MSc Thesis

Uncertainty in global hydrological modeling related to meteorological forcing and river routing characteristics.

By Lex van Dolderen

Utrecht University
Department of Physical Geography
Geographical Hydrology

First corrector: Dr. Rens van Beek
Second corrector: Prof. Marc Bierkens

Contents

Abstract

Global hydrological models (GHM) can be used in studies on a range of topics related to land surface hydrology. With growing computational power and availability of better input data, efforts have to be made to improve such models in order to use their full potential. Two factors that determine the quality of GHM simulations are meteorological forcing and parameterization of flow characteristics. Question is how model results respond to changes in these features.

In the first part of this study two sources of meteorological input are compared, i.e. the reanalysis products ERA-Interim and MERRA, using a reference dataset (CRU TS2.1). Both the actual variable fields, like precipitation, and GHM output based on the products are analyzed. The model that is used is PCR-GLOBWB. It solves the land surface water balance and has a river routing module that simulates discharge and flooding. One of the objectives is to find out which of the reanalysis products is most suitable for further use in this study and in other future research. This turns out to be ERA-Interim, mainly because MERRA suffers from a problem with cloud cover in tropical regions.

The second part of the study consists of a sensitivity analysis, in which the influence of changes in channel and floodplain properties on river discharge and flooding is investigated. PCR-GLOBWB is run with 36 different sets of parameters. The resulting discharge and flood extent about the possibilities of improving the model by adjusting them. Comparing the influence of forcing and parameter variations on model performance will tell something about what the focus should be on in future studies on model improvements.

1. Introduction

A global hydrological model (GHM) can serve as a tool for research in hydrology, water management and other scientific fields related to water. Examples of applications of this type of modeling are assessments of continental runoff (Nijssen et al., 2001; Fekete et al., 2005), flood risk (Winsemius et al., 2013) and water availability (Wada et al., 2011). There are several well-known GHMs built by research groups throughout the world, e.g. VIC (Liang et al., 1994), MAC-PDM (Arnell, 1999), WATERGAP (Döll et al., 2003), WASMOD-M (Widen-Nilsson, 2007). Some of these models focus purely on the water balance, while others combine it with the land surface energy balance. Some models also include a river routing module to simulate river discharge and flooding. Apart from studies on a global scale, GHMs can be very useful in studies at basin scale, especially when it comes to river discharge and flooding. For certain areas in the world it is often not feasible to make higher resolution models for case studies. For example, simulations can be used by governments, re-insurance companies or NGOs to determine flood risk and making policy on reducing it. Considering the potential of GHMs for this purpose it is important that a model produces reliable and accurate data, both quantitatively and qualitatively. Biases in model results should be acceptable and variability within and between simulations has to be realistic.

When looking at the processes and input of a GHM, factors that determine the quality of simulations can be divided in a few groups. One of the fundamentals is meteorological forcing. Changes in variables like the amount of precipitation that contributes to runoff or temperature that controls the release of melt water, have immediate consequences for the output of the model. Another important factor is the parameterization of a GHM. Depending on the complexity of the model and its purposes a number of parameters need to be set. Examples are vegetation types, which determines interception and evapotranspiration, or soil properties for the calculation of infiltration and groundwater flow.

Global hydrological models are still a relatively new resource for hydrological computations and there is room for improvement in all facets. For instance, efforts are made to increase spatial resolution, which is enabled by growing computational power. Another area where progress is made is in the field of meteorological modeling, playing an important role being the source of forcing.

There are two main sources of meteorological input, depending on whether the interest is in past or future hydrology. GCMs (General Circulation Models) are generally used in studies on future hydrology and climate change. This type of models simulates atmospheric and oceanic water and energy transport based on thermodynamics and the rotation of the planet, producing a variety of variable fields. Reanalysis products on the other hand are primarily used for meteorological variables from the past. Reanalysis consist of modeled data, similar to GCMs, but assimilated with observations. This way their data are more reliable, with a better chance of realistic reproduction of hydrology. In this study two reanalysis products are compared. This is done by comparing the actual variables, as well as the influence they have on GHM results.

