Sosa, J., Sampson, C., Smith, A., Neal, J., & Bates, P. (2020). A toolbox to quickly prepare flood inundation models for LISFLOOD-FP simulations. *Environmental Modelling and Software*, *123*, [104561]. https://doi.org/10.1016/j.envsoft.2019.104561

A toolbox to quickly prepare flood inundation models for LISFLOOD-FP simulations

Jeison Sosa (1), Chris Sampson (1,2), Andy Smith (1,2), Jeff Neal (1,2), Paul Bates (1,2)

(1) School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, UK
(2) Fathom, Engine Shed, Temple Meads, Bristol, BS1 6QH, UK

Corresponding author
Jeison Sosa
School of Geographical Sciences
University of Bristol
Bristol, UK
Tel: +44 7460 807038
E-mail: j.sosa@bristol.ac.uk

Abstract

Hydrodynamic floodplain inundation models have been popular for many years and used extensively in engineering applications. Continental scale flood studies are now achievable using such models due to the development of terrain elevation, hydrography and river width datasets with global coverage. However, deploying flood models at any scale is time-consuming since input data needs to be processed from different sources. Here we present LFPtools, which is an open-source Python package which encompasses most commonly used methods to prepare input data for large scale flood inundation studies using the LISFLOOD-FP hydrodynamic model. LFPtools performance was verified over the Severn basin in the UK where a 1 km flood inundation model was built within 1.45 mins. Outputs of the test case were compared with the official flood extent footprint of a real event and satisfactory model performance was obtained: Hit rate=0.79, False alarm ratio=0.24 and Critical success index=0.63.

Keywords

Large-scale, continental-scale, modelling, toolbox, hydraulics, flood, LISFLOOD-FP, Python

Highlights

- LFPtools provides data processing methods to deploy LISFLOOD-FP models.
- LFPtools is written in way that more complex methods can be easily added.
- LFPtools can be used within a sensitivity analysis framework.
- LFPtools is intended for both non-specialist and experienced flood modellers.

Software availability

The toolbox developed in this research is written in Python and built on top of GDAL (https://www.gdal.org), Cython (http://cython.org/), Pandas (https://pandas.pydata.org/), Numpy (http://www.numpy.org/), xarray (http://xarray.pydata.org) and TauDEM (http://hydrology.usu.edu/taudem/). Code and installation instruction are available at https://github.com/jsosa/LFPtools. The toolbox is distributed under the 3-Clause BSD license.

1 Introduction

Hydrodynamic models designed to simulate floodplain inundation have been popular for many years and are widely used in engineering applications. These models, such as TUFLOW (Syme, 1991), JFLOW (Bradbrook et al., 2004), TRENT (Villanueva and Wright, 2006) and LISFLOOD-FP (Bates et al., 2010), route water through channels and floodplains following shallow water flow theory.

Global to continental scale flood studies are being used for insurers, multi-national corporations, NGOs and national governments. They have been made possible as a result of the appearance of global coverage datasets of terrain elevation (Farr et al., 2007; Tadono et al., 2015; Yamazaki et al., 2017; Rizzoli et al., 2017, Wessel et al., 2018), hydrography (Lehner et al., 2008; Yamazaki et al., 2019) and river width (Andreadis et al., 2013; Yamazaki et al., 2014; Allen and Pavelsky, 2018). These data sets, coupled with the parallel development of efficient two-dimensional flood models (Bates et al., 2010, Neal et al., 2012; Sanders et al., 2010) and advances in computational power (Neal et al., 2018; Lamb et al., 2009), have led to the implementation of flood inundation studies in data-sparse areas around the world at very high resolutions (10^2-10^3 m). As consequence, a variety of applications involving flood hydrodynamic variables —flood extent, water depth, flow velocity, flow discharge— have been explored (Winsemius et al., 2013; Sampson et al., 2015; Wing et al., 2018; Dottori et al., 2017; Alfieri et al., 2018; Schumann et al., 2016; Lu et al., 2016)

Building a flood model can be time-consuming since input data need to be processed from a variety different sources and adapted to a particular user's problem. The increasing quantity, complexity and resolution of useful datasets imparts an ever-growing burden of knowledge on model developers. Furthermore, the frequent update cycles of some datasets can cause module builds to go out of date quickly. Therefore, developing a flood inundation model requires a high level of skill in handling geographical information using Graphical User Interface (GUI) driven software packages such as ArcGIS and QGIS. These present a workable solution for the treatment of data, but typically only at small-scales due to their high demands for computing resource and user intervention. Instead, at continental-scale command line interface (CLI) software packages are the best candidates for the preparation of flood inundation models since they provide robustness and computational efficiency. CLI packages can also be simpler and more streamlined than general GIS software, providing only the functionality that users need and thus making sophisticated flood inundation modelling more accessible to specialist users.

In this paper we present LFPtools, a Python CLI package which attempts to encompass the most commonly used methods to prepare input data for flood inundation studies using LISFLOOD-FP (Sampson et al., 2015; Schumann et al., 2013; Hawker et al., 2018) a widely used flood inundation model. Among the capabilities LFPtools can provide are: DEM upscaling, bank elevation estimation, bed elevation estimation, river width subtraction and interpolation, elevation smoothing algorithms, continent basin splitting, and more. Whilst the software has been built specifically for the LISFLOOD-

FP model, many of the operations it encodes are useful for a wide range of other flood inundation models, especially those operating on regular grids. LFPtools can act as an intermediate platform to streamline the preparation of local, continental or global flood inundation studies in different fields by bringing ease of use to non-expert users and efficiency to expert ones. For example, new experimental studies on hydrological-hydrodynamic modelling, sensitivity analysis (SAFE Toolbox Pianosi et al., 2015; SALib Herman et al, 2017) will be achievable more straightforwardly. LFPtools is open-source and presents a series of tools to estimate the variables required for flood inundation modelling in rapid and automated manner. As open-source, users can revise the code, modify or add new methods easily and transparently. The tools were verified over the Severn basin where a 1 km flood inundation model was built in under 2 minutes on a standard laptop (1.6 GHz Intel Core i5; 8 GB 1600 MHz DDR3).

2 The flood model LISFLOOD-FP

LISFLOOD-FP (Bates et al., 2010) is a floodplain inundation model which solves the Saint-Venant equations at very low computational cost by neglecting the flow advection term, as this is unimportant for typical gradually varying and subcritical floodplain flows. The implementation of LISFLOOD-FP Sub-Grid (Neal et al., 2012) extends the two-dimensional model for application to large domain areas where channels may be smaller than typical grid resolutions by treating river and floodplain channel networks as sub-grid scale features. Sub-grid topographic information such as realistic river width estimates is important since it increases model accuracy in terms of water level simulation, wave propagation speed, and inundation extent (Yamazaki et al., 2011; Neal et al., 2012).

