel C. Winsemius, Paul D. Bates, and Marc F. P. Bierkens supervised the research and provided important advice. Jannis M. Hoch designed and executed the research, and also prepared the manuscript as well as code, with contribution from all co-authors.

Competing interests. Author Jeffrey C. Neal is a topical editor of the journal.

Acknowledgements. The authors declare that they have no conflict of interest. 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.

References

Alcamo, J., Döll, P., Kasper, F. and Siebert, S.: Global change and global scenarios of water use and availability: An Application of WaterGAP1.0., 1997.

De Almeida, G. A. M. and Bates, P. D.: Applicability of the local inertial approximation of the shallow water equations to flood modeling, Water Resour. Res., 49(8), 4833–4844, doi:10.1002/wrcr.20366, 2013.

Bates, P. D. and de Roo, A.: A simple raster-based model for flood inundation simulation, J. Hydrol., 236(1–2), 54–77, doi:10.1016/S0022-1694(00)00278-X, 2000.

Bates, P. D., Horritt, M. S. and Fewtrell, T. J.: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, J. Hydrol., 387(1), 33–45, doi:10.1016/j.jhydrol.2010.03.027, 2010.

Baugh, C. A., Bates, P. D., Schumann, G. J.-P. and Trigg, M. A.: SRTM vegetation removal and hydrodynamic modeling accuracy, Water Resour. Res., 49, 5276–5289, doi:10.1002/wrcr.20412, 2013.

van Beek, L. P. H.: Forcing PCR-GLOWB with CRU data, Department of Physical Geography, Utrecht University. [online] Available from: http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pdf, 2008.

van Beek, L. P. H. and Bierkens, M. F. P.: The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification. [online] Available from: http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf, 2008.

van Beek, L. P. H., Wada, Y. and Bierkens, M. F. P.: Global monthly water stress: 1. Water balance and water availability, Water Resour. Res., 47(7), W07517, doi:10.1029/2010WR009791, 2011.

Beven, K. J., Cloke, H. L., Pappenberger, F., Lamb, R. and Hunter, N. M.: Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface, Sci. China Earth Sci., 58(1), 25–35, doi:10.1007/s11430-014-5003-4, 2015.

Biancamaria, S., Bates, P. D., Boone, A. and Mognard, N. M.: Large-scale coupled hydrologic and hydraulic modelling of the Ob river in Siberia, J. Hydrol., 379(1–2), 136–150, doi:10.1016/j.jhydrol.2009.09.054, 2009.

Bierkens, M. F. P.: Global hydrology 2015: State, trends, and directions, Water Resour. Res., 51(7), 4923–4947, doi:10.1002/2015WR017173, 2015.

Castro Gama, M., Popescu, I., Mynett, A., Shengyang, L. and Van Dam, A.: Modelling extreme flood hazard events on the middle Yellow River using DFLOW-flexible mesh approach, Nat. Hazards Earth Syst. Sci. Discuss., 1(6), 6061–6092, doi:10.5194/nhessd-1-6061-2013, 2013.

Ceola, S., Laio, F. and Montanari, A.: Satellite nighttime lights reveal increasing human exposure to floods worldwide, Geophys. Res. Lett., 41(20), 7184–7190, doi:10.1002/2014GL061859, 2014.

Clarke, R. T., Mendiondo, E. M. and Brusa, L. C.: Uncertainties in mean discharges from two large South American rivers due to rating curve variability, Hydrol. Sci. J., 45(2), 221–236, doi:10.1080/02626660009492321, 2000.

CSDMS: CSDMS Basic Model Interface (version 1.0), [online] Available from: https://csdms.colorado.edu/wiki/BMI_Description (Accessed 25 August 2016), 2016.

CSDMS: CSDMS Web Modeling Tool, [online] Available from: https://csdms.colorado.edu/wmt/ (Accessed 2 February 2017), 2017.

Dadson, S. J., Ashpole, I., Harris, P., Davies, H. N., Clark, D. B., Blyth, E. and Taylor, C. M.: Wetland inundation dynamics in a model of land surface climate: Evaluation in the Niger inland delta region, J. Geophys. Res. Atmos., 115(23), 1–7, doi:10.1029/2010JD014474, 2010.

Deltares: Delft3D Flexible Mesh Technical Manual. [online] Available from:

https://content.oss.deltares.nl/delft3d/manuals/D-Flow_FM_Technical_Reference_Manual.pdf, 2017a.

Deltares: Delft3D Flexible Mesh User Manual, [online] Available from: https://content.oss.deltares.nl/delft3d/manuals/D-Flow_FM_User_Manual.pdf (Accessed 3 February 2017b), 2017.

Döll, P., Kaspar, F. and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, J. Hydrol., 270(1–2), 105–134, doi:10.1016/S0022-1694(02)00283-4, 2003.

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. and Feyen, L.: Development and evaluation of a framework for global flood hazard mapping, Adv. Water Resour., 94, 87–102, doi:10.1016/j.advwatres.2016.05.002, 2016.

FAO: FAOSTAT, [online] Available from: http://www.fao.org/faostat/en/#home (Accessed 13 January 2017), 2017.

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. and Alsdorf, D.: The Shuttle Radar Topography Mission, Rev. Geophys., 45(2), RG2004, doi:10.1029/2005RG000183, 2007.

de Graaf, I. E. M., Sutanudjaja, E. H., van Beek, L. P. H. and Bierkens, M. F. P.: A high-resolution global-scale groundwater model, Hydrol. Earth Syst. Sci., 19(2), 823–837, doi:10.5194/hess-19-823-2015, 2015.

