LISFLOOD at 30 arc sec spatial resolution

Description of the input layers at 30 arc sec spatial resolution and a first model comparison with the native 5 km spatial resolution for a case study in the Elbe basin, Germany

Internship Report

Jonathan Bänziger

Contact: jonathan_baenziger@hotmail.com

December 2018

Abstract

A new high resolution 30 arc sec (~1 km at the Equator) data set describing static maps and meteorological forcing was developed for the hydrological model LISFLOOD. LISFLOOD is the underlying rainfall-runoff-routing model of the European Flood Awareness System (EFAS) which operationally monitors and forecasts floods across Europe. There is a general expectation that moving towards spatial hyperresolution will allow for a better representation of the heterogeneity in topography, soils, and vegetation and their effects on hydrological dynamics. Higher spatial model resolutions might further allow for a better discretization especially for locations with small upstream drainage areas than coarser resolutions. This study compared the model performance between the native 5 km LISFLOOD setup and the newly 30 arc sec setup with regard to discharge simulations for a case study in the Elbe basin, Germany. The comparison was done twice: for an un-calibrated and calibrated model scenario.

Static maps describing topographic and river channel characteristics were newly derived on 30 arc sec, whereas other maps such as land use and soil characteristics were only resampled from the native 5 km LISFLOOD setup to the higher resolution due to the lack of information about their sources and/or the time constraint of this project. Additionally, the new high resolution data set was developed on the geographic projection WGS84 rather than the native ETRS-LAEA projection of the 5 km data set in order to have a consistent framework with the Global Flood Awareness System (GloFAS) which is the equivalent of EFAS on the global scale. Until now, the 30 arc sec setup neglects sub-grid information on channel length and channel gradient and hence, for comparison reasons, the 5 km setup was employed the same way. Nevertheless, to investigate the model performance from the optimal native 5 km LISFLOOD setup, the sub-grid information on channel length and channel gradient was included in the 5 km[++] setup.

In case LISFLOOD run un-calibrated, the results suggested that there is a drop in model performance if sub-grid information on channel length and gradient is neglected. Even the 5 km setup performed slightly better regarding the Kling-Gupta Efficiency than the newly 30 arc sec setup for all validation stations. This could be due to the used standard parameter set which was more or less manually tuned for the 5 km LISFLOOD setup, the resampling of certain static maps from the native 5 km to the 30 arc sec grid, the omission of increasing the temporal scale or the large-scale model parameterization which does not account for small-scale processes. The calibrated model outputs gave a different picture where the 5 km and 5 km[++] setups showed better agreements to the observations in the downstream part, while the 30 arc sec data set yielded an improvement for the upstream area. Because the parameterization of LISFLOOD was originally designed for large-scale catchment processes the impacts of spatial resolution become less significant for locations with large upstream drainage areas.

The newly derived 30 arc sec data set is by no means complete and ideal. Static maps such as land use and soil characteristics have to be derived from high resolution sources in order to include more details in the 30 arc sec grid. Because the spatial and temporal scales are linked, future studies should also focus on increasing the temporal resolution of LISFLOOD, also for model calibration. Furthermore, new parameterization techniques should be developed to account for small-scale processes which take on greater significance while moving towards hyperresolution.

Preface

This report documents the procedure and outcomes of a five-month traineeship at the Joint Research Centre (JRC) of the European Commission in Ispra, Italy. For the Master's programme Earth Surface and Water at the University of Utrecht in the Netherlands, a five-month traineeship can be done for which 30 ECTS are allocated. The general goal of the traineeship is to enable students to gain practical experiences by working for instance in a company. The traineeship was undertaken from the 16th of May 2018 to the 16th of October 2018 and was supervised internally by Dr. Vera Thiemig and externally by Prof. Dr. Steven de Jong (Utrecht University). This report serves as proof of performance for both the JRC and the University of Utrecht.

My thanks goes to the whole Disaster Risk Management Unit of the JRC who welcomed and affiliated me in a very friendly and straightforward manner. Further, I would like to thank especially Vera Thiemig, Lorenzo Alfieri, Valerio Lorini, Goncalo Gomes, Peter Salamon, Domenico Nappo, Juliana Disperati, and Francesco Dottori for their support during this internship.

Table of Contents

1 Introduction

The European Flood Awareness System (EFAS) developed by the Joint Research Centre (JRC) of the European Commission provides complementary, added-value information (e.g. probabilistic, medium range flood forecast, flash flood indicators or impact forecasts) to the relevant national and regional authorities. Furthermore, EFAS keeps the Emergency Response Coordination Centre (ERCC) informed about ongoing and possibly upcoming flood events across Europe. The prime aim of EFAS is to support preparatory measures before major flood events strike, particularly in trans-national river basins and throughout Europe in general (Bartholmes *et al.* 2009). EFAS is the first operational European system monitoring and forecasting floods across the whole of Europe. Since 2012, EFAS is fully operational as part of the Copernicus Emergency Management Service.

LISFLOOD is the underlying hydrological rainfall-runoff-routing model of EFAS. It is capable of simulating the hydrological processes that occur in a catchment (van der Knijff *et al.* 2010, Burek *et al.* 2013a). Currently the model runs on a 5 km grid resolution with daily time steps. The ETRS89 Lambert Azimuthal Equal Area (ETRS-LAEA) is used as a single projected coordinate reference system for the entire European domain (Burek *et al.* 2014).

For earth system models in general, there is a drive towards hyperresolution scales which are in the order of 1 km or less (Beven *et al.* 2014). It is expected that using a higher spatial model resolution allows for a better representation of the heterogeneity in topography, soils, and vegetation and their effects on hydrological dynamics (Wood *et al.* 2011). Moreover, a poor representation of the upstream contributing area affects the simulation of streamflow (Li *et al.* 2013). Higher spatial model resolutions might thus allow for a better discretization especially for locations with small upstream drainage areas than coarser resolutions (Arora *et al.* 2001, Li *et al.* 2013). Beside moving towards a higher spatial model resolution, the JRC is planning to have the EFAS model nested into the Global Flood Awareness System (GloFAS, Alfieri *et al.* 2013) in order to have a more interactive and consistent framework. GloFAS is the equivalent to EFAS on the global scale and runs LISFLOOD on spherical coordinates (lat-lon) with a spatial resolution of 0.1°. This study investigates the benefits and affiliated costs of increasing the spatial resolution of EFAS from 5 km to 30 arc sec or 0.00833° (~1 km at the Equator) based on a case study in the Elbe basin, Germany.

This internship report is structured as followed: The introduction is followed by a brief description on the hydrological model LISFLOOD. Then a technical description is provided on how the high resolution data (static input layers as well as the meteorological forcing) was produced, including the Python scripts that were developed for the map productions (attached in the digital appendix). Thereafter a hydrological comparison is presented, which was performed between the outcomes of the model runs with the native 5 km spatial resolution and the 30 arc sec spatial resolution (hereafter referred to as 30" data set), by comparing the model output against observed discharges. This comparison was done twice: for an un-calibrated and calibrated scenario. Finally, the study is rounded off by the conclusions.

Note: This report does not aim to replace any former technical report nor to present a comprehensive description of all the input layers used in LISFLOOD. The objective is to provide a description on how the 30" data set was derived and how it can be further developed in the future. For more detailed documentation about the input layers and the LISFLOOD model itself, this report refers to the LISFLOOD

manual (Burek *et al.* 2013a) and the technical report describing the input layers for LISFLOOD (Burek *et al.* 2014).

2 LISFLOOD – Model description

This section shortly describes the hydrological model LISFLOOD (Burek *et al.* 2013a). LISFLOOD was developed by the JRC with the objective to acquire a tool that can be used in large and trans-national catchments for flood forecasting. The model is grid-based and designed to be applied across a wide range of spatial and temporal scales (Burek *et al.* 2013a). However, LISFLOOD was primarily developed for the simulation of large river basins and thus small-scale processes are often simulated in a simplified way (van der Knijff *et al.* 2010).

As shown in Figure 2-1, the model consists of a two-layer soil water balance sub-model, sub-models for the simulation of groundwater and subsurface flow (using two parallel interconnected linear reservoirs), a sub-model for the routing of surface runoff to the nearest river channel, and a sub-model for the routing of channel flow (not shown in Figure 2-1) (Burek *et al.* 2013a). Since a couple of years, a third soil layer has been introduced (also not shown in Figure 2-1). LISFLOOD is able to model interception, evapotranspiration, infiltration, surface runoff, water uptake by vegetation, leaf drainage, snow melt, preferential flow, drainage to the groundwater, sub-surface and groundwater flow, and flow through river channels. Snowmelt is simulated by a simple degree-day factor method but takes accelerated snowmelt into account when it rains (Thielen *et al.* 2009). Infiltration is estimated using the widely-used Xinanjiang approach (Zhao and Liu 1995, Todini 1996). Adding a fraction of the water that is available for infiltration directly to the groundwater accounts for preferential flow (van der Knijff *et al.* 2010). Note that capillary rise and deep groundwater systems are not simulated. Water in each channel pixel is routed through the channel network by using the kinematic wave approximation. The model is also able to apply the full dynamic wave routing which in turn requires some additional channel cross-section information. Furthermore, special structures such as lakes, reservoirs and polders can be taken into account by stating their location, size and in- and outflow boundary conditions.
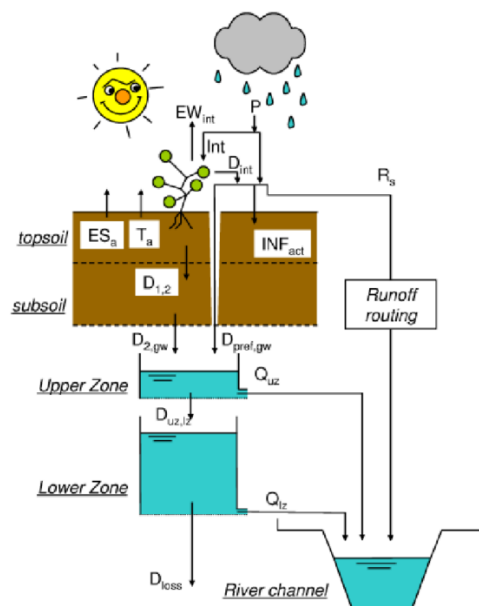

Figure 2-1 Overview of the LISFLOOD model.