Hydrodynamics in LISFLOOD-FP are solved using a momentum equation derived from the quasi-linearized one-dimensional form of the Saint-Venant equation described in Eq. (1) where q is the flow per unit width, h is the flow depth, z is the bed elevation, g is the acceleration due to gravity, n is the Manning's friction coefficient and R is the hydraulic radius which for wide shallow flows can be approximated with the flow depth h.

$$\frac{\delta q}{\delta t} + \frac{gh\delta(h+z)}{\delta x} + \frac{gn^2q^2}{R^{4/3}h} = 0 \qquad (1)$$

The final form of the unit flow at the next time step is obtained by discretising Eq. (1) with respect to the time step Δt as described in Eq. (2):

$$q_{t+\Delta t} = \frac{q_t - gh_t\Delta t\frac{\partial(h_t+z)}{\partial x}}{(1+gh_t\Delta t n^2 q_t/{h_t}^{10/3})} \qquad (2)$$

The model has been widely used for different applications at small and large scales (Wilson et al., 2007; Biancamaria et al., 2009; Neal et al., 2012; Schumann et al., 2013; Schumann et al., 2016; Alfieri et al., 2014; Sampson et al., 2015; Wing et al., 2018) due its computational speed which is mainly given by neglecting the flow advection in the shallow water equation but also by employing a highly efficient finite difference numerical solution scheme (de Almeida et al., 2012; de Almeida and Bates, 2013).

The reader is advised to consult the user manual (Bates et al., 2013) for more information on technical aspects.

3 Capabilities and features of LFPtools

LFPtools is written in Python and built on top of well-known open-source libraries: GDAL, Cython, Pandas, Numpy and xarray. The TauDEM toolbox (Tarboton, 2005) is also required for some functionalities. The library handles I/O operations via well-known file formats such as ESRI Shapefiles and GeoTIFF.

3.1 Floodplain elevations

Floodplain elevations define the grid output resolution. Those elevations can be obtained directly using a Digital Elevation Model (DEM) as-is (i.e. at native resolution). Alternatively, if the native DEM contains noise, usually derived from instrument error, upscaling the native data will reduce that noise in a coarser floodplain elevation grid, but may also smooth or loose important small scale elevation features (Neal et al., 2012; Hawker et al., 2018).

lfp-rasterresample is the program included in the library to upscale DEMs. The program can handle arrays of any size since it never loads entire arrays on memory but instead it loads a small portion of the array corresponding to the aggregation kernel to be upscaled. The program receives three inputs: a high-resolution DEM, a target resolution mask and a searching window threshold. Only cells with mask=1 will be considered for calculation. The upscaling method is described as follows:

1. A user-defined threshold is applied to a centre cell of the target mask to lump together high-resolution values.
2. A modified z-score (Iglewicz and Hoaglin, 1993; based on the median absolute deviation) is calculated for every DEM cell in the kernel. z-score values larger than 3.5 are identified as outliers and subsequently removed from the aggregation kernel.
3. In the aggregation kernel, different reduction algorithms can be applied (e.g., mean, min, meanmin). 'meanmin' is an interesting reduction method which averages the minimum and mean values from the kernel and emphasises topographic valleys in the calculation. Important to mention that more reduction algorithms can be easily added in the source code by users should they be required.

Step 2 is important to consider since native DEMs might present irregularities in some places. For example, in development testing a disagreement was found in the aggregation kernel for a target cell in the Seine River using the native ~90 m resolution MERIT DEM. In particular, some strong negative values (~-10 m) were found in an area where the typical topographic elevation was ~30 m (See Fig. 1). The automatic detection algorithm in step 2 prevents inclusion of these values before step 3.

Different aggregation methods from Step 3 are compared for a small part of the River Thames using the toolbox in Fig. 2.

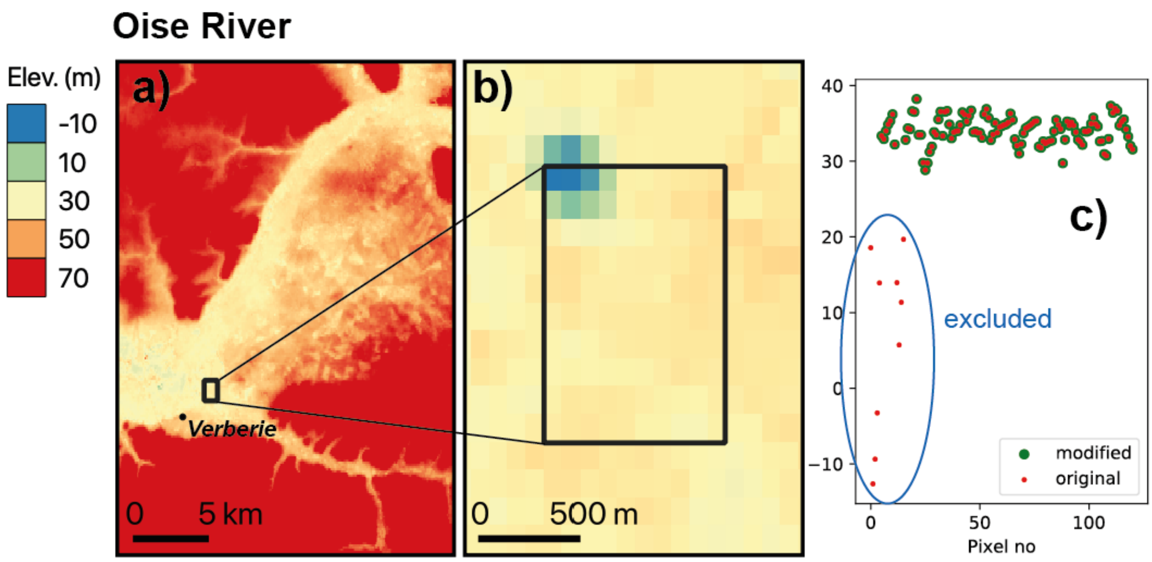

Figure 1: Outlier detection procedure: a) original 90 m resolution DEM and aggregation kernel (in black), b) zoom-in at aggregation kernel (area ~1 km$_2$) and c) automatic detection of outliers in kernel (in green) points retained for upscaling and (in red) all points.

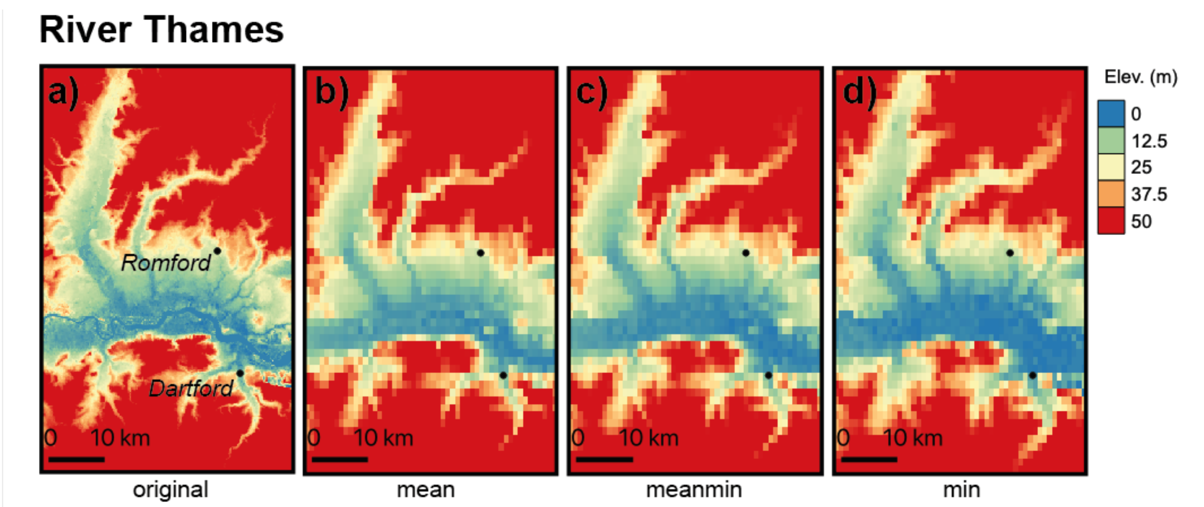

Figure 2: Upscaling methods comparison at 1 km resolution: a) original 90 m resolution DEM, b) 'mean' aggregation, c) 'meanmin' aggregation and d) 'min' aggregation