Gupta, H. V, Kling, H., Yilmaz, K. K. and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria : Implications for improving hydrological modelling, J. Hydrol., 377(1–2), 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.

Hardy, R. J., Bates, P. D. and Anderson, M. G.: The importance of spatial resolution in hydraulic models for floodplain environments, J. Hydrol., 216(1), 124–136, doi:http://dx.doi.org/10.1016/S0022-1694(99)00002-5, 1999.

Harris, I., Jones, P. D., Osborn, T. J. and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, Int. J. Climatol., 34(3), 623–642, doi:10.1002/joc.3711, 2014.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S.: Global flood risk under climate change, Nat. Publ. Gr., 3(9), 816–821, doi:10.1038/nclimate1911, 2013.

Hoch, J. M., Haag, A. V., Dam, A. van, Winsemius, H. C., Beek, L. P. H. van and Bierkens, M. F. P.: Assessing the impact of hydrodynamics on large-scale flood wave propagation - a case study for the Amazon Basin, Hydrol. Earth Syst. Sci., 21(1), 117–132, doi:10.5194/hess-21-117-2017, 2017.

Jongman, B., Hochrainer-Stigler, S., Feyen, L., Aerts, J. C. J. H., Mechler, R., Botzen, W. J. W., Bouwer, L. M., Pflug, G., Rojas, R. and Ward, P. J.: Increasing stress on disaster-risk finance due to large floods, Nat. Clim. Chang., 4(4), 1–5, doi:10.1038/NCLIMATE2124, 2014.

Kållberg, P., Berrisford, P., Hoskins, B. J., Simmons, A. J., Uppala, S., Lamy-Thepaut, S. and Hine, R.: ERA-40 atlas, ERA-40 Proj. Rep. Ser. 19, Eur. Cent. for Medium Range Weather Forecasts, Reading, UK, 2005.

Karssenberg, D., Schmitz, O., Salamon, P., de Jong, K. and Bierkens, M. F. P.: A software framework for construction of process-based stochastic spatio-temporal models and data assimilation, Environ. Model. Softw., 25(4), 489–502, doi:10.1016/j.envsoft.2009.10.004, 2010.

Kernkamp, H. W. J., van Dam, A., Stelling, G. S. and de Goede, E. D.: Efficient scheme for the shallow water equations on unstructured grids with application to the Continental Shelf, Ocean Dyn., 61, 1175–1188, doi:10.1007/s10236-011-0423-6, 2011.

Kim, J., Warnock, A., Ivanov, V. Y. and Katopodes, N. D.: Coupled modeling of hydrologic and hydrodynamic processes including overland and channel flow, Adv. Water Resour., 37, 104–126, doi:10.1016/j.advwatres.2011.11.009, 2012.

Li, H., Beldring, S. and Xu, C.-Y.: Stability of model performance and parameter values on two catchments facing changes in climatic conditions, Hydrol. Sci. J., 60(7–8), 1317–1330, doi:10.1080/02626667.2014.978333, 2015.

Lian, Y., Chan, I.-C., Singh, J., Demissie, M., Knapp, V. and Xie, H.: Coupling of hydrologic and hydraulic models for the

Illinois River Basin, J. Hydrol., 344(3–4), 210–222, doi:10.1016/j.jhydrol.2007.08.004, 2007.

Liang, X., Lettenmaier, D. P., Wood, E. F. and Burges, S. J.: A simple hydrologically based model of land surface water and energy fluxes for general circulation models, J. Geophys. Res. Atmos., 99(D7), 14415–14428, doi:10.1029/94JD00483, 1994.

Mahe, G., Bamba, F., Soumaguel, A., Orange, D. and Olivry, J. C.: Water losses in the inner delta of the River Niger: water balance and flooded area, Hydrol. Process., 23, 3157–3160, doi:10.1002/hyp.7389, 2009.

Meade, R. H., Rayol, J. M., Conceição, S. C. Da and Natividade, J. R. G.: Backwater Effects in the Amazon River of Basin, Environ. Geol. Water Sci., 18(2), 105–114, doi:10.1007/BF01704664, 1991.

Moussa, R. and Bocquillon, C.: Criteria for the choice of flood-routing methods in natural channels, J. Hydrol., 186(1–4), 1–30, doi:http://dx.doi.org/10.1016/S0022-1694(96)03045-4, 1996.

Muis, S., Verlaan, M., Winsemius, H. C., Aerts, J. C. J. H. and Ward, P. J.: Corrigendum: A global reanalysis of storm surges and extreme sea levels, Nat. Commun., 7, 12913 [online] Available from: http://dx.doi.org/10.1038/ncomms12913, 2016.

Munich Re: Munich Reinsurance Company Geo Risks Research: NatCatSERVICE Database, 2013.

Neal, J. C., Schumann, G. J.-P. and Bates, P. D.: A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas, Water Resour. Res., 48(11), 1–16, doi:10.1029/2012WR012514, 2012a.

Neal, J. C., Villanueva, I., Wright, N., Willis, T., Fewtrell, T. and Bates, P. D.: How much physical complexity is needed to model flood inundation?, Hydrol. Process., 26(15), 2264–2282, doi:10.1002/hyp.8339, 2012b.