3 Description of input layers

Since moving towards spatial hyperresolution also means an increase in computational costs, a basin rather than the entire European domain (Figure 3-1) was preferred for testing the 30" data set. Due to the good availability of data, the Elbe basin was chosen as study area.

LISFLOOD requires over 80 static input layers describing for instance topography, soil characteristics, or channel geometry. Not all 30" input layers, however, were derived from scratch due to a) the lack of information regarding the source of certain input layers of the native 5 km setup and b) the time constraint of five months given by the length of this internship. As a result of that, certain input layers such as land use for instance were derived by resampling (i.e. downscaling) the 5 km land use layer from the native LISFLOOD setup to the 30" grid instead of deriving the 30" land use layer by upscaling the source land use layer which would be a high resolution (< 1 km) remote sensing map. By downscaling the 5 km layer from the native LISFLOOD setup, no further information will be included in the 30" grid. Table 3-1 provides an overview about which input layers (thematically grouped) were simply resampled from the native 5 km LISFLOOD setup and those which were derived newly on 30". The 30" layers generated by resampling the 5 km layers cover the entire European domain while others, such as the local drain direction (LDD) or digital elevation model (DEM), were only produced for the Elbe catchment so far.

Table 3-1 30" thematic input layers and their spatial domain and source resolution.

Thematic layers	Spatial domain		Source resolution
	Europe	Elbe	
Topography (LDD, DEM, etc.)		X	30"
River channel (width, depth, etc.)		X	30"
Land use (forest, water, etc.)	X	X	5km
Soil characteristics (lambda, ksat, etc.)	X	X	5km
Meteorological forcing (pr, ta, etc.)	X	X	5km

3.1. LISFLOOD version and NetCDF raster format

LISFLOOD is implemented in the PCRaster Environmental Modelling language (Wesseling *et al.* 1996). Besides a number of valuable features, e.g. build-in functions, the PCRaster format lacks to deal with spatial-temporal data in a structured and space-saving way. For instance, a time series of spatial data such as precipitation will be stored as single raster file per each time step. The Network Common Data Form (NetCDF) is capable of storing spatial-temporal data as a three dimensional array (two coordinate dimensions and one time dimension) in one single file. Recent LISFLOOD versions are able to read and write NetCDF rasters. The 30" input layers for LISFLOOD (static maps and meteorological forcing) are thus provided as NetCDFs. It should be noted that LISFLOOD reads the input NetCDF raster by means of indexing and not by variable names. The structure of the NetCDF raster containing e.g. the variable pr (precipitation) should hence be like:

NetCDF_raster [x/lon,   y/lat,   coor. sys,   time,   pr]

Index   0   1   2   3   4

Importantly, the first two variables in the NetCDF raster represent the coordinates x/lon and y/lat, respectively, and the precipitation variable pr is sorted in the last position.

The raster extent and raster resolution are internally calculated in LISFLOOD based on the first two values of both coordinate arrays, i.e. lon[0], lon[1], lat[0], and lat[1]. The first values of each coordinate variable have to represent the upper left corner of the raster! It is thus crucial to have the exact and correct values in the coordinate arrays. LISFLOOD further rounds the extracted values from the coordinate variables for unknown purposes. Since LISFLOOD was initially implemented on the LAEA projection with a 5 km grid resolution, rounding the coordinate values does not pose any problems as there are no fractional digits. This rounding in LISFLOOD, however, does pose a problem if coordinates values may have some fractional digits which is the case in the 30" data set. Hence, the rounding commands in the LISFLOOD code were removed. The spatial resolution of the 30" data set is 0.00833°.

3.2. Projection, raster extent and mask

Likewise GloFAS, the 30" data set uses the World Geodetic System of 1984 (WGS84) as spatial reference (EPSG: 4326). In order to produce the input layers for the 30" data set, the native 5 km LAEA projection maps were re-projected to the aforementioned WGS84 reference system. As a result of the conversion from planar to spherical coordinates, the European domain (Figure 3-1) from the native 5 km setup is not rectangular anymore. Figure 3-1 shows the map extent for the European domain as well as the Elbe catchment where the 30" data set was employed. Running LISFLOOD only for a catchment is possible by means of implementing a Boolean mask map.
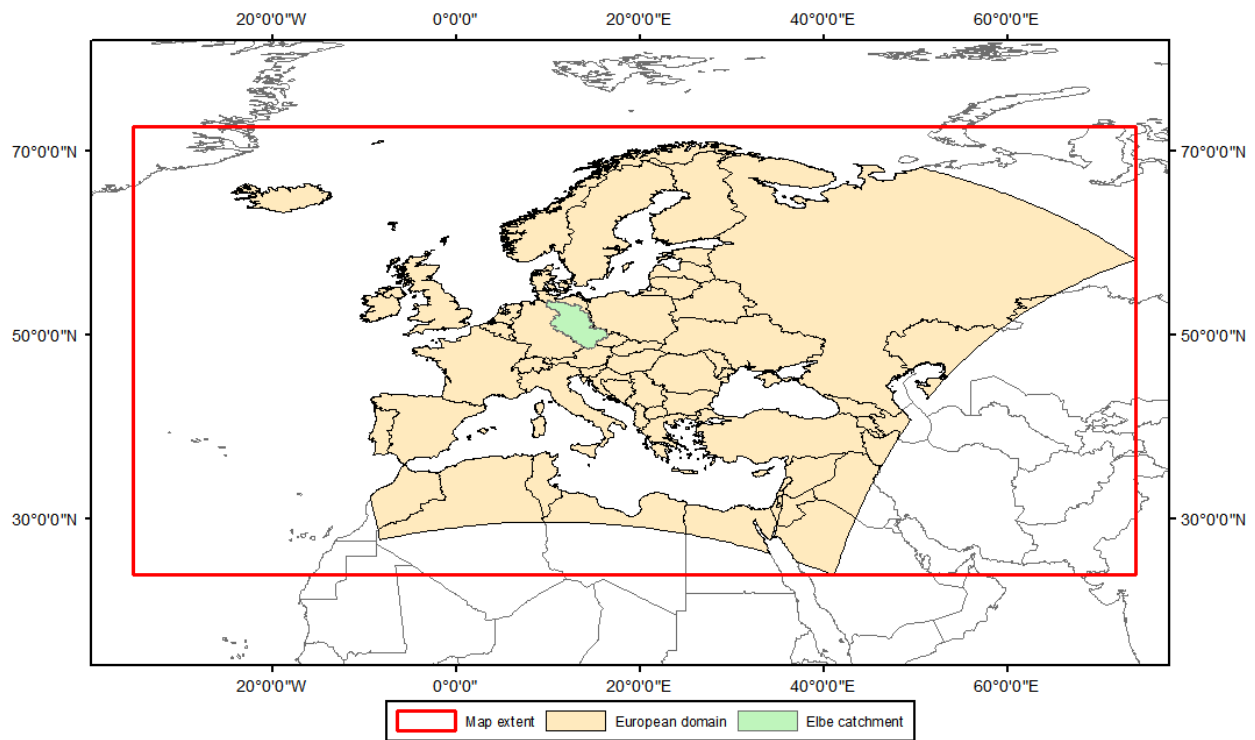

Figure 3-1 Map extent for the European domain and the Elbe catchment.

5

The map extent is given in Table 3-2. Note that the negative value for the left side of the raster denotes West of the Prime Meridian.

Table 3-2 Map extent

Extent	
Top	72.66664
Left	-35.12499
Right	73.95636
Bottom	23.92781

3.3. LISFLOOD options

LISFLOOD can be run in different modes, depending on the user's needs. Subsequently, some additional insights to the description given in Burek *et al.* (2013a) are provided:

1) The *groundwatersmooth* option is meant to work in combination with the *water use* option (requires additional maps wregion and energyconsumptionuse described in Chapter 3.4.7). It smooths out sudden groundwater level changes due to water abstractions. Either both options are switched on or off.
2) Option *indicator* (requires population map, not described in this report because it was not used) produces some water indices but does not really affect the output.
3) Option *transientlandusechange* (requires population map as well) is used for climatic simulations of several years where land use changes over the years.
4) Option *InitLISFLOOD* should be run without any report statements, such as repdischarge.tss, because the generated maps lzavin (groundwater inflow) and avgdis (average discharge) will contain otherwise a time variable in their files and LISFLOOD would not able to read these maps for the normal run.
5) *GridSizeUserDefined* option needs to be active if the maps are provided in spherical coordinates such as lat-lon.

3.4. Static maps

A major difference between the data set described in this report and the one described by Burek *et al.* (2014) is the spatial reference of the maps. By using lat-lon coordinates, the true cell area of the raster pixels varies especially with latitudes. This fact needs to be taken into account when deriving certain maps such as upstream area or DEM gradient where true cell area or cell length is required, respectively.

All subsequently described static maps were derived by scripts written in Python using Gdal, Numpy, Xarray, or PCRaster libraries. The scripts are provided in the digital appendix. PCRaster and NetCDF format maps are henceforth referred to as PCR and NC maps, respectively.