When it comes to the parameterization the focus lies on parameters that are directly related to river flow. After local runoff amounts are calculated, channel and floodplain characteristics determine the speed with which the runoff is discharged downstream and how high water levels rise. The parameters that represent these features are also related to river regulation. Additionally, this makes them interesting from the perspective of designing measures, e.g. for flood reduction. Human influences on river flow consist of changing embankment levels, altering hydraulic properties of both channels and floodplains, canalizing and building dams to create reservoirs (Vörösmarty and Sahagian, 2000). Where the latter two are local phenomena that take adaptations in the drainage network to represent them, friction and embankment levels can be altered on a global scale in a general way.

The influence of changes in the flow characteristics is determined by means of a sensitivity analysis, varying three parameters, i.e. channel friction, floodplain friction and channel depth. The analysis is done by running the GHM PCR-GLOBWB (PCRaster Global Water Balance, van Beek and Bierkens (2008)) with different parameter combinations. The meteorological forcing is retrieved from the reanalysis that gives the best results in the first part of this study. Finally, the analyses of the influence of the meteorological forcing and the parameterization can be compared to see where the focus of future model improvements should lie.

The main research questions of this study are:

- Which reanalysis product, MERRA or ERA-Interim, is best for future use in hydrological studies?
    - Which reanalysis compares best with climate data?
    - Which reanalysis gives the most realistic discharge when used as GHM input?
- How does river regulation influence discharge and flooding in global hydrological modeling?
    - How do discharge patterns respond to changes in parameters related to river regulation?
    - How does river flooding respond to changes in parameters related to river regulation?
    - Is there a parameterization that improves the overall modeling performance in discharge reproduction compared to the current standard settings?
- How do the influences of meteorological forcing and parameterization compare?
    - On which of the facets should be focused to make improvements in global hydrological modeling?

2. Literature review and dataset description

2.1 Similar studies

Looking at earlier research that is similar to this, the assessment that Candogan Yossef et al. (2011) carried out is very interesting. They studied the modeling skill of PCR-GLOBWB in reproducing discharge extremes for 20 large river basins, looking both at hydrographs and floodings. Although there were biases in the results the overall skill was satisfactory.

Within the framework of her PhD, Sperna Weiland (2011) did research on a variety of topics related to global hydrological modeling. One of the studies is very similar to this, as it also covers uncertainties in model forcing and parameterization. The inputs that are considered and the method of analysis is however somewhat different. Amongst others a larger set of parameters is used to make variations. The study shows that the parameterization has only minor influence on the hydrographs that the model produces. None of the applied parameter sets resulted in consistently improved simulations with different forcing datasets and for all river basins.

2.2 Reanalysis products

Reanalysis is a scientific method that combines numerical modeling and observations in order to create a dataset of weather and climate variables that covers a great area, up to the global scale. The observations often have various sources, for example weather balloons, airplanes, ships, ground-based stations and satellites. Reanalysis products contain a wide variation in variables, from temperature, humidity, wind speed and air pressure at different altitudes to albedo, snow characteristics, cloud cover and incoming and outgoing radiation. There are several well-known reanalysis products developed in the last couple of decades. The most commonly used products are: NCEP/NCAR (Kalnay et al., 1996), ERA-40 (Uppala et al., 2005), JRA-25 (Onogi et al., 2007), ERA-Interim (Dee et al., 2011) and MERRA (Rienecker et al., 2011). The reanalysis products that are used in this study are ERA-40, ERA-Interim and MERRA. In the next sections these three products are described, mainly focusing on the latter two, as one of these is intended to replace ERA-40.

ERA-40

The ERA-40 reanalysis product is developed by the European Centre for Medium-range Weather Forecasts (ECMWF) and available for the period 1957-2002 on 0.5° spatial resolution. It has been used as input for PCR-GLOBWB in combination with climatological datasets CRU TS2.1 (New et al., 2000, 2002; Mitchell and Jones, 2005) and CRU CLIM 1.0 (New et al., 1999). These are used for additional correction of the reanalysis. The CRU products have the same spatial resolutions as ERA-40. Their monthly fields are downscaled temporally to daily resolution. Other examples of ERA-40 related research are on the subject of crop yield (Challinor et al., 2005), drought variability (Bordi et al., 2006) and soil moisture fluctuations (Wang et al., 2010).