3.2 Channel widths

LISFLOOD-FP Sub-Grid needs several input variables to run a flood simulation, one of which is river width estimates at every cell in the river network. With the appearance of global river width data sets based on remote sensing techniques (GWD-LR Yamazaki et al., 2014; GRWL Allen and Pavelsky 2018) and empirical formulations (Andreadis et al., 2013) it is now feasible to use these data sets as width sources in flood studies for data-sparse regions.

Global river width databases may have some degree of geolocation shift in relation to the corresponding rivers extracted from hydrography databases making them difficult to use in their native format. This problem may appear if these databases are derived from different sources or due to resolution dissimilarity; for example, DEM derived river networks and remotely sensed open water locations. Commonly, a nearest neighbour function in a searching window is used to assign the nearest value from a river width database to a river cell in a flood study. However, there might be cases where the searching window is too small and no width values are found, in this case increasing the window size is not an appealing option since it might result in an incorrect river width assignment from a tributary. Instead, it is advisable to use an interpolation with values already assigned. It is important to note that leaving a river cell with no width assigned is a critical issue since LISFLOOD-FP Sub-Grid cannot perform calculations on river cells with zero width.

LFPtools includes a routine (*lfp-getwidths*) to automatically assign width values to river cells, it works in the following way:

1. River cell widths are assigned based on the nearest-neighbour within a searching window.
2. If no width value is assigned from the source database, the missing value is automatically interpolated with values already assigned.

Fig. 3 shows an example of three river cells with widths unassigned due to the searching window size problem. Fig. 3a shows a river reach (blue) at ~1 km, red dots are centroids of river cells and the black solid line is river vector from the GRWL database (~30 m). From the figure only three points (A, B, C) were not able to find an appropriate width value in their neighbourhood (red dash line), those values were automatically calculated by interpolation in *lfp-getwidths* see Fig. 3b

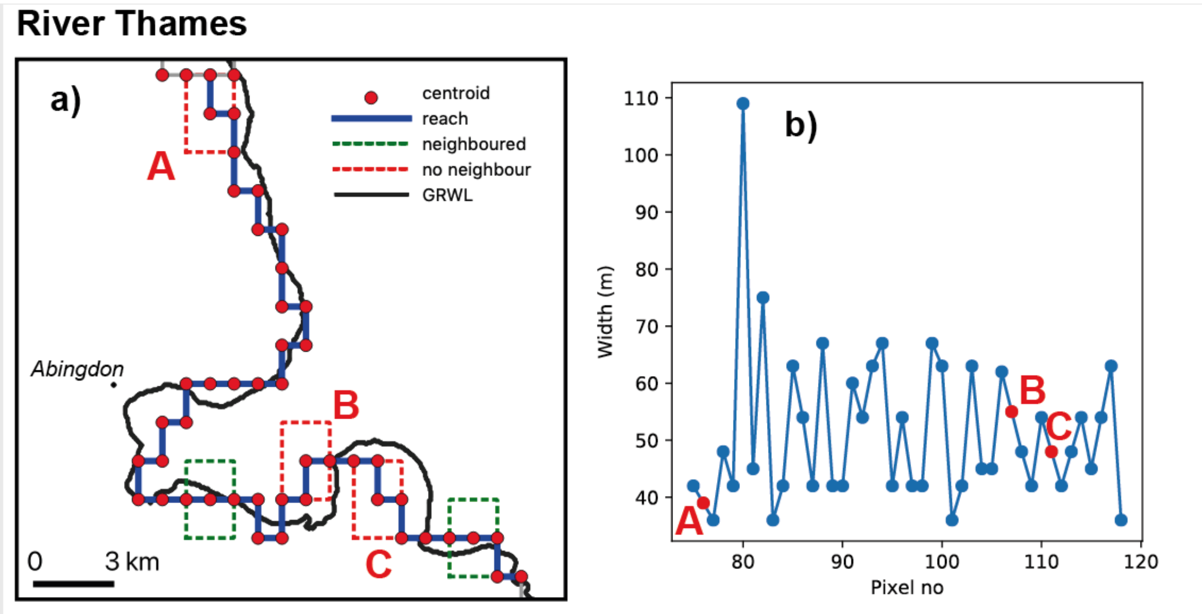

Figure 3: River widths assignment: **a)** Example showing three river cells unassigned due to small size in searching window at locations A, B and C and **b)** (in blue) width values that yield in the searching window (in red) width values interpolated.

3.3 Bank elevations

The LISFLOOD-FP Sub-Grid uses the DEM elevation as the bank height elevations, which when combined with the channel bed elevation defines the channel bankfull depth. It is therefore recommended to recalculate the bank height elevations to get better estimates because of the critical role this value plays in flooding simulations.

If a native resolution DEM is used, bank height elevations are self-defined. However, if a coarser resolution model is created, high-resolution cell aggregation is required. *lfp-getbankelevs* reads a target river network mask (mask=1 will be considered for calculation), a high-resolution DEM, and a searching window threshold to aggregate cells and apply a reduction algorithm (nearest, mean, min, meanmin). Resulting elevations might contain irregularities that may result in model instabilities caused by local supercritical flows and flow blocking effects if the channel bed follows the banks. Those irregularities can be solved by applying a smoothing algorithm along the river.

LFPtools includes a routine (*lfp-fixelevs*) which includes two approaches to deal with this problem:

1. Adjust bank heights by minimising the amount of modifications following the method developed by Yamazaki et al., (2012). This algorithm removes all the pits in the spaceborne DEM caused by

vegetation canopies, sub-pixel sized structures, and random radar speckles while minimizing the amount of modification required for removing the pits.

2. Apply a weighted local regression (LOWLESS) (Cleveland, 1979) in the downstream direction as in Schumann et al., (2013).