As already mentioned in Chapter 3, due to time constraints, not all input data were created from scratch. Input maps characterizing land use and land use change (Chapter 3.4.5), soil properties and leaf

area index (Chapter 3.4.6), and water demand (Chapter 3.4.7) for instance were simply resampled and re-projected from the 5 km setup to 30". To do so, the gdalwarp utility from the Gdal library was applied. Gdal supports NetCDF but generates variables with predefined labels and positions. Hence, the outputs needed some modifications in order to be read by LISFLOOD (see LISFLOOD format convention in Chapter 3.1). For that purpose, a script was written which restructures and renames the variables in the initially generated NC maps. All generated static maps were modified afterwards by this script (see modify_static_netcdfs.py in the digital appendix).

A detailed description on how each map was created is provided in a structured way by thematic groups in the following sub-sections.

3.4.1. Pixel area and pixel length

a) Pixel area and pixel length (pixarea_leng.py)

Because the Earth is spherical, the meridians of longitude are approaching when moving towards the poles. In that sense the true area becomes smaller at higher latitudes. The distance between the latitudinal lines changes only marginally and can thus be neglected. Little documentation is found in literature regarding formulas calculating the area of a spherical quadrangle bounded by latitudinal and longitudinal lines. The formulas to calculate the pixel area and length were taken from [https://gis.stackexchange.com/questions/127165/more-accurate-way-to-calculate-area-of-rasters?utm _medium=organic&utm_source=google_rich_qa&utm_campaign=google_rich_qa, Stand: August 2018] and [https://github.com/wri/geodesy, Stand: August 2018]. The script pixarea_leng.py generates pixel area and pixel length as NC maps. Note that the variable positions in the NC maps need to be re-structured to fit the LISFLOOD format convention as stated at the beginning of Chapter 3.4.

3.4.2. Topography

A DEM is required to generate the slope gradient and elevation standard deviation maps. The DEM used here is the 30" void-filled elevation (un-conditioned) from HydroSHEDS (Lehner *et al.* 2008).

b) Local drain direction (LDD)

The LDD is derived from the HydroSHEDS drainage direction map (DIR) on 30" (Lehner *et al.* 2008). The DIR on 30" is a result from a DEM upscaling process including stream burning and sink filling. Further documentation about the DIR can be found in Lehner *et al.* (2008). The DIR is then converted into a LDD PCR map. Note that PCRaster uses a different code for the 8 point pour algorithm than for instance HydroSHEDS which in turn has the same flow direction implementation than ESRI's. The LDD PCR map is afterwards converted to a NC map.

c) Upstream area (uparea.py)

The generated LDD and pixel area maps are needed to obtain the upstream area. PCRaster provides a build-in function called 'accuflux' that can be used to calculate for each raster cell the corresponding upstream area. The function accumulates all pixel areas flowing into each downstream cell.

d) Gradient (gradient.py)

On planar coordinates where each raster cell represents the same true area, the calculation of the slope gradient is straightforward being the fraction between the steepest downhill decent and the horizontal length scale. On spherical coordinates, this calculation becomes more complex since the true area of the raster cells becomes smaller with higher latitudes. This means that the horizontal length scale changes spatially. The procedure used here is taken from [https://gis.stackexchange.com/questions/14750/using-srtm-global-dem-for-slope-calculation, Stand: August 2018] and follows the suggestions from Snyder (1987):

1) Create a latitude grid and compute its cosine
2) Project both the DEM and the cosine of the latitude using a Mercator projection in which scale is true at the Equator
3) Compute the slope of the projected DEM as percentage
4) Divide this slope by the projected cosine(latitude) grid
5) Reproject the slope grid back to WGS84

e) Elevation standard deviation (raster_upscaling.py)

LISFLOOD requires information about the standard deviation of surface elevation on a sub-grid level which is used for modelling snow accumulation and melt. This information needs to be extracted from a source having a higher resolution than the resolution on which the data set is being developed. HydroSHEDS not only provides the used 30" DEM but also a 15" and 3" DEM from which such sub-grid information can be extracted. The 15" DEM is available as a single file whereas the 3" DEM is divided into multiple tiles. For time-saving reasons, the 15" DEM has been used to derive the standard deviation of surface elevation in each 30" grid. This is of course unfavorable since the 15" DEM has only four raster cells within each 30" cell. This should be changed at the next opportunity where the 3" DEM is used instead. The script provided is able to retrieve the standard deviation from any given higher resolution raster.

3.4.3. River channel

f) Channel bottom width (chanbw_Step1.py and chanbw_Step2.py)

Two steps are required to obtain the channel bottom width. First, channel bottom width is related to upstream area with an empirical relationship. LISFLOOD is then run preferably for a long term run and the resulting daily average discharge is used in a second step to retrieve the final channel bottom width through the equation from Pistocchi et al. (2006).

Step 1:

$$chanbw = upArea * 0.0032 \tag{3.1}$$

Step 2:

$$chanbw = 7.0825 * avgQ^{0.539} \tag{3.2}$$

with *chanbw* being the channel bottom width [m], *upArea* the upstream area [km^2], and *avgQ* the long-term daily average discharge [m^{3}s^{-1}] resulting from the initial LISFLOOD run.

g) Channel bankful depth (chanbnkf_Step1.py and chanbnkf_Step2.py)

Simultaneously with the channel bottom width, the bankful depth is extracted in the two aforementioned steps as well. First step relates bankful depth to upstream area. The second step uses the Manning's equation to get a more appropriate estimate of the channel bankful depth.

Step 1:

$$chanbnkf = 0.27 * upArea^{0.33} \tag{3.3}$$

Step 2:

$$chanbnkf = 1.004 * n^{3/5} \, Q^{3/5} W^{-3/5} S_o^{-3/10} \tag{3.4}$$

with $Q = avgQ * 2.0$

where *chanbnkf* denotes channel bankful depth [m], *upArea* the upstream area [km^2], *n* Manning's roughness coefficient [-], *Q* bankful discharge [m^{3}s^{-1}], *avgQ* long-term daily average discharge [m^{3}s^{-1}], *W* channel bottom width [m], and S_o the channel slope [m m^{-1}].

h) Floodplain width (chanflpn_Step1.py and chanflpn_Step2.py)

The floodplain width is given by,

$$chanflpn = chanbw * 3.0 \tag{3.5}$$

where *chanflpn* denotes floodplain width [m] and *chanbw* the channel bottom width [m]. Because the floodplain width is linked to the channel bottom width, this calculation needs to be repeated.

i) Channel map (chan.py)

The channel map (chan) is a Boolean map where cells have a value of 1 at channel pixels and 0 at all other pixels. Since LISFLOOD treats all cells as river cells, the channel map has for the whole Elbe catchment a raster value of 1.

j) Channel side slope (chans.py)

The channel side slope map (chans) is a spatially uniform map having a value of 1.0 [m m⁻¹] for all river cells. The channel side slope belongs to the characteristics describing the channel cross-section geometry. If known, this can be a map with spatially varying channel side slopes. In its simplest form, a spatially uniform value for all channels is assigned.

k) Channel Manning's (chanman.py)

The Manning's roughness coefficient n directly influences travel times and peak amplitudes where a small n results in smooth surfaces and thus faster travel times and higher peaks, and a large n in rougher surfaces and hence slower travel times and lower peaks (Burek *et al.* 2014).

A regression function linking Manning's n to upstream area and the DEM is used to obtain the base map. A minimum value of 0.025 is used for flatland rivers with large upstream areas. A maximum of 0.015 is added for flatland rivers and small upstream areas (upstream area dependent) and another maximum of 0.030 is added if in mountainous areas (elevation dependent) (Burek *et al.* 2014).

$$Manning's\ n = 0.025 + 0.015 * \min\left(\frac{50}{upArea}, 1\right) + 0.030 * \min\left(\frac{DEM}{2000}, 1\right) \tag{3.6}$$

with *upArea* being the upstream area [km²], and *DEM* the digital elevation model [m].

l) Channel length and channel gradient

Channel length and channel gradient are sub-grid-scale topographic parameters derived by upscaling the LDD (see Burek *et al.* 2014). Even though the used 30" LDD from HydroSHEDs has been upscaled as well, HydroSHEDs does not provide this sub-grid information. Therefore, for the 30" data set, the channel length and channel gradient maps were not produced. Instead, the pixel length (Section a) and slope gradient from the DEM (Section d) were used as channel length and channel gradient maps, respectively.

3.4.4. Reservoir, lake, and outlet locations

m) Reservoirs, lakes, outlets (latlon_grids.py and shift_pointdata.py)

Reservoirs and lakes affect the hydrological cycle and therefore describing these features at the exact location is substantial. When model outputs are validated against observations at gauge stations, it is obvious that the location of the outlet (i.e. where LISFLOOD is writing the output to disk) should match the location of the gauge station. Depending on the spatial reference and the local drainage direction used, the locations of the reservoirs, lakes, and outlets/gauges may need to be shifted from their original position. This requires that the locations of these features are visually inspected by e.g. open street map and if necessary are corrected. For gauge stations, one could examine the discharges or compare the upstream area in order to verify if a gauge station lies on the main stream or on a small tributary. For the 30" data set in the Elbe catchment an automated shifting algorithm has been applied

to shift reservoirs, lakes, and outlet/gauge locations to the main river. For each e.g. gauge station the algorithm extracts first the underlying upstream area. If the extracted upstream area is higher than a given threshold, then this raster cell is assumed to represent the true position and no shifting will take place. However, in case the extracted upstream area is lower that the given threshold, an initial 3x3 window (red square in Figure 3-2) is generated. If at least one extracted upstream area value within this 3x3 window is higher than the threshold, the gauge station will be shifted towards the raster cell having the highest upstream area. Provided that all upstream areas within the 3x3 window are smaller than the threshold, the algorithm generates subsequently a 5x5 window and repeats the procedure. This continues until at least one raster cell contains an upstream area that exceeds the given threshold. Figure 3-2 shows an example where a 3x3 window is sufficient to obtain raster cells having larger upstream areas than the threshold and the gauge station (yellow dot) is snapped to the cell having the highest upstream area within the 3x3 window (shaded raster cell).