O’Loughlin, F. E., Paiva, R. C. D., Durand, M., Alsdorf, D. E. and Bates, P. D.: A multi-sensor approach towards a global vegetation corrected SRTM DEM product, Remote Sens. Environ., 182, 49–59, doi:10.1016/j.rse.2016.04.018, 2016.

Paiva, R. C. D., Collischonn, W. and Tucci, C. E. M.: Large scale hydrologic and hydrodynamic modeling using limited data and a GIS based approach, J. Hydrol., 406(3–4), 170–181, doi:10.1016/j.jhydrol.2011.06.007, 2011.

Paiva, R. C. D., Collischonn, W. and Buarque, D. C.: Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon, Hydrol. Process., 27(3), 333–346, doi:10.1002/hyp.8425, 2013.

Pappenberger, F., Dutra, E., Wetterhall, F. and Cloke, H. L.: Deriving global flood hazard maps of fluvial floods through a physical model cascade, Hydrol. Earth Syst. Sci., 16(11), 4143–4156, doi:10.5194/hess-16-4143-2012, 2012.

Peckham, S. D., Hutton, E. W. H. and Norris, B.: A component-based approach to integrated modeling in the geosciences: The design of CSDMS, Comput. Geosci., 53, 3–12, doi:10.1016/j.cageo.2012.04.002, 2013.

Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J. and Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia, Remote Sens. Environ., 112(9), 3469–3481, doi:10.1016/j.rse.2008.03.018, 2008.

Rudorff, C. M., Melack, J. M. and Bates, P. D.: Flooding dynamics on the lower Amazon floodplain: 1. Hydraulic controls on water elevation, inundation extent, and river-floodplain discharge, Water Resour. Res., 50(1), 619–634, doi:10.1002/2013WR014091, 2014a.

Rudorff, C. M., Melack, J. M. and Bates, P. D.: Flooding dynamics on the lower Amazon floodplain: 2. Seasonal and interannual hydrological variability, Water Resour. Res., 50(1), 635–649, doi:10.1002/2013WR014714, 2014b.

Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L. and Freer, J. E.: A high-resolution global flood hazard model, Water Resour. Res., 51, 7358–7381, doi:10.1002/2015WR016954, 2015.

Schumann, G. J.-P., Neal, J. C., Voisin, N., Andreadis, K. M., Pappenberger, F., Phanthuwongpakdee, N., Hall, A. C. and Bates, P. D.: A first large-scale flood inundation forecasting model, Water Resour. Res., 49(10), 6248–6257, doi:10.1002/wrcr.20521, 2013.

Tayefi, V., Lane, S. N., Hardy, R. J. and Yu, D.: A comparison of one- and two-dimensional approaches to modelling flood inundation over complex upland floodplains, Hydrol. Process., 21(23), 3190–3202, doi:10.1002/hyp.6523, 2007.

Trigg, M. A., Wilson, M. D., Bates, P. D., Horritt, M. S., Alsdorf, D. E., Forsberg, B. R. and Vega, M. C.: Amazon flood wave hydraulics, J. Hydrol., 374(1–2), 92–105, doi:10.1016/j.jhydrol.2009.06.004, 2009.

Trigg, M. A., Birch, C. E., Neal, J. C., Bates, P. D., Smith, A., Sampson, C. C., Yamazaki, D., Hirabayashi, Y., Pappenberger, F., Dutra, E., Ward, P. J., Winsemius, H. C., Salamon, P., Dottori, F., Rudari, R., Kappes, M. S., Simpson, A. L., Hadzilacos, G. and Fewtrell, T. J.: The credibility challenge for global fluvial flood risk analysis, Environ. Res. Lett., 11(9), 94014, doi:10.1088/1748-9326/11/9/094014, 2016.

Visser, H., Bouwman, A., Ligtvoet, W. and Petersen, A. C.: A statistical study of weather-related disasters: past, present and future, PBL Netherlands Environ. Assess. Agency, Hague/Bilthoven, Netherlands, 2012.

Voinov, A. and Shugart, H. H.: “Integronsters”, integral and integrated modeling, Environ. Model. Softw., 39, 149–158, doi:10.1016/j.envsoft.2012.05.014, 2013.

Wanders, N. and Wada, Y.: Human and climate impacts on the 21st century hydrological drought, J. Hydrol., 526(0), 208–220, doi:http://dx.doi.org/10.1016/j.jhydrol.2014.10.047, 2015.

Ward, P. J., Jongman, B., Salamon, P., Simpson, A., Bates, P. D., de Groeve, T., Muis, S., de Perez, E. C., Rudari, R., Trigg, M. A. and Winsemius, H. C.: Usefulness and limitations of global flood risk models, Nat. Clim. Chang., 5(8), 712–715, doi:10.1038/nclimate2742, 2015.

Winsemius, H. C., van Beek, L. P. H., Jongman, B., Ward, P. J. and Bouwman, A.: A framework for global river flood risk assessments, Hydrol. Earth Syst. Sci., 17(5), 1871–1892, doi:10.5194/hess-17-1871-2013, 2013.

Winsemius, H. C., Aerts, J. C. J. H., van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., Kwadijk, J. C. J., Ligtvoet, W., Lucas, P. L., van Vuuren, D. P. and Ward, P. J.: Global drivers of future river flood risk, Nat. Clim. Chang., 6(4), 381–385 [online] Available from: http://dx.doi.org/10.1038/nclimate2893, 2016.

Wood, E. F., Lettenmaier, D. P. and Zartarian, V. G.: A land-surface hydrology parameterization with subgrid variability for general circulation models, J. Geophys. Res., 97(D3), 2717, doi:10.1029/91JD01786, 1992.