Both methods are compared for the main channel of the River Thames, UK in Fig. 4b

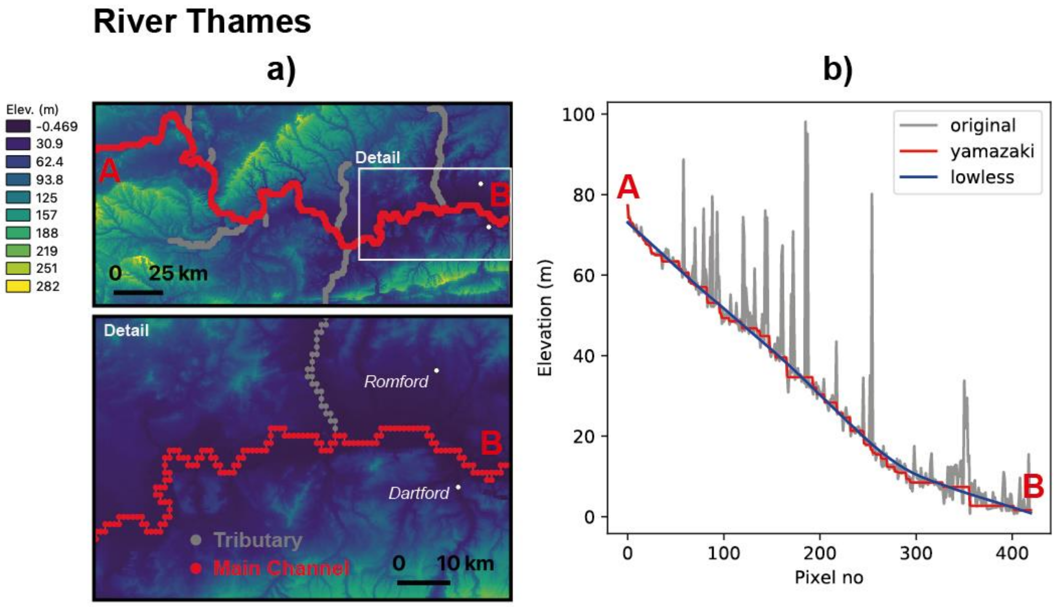

Figure 4: Smoothing method available in LFPtools. These methods were applied to the main channel of the River Thames: **a)** (in red) main channel of the River Thames and (in grey) tributaries, **b)** (in grey) original elevation extracted by the nearest-neighbour (in red) Yamazaki's method (in blue) Locally weighted smoothing

3.4 River depths

Standard LISFLOOD-FP Sub-Grid treats river cross-sections as rectangular. Due to this fact channel depths may differ from in-situ river depth surveys. With some calibration this approximation works very well at large scales producing reasonable results in most places as long as accurate estimations of bank heights and widths are used. Unlike bank heights and river widths that can be determined from satellite data, river depths need to be approximated. Two approaches have been proposed to achieve this goal and are included in the *lfp-getdepths* tool — a simple empirical power law formulation (Neal et al., 2012) and the Manning's equation (Sampson et al., 2015). A user-defined raster (e.g., survey data on river bathymetry) can also be used to assign depths to cells if none of the previous methods are used.

Power law relationship

Leopold and Maddock (1953) derived a series of power law relationships given by Eq. (5), (6) and (7) where W is water-surface width, Q is discharge, D is mean depth and V is mean velocity

$$W = aQ^b \quad (3)$$

$$D = cQ^f \quad (4)$$

$$V = kQ^m \quad (5)$$

It is straightforward to equate Eq. (3) and (4) to obtain Eq. (6)

$$D = \left(\frac{c}{a^{f/b}}\right) W^{f/b} \quad (6)$$

where (a, b, c, f) are empirical values depending on the geomorphology of the bed. Sometimes it is preferred to use only one pair of constants (r,p) as in Eq. (7). See Hey and Thorne (1986) for empirical values for gravel-bed rivers in the UK.

$$D = rW^p \quad (7)$$

Manning's equation

The Manning's equation for a rectangular channel is described by Eq. (8) where A is the cross-section area expressed as $A = WD$ with W width and D depth, R is the hydraulic radius $R = A/(W + 2D)$, S is the channel cell slope —it can be calculated via *lfp-slopes* or directly extracted from an external data set (Cohen et al., 2018)— n is the Manning's coefficient and Q_{bf} is the bankfull flow.

$$Q_{bf} = \frac{AR^{2/3}S^{1/2}}{n} \quad (8)$$

The Manning's equation considers bankfull flow Q_{bf} as a known variable, however it is not always the case. If not measured in the field, bankfull flow is usually estimated by fitting a statistical distribution on the annual flow peaks of a streamflow time series where bankfull conditions occur at return periods of 1.5-2 years (Scheneider et al., 2011). Fig. 5 shows the aforementioned procedure for the Kingston gauging station from the National River Flow Archive (NRFA) on the River Thames, UK.

A comparison between the Power law relationship and Manning's equation is presented for the River Thames in Fig. 6. Bankfull flow (yellow dots) was obtained by subtracting the 2-year return period in a Pearson Type III distribution fitted on the annual maxima time series derived by means of a 24-year streamflow reanalysis from the European Forecasting Awareness System (EFAS) (Thielen et al., 2009). River width estimates used in Eq. (7) were obtained from the GRWL database using *lfp-getwidths*. At locations where no-bankfull width is available, the nearest bankfull value was assigned. Fig 6c shows (in grey) bank elevations after smoothing in the main channel, (in blue) bed elevations (i.e., bank elevation - depth) using the Manning's Eq. (8) and (in red) using the power law relationship Eq. (7). A zoom for the downstream section is shown in Fig 6d and reveals considerable differences in the delta area.

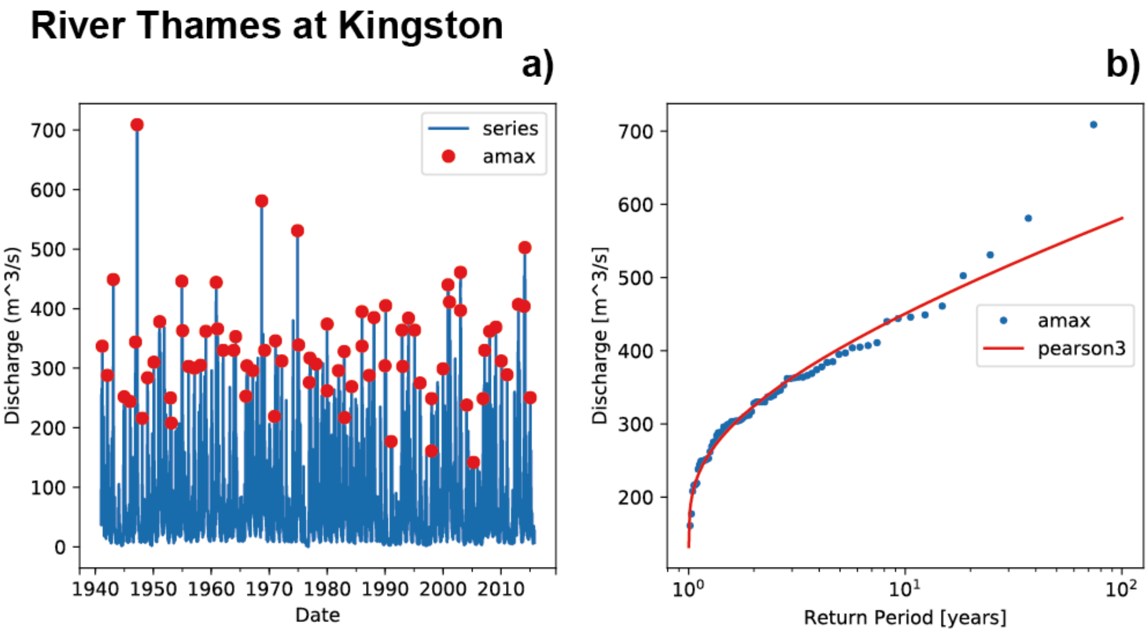

Figure 5: Observed river discharge in the River Thames at Kingston Station. Bankfull was estimated by fitting a statistical distribution on the annual maxima and retrieving the discharge value for the 2-yr return period: **a)** annual maxima between 1940-2015 (red dots). **b)** Pearson Type III distribution fitted on the annual maxima (red line), here the distribution parameters were estimated via L-moments. This figure was generated by using the *hydroutils* library (Sosa, 2018).