Different thresholds were used for the outlets representing gauge locations and for the reservoirs. In the Elbe catchment there is only one lake whose original position is already correct. An upstream area threshold of 500 km^2 was used for the outlets/gauges whereas a threshold of 4 km^2 was taken for the reservoirs. The reason for the lower threshold for the reservoirs is that some reservoirs are actually lying on small tributaries and not necessarily on the main river. Note that this procedure to shift the outlets and reservoirs has been tested for the Elbe catchment only and not for a larger area with multiple catchments where different thresholds per catchment might be necessary.

Figure 3-2 Shifting algorithm employed in the 30" data set for outlet locations representing gauge stations and reservoir locations.

3.4.5. Land use and land use change

n) Land use and land use change (proj_and_resamp.py and proj_and_resamp_stackedNetCDF.py)

LISFLOOD accounts for sub-grid variability in land cover for open water, sealed area, forest, paddy rice irrigated area, crop irrigated area, and 'other'. These maps display a percentage representing the frequency of its class within each upscaled raster cell. This means that in order to obtain these maps, an

upscaling procedure needs to be performed from an original high resolution land use map to a coarser grid. Clearly, for each coarse raster cell (e.g. 5 km) the fractions of all land use classes sum to 100%. Unfortunately, only incomplete documentation was found about the original high resolution land use map and its upscaling technique. Additionally, there are two more maps (fracgwused, fracncused) whose source is not known. Due to the lack of information, these maps were resampled and re-projected from the 5 km setup to 30". However, in case a high resolution land use map is available, the script raster_upscaling.py in Section e) can be used to derive the fractions of each class for any coarser raster.

3.4.6. Soil characteristics and leaf area index (LAI)

o) Soil characteristics and LAI (proj_and_resamp.py and proj_and_resamp_stackedNetCDF.py)

Only incomplete documentation was found about the origin of the soil characteristics and LAI maps. Note that there are four LAI maps (forest, paddy rice irrigated, crop irrigated, and "other")! Pedotransfer functions are used to derive the required soil characteristics. However, only for two out of three implemented soil layers these functions are described. The maps were thus only resampled and re-projected from the 5 km setup to 30".

3.4.7. Water demand / water use

p) Water demand (proj_resamp_stackedNetCDF.py)

LISFLOOD provides an option to include water withdrawal from the domestic, energy, industrial, and agriculture sectors. Hence, maps representing the yearly total water withdrawal demand for the four sectors are required. Water is then withdrawn only from discharge in the river network for the corresponding raster cell. If the water demand is larger than the available amount of water within this grid cell, water is taken from downstream moving along the LDD (Burek *et al.* 2013a). However, no documentation was found about the origin of the water demand maps concerning the domestic, energy, industrial, and agriculture sectors. The maps were resampled and re-projected from the 5 km setup to 30".

q) Eflow (eflow.py)

The water use option requires a map called eflow describing the 10th percentile of discharge from a preferably long-term run, e.g. repeat the same long-term run used to obtain channel bottom width (Section f) and channel bankful depth (Section g) while integrating directly these new channel maps. For that purpose, LISFLOOD runs on as natural conditions as possible with disabled water use option and without reservoirs. Reservoirs are excluded by keeping the *reservoir* option active but using a reservoir map (resZero) containing only zero's.

Numpy or Xarray libraries have functions available to calculate percentiles for three dimensional arrays (two coordinate dimensions and one time dimension). If the three dimensional discharge array is small

enough to fit into memory, the calculations can be done by using the Numpy percentile function. However, since a long-term discharge time series is favored, the discharge array becomes easily far too big and using the Numpy function will result in memory errors. Hence, the Xarray library should be used as it supports reading the data in chunks. Depending on the size of the discharge array, the Xarray quantile function is either applied directly on the entire three dimensional array or one has to loop over each raster cell and calculate the percentile on the extracted one dimensional array along the time axis.

Either way, this may be quite time consuming especially when the discharge array also accommodates a lot of nodata values (e.g. when LISFLOOD runs with a mask map). Therefore, the script employed here crops firstly the three dimensional discharge array to the smallest possible extent and excludes most of the redundant nodata values in the calculations.

r) Water regions (wregions.py)

Water use is restricted to water regions (wregion) representing sub basins within a country (regions should not cross borders). The idea is to have the regions matched to ''drinkingwater-managed-areas''. No water is taken from downstream of a water region (Burek *et al.* 2013a). Sub basins are created at river confluences and receive unique ID's. Besides the LDD, a PCR map delineating the countries is required.

Instead of deriving water regions at river confluences, the river basin districts from the Water Framework Directive can be used which might match the "drinkingwater-managed-areas" better. However, an initial comparison between these two water regions inputs did not reveal any differences in simulated discharges and model run times whatsoever.

s) Energyconsumptiveuse (energyconsumptiveuse.py)

Power plants located near large rivers, lakes, or oceans often use an inefficient open-cooling-tower technique resulting in large water losses due to evaporation/cooling. Power plants at smaller rivers are likely to have a 'closed' cooling system which is more efficient. For rivers having a larger upstream drainage area than 20,000 km^2, an 11 km zonation was applied. A higher energy consumption is then allocated to this zonation than to the remaining area.

3.4.8. Stand-alone maps

t) Lzthreshold, gwbodies, lakemask, and lusemask (lzthreshold.py and boolean_maps.py)

The map lzthreshold describes the threshold value below which there is no outflow from the lower zone to the channel and has a spatially uniform value for the entire domain. The locations of groundwater bodies and lakes are represented as Boolean maps in gwbodies and lakemask, respectively. The latter is used to calculate evaporation from open water. The land use mask (lusemask) map is also a Boolean mask but contains for the entire model domain a true value (1). These maps were resampled and re-projected from the 5 km setup to 30".

3.5. Meteorological forcing

u) Meteorological forcing (forcing_e0.py, forcing_es.py, forcing_et.py, forcing_pr.py, and forcing_ta.py)

LISFLOOD requires meteorological forcing such as precipitation rate, average temperature, potential evaporation rate for free water surface and bare soil, and potential evapotranspiration rate for reference crop. The meteorological forcing is derived by interpolating observed precipitation and temperature and with a supplement model of LISFLOOD called LISVAP (Burek *et al.* 2013b) which produces the potential evaporation/evapotranspiration rate maps. Regarding precipitation and temperature, the former interpolation technique Inverse Distance Weighting (IDW) has been recently replaced by SPHEREMAP (Willmott *et al.* 1985). SPHEREMAP works similar to IDW but takes into account the spatial clustering of the observation stations.

Also the meteorological forcing was resampled and re-projected from the 5 km setup to 30". Unlike the static maps, this procedure becomes a bit more complex because temporal data stored as NetCDF cannot be easily re-projected. The provided scripts extract the spatial two dimensional array at each time step, converts it to a tiff, re-project the tiff to WGS84, and puts the new array back into a NetCDF raster.

4 Hydrological comparison between 30" and native 5 km resolution

Because the 30" data set omits the sub-grid information on channel length and channel gradient (Chapter 3.4.3), the 5 km setup neglects this information as well in order to ensure consistency. Therefore, the pixel length and slope gradient from the DEM were used as channel length and channel gradient maps in the 5 km setup, respectively. Nevertheless, a third setup called 5 km[++] was created including the sub-grid parameters channel length and channel gradient to also examine the optimal setup from the native 5 km LISFLOOD resolution.

A LISFLOOD run was performed for all three setups (5 km, 5 km[++], and 30") for the period 2001 to 2005. Firstly, the model run un-calibrated, using uniform default values for all calibration parameters: time constants for water in the upper zone and in the lower zone, maximum rate of percolation going from the upper to the lower zone, maximum loss rate out of the lower response box, power coefficient in the Xinanjiang distribution function, power coefficient related to preferential flow and soil moisture storage, snow melt coefficient, and multiplier for the channel Manning's *n* (Table 4-1). See Burek *et al.* (2013a) for further documentation of these parameters. The simulated discharges were validated against observations at six stations for the un-calibrated runs (Figure 4-1).

Secondly, the aforementioned eight parameters were calibrated at Station 172, Station 169, Station 349, and Station 7 (Figure 4-1) for the period 2005 to 2010. Station 168 and Station 165 were deliberately excluded from the calibration procedure due to time constraints. Subsequently, the model run with the calibrated parameters for the period 2001 to 2005 and the resultant simulated discharges were validated against observations at the four mentioned stations. The calibration was performed by using the Non-dominated Sorting Genetic Algorithm-II (NSGA-II) multi-objective genetic algorithm (Deb *et al.* 2002).

Table 4-1 Calibration parameters, used value ranges in the calibration as well as their default values.