ERA-Interim

The ERA-Interim reanalysis product (Dee et al., 2011) is the interim product between ERA-40 and a future reanalysis that will span the entire twentieth century (ERA-CLIM). The dataset covers the period from 1979 and is updated with a lag of about a month. There have been some improvements on the data assimilation and observation with respect to ERA-40. In the paper of Dee et al. (2011) the average daily precipitation of ERA-Interim and ERA-40 are compared. The most pronounced differences between the two is the higher precipitation on Europe (1-2 mm/day), the eastern part of the USA (ca. 1 mm/day), Central Africa (ca. 2 mm/day) and some spots in the northern part of South-America (up to ca. 6 mm/day). Of course there are locations where ERA-Interim precipitation is smaller than that of ERA-40, but the only significant difference is located over the Pacific Ocean, which is not of interest to this study.

ERA-Interim has been available for a relatively short period, so the number of publications on it is little. Betts et al. (2009) made a comparison between ERA-Interim, ERA-40 and observations for three major river basins, being the Amazon, Mississippi and Mackenzie, mainly looking at the effect of cloud cover. They conclude that ERA-Interim is an improvement for these three catchments, because of the new humidity analysis and assimilation system.

Szczypta et al. (2010) have investigated the performance of ERA-Interim for France and compared parameters like precipitation, temperature and air humidity to those of a high-resolution product called SAFRAN (Durand et al., 1993). They show that ERA-Interim underestimates precipitation over France by 26%. It is interesting to see if the results of this study are similar.

Balsamo et al. (2010) validated the precipitation of ERA-Interim over the USA by comparing it to different precipitation datasets, but primarily to PRISM (USDA). The correlation of monthly average values of precipitation turned out to be high (0.85) and even higher (0.9) when rescaled with the GPCP dataset. On the downside, the rescaling makes the average annual bias to PRISM increase from -0.013 mm/day to +0.101 mm/day. The correlation of daily precipitation amounts is 0.560 (0.575 for rescaled values).

MERRA

The Modern-Era Retrospective Analysis for research and Application (MERRA) (Rienecker et al., 2011) has been developed by NASA's Global Modelling and Assimilation Office. This reanalysis spans the same period as ERA-Interim, 1979-present and has a spatial resolution of 0.5° latitude x 0.67° longitude. The paper of Rienecker et al. (2011) gives an introduction to the model and compares it to ERA-Interim. The data that is used in the assimilation process is almost the same for both products. The differences between the two are mainly caused by the use of different models and different assimilation methods. The paper also mentions that changes in observing systems have a negative effect on the quality and consistency of reanalyses. If these changes are not processed in the right way, sudden jumps or trends in climate variables may occur in a dataset that are not actually happening. Especially precipitation, observed by different satellite systems, is vulnerable to these fluctuations. This phenomenon is mentioned as one of the most important points that should be improved in future reanalysis products.

Reichle et al. (2011) developed a supplemental and improved set of land surface hydrological fields called MERRA-Land. GPCP-corrected precipitation data are used to rerun the land component of the MERRA model. The skill of MERRA-land in reproducing fields of soil moisture, runoff and snow cover is compared with that of MERRA and ERA-Interim for nine large river basins. Overall, MERRA-Land turns out to be an improvement compared to MERRA, although this is not always significant for every parameter and for every basin.

Bosilovich et al. (2011) describe the quantification of the global energy and water budget in MERRA. Especially the analysis of precipitation in this paper is of great interest to this study. Annual average fields of precipitation from MERRA (and other reanalysis products) are compared to the GPCP dataset. There are several regions that show deviating values, most prominently South-America and Central Africa. The precipitation of these regions is underestimated significantly in MERRA compared to both GPCP and other reanalyses, including ERA-40 and ERA-Interim. They also looked at the effect of changing observing systems, in this case the introduction of the Advanced Microwave Sounding Unit (AMSU) where before that the Special Sensor Microwave/Imager (SSM/I) was used for observing, amongst others cloud liquid water and precipitation. This introduction causes a significant difference in annual average precipitation and evaporation, both positive and negative, between the first decade of this century and the period before that, for several regions in the world.