World Resources Institute: Aqueduct Global Flood Analyzer, [online] Available from: http://floods.wri.org (Accessed 17 June 2017), 2017.

Yamazaki, D., Kanae, S., Kim, H. and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resour. Res., 47(4), 1–21, doi:10.1029/2010WR009726, 2011.

Yamazaki, D., Baugh, C. A., Bates, P. D., Kanae, S., Alsdorf, D. E. and Oki, T.: Adjustment of a spaceborne DEM for use in floodplain hydrodynamic modeling, J. Hydrol., 436–437, 81–91, doi:10.1016/j.jhydrol.2012.02.045, 2012.

Yamazaki, D., O’Loughlin, F. E., Trigg, M. A., Miller, Z. F., Pavelsky, T. M. and Bates, P. D.: Development of the Global Width Database for Large Rivers, Water Resour. Res., 50, 2108–2123, doi:10.1002/2012WR013085.Received, 2014.

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O’Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S. and Bates, P. D.: A high-accuracy map of global terrain elevations, Geophys. Res. Lett., 44(11), 5844–5853, doi:10.1002/2017GL072874, 2017.

Yossef, N. C., van Beek, L. P. H., Kwadijk, J. C. J. and Bierkens, M. F. P.: Assessment of the potential forecasting skill of a global hydrological model in reproducing the occurrence of monthly flow extremes, Hydrol. Earth Syst. Sci., 16, 4233–4246, doi:10.5194/hess-16-4233-2012, 2012.

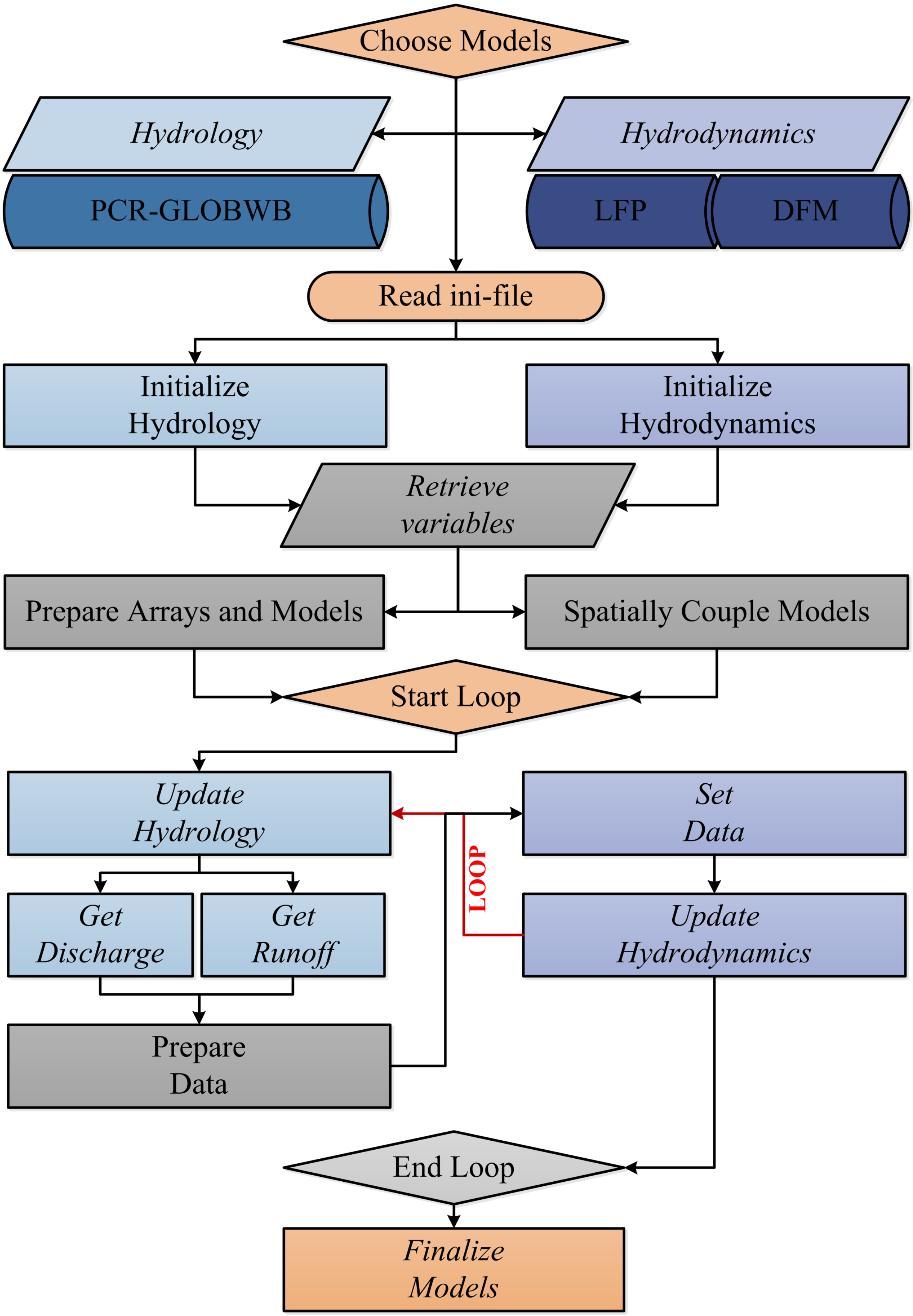

Figure 1: Flow diagram of steps executed in GLOFRIM as well as model currently available within the framewok; all steps in italic are ~~taken by using~~executed by employing the Basic Model Interface (BMI)