Parameter name	CalibrationRange		Default value	Unit
	minValue	maxValue		
UpperZoneTimeConstant	3	40	10	days
LowerZoneTimeConstant	40	500	100	days
GwPercValue	0.01	2	0.8	mm day^{-1}
GwLoss	0	0.5	0	-
b_Xinanjiang	0.01	1	0.5	-
PowerPrefFlow	0.5	8	4	-
SnowMeltCoef	2.5	6.5	4	mm °C^{-1} day^{-1}
CalChanMan1	0.1	15	3	-

The Kling-Gupta efficiency (KGE, Gupta *et al.* 2009) was used to validate the model performance at the respective stations. KGE measures the model's skill by considering variability, bias and correlation. The components variability and bias compare simulated and observed standard deviations and means, respectively. Further differences between simulated and observed, such as timing of the peaks, and shapes of the rising and falling limb of the hydrographs, are analyzed by the correlation coefficient

(Gupta *et al.* 2009). These three components are then combined and yield the KGE in which a value of one is ideal. KGE is defined as:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha - 1)^2 + (\beta - 1)^2},$$

(4-1)

with

$r = $ *Pearson product-moment correlation coefficient*

$\alpha = \sigma_S / \sigma_O$

$\beta = \mu_S / \mu_O$

where σ_S and σ_O denote the standard deviation of the simulation and observation, respectively. μ_S and μ_O are the simulated and observed mean.
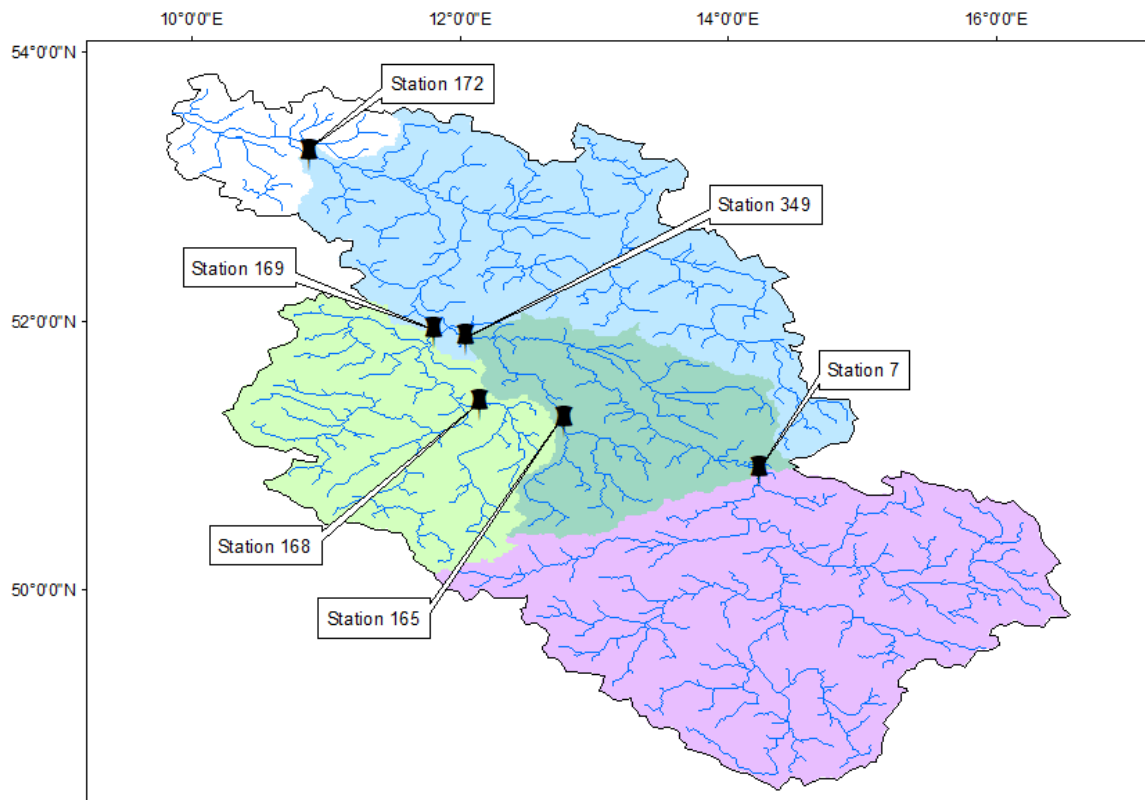

Figure 4-1 Used observation stations for calibration and validation in the Elbe catchment. Only Station 172, Station 169, Station 349, and Station 7 were used for calibration though. Corresponding to these four stations, the sub basins are shown.

4.1. Results

The presented results for all stations in the following tables and figures are sorted in ascending order based on their upstream drainage area.

4.1.1. River network structure

Figure 4-2 illustrates the river network structure derived from the two different spatial resolutions in the south-east part of the study area. For comparison reasons, the 5 km river network was reprojected from the native LAEA projection to WGS84. The higher spatial resolution and hence the larger amount of raster cells results in a smoother river network sinuosity. The river network resulting from the 5 km spatial resolution is more cornered due to the large pixel sizes.
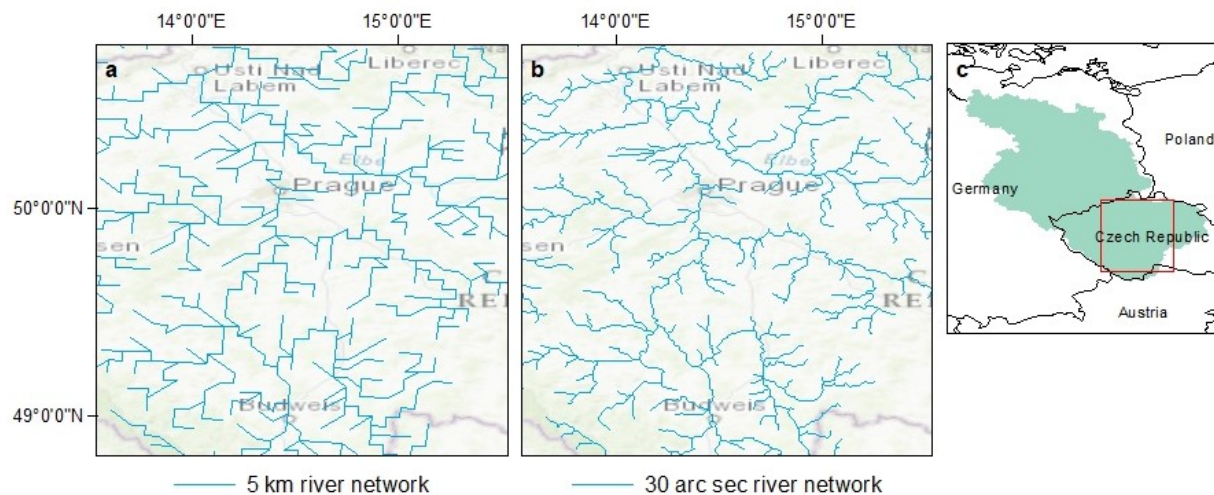

Figure 4-2 Resultant river networks derived on 5 km (a) and 30 arc seconds (b) spatial resolution. The overview map (c) shows the location in the Elbe basin. Note the 5 km river network was reprojected from the native ETRS-LAEA projection to WGS84 for comparison reasons.

4.1.2. Un-calibrated runs

Observed and simulated discharges from the un-calibrated model runs are presented in Figure 4-3. In general, the model captures the timing of the hydrographs pretty well but over predicts discharges at almost all stations except station 165. The flood event in late August 2002 is captured but highly over predicted, especially downstream resulting in discharges over 8,000 m^3s^{-1} (stations 349 and 172). By visually inspecting the simulated discharges from the 30" data set and the two 5 km data sets, no distinctive differences between them can be observed. At the stations 349 and 172, the 5 km^{++} data set seems to capture the observed hydrograph better than the two other 5 km and 30" data sets.