3. Methodolgy

3.1 Comparison of meteorological forcing data

A comparison of the two products has to determine which one gives the best results. The ultimate way of testing this is to put the reanalysis data in a hydrological model, in this case PCR-GLOBWB, and look at the performance with respect to observations. Before the hydrological model is run with the reanalysis data, monthly fields of precipitation, potential evapotranspiration and temperature are analyzed. The results of this analysis give a first view on what can be expected when the data is used in a model run. The CRU dataset is used as reference, as it is considered to be a reliable representation of reality. The analysis is done both in a quantitative and a qualitative way in terms of respectively bias and correlation. The bias is calculated by averaging all the differences between the fields of the reanalysis products and CRU, resulting in a shortage or excess in meter per day for precipitation and evapotranspiration and degrees Celsius for temperature for each cell of the grid. The correlation coefficient between the two reanalyses and CRU is only calculated for precipitation, being the driving variable for hydrological modeling. Same as bias, it is determined for the entire global land surface and because monthly values are used it tells something about the similarity of seasonal patterns in precipitation.

PCR-GLOBWB

The PCR-GLOBWB model is used in the remainder of the study, starting of with the comparison of ERA-Interim and MERRA. The model has been set up by Van Beek and Bierkens (2008), with the intention to build a model taking into account as many hydrological processes as possible and that can be used for all kinds of assessments where other GHMs often have a limited number of applications. As the name already indicates PCR-GLOBWB focuses on the terrestrial water cycle and does not regard the land surface energy balance. The model has a spatial resolution of 0.5°x0.5° and is run with daily time steps. Land surface properties, like vegetation cover and soil type, are embedded in the model by means of sub-grid variability. In the first place, PCR-GLOBWB calculates the storage of water in and transport between several hydrological compartments, being two soil layers, a groundwater layer, snow cover, glaciers, canopy interception and open water. As mentioned before, the input of the model consists of a time series of three meteorological fields, i.e. precipitation, potential evapotranspiration and temperature. In most of the earlier research using PCR-GLOBWB the CRU TS 2.1 monthly data set (Mitchell and Jones, 2005) was used, downscaled to daily fields with ERA-40 reanalysis. These datasets need to be replaced as ERA-40 covers the period 1957-2002 and is no longer updated. The new generation reanalyses ERA-Interim and MERRA are candidates to take over from ERA-40.

For each time step of the model run, the water balance is solved for each cell in the grid, resulting in a certain amount of runoff. PCR-GLOBWB contains a routing module that calculates river discharge from runoff. With the use of the kinematic wave approximation, based on Manning's surface roughness coefficient, the development of discharge along a global drainage network is simulated. Lakes, reservoirs and floodplains are included in the model, enabling the effects of river regulation and flooding. The primary modeling result is a series of maps, one for each time step, with discharge on a 0.5° resolution grid. The combination of these maps gives the hydrograph for each grid cell for the period of interest. Simulated flooding results are represented by two values, i.e. the fraction of the cell area that is flooded and the flood depth.

Forcing PCR-GLOBWB

With the results of the comparison of the reanalysis data in mind, PCR-GLOBWB is run with both meteorological forcing sets and the CRU control set. The focus of this part of the study is fully on the resulting discharge, because it is in this context the most significant feature of land surface hydrology. Again, monthly discharge data is used, which is a small enough temporal scale to determine the basic performance of the model forced with the two different reanalysis datasets and large enough not to use the routing module, which saves computation time.

The reference data that is used for the analysis of the discharge data comes from the long-term inventory Global Runoff Data Centre (GRDC). It consists of 3561 files containing a summary of the monthly discharge measured by stations all over the world, for varying periods throughout the last century. For each calendar month minimum, mean and maximum discharge values of the covered period are provided and additionally the same statistics are covered for each year. Combining these gives a generalized hydrograph that can be adjusted according to the yearly conditions. As the distribution of stations over the globe is not uniform and there is a lot of overlap within river basins, a selection of 65 stations is made that is a fair representation of the overall global river runoff. A list of these stations is included in appendix A.

The simulated and observed river runoff are compared, again looking at both bias and correlation. In the quantitative analysis average discharge and generated runoff per unit area are used. The second is calculated by dividing the average discharge by the area upstream of the station. This makes that any error in area between the model and the observation data is compensated for and the value no longer depends on the size of the basin. To get a good overview of the differences between the results of the two model runs, both discharge and generated runoff are plotted in regression charts. Finally, the correlation coefficient is calculated to determine the resemblance between the simulated and the observed discharge signal.