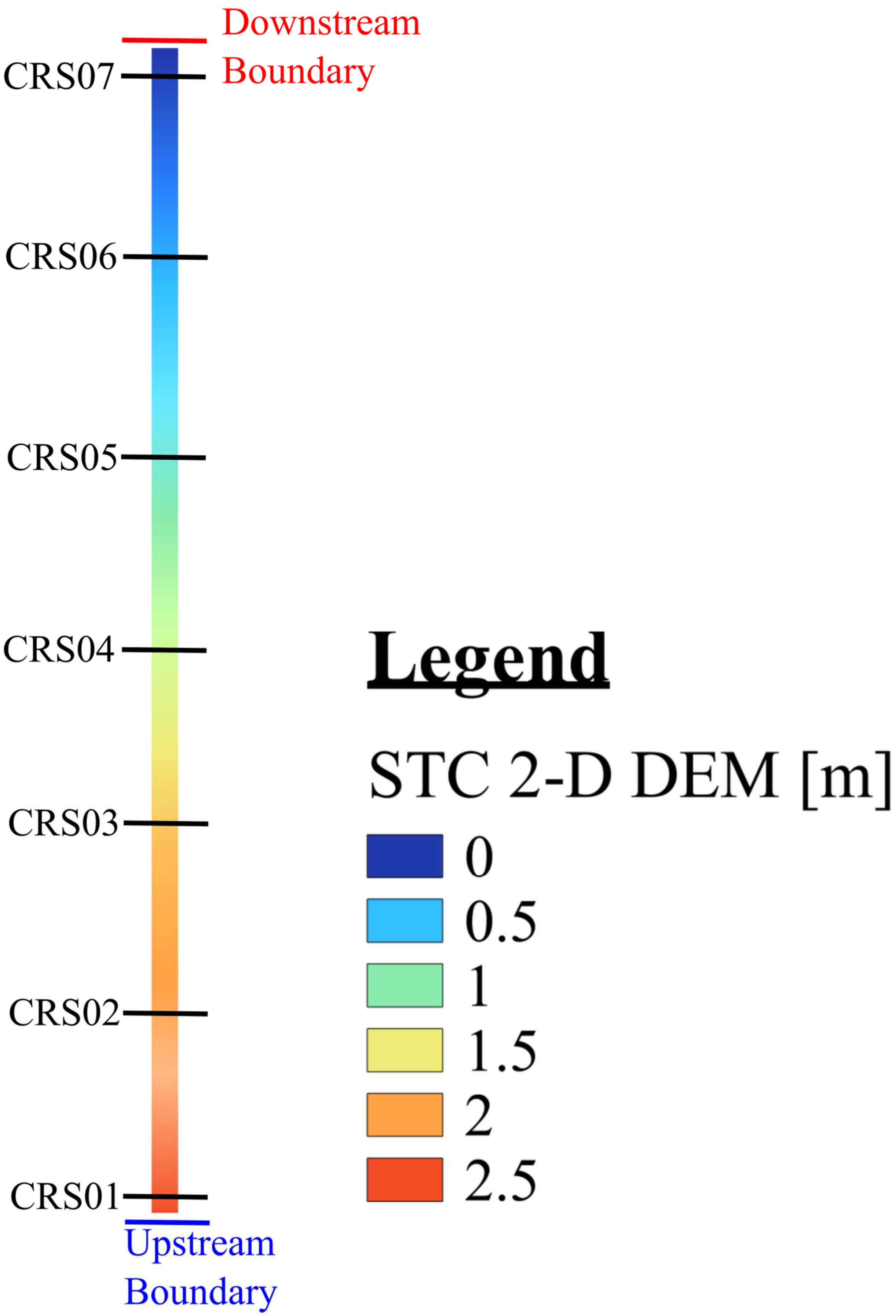

Figure 2: DEM of the 2-D synthetic test case for LFP and DFM.