Station 168
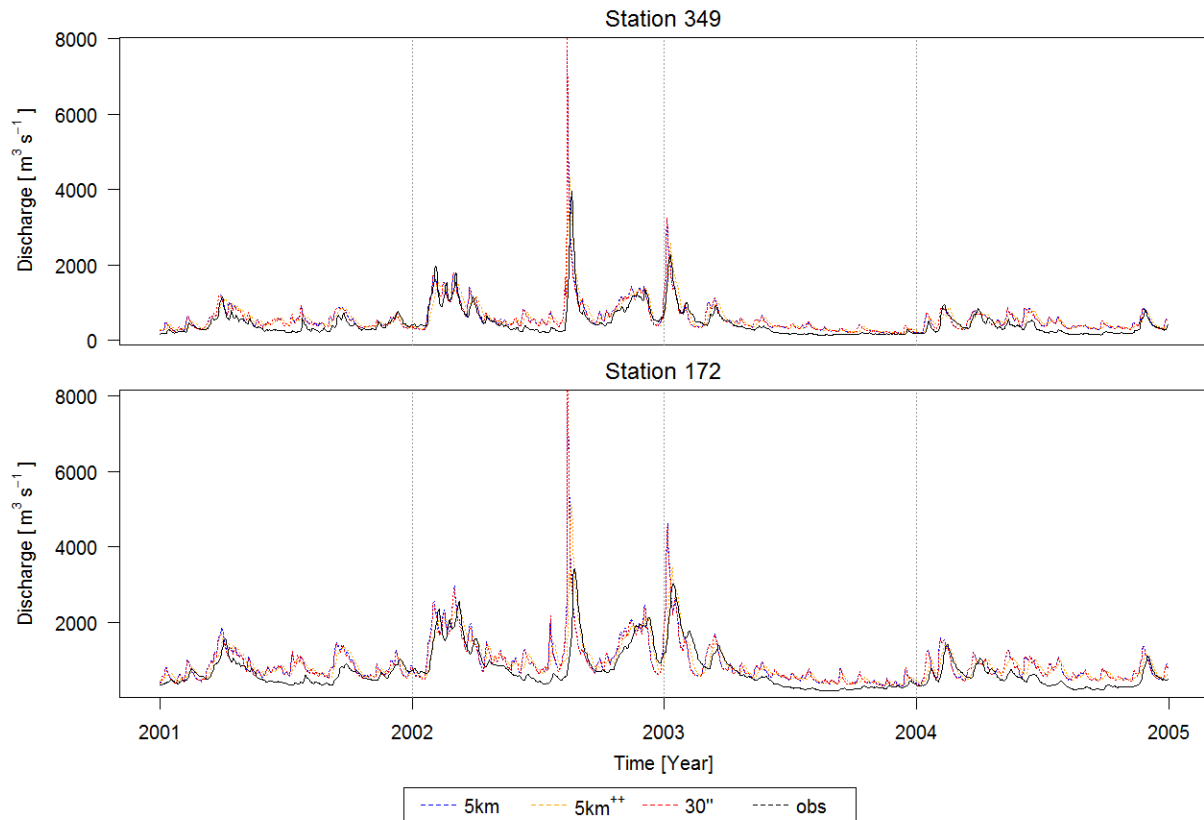
Station 165

Station 169
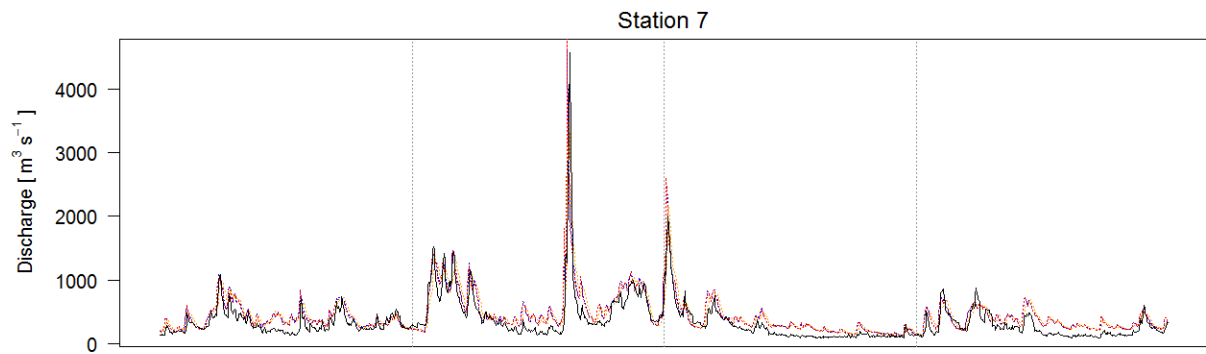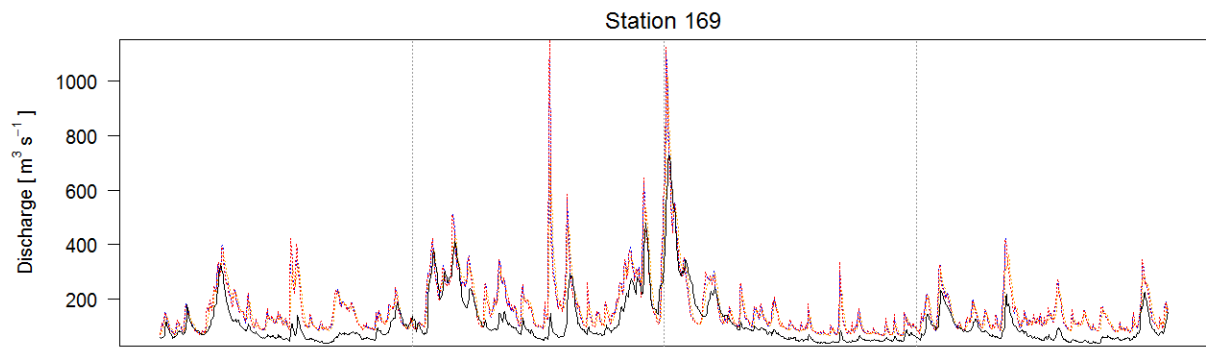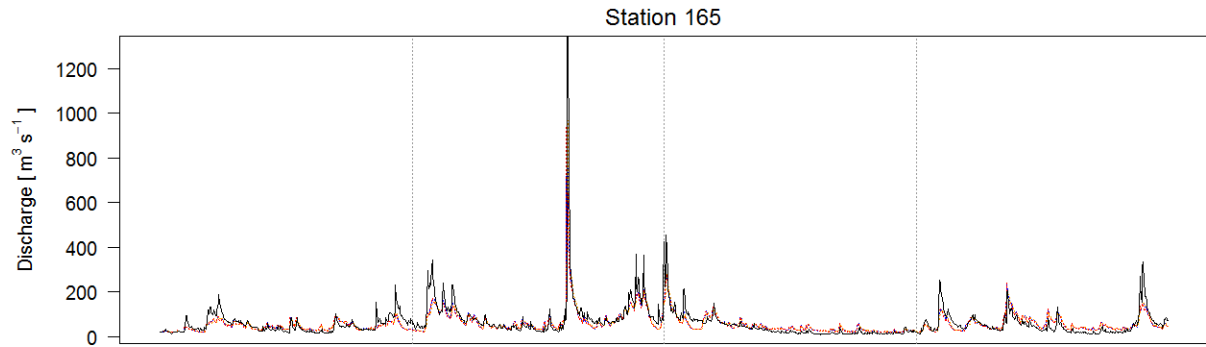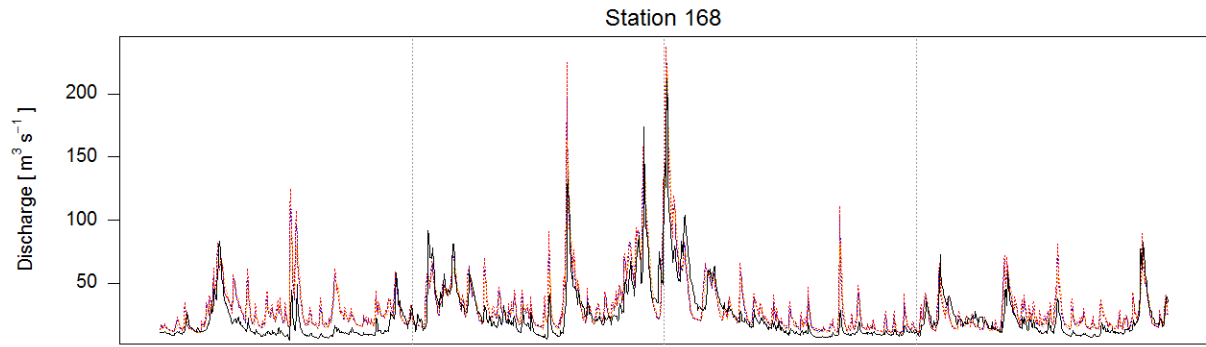
Station 7

Figure 4-3 Observed and simulated discharges (un-calibrated model runs) in the Elbe basin.

The downstream stations (stations 349 and 172) are indeed better represented by the 5 km[++] as Table 4-2 shows. KGE are in the order of ~0.15 higher than the ones resulting from the 5 km and 30" data sets. In fact, the 5 km[++] setup not only performs better in the downstream part, but also in the upstream part. Comparing the 5 km and 30" setups, the 5 km results in slightly better KGE. Especially in the upstream part (stations 168 and 165), the difference in KGE becomes more distinctive.

Table 4-2 Kling-Gupta efficiency (KGE) obtained to evaluate the un-calibrated model outputs.

Stations	Spatial resolutions		
	5km	5km[++]	30"
168	0.67	0.71	0.61
165	0.62	0.64	0.57
169	0.42	0.48	0.41
7	0.67	0.69	0.67
349	0.60	0.71	0.57
172	0.53	0.71	0.53

19

4.1.3. Calibrated runs

Observed and simulated discharges from the calibrated model runs are shown in Figure 4-4. Clearly, the simulated hydrographs from the calibrated model runs capture the observations better than the un-calibrated ones in Figure 4-3. The over prediction of the flood event in late August 2002 remains though. For all stations, low flow is slightly under estimated and the peaks in some occasions over predicted. Likewise Figure 4-3, distinctive differences between the 5 km and 30" data sets are hardly recognizable. All three setups capture the timing of the hydrographs fairly well.
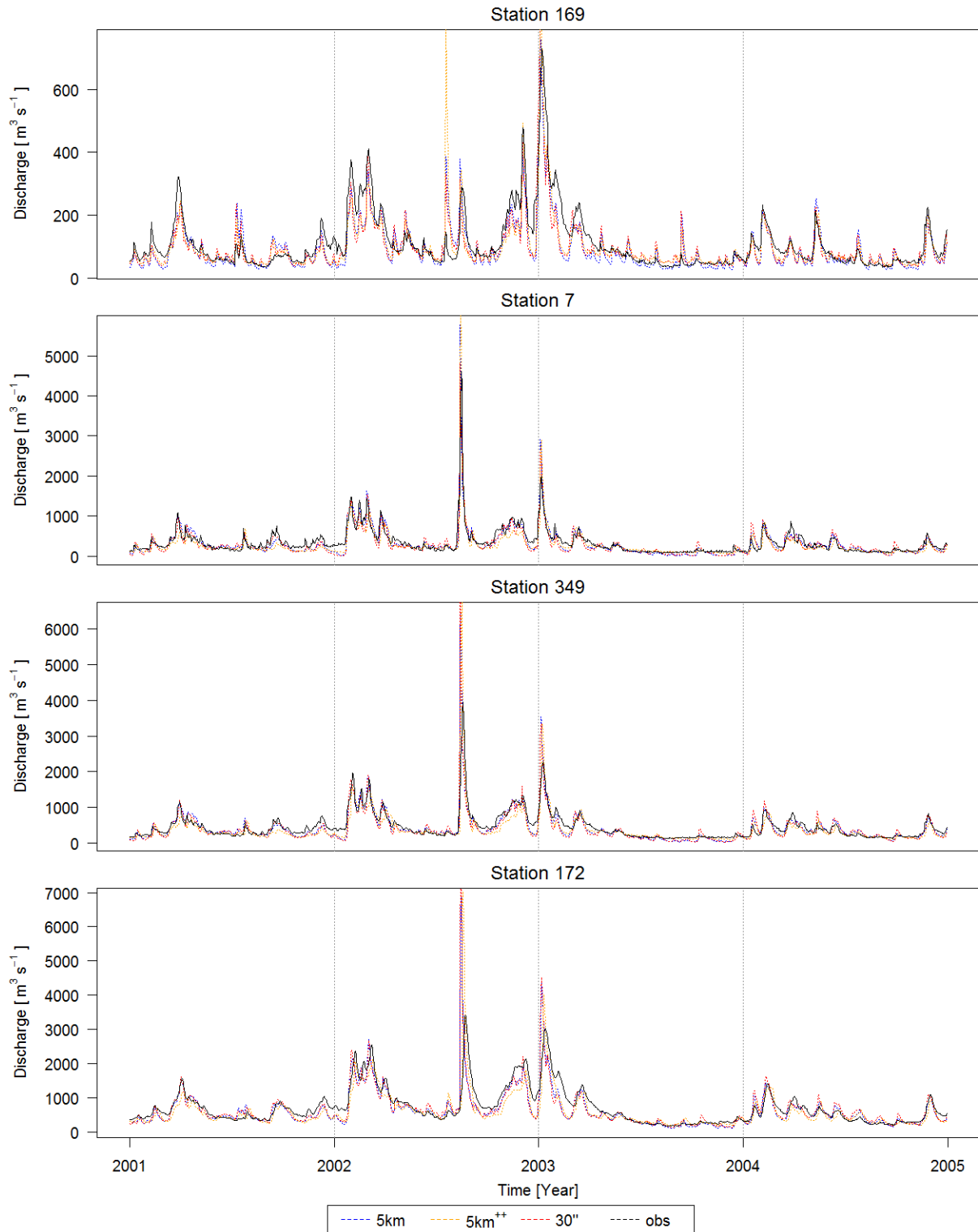
Figure 4-4 Observed and simulated discharges (calibrated model runs) in the Elbe basin.