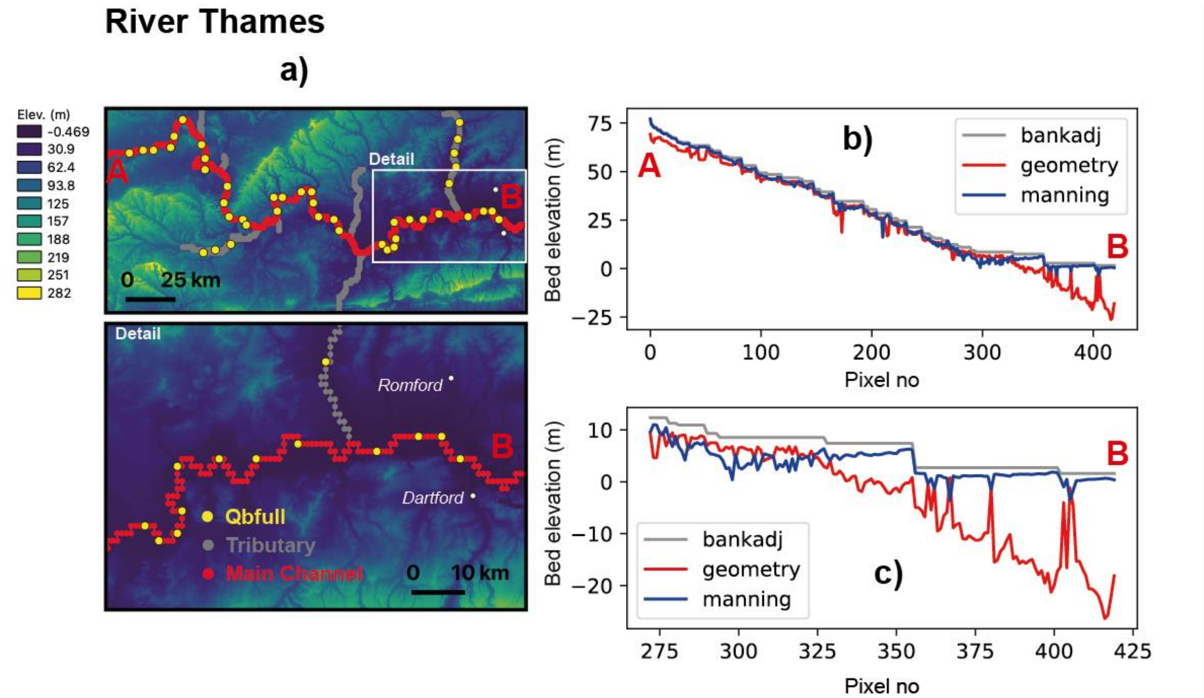

Figure 6: River depth estimation using hydraulic geometry equations and Manning's equation: **a)** River Thames (in red) tributaries (in grey), **b)** depth estimation via hydraulic geometry (in red) and Manning's equation (in blue) for the lower part of the River Thames and **c)** zoom-in delta area of the River Thames

3.5 Continental tools

The library includes two programs designed to automate delineation of basins within large regions *lfp-prepdata* and *lfp-split*.

lfp-prepdata incorporates a subroutine to clip global data sets of DEM, hydrography and river width based on a user-defined extent. Thereafter, a user-defined threshold is applied to the flow accumulation area (or upslope drainage area) to define a river network. The TauDEM toolbox (Tarboton, 2005) is used to generate a network topological connectivity for the whole area and to delineate basins within the region (NNN_Tree.csv, NNN_Coord.csv and NNN_Rec,csv in Fig. 7). The routine also includes a function to convert D8 connected river networks to D4 connectivity based on the flow directions map given by the hydrography. *lfp-split* breaks up the region into individual basins with a basin-number associated. Folders are created with a basin-number and each of them contains clipped data associated with that basin. After basin required data is split in this way the tools described in Sections 3.1-3.4 can be applied. Fig. 7 shows a flowchart describing how the tools can connect to each other to automatically build models at continental-scale.