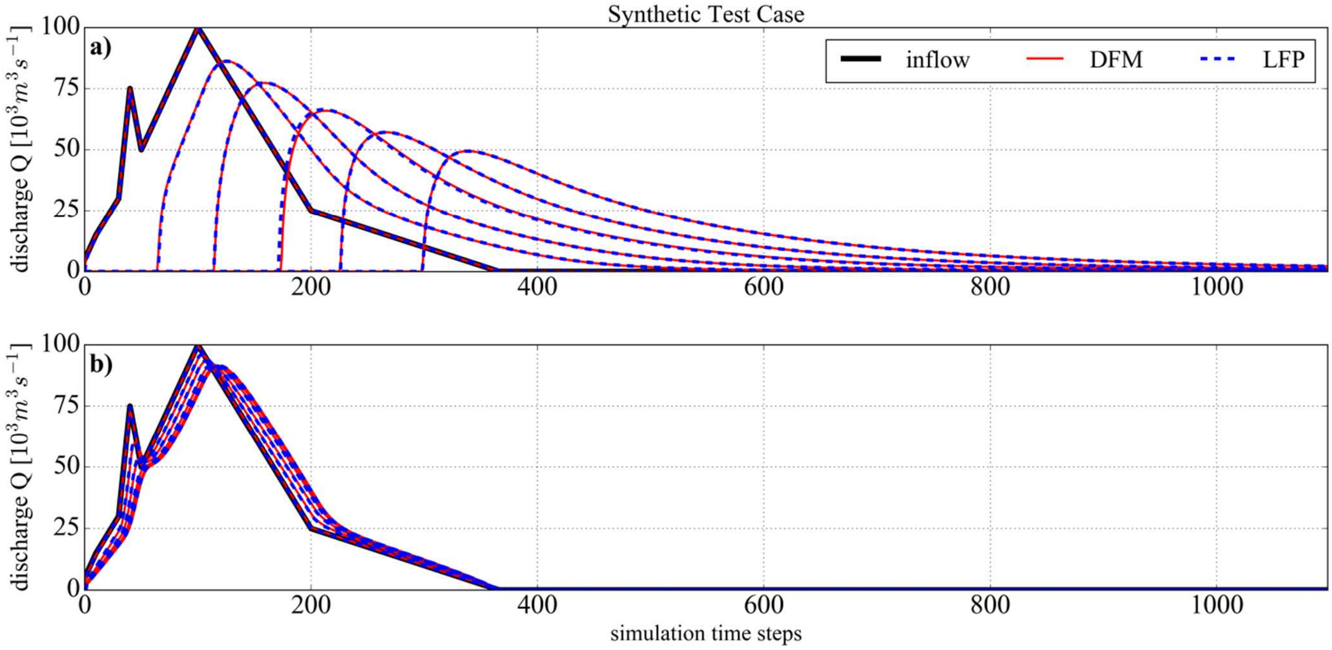

Figure 3: Simulated discharge of (a) 2-D and (b) 1-D synthetic test case

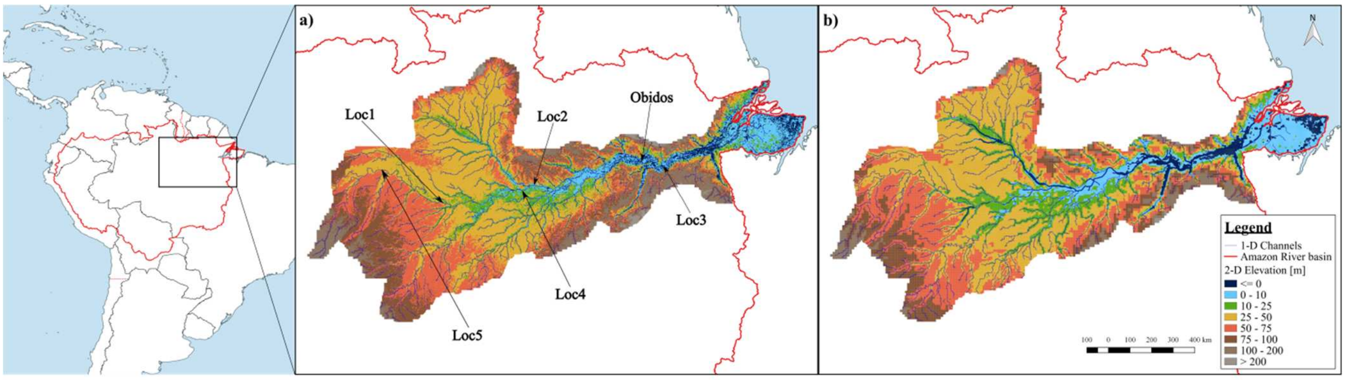

Figure 4: Digital Elevation Model as well as 1-D channel network as used in LFP (a) and DFM (right); discharge was benchmarked and validated at Óbidos while water levels were compared at three locations throughout the domain