The resultant KGE from the calibrated model runs are presented in Table 4-3. It can be seen that the 30"
data set performs best for the catchments having smaller upstream areas. That is, a KGE of 0.79 and
0.84 at the stations 169 and 7, respectively. For the catchments with larger upstream areas, the 5 km
and 5 km^{++} perform better for the stations 349 and 172, respectively. Comparing with the KGE from the
un-calibrated model runs which indicated a clear trend with the 5 km^{++} setup performing overall the
best (Table 4-2), the KGE from three setups do not exhibit a similar trend when calibrated. Considering
all catchments/stations, the 5 km setup yields the best KGE ranging from 0.75 to 0.83.

Table 4-3 Kling-Gupta efficiency (KGE) obtained to evaluate the calibrated model outputs.

Stations	Spatial resolutions		
	5km	5km^{++}	30"
169	0.76	0.78	0.79
7	0.83	0.78	0.84
349	0.76	0.75	0.65
172	0.75	0.78	0.70

4.1.4. Run times

Table 4-4 shows the total runtimes [seconds] for the model runs using the best calibration parameter set
over the whole validation and calibration period (2000 to 2010). Further, the number of pixels in the sub
basins for both resolutions are provided. On average, the high resolution 30" data set runs 18 times
longer than the coarser 5 km setup. Table 4-4 further reveals that there are some considerable
differences in run times between the sub basins within the 30" data set. This behavior is not observed
within the 5 km setup whose run times hardly deviate among the sub basins. While moving towards 30",
the number of pixels rises sharply by almost one and a half order of magnitude.

Table 4-4 Total runtime [seconds] for the period 2000 to 2010 (validation & calibration period) and number of pixels in the sub basins.

Stations	Spatial resolutions			
	5km		30"	
	runtime	pixel count	runtime	pixel count
169	784	950	7,267	43,570
7	798	2,056	19,513	90,614
349	759	738	6,917	34,540
172	725	1,513	22,442	66,639

4.2. Discussion

4.2.1. Un-calibrated runs

Regarding the KGE, the 5 km and 5 km[++] LISFLOOD setups show a better model performance than the high resolution 30" data set. Especially by including sub-grid information on channel gradient and channel length, the model performance is considerably improved (see Table 4-2). This implies that also on the higher spatial resolution of 30" the omission of sub-grid variability in channel length and gradient results in the loss of valuable information and causes a drop in model performance. Even when the channel gradient and channel length are excluded, the 5 km setup performs slightly better than the 30" data set, notably in the upstream area (stations 168 and 165). Further analyses are needed to track down these causes. Possible reasons could be:

a) Beside the calibration parameters (Table 4-1), LISFLOOD uses even more parameters e.g. related to evapo(transpi)ration and interception, snow and frost, infiltration, groundwater, and routing. This standard parameter set was adopted as it stands in the 30" setup. However, this standard parameter set was developed and more or less manually tuned for the 5 km LISFLOOD setup.

b) Not all static maps were produced on a higher resolution. Soil characteristics, land use and land use change, to name a few, were downscaled from 5 km to 30". Of course, all static maps (and also the meteorological forcing) are preferably derived by upscaling an original high resolution source map to the desired 30" resolution. By downscaling the 5 km grid to the 30", no additional information is included in the higher resolution 30" grid. However, LISFLOOD takes the spatial sub-grid variability of land use cover into account (Chapter 3.4.5) which is derived from an original high resolution land use map. By including sub-grid information and thus representing spatial heterogeneity, the model becomes less dependent on spatial resolution (Schumann *et al.* 2018).

c) The temporal resolution was not increased. River flow is specified as a function of distance *x* along the channel and time *t* by the kinematic wave approximation (Chow *et al.* 1988). By moving towards spatial hyperresolution one should not neglect the temporal scale because the spatial and temporal scales are linked (Melsen *et al.* 2016a). This study should hence further be expanded by increasing the temporal model resolution. A possible testing setup could look like:

   - 5 km   -   24h
   - 5 km   -   6h
   - 30"   -   24h
   - 30"   -   6h

d) The model structure of LISFLOOD. Moving towards higher spatial model resolutions does not necessarily mean an improvement of the model outputs. This depends whether small-scale processes are actually simulated in the model and if their descriptions are only conceptual or physically based (Beven *et al.* 2015). Furthermore, Mateo *et al.* (2017) noted that the assumption of a single-downstream river network may not be true on higher spatial resolutions and highlighted the importance of multiple pathways and river bifurcations. LISFLOOD for

instance uses a single-downstream-grid river channel and is thus not able to simulate river bifurcations. This issue should be addressed in future model development. Also the applied river routing scheme might have to be revised when it comes to modelling river flow on higher resolutions. Meteo *et al.* (2017) emphasized the importance of floodplain flow and connectivity when spatial resolution is increased. In order to realistically represent floodplain flow, inertia and advection terms may no longer be neglected as done by the kinematic wave approximation.

Sutanudjaja *et al.* (2018) developed and compared the hydrological model PCR-GLOBWB (van Beek and Bierkens 2009) on 5 arc min and 30 arc min and found an improvement in the model performance with the higher-resolution model. The higher-resolution model was able to capture the seasonality, inter-annual anomalies, and general discharge characteristics better than the model on 30 arc min.

4.2.2. Calibrated runs

The results show that when the three implemented setups are calibrated the high resolution 30" data set performs better in the upstream area (stations 169 and 7) than the two 5 km setups. On the other hand, the high resolution data set yields an inferior model performance for the downstream part of the Elbe catchment. Unlike the un-calibrated runs, none of the three implemented setups outperforms the others when they are calibrated (Table 4-3). The coarse resolution 5 km and 5 km++ data sets produce robust discharge estimates. Since the parameterization of LISFLOOD is designed for large-scale catchment processes, the coarser grids (5 km and 5 km++) are assumed to perform better downstream (large upstream areas). An improvement of the streamflow simulation with the high resolution 30" data set is expected in particular for locations with small upstream drainage areas. This is in line with the findings obtained in this study.

Comparing the results in this study to findings given in literature is difficult due to the use of a) different hydrological models and hence different parameterization schemes for land surface and routing, b) different effective parameters which are calibrated, and c) different objective functions and target state variables (e.g. discharge) on which calibration is performed.

A lot of studies focus on parameter transferability when calibration is performed at different temporal and spatial resolutions (e.g. Haddeland *et al.* 2002, Troy *et al.* 2008, and Melsen *et al.* 2016b). Little is found in literature about the differences when a hydrological model is calibrated on both coarser and higher spatial resolutions. Vázquez *et al.* (2002) calibrated and validated the hydrological model MIKE-SHE on three different spatial resolutions (300 m, 600 m, and 1,200 m) in the Gete catchment in Belgium. The results revealed that model performance was overall best when the model was calibrated at a spatial resolution of 600 m. They further analyzed the model performance for high and low flow only. The 600 m grid outperformed the two other grids for high flows as well. The present study could be expanded by further analyzing the high and low flows independently, in order to investigate which of the three implemented setups (5 km, 5 km++, and 30") yield the best model performance for peak discharges only. Kumar *et al.* (2013) tested the sensitivity of daily streamflow simulations to different spatial resolutions (2, 4, 8, and 16 km) by applying the mesoscale hydrologic model (mHM) in the Neckar river basin, Germany. The model performances among the employed spatial resolutions were comparable when the effective parameters were calibrated at the respective spatial resolution. Parameter transferability from coarser to finer resolutions introduced degraded model performance.

Melsen *et al.* (2016a) identified that over the years the spatial resolution of global hydrological models (GHMs) has steadily increased whereas the calibration and validation time interval has remained unchanged. They argue that in order to resolve the processes that are relevant at the applied spatial resolution, the calibration and validation time interval should keep pace with the increase in spatial resolution. It would hence be beneficial to investigate the model performance of LISFLOOD when calibration is performed on both higher spatial and temporal resolution.