In advance of the analysis it is hard to determine what criteria will be decisive in the process of picking one of the reanalysis products to be the future standard input. There may be regional differences that do well for one and bad for the other and vice versa. Only in case of significant differences in quality the choice will be easy, otherwise all pros and cons need to be considered.

3.2 Free flooding versus no flooding

When the new standard meteorological forcing is chosen the scope of the study shifts towards the main interest, i.e. the influence of river regulation on discharge and flooding. First, the model is run twice with different routing settings, one run with flooding enabled and one in which all discharge stays within the river channels, respectively representing a natural situation and a situation with extreme regulation with (infinitely high) embankments. This is done in an attempt to make a provisional distinction between regulated and non-regulated rivers.

3.3 Scenarios and analysis

In the no-flooding scenario the retentive function of the floodplain is disabled, which makes that flood peaks are better preserved while moving downstream and average flow velocities are higher compared to the first situation. Therefore water levels rise earlier and faster downstream of a rainfall event and last for a shorter period. To rule out any effects of the extremer discharge peaks on calculated model efficiency, 5-day average discharge values are analyzed along with daily values.

Overall the average discharge should be slightly higher for the no-flooding scenario, as the open water surface area is smaller which causes a decrease in evaporation. Although the no-flooding scenario is quite extreme, it is expected that some rivers perform better under these circumstances, especially those which are located in wealthy, densely populated areas that have high levels of river regulation. It is also possible that an improvement in performance is not exclusively related to river regulation, for instance if other parameters such as tortuosity or friction deviate from reality. It is therefore important to look at the analysis results critically.

The same selection of 65 stations is used in this analysis. The focus is now more on the resemblance between the simulated and observed hydrograph and less on the quantity of discharge, as this is mainly a result of the meteorological forcing of the model. Next to the correlation the Nash-Sutcliffe model efficiency coefficient (NSE) is calculated. It is a measure similar to the correlation, but more suitable to use in the comparison of hydrographs. It uses the following equation:

$$NSE = 1 - \frac{\sum_{t=1}^{T} \left(Q_o^t - Q_s^t\right)^2}{\sum_{t=1}^{T} \left(Q_o^t - \overline{Q_o}\right)^2},$$

where Q^t_o and Q^t_s are respectively observed and simulated discharge at a given time step *t* in range [1,T]. An NSE of 1 indicates perfect resemblance between the observed and simulated hydrograph, NSE=0 means the model is as good a reproduction as the horizontal line based on the average of all observations and negative NSE means it is worse than that.

A disadvantage of the NSE measure is that it is very sensitive to bias and as the model outcome is biased to a considerable degree for most of the stations, this makes the NSE less applicable. To deal with this problem the NSE is calculated twice for each station, once with and once without a bias correction. The bias corrected discharge $Q^t_{s,b}$ is calculated in the following way:

$$Q^t_{s,b} = \frac{Q^t_s}{\sum_{t=1}^{T} Q^t_s / \sum_{t=1}^{T} Q^t_o}.$$

The bias correction can be justified for the same reason it is unnecessary to do a quantitative analysis in this part of the study. The combination of both NSE values gives a meaningful insight in the performance of the model and the influence that errors in the meteorological input have on it.

With the results of this analysis a new selection of several basins is made for the final part of the study. Having in mind the purpose of this last part, there are three criteria that the selection and the basins in it should meet:

1) There is room for improvement in the performance of the model in the basins, i.e. NSE is not too close to 1.
2) The bias is not too large, here a threshold of -/+30% is set such that most of the effects of major quantitative flaws in the meteorological forcing are ruled out.
3) The selection as a whole should be a good representation of all global river basins, meaning spread over the globe and containing different types and sizes.

In order to fulfill the third criterion the other two may be overruled.

3.4 Parameter variations

In the last part of the study the influence of river regulation on discharge and flooding is looked at in more detail. The routing section of the model is run with different sets of parameters. The parameters that are changed are floodplain surface roughness, channel roughness and channel depth. In the coming chapter the variations and the expected consequences for the modeling results are discussed. Finally, the analysis of the discharge and flooding data is described.