LFPtools flowchart

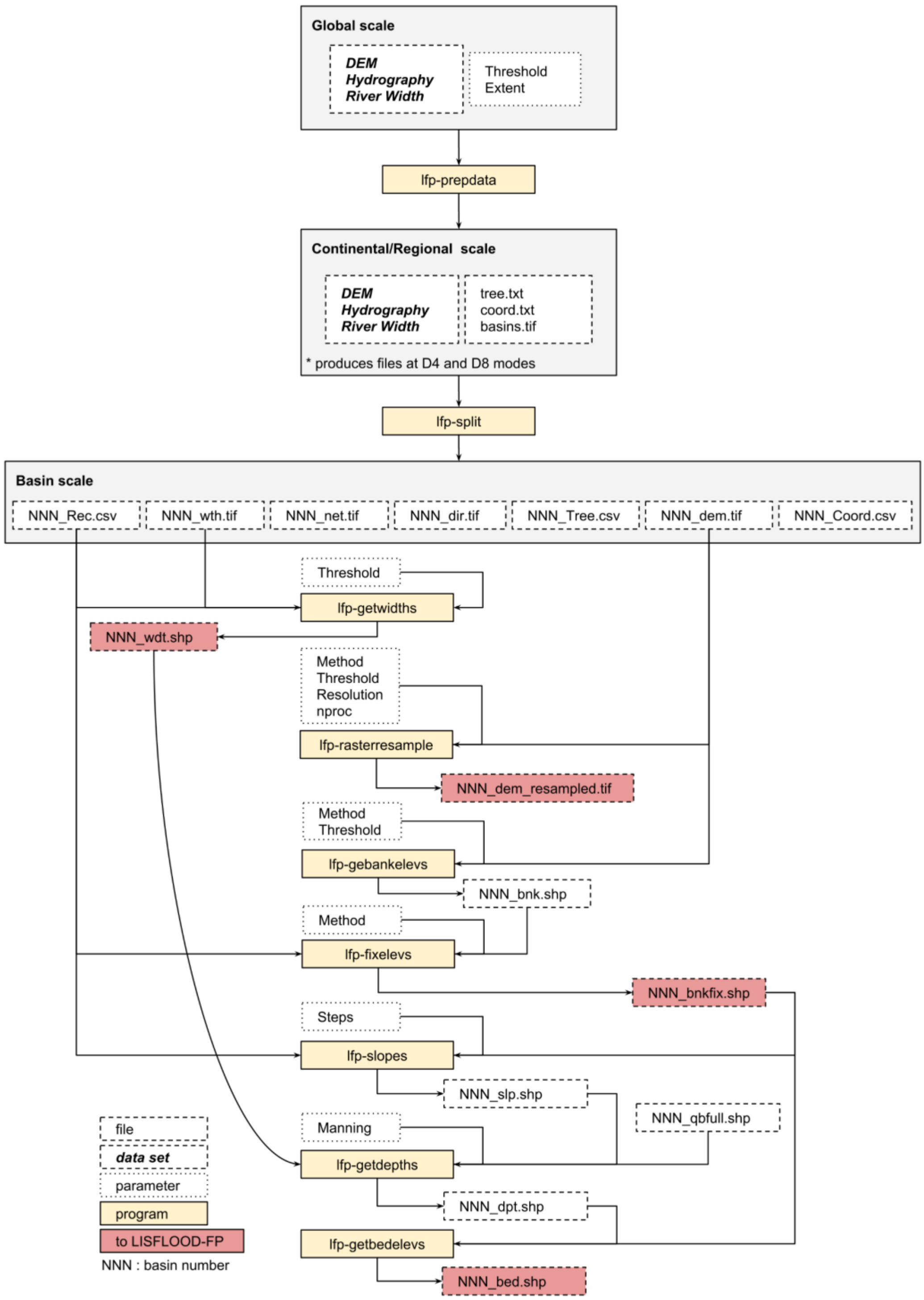

Figure 7: Flowchart using LFPtools for continental-scale studies. Command-line tools are presented in yellow boxes, white dashed boxes represent input data sets and white dotted boxes free parameters. Outputs to LISFLOOD-FP are coloured in red.

3.6 Usage

In order to facilitate the use of the tools LFPtools can be called via command-line, however if preferred it can also be imported as a Python module. All tools can be invoked via the command line by typing the name of the tool followed by the -i keyword and the name of the configuration file:

```
$ lfp-getwidths -i config.txt
```

where the configuration file 'config.txt' is a text file containing a [tool-name] header followed by variable=argument entries. Input variable descriptions are specified when typing the name of the tool in the command-line followed by the –h keyword: $ lfp-getwidths -h

LFPtools can be imported as a Python module as follows:

```
import lfptools as lfp
```

An overview of tools with a brief description is given in Table 1.

Program	**Description**
lfp-depths	Get estimates of depth
lfp-fixelevs	Smooth elevations
lfp-getbankelevs	Retrieve bank elevations
lfp-slopes	Estimate slopes in a river network
lfp-getwidths	Retrieve river widths
lfp-rasterresample	Upscale a high-resolution DEM into a user-defined resolution
lfp-split	Breaks up a study area in individual basins with a basin number associated
lfp-prepdata	Clip global data sets given a user-defined extent and threshold. The threshold is used to define a river network based on the upslope area

Table 1: Summary of programs in LFPtools

4 A flood inundation model for the Severn River in England, UK

LFPTools was used to build a flood inundation model for the Severn river basin in the UK. A one-month simulation (April 1998) was undertaken in order to capture an observed flood event that happened during this period. An additional one month 'warm-up' period was included to bring the model into a hydraulic steady state condition prior to the commencement of the April 1998 period. The model was built from LIDAR-based terrain data (at 90 m resolution) where the floodplain terrain was upscaled to 1 km resolution using the 'mean' aggregation method and removing outliers. Bank heights were defined using the 'nearest neighbour' method. River channels were explicitly represented using HydroSHEDS (Lehner et al., 2008) as input hydrography at 1 km resolution. Channel widths were retrieved from the GRWL database while river depths were estimated through the hydraulic geometry method (Eq. 5) with $r = 0.12$ and $p = 0.78$. The model was forced using daily gauged flows from the UK National River Flow Archive (NRFA) for the simulation period mentioned before. Data sources used in this study are briefly described in Table 2.

Data set	**Description**	**Source**
LIDAR DTM	Composite at 1 m resolution	Data available at data.gov.uk
HydroSHEDS	Hydrography at 1 km resolution	Lehner et al., 2018. Data available at hydrosheds.org
GRWL	Landsat-based global river width database at 30 m resolution	Allen and Pavelsky, 2018. Data available at https://zenodo.org/record/1297434
NRFA	Streamflow data from gauge stations	Data available at nrfa.ceh.ac.uk
Recorded Flood Outlines for UK	Records of historic flooding from rivers, the sea, groundwater and surface water	Data available at data.gov.uk

Table 2: Data sets used to build the flood inundation model in the Severn river basin

Resulting water depths from LISFLOOD-FP at 1 km resolution were subsequently downscaled onto 90 m resolution using an algorithm similar to Schumann et al., 2014**.** In particular, the algorithm takes water surface elevation (WSE) at 1 km resolution and subtracts its corresponding 90 m DEM values. From this arithmetic operation, a grid at 90 m resolution is created with positive values representing the water depth (wet cells) whilst negative values (dry cells) are replaced with nodata values.

The performance of flood model in the Severn river basin in terms of flood extent was quantified using three scores: Hit rate (H), Falsa alarm ratio (F) and Critical success index (C). *H* tests the tendency of the model towards underprediction and can range from 0 (none of the wet benchmark data is wet model data) to 1 (all of the wet benchmark data are wet model data). *F* examines the tendency of the model towards overprediction and can range from 0 (no false alarms) to 1 (all false alarms). *C* accounts for both overprediction and underprediction and can range from 0 (no match between modelled and benchmark data) to 1 (perfect match between modelled and benchmark data). A detailed explanation of these scores is available in Wing et al., 2017.

Simulated water depth results for the 15_{th} April 1998 are shown in Fig. 8. From the figure is clear that in most places water remains in the channel and where water elevations exceed bankfull heights water spreads onto the floodplains. Simulated water depth on the 15_{th} April 1998 were compared with the official event footprint from the English Environment Agency (EA) and the ‘Agreement’ between both flood extents are presented in the Fig. 8 right-hand panel. The ‘Agreement’ in Fig. 8 refers to areas in the map where the EA flood extent and the simulated flood extent overlap each other. In terms of flood extent, the model obtained satisfactory comparison scores against observations: H=0.79, F=0.24 and C=0.63. Example files are available at the LFPtools web repository.