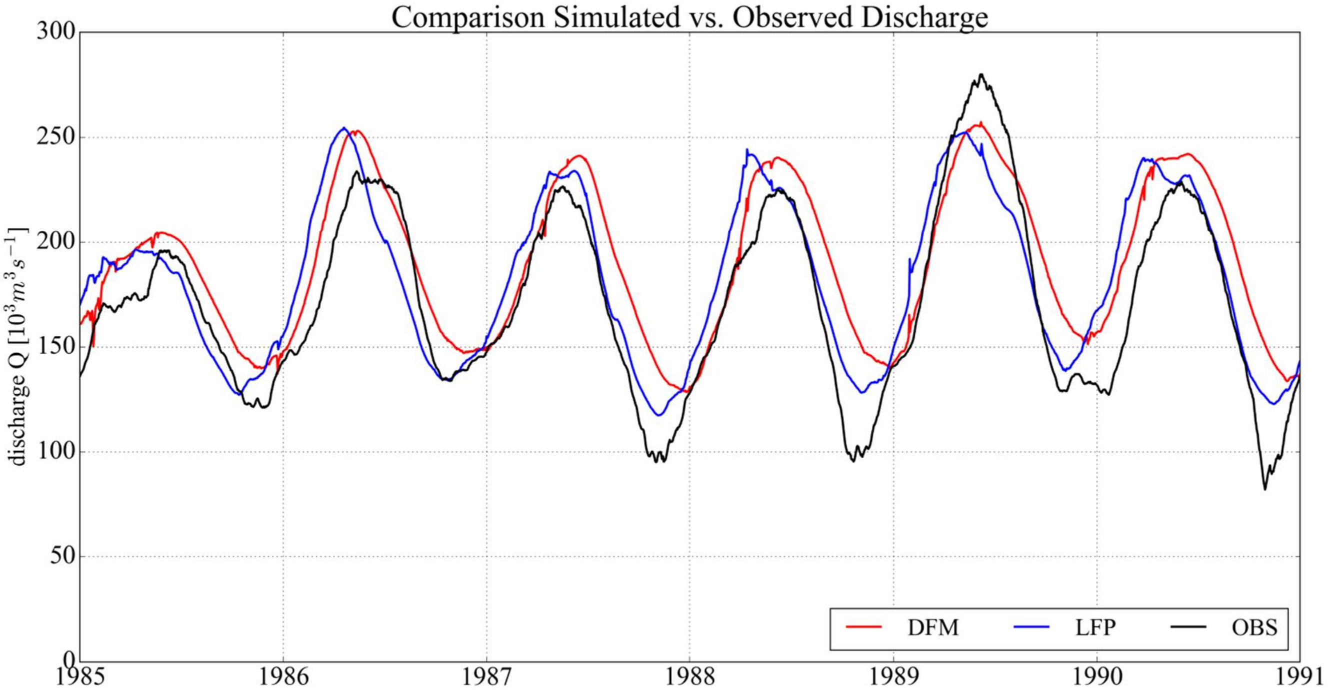

Figure 5: Observed discharge from the Global Discharge Data Centre (GRDC) as well as simulated discharge from both DFM and LFP at Óbidos

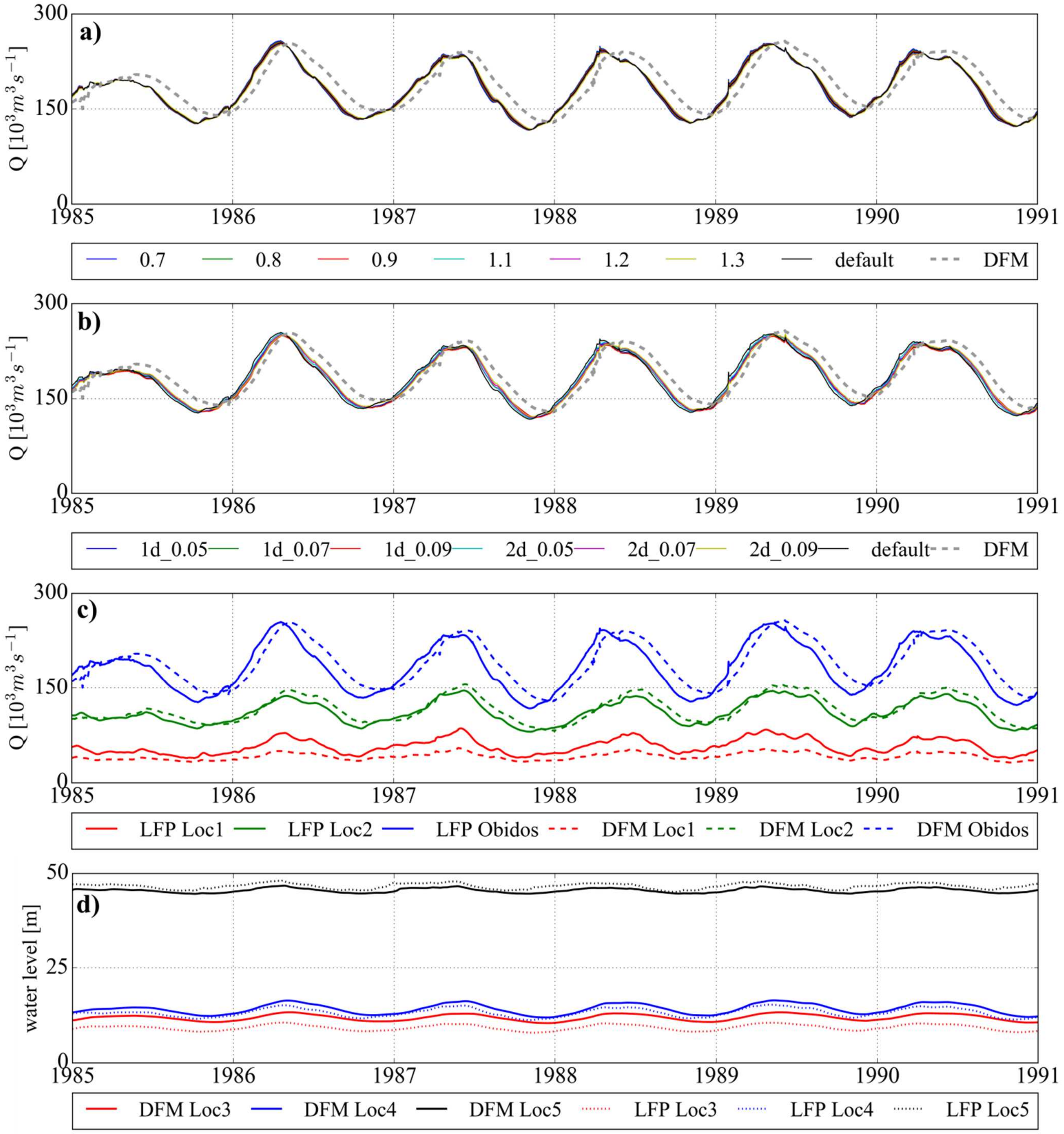

Figure 6: Results of sensitivity analysis of (a) the meandering coefficient and (b) both 1-D and 2-D surface roughness coefficients in LFP. Since the D4 system in LFP can both decrease and increase effective river dimension, the dimensionless meandering coefficient was not only reduced from default (1.0) to 0.09, 0.08, and 0.07, but also increased to 1.1, 1.2, and 1.3. As default Manning's surface roughness is already low (0.03 s $m^{-1/3}$), coefficients were increased to 0.05, 0.07, and 0.09;(c) Comparison of simulated discharge across basin to assess impact of spatial resolution on simulated discharge; to that end two additional observations upstream of Obidos were introduced, Loc1 (most upstream) and Loc2 (intermediate upstream); (d) Comparison of simulated water depth at three different locations (Loc3, Loc4, and Loc5) randomly picked within the domain