In addition to the aforementioned trend, there has been as well an increase in the parameterization complexity of land surface models (LSMs) over the past decades, whereas the procedure to estimate the effective parameters required for the parameterization remained unchanged (Samaniego *et al.* 2017). They argue that state-of-the-art LSMs are able to describe comprehensively the exchange of energy and matter between the atmosphere, the vegetation, and the land surface but they still use simple pedotransfer functions to estimate fundamental soil properties. Therefore, the focus should not only be on moving towards hyperresolution but also on the development of fine-scale process parameterizations (Beven and Cloke 2012, Beven *et al.* 2015). Even if the spatial model resolution is increased, the issue of sub-grid variability might not be resolved since many of the relevant processes take place on even smaller scales (Melsen *et al.* 2016a). Hence, "the challenge is not hyperresolution but an appropriate scale-dependent sub-grid parameterization that recognizes the epistemic uncertainties in knowing and representing the characteristics of a grid element, and the nonlinearity and hysteresis in its dynamics" (Beven and Cloke 2012, p.1).

4.2.3. Run times

Counting the amount of raster pixels from the two spatial domains reveals that the 30" resolution exhibits 43 times the amount of pixels of the 5 km domain. The total amount of raster pixels in the entire Elbe catchment at 30" spatial resolution is 250,072 and 5,735 at 5 km spatial resolution. Table 4-4 indicates that the run times from the 5 km setup are almost unaffected by the size of the sub basin domain. By moving towards higher spatial resolution and hence increased amounts of pixels in the sub basin domain, a dependency emerges between the model run time and the size of the model domain. Station 7 exhibits the largest sub basin but its model run time is shorter than compared with station 172 for both model setups (5 km and 30") (Table 4-4). This might be attributed to the fact that station 172 is the most downstream sub basin and receives additional discharge inflows from the stations 169 and 349 which translates in greater computational costs. Further analyses are needed to investigate the relationship between the spatial model resolution, the size of the model domain and the model run time.

5 Conclusions

This study investigated the benefits and affiliated costs of increasing the spatial resolution of LISFLOOD from 5 km to 30" or 0.00833° (~1 km at the Equator) based on a case study in the Elbe basin, Germany. There is a general expectation that moving towards spatial hyperresolution will allow for a better representation of the heterogeneity in topography, soils, and vegetation and their effects on hydrological dynamics. The high resolution grid is expected to allow for a better discretization in particular for locations with small upstream drainage areas.

The comparison between the 30" and 5 km resolution setup was done for an un-calibrated and calibrated model scenario. Until now, the 30" setup neglects sub-grid information on channel length and channel gradient and hence, for comparison reasons, the 5 km setup was employed the same way. Nevertheless, to investigate the model performance from the optimal native 5 km LISFLOOD setup, the sub-grid information on channel length and channel gradient was included in the 5 km^{++} setup. In case LISFLOOD run un-calibrated, the results suggested that there is a drop in model performance if sub-grid information on channel length and gradient is neglected. Even the 5 km setup performed slightly better regarding the Kling-Gupta Efficiency than the newly 30" setup for all validation stations. This could be due to the used standard parameter set which was more or less manually tuned for the 5 km LISFLOOD setup, the resampling of certain static maps from the native 5 km to the 30" grid, the omission of increasing the temporal scale, or the large-scale model parameterization which does not account for small-scale processes. The calibrated model outputs gave a different picture where the 5 km and 5 km^{++} setups showed better agreements to the observations in the downstream part, while the 30" data set yielded an improvement for the upstream area. Because the parameterization of LISFLOOD was originally designed for large-scale catchment processes the impacts of spatial resolution become less significant for locations with large upstream drainage areas.

The newly derived 30" data set is by no means complete and ideal. Static maps such as land use and soil characteristics have to be derived from high resolution sources in order to include more details in the 30" grid. Because the spatial and temporal scales are linked, future studies should also focus on increasing the temporal resolution of LISFLOOD, also for model calibration. Furthermore, new parameterization techniques should be developed to account for small-scale processes which take on greater significance while moving towards hyperresolution.

References

Alfieri, L., Burek, P., Dutra, E., Krzeminski, B., Muraro, D., Thielen, J., & Pappenberger, F. (2013). GloFAS-global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, *17*(3), 1161.

Arora, V., Seglenieks, F., Kouwen, N., & Soulis, E. (2001). Scaling aspects of river flow routing. Hydrological processes, 15(3), 461-477.

Bartholmes, J. C., Thielen, J., Ramos, M. H., & Gentilini, S. (2009). The european flood alert system EFAS–Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrology and Earth System Sciences*, *13*(2), 141-153.

Beven, K. J., & Cloke, H. L. (2012). Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al. Water Resources Research, 48(1).

Beven, K., Cloke, H., Pappenberger, F., Lamb, R., & Hunter, N. (2015). Hyperresolution information and hyperresolution ignorance in modelling the hydrology of the land surface. *Science China Earth Sciences*, *58*(1), 25-35.

Burek, P., van der Knijff, J., & de Roo, A. (2013a). LISFLOOD – Distributed Water Balance and Flood Simulation Model. JRC technical reports. Revised user manual. Available under [https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/lisflood-distributed-water-balance-and-flood-simulation-model-revised-user-manual-2013, Stand: 10.03.2018]

Burek, P., van der Knijff, J., & Ntegeka. V. (2013b). LISVAP – Evaporation pre-processor for the LISFLOOD Water Balance and Flood Simulation Model. JRC technical reports. Revised user manual. Available under [https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/lisvap-evaporation-pre-processor-lisflood-water-balance-and-flood-simulation-model, Stand: 10.03.2018]

Burek, P., Bianchi, A., & Gentile, A. (2014). A Pan-European Data Set for hydrological modelling. JRC technical reports. Description of the input layers for the hydrological LISFLOOD model. Draft.

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). Applied hydrology, 572 pp. Editions McGraw-Hill, New York.

Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. A. M. T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE transactions on evolutionary computation, 6(2), 182-197.

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. Journal of Hydrology, 377(1-2), 80-91.

Haddeland, I., Matheussen, B. V., & Lettenmaier, D. P. (2002). Influence of spatial resolution on simulated streamflow in a macroscale hydrologic model. Water Resources Research, 38(7).

Kumar, R., Samaniego, L., & Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. Water Resources Research, 49(1), 360-379.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union*, *89*(10), 93-94.

Li, H., Wigmosta, M. S., Wu, H., Huang, M., Ke, Y., Coleman, A. M., & Leung, L. R. (2013). A physically based runoff routing model for land surface and earth system models. Journal of Hydrometeorology, 14(3), 808-828.

Mateo, C. M. R., Yamazaki, D., Kim, H., Champathong, A., Vaze, J., & Oki, T. (2017). Impacts of spatial resolution and representation of flow connectivity on large-scale simulation of floods. Hydrology and Earth System Sciences, 21(10), 5143-5163.

Melsen, L. A., Teuling, A. J., Torfs, P. J., Uijlenhoet, R., Mizukami, N., & Clark, M. P. (2016a). HESS Opinions: The need for process-based evaluation of large-domain hyper-resolution models. Hydrology and Earth System Sciences, 20(3), 1069-1079.

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Clark, M., & Uijlenhoet, R. (2016b). Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin. Hydrology and Earth System Sciences, 20(6), 2207-2226.

Pistocchi, A., & Pennington, D. (2006). European hydraulic geometries for continental SCALE environmental modelling. *Journal of Hydrology*, *329*(3-4), 553-567.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., ... & Attinger, S. (2017). Toward seamless hydrologic predictions across spatial scales. Hydrology and Earth System Sciences, 21(9), 4323-4346.

Schumann, G., Bates, P. D., Apel, H. & Aronica, G. T. (2018). Global Flood Hazard Mapping, Modeling, and Forecasting. In Global Flood Hazard (eds G. J. Schumann, P. D. Bates, H. Apel and G. T. Aronica).

Snyder, J. P. Map Projections: A Working Manual; USGS series, Professional Paper 1395; Geological Survey (US), 1987.

Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., ... & Jong, K. (2017). PCR-GLOBWB 2: a 5 arc-minute global hydrological and water resources model. Geosci. Model Dev. Discuss, 2017, 1-41.

Thielen, J., Bartholmes, J., Ramos, M. H., & Roo, A. D. (2009). The European flood alert system–part 1: concept and development. Hydrology and Earth System Sciences, 13(2), 125-140.

Todini, E. (1996). The ARNO rainfall—runoff model. Journal of hydrology, 175(1-4), 339-382.

Troy, T. J., Wood, E. F., & Sheffield, J. (2008). An efficient calibration method for continental-scale land surface modeling. Water Resources Research, 44(9).

van Beek, L. P. H., & Bierkens, M. F. P. (2009). The global hydrological model PCRGLOBWB: conceptualization, parameterization and verification. Technical report, Department of Physical Geography, Utrecht University, The Netherlands.

van der Knijff, J. M., Younis, J., & De Roo, A. P. J. (2010). LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, *24*(2), 189-212.

Vázquez, R. F., Feyen, L., Feyen, J., & Refsgaard, J. C. (2002). Effect of grid size on effective parameters and model performance of the MIKE-SHE code. Hydrological processes, 16(2), 355-372.

Wesseling, C. G., Karssenberg, D. J., Burrough, P. A., & van Deursen, W. P. (1996). Integrating dynamic environmental models in GIS: the development of a Dynamic Modelling language. *Transactions in GIS*, *1*(1), 40-48.

Willmott, C. J., Rowe, C. M., & Philpot, W. D. (1985). Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *The American Cartographer*, *12*(1), 5-16.

Wood, E. F., Roundy, J. K., Troy, T. J., Van Beek, L. P. H., Bierkens, M. F., Blyth, E., ... & Gochis, D. (2011). Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resources Research*, *47*(5).

Zhao, R.J., & Liu, X.R., (1995). The Xinanjiang model. In: Singh, V.P. (ed.), *Computer Models of Watershed Hydrology*, 215-232.