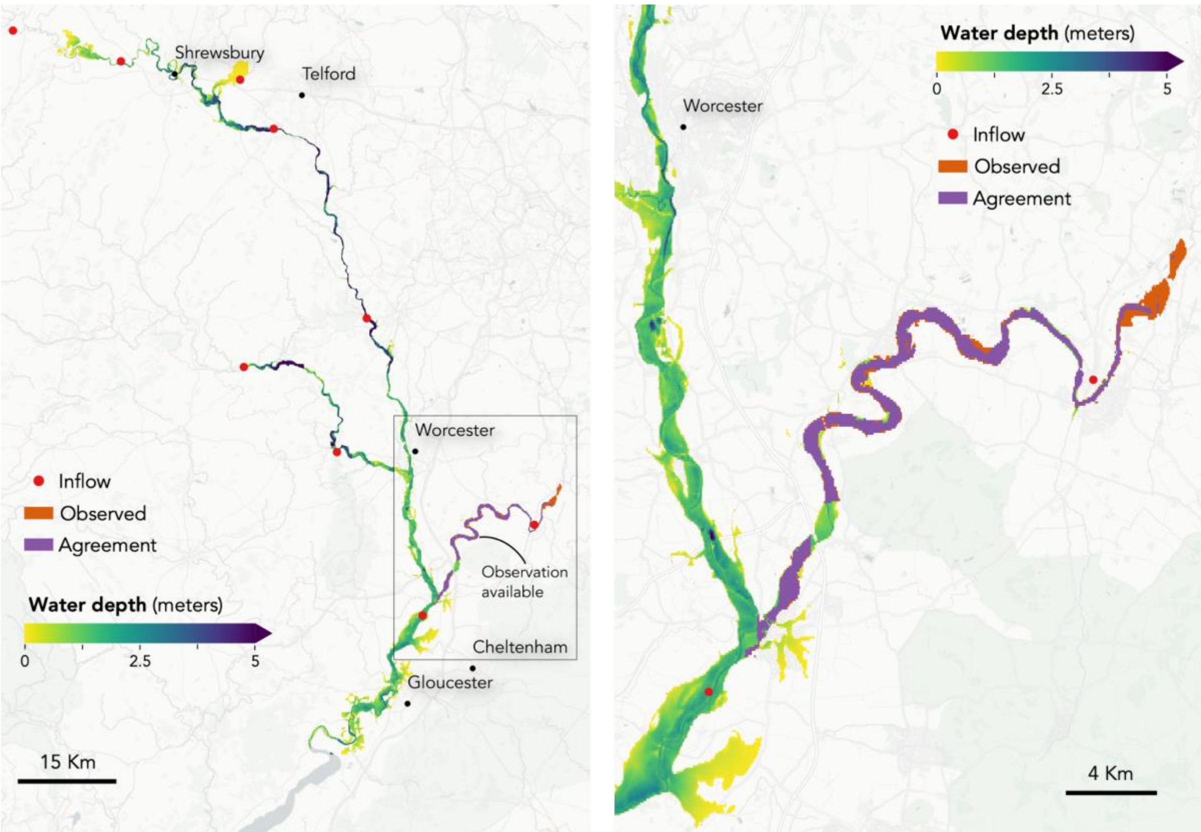

Figure 8: Flood inundation model prepared for the Severn basin in England, UK during the flood event of April-1998. The event was compared with official footprint of the event (orange). The agreement between the model and the output is also shown (purple). Note that the observed data only cover limited portions of the model domain which are not contiguous. In areas with no observed

data we simply plot the modelled water depth. Also, the moderately low *Hit Rates* occur since the observed flood extent area is upstream of the inflow point (East of the domain in the right-hand panel), hence, no forcing data is available to predict water depths in that area.

5 Conclusions

A Python CLI package has been developed to help prepare input data for flood studies carried out using LISFLOOD-FP. The package encompasses the most frequently used methods for flood inundation modelling data preparation, and also facilitates the addition of new ones if desired. LFPtools can be thought of as a platform to streamline the preparation of flood inundation studies in different fields by bringing ease of use to non-expert users and efficiency to expert ones. It is built on top of the state-of-the-art Python libraries to handle large sets of data and it is in active development. It is important to mention that these tasks could be done in a GIS package, but only with quite extreme difficulty and for small data arrays. The tasks performed by LFPTools are generic for structured grids and can be used to prepare input data sets for any hydraulic model.

LFPtools programs were verified in the UK's Severn basin on a model built at 1 km resolution using publicly available data sets only. The test basin was used to simulate the event of April 1998 and results are presented in Fig. 8. From the figure it is clear that most of the water is kept in channels with some places inundated suggesting a normal hydrodynamic behaviour. After comparison, the model obtained satisfactory scores against the official event footprint: H=0.79, F=0.24 and C=0.63. It is important to mention that the Severn scenario was used only to broadly test the tools and not to simulate the real event to an engineering standard.

The Severn river basin used in this study is only a small example on how the tools can be employed and the tools have been designed so they can be integrated within a framework to build continental to global scale studies. For example, LFPtools can be used within a modelling framework to build a continental-scale flood hindcast or reanalysis, a modelling framework of continental-scale flood extent for an early warning system or even within a framework to predict flood inundation variables (flood extent, water depth, etc) in a climate change context.

Global to continental scale models are being used by insurers, multi-national corporations, NGOs and national governments to tackle problems such as rapid flood disaster response, urban planning and climate change adaptation. Thus, flood models at such scales are important decision making tools and building them demands great effort to research scientists. We envisage that this innovative set of tools will help to significantly reduce these costs.

Acknowledgments

J.S. received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 676027. P.B. was supported by a Leverhulme Research Fellowship and a Wolfson Research Merit Award from the Royal Society. The authors would like to thank Gemma Coxon for providing the NRFA data.

References

Alfieri, L., Dottori, F., Betts, R., Salamon, P., Feyen, L., 2018. Multi-Model Projections of River Flood Risk in Europe under Global Warming. Climate 6, 6. https://doi.org/10.3390/cli6010006

Alfieri, L., Salamon, P., Bianchi, A., Neal, J., Bates, P., Feyen, L., 2014. Advances in pan-European flood hazard mapping. Hydrol Process 28, 4067–4077. https://doi.org/10.1002/hyp.9947

Allen, G.H., Pavelsky, T.M., 2018. Global extent of rivers and streams. Science. https://doi.org/10.1126/science.aat0636

Andreadis, K.M., Schumann, G.J.-P., Pavelsky, T., 2013. A simple global river bankfull width and depth database: Data and Analysis Note. Water Resour Res 49, 7164–7168. https://doi.org/10.1002/wrcr.20440

Bates, P., Trigg, M., Neal, J., Dabrowa, A., 2013. LISFLOOD-FP user manual. School of Geographical Sciences, University of Bristol, UK.

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

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

Bradbrook, K.F., Lane, S.N., Waller, S.G., Bates, P.D., 2004. Two dimensional diffusion wave modelling of flood inundation using a simplified channel representation. Intl. J. River Basin Management 2, 211–223. https://doi.org/10.1080/15715124.2004.9635233

Cleveland, W.S., 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of the American Statistical Association 74, 829–836. https://doi.org/10.1080/01621459.1979.10481038

Cohen, S., Wan, T., Islam, M.T., Syvitski, J.P.M., 2018. Global river slope: A new geospatial dataset and global-scale analysis. J Hydrol 563, 1057–1067. https://doi.org/10.1016/j.jhydrol.2018.06.066

de Almeida, G.A.M., Bates, P., 2013. Applicability of the local inertial approximation of the shallow water equations to flood modeling. Water Resour Res 49, 4833–4844. https://doi.org/10.1002/wrcr.20366

de Almeida, G.A.M., Bates, P., Freer, J.E., Souvignet, M., 2012. Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. Water Resour Res 48. https://doi.org/10.1029/2011WR011570

Dottori, F., Kalas, M., Salamon, P., Bianchi, A., Alfieri, L., Feyen, L., 2017. An operational procedure for rapid flood risk assessment in Europe. Nat Hazards Earth Syst Sci 17, 1111–1126. https://doi.org/10.5194/nhess-17-1111-2017

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., Alsdorf, D., 2007. The Shuttle Radar Topography Mission. Rev Geophys 45. https://doi.org/10.1029/2005RG000183

Hawker, L., Rougier, J., Neal, J., Bates, P., Archer, L., Yamazaki, D., 2018. Implications of Simulating Global Digital Elevation Models for Flood Inundation Studies. Water Resour Res 54, 7910–7928. https://doi.org/10.1029/2018WR023279

Herman, J., Usher, W., 2017. SALib: An open-source Python library for Sensitivity Analysis. The Journal of Open Source Software 2, 97. https://doi.org/10.21105/joss.00097

Hey, R.D., Thorne, C.R., 1986. Stable Channels with Mobile Gravel Beds. J Hydraul Eng 112, 671–689. https://doi.org/10.1061/(ASCE)0733-9429(1986)112:8(671)

Iglewicz, B. and Hoaglin, D.C., 1993. How to detect and handle outliers (Vol. 16). Asq Press.