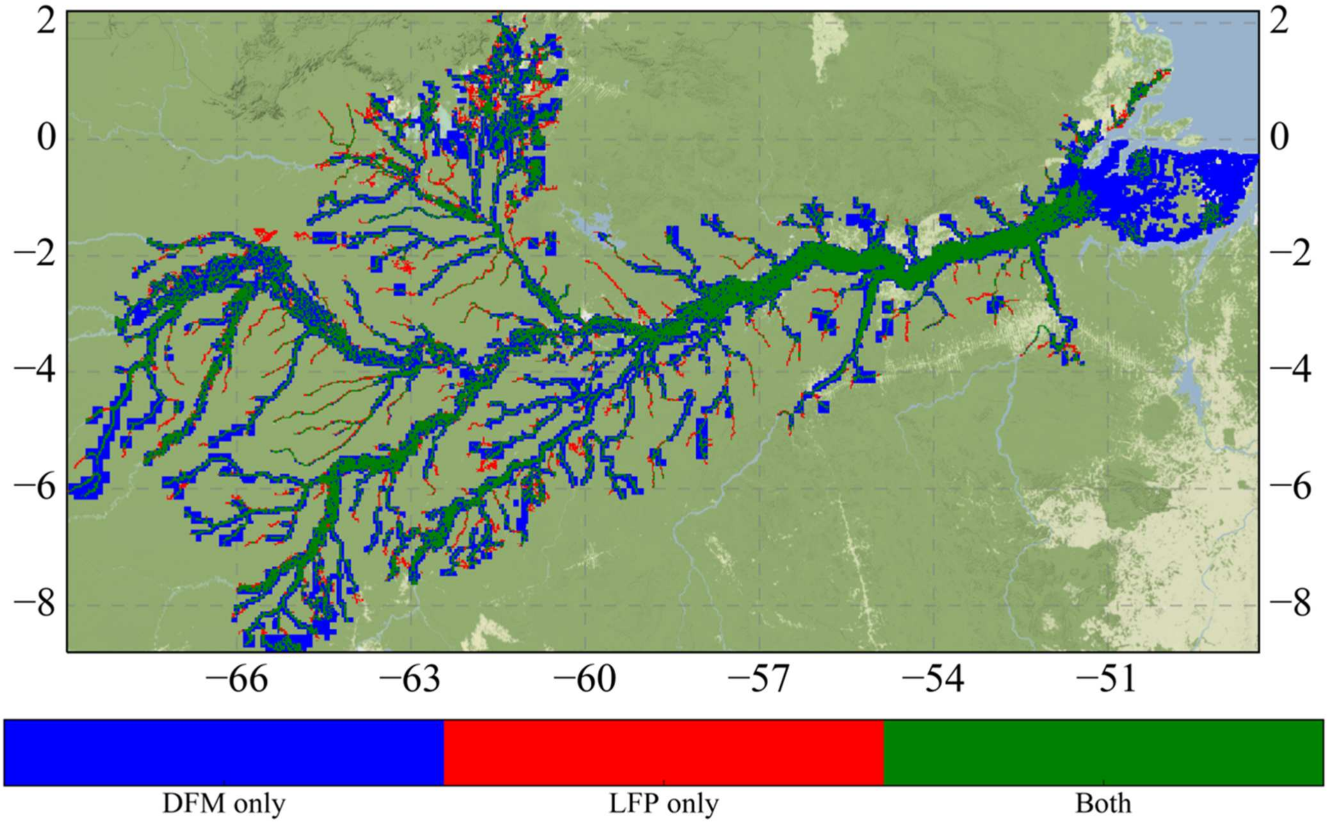

Figure 7: Benchmarking simulated inundation extent by DFM and LFP.

Table 1: Run times of different set-ups in synthetic test case

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

Table 2: Overview of key properties of hydrodynamic schematizations coupled to PCR-GLOBWB in this study

	2-D cells	**1-D cells**	**Smallest cell size**	**Largest cell size**
DFM	41,207	12,185	2 x 2 km	10 x 10 km
LFP	174,982	17,119	2 x 2 km	2 x 2 km

Table 3: Results of Pearson's coefficient r, root mean square error RMSE, and Kling-Gupta-Efficiency KGE obtained to benchmark discharge as well as run times of coupled runs

	r	**RMSE**	**KGE**	**Run time**

DFM	0.92	25,289 m^3	0.76	7 h
LFP	0.89	22,291 m^3	0.82	6 h

Table 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 arcsec spatial resolution

	Loc3	**Loc4**	**Loc5**
Input elevation	4.0	7.0	44.5
Model elevation LFP	-0.2	2.4	37.4
Model elevation DFM	0.5	4.9	42.5
Cell area LFP	~4 x 10^6		
Cell area DFM	7,7 x 10^6	7,7 x 10^6	30,9 x 10^6

Table 5: Resulting benchmarking indicators for inundation extent

	H	**F**	**C**
LFP / DFM	0.85	0.50	0.46