Lamb, R., Crossley, M., Waller, S., 2009. A fast two-dimensional floodplain inundation model. Proceedings of the Institution of Civil Engineers - Water Management 162, 363–370. https://doi.org/10.1680/wama.2009.162.6.363

Lehner, B., Verdin, K., Jarvis, A., 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos, Transactions American Geophysical Union 89, 93. https://doi.org/10.1029/2008EO100001

Leopold, L.B., Maddock Jr., T., 1953. The hydraulic geometry of stream channels and some physiographic implications (Report No. 252), Professional Paper. Washington, D.C.

Lu, X., Zhuang, Q., Liu, Y., Zhou, Y., Aghakouchak, A., 2016. A large-scale methane model by incorporating the surface water transport: Development of a Methane Model. J Geophys Res Biogeosci 121, 1657–1674. https://doi.org/10.1002/2016JG003321

Neal, J., Dunne, T., Sampson, C., Smith, A., Bates, P., 2018. Optimisation of the two-dimensional hydraulic model LISFOOD-FP for CPU architecture. Environ Model Softw 107, 148–157. https://doi.org/10.1016/j.envsoft.2018.05.011

Neal, J., Schumann, G., Bates, P., 2012. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas. Water Resour Res 48. https://doi.org/10.1029/2012WR012514

Pianosi, F., Sarrazin, F., Wagener, T., 2015. A Matlab toolbox for Global Sensitivity Analysis. Environ Model Softw 70, 80–85. https://doi.org/10.1016/j.envsoft.2015.04.009

Rizzoli, P., Martone, M., Gonzalez, C., Wecklich, C., Borla Tridon, D., Bräutigam, B., Bachmann, M., Schulze, D., Fritz, T., Huber, M., Wessel, B., Krieger, G., Zink, M., Moreira, A., 2017. Generation and performance assessment of the global TanDEM-X digital elevation model. Isprs J Photogramm 132, 119–139. https://doi.org/10.1016/j.isprsjprs.2017.08.008

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

Sanders, B.F., Schubert, J.E., Detwiler, R.L., 2010. ParBreZo: A parallel, unstructured grid, Godunov-type, shallow-water code for high-resolution flood inundation modeling at the regional scale. Adv Water Resour 33, 1456–1467. https://doi.org/10.1016/j.advwatres.2010.07.007

Schneider, C., Flörke, M., Eisner, S., Voss, F., 2011. Large scale modelling of bankfull flow: An example for Europe. J Hydrol 408, 235–245. https://doi.org/10.1016/j.jhydrol.2011.08.004

Schumann, G.J.-P., Andreadis, K.M., Bates, P.D., 2014. Downscaling coarse grid hydrodynamic model simulations over large domains. J Hydrol 508, 289–298. https://doi.org/10.1016/j.jhydrol.2013.08.051

Schumann, G.J.-P., Neal, J.C., Voisin, N., Andreadis, K.M., Pappenberger, F., Phanthuwongpakdee, N., Hall, A.C., Bates, P.D., 2013. A first large-scale flood inundation forecasting model: Large-Scale Flood Inundation Forecasting. Water Resour Res 49, 6248–6257. https://doi.org/10.1002/wrcr.20521

Schumann, G.J.-P., Stampoulis, D., Smith, A.M., Sampson, C.C., Andreadis, K.M., Neal, J.C., Bates, P.D., 2016. Rethinking flood hazard at the global scale. Geophys Res Lett 43, 10,249-10,256. https://doi.org/10.1002/2016GL070260

Sosa, J., 2018. Hydroutils. https://doi.org/10.5281/zenodo.1408076

Syme, W.J., 1991. Dynamically Linked Two-dimensional/One-dimensional Hydrodynamic Modelling Program for Rivers, Estuaries & Coastal Waters (MEngSc thesis). University of Queensland, Australia.

Tadono, T., Takaku, J., Tsutsui, K., Oda, F., Nagai, H., 2015. Status of “ALOS World 3D (AW3D)” global DSM generation. Proceeding 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3822–3825. https://doi.org/10.1109/IGARSS.2015.7326657

Tarboton, D.G., 2005. Terrain analysis using digital elevation models (TauDEM).

Thielen, J., Bartholmes, J., Ramos, M.-H., de Roo, A., 2009. The European Flood Alert System – Part 1: Concept and development. Hydrol Earth Syst Sci 13, 125–140. https://doi.org/10.5194/hess-13-125-2009

Villanueva, I., Wright, N.G., 2006. Linking Riemann and storage cell models for flood prediction. Proceedings of the Institution of Civil Engineers - Water Management 159, 27–33. https://doi.org/10.1680/wama.2006.159.1.27

Wessel, B., Huber, M., Wohlfart, C., Marschalk, U., Kosmann, D., Roth, A., 2018. Accuracy assessment of the global TanDEM-X Digital Elevation Model with GPS data. Isprs J Photogramm 139, 171–182. https://doi.org/10.1016/j.isprsjprs.2018.02.017

Wilson, M., Bates, P., Alsdorf, D., Forsberg, B., Horritt, M., Melack, J., Frappart, F., Famiglietti, J., 2007. Modeling large-scale inundation of Amazonian seasonally flooded wetlands. Geophys Res Lett 34. https://doi.org/10.1029/2007GL030156

Wing, O.E.J., Bates, P.D., Sampson, C.C., Smith, A.M., Johnson, K.A., Erickson, T.A., 2017. Validation of a 30 m resolution flood hazard model of the conterminous United States. Water Resour Res 53, 7968–7986. https://doi.org/10.1002/2017WR020917

Wing, O.E.J., Bates, P.D., Smith, A.M., Sampson, C.C., Johnson, K.A., Fargione, J., Morefield, P., 2018. Estimates of present and future flood risk in the conterminous United States. Environ Res Lett 13, 034023. https://doi.org/10.1088/1748-9326/aaac65

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

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

Yamazaki, D., Kanae, S., Kim, H., Oki, T., 2011. A physically based description of floodplain inundation dynamics in a global river routing model. Water Resour Res 47. https://doi.org/10.1029/2010WR009726

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

Yamazaki, D., O'Loughlin, F., Trigg, M.A., Miller, Z.F., Pavelsky, T.M., Bates, P.D., 2014. Development of the Global Width Database for Large Rivers. Water Resour Res 50, 3467–3480. https://doi.org/10.1002/2013WR014664

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P.D., Allen, G., Pavelsky, T., 2019. MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets. Water Resour. Res. 2019WR024873. https://doi.org/10.1029/2019WR024873