Channel depth

One of the most tangible forms of river regulation are embankments. The hydrograph of a river is influenced by raising the critical discharge that a channel can transport before it floods. Apart from the local consequences of raising flood security levels, embankments cause discharge peaks to remain higher and arrive faster downstream.

As there is no global dataset for embankment levels, this feature is implemented in the PCR-GLOBWB model in a simplified manner. Instead of introducing a separate parameter for embankments, channel depth is adjusted to raise the critical discharge of a river. The standard channel depth used in the model is calculated with the use of the bankfull discharge. The channel depth adjustment is generalized by applying one value globally. This makes that the reference level is 0 meter additional channel depth. The alternative values that are chosen for the assessment are 1, 2.5 and 5 meter. It can be assumed that human operations only lead to an increase in embankment levels, so it is not necessary to take negative values into account.

Roughness parameters

The roughness is implemented into the model with two Manning's values (n_M), one for channels and one for floodplains. With respect to human influences on river flow, floodplain roughness is more subject to changes than channel roughness. Floodplain roughness is mainly determined by vegetation, which properties can be affected by cutting or (re)planting. Channel roughness is mostly derived from small scale morphology and to a limited extent from composition of the river bed material and vegetation. Channel roughness can be influenced by dredging and canalizing, in most cases leading to a lower friction.

Type of Channel and Description	Minimum	Normal	Maximum
Natural streams - minor streams (top width at floodstage < 100 ft)			
1. Main Channels			
a. clean, straight, full stage, no rifts or deep pools	0.025	0.030	0.033
b. same as above, but more stones and weeds	0.030	0.035	0.040
c. clean, winding, some pools and shoals	0.033	0.040	0.045
d. same as above, but some weeds and stones	0.035	0.045	0.050
e. same as above, lower stages, more ineffective slopes and sections	0.040	0.048	0.055
f. same as "d" with more stones	0.045	0.050	0.060
g. sluggish reaches, weedy, deep pools	0.050	0.070	0.080
h. very weedy reaches, deep pools, or floodways with heavy stand of timber and underbrush	0.075	0.100	0.150
2. Mountain streams, no vegetation in channel, banks usually steep, trees and brush along banks submerged at high stages			
a. bottom: gravels, cobbles, and few boulders	0.030	0.040	0.050
b. bottom: cobbles with large boulders	0.040	0.050	0.070
3. Floodplains			
a. Pasture, no brush			
1.short grass	0.025	0.030	0.035
2. high grass	0.030	0.035	0.050
b. Cultivated areas			
1. no crop	0.020	0.030	0.040
2. mature row crops	0.025	0.035	0.045
3. mature field crops	0.030	0.040	0.050
c. Brush			
1. scattered brush, heavy weeds	0.035	0.050	0.070
2. light brush and trees, in winter	0.035	0.050	0.060
3. light brush and trees, in summer	0.040	0.060	0.080
4. medium to dense brush, in winter	0.045	0.070	0.110
5. medium to dense brush, in summer	0.070	0.100	0.160
d. Trees			
1. dense willows, summer, straight	0.110	0.150	0.200
2. cleared land with tree stumps, no sprouts	0.030	0.040	0.050
3. same as above, but with heavy growth of sprouts	0.050	0.060	0.080
4. heavy stand of timber, a few down trees, little undergrowth, flood stage below branches	0.080	0.100	0.120
5. same as 4. with flood stage reaching branches	0.100	0.120	0.160

Table 1: Manning's hydraulic roughness values for channels and floodplains. Selection of relevant properties from (Chow, 1959)

A higher n_M reduces the velocity of water that flows over a surface, as there is more friction. A logical consequence of the deceleration is a rise of the water level, which can lead to more flooding. Chow (1959) made a classification of vegetation and surface properties and n_M values that apply to those situations Table 1 shows the classes that are relevant here. For the current standard parameterization n_M for floodplains is 0.1, which belongs to the class *"medium to dense brush, in summer"*, and for channels it is 0.04, corresponding to *"clean, winding, some pools and shoals"*. These are both moderate classes, assumed to be a good representation of the average situation. For the purpose of this study two values are chosen for each parameter, one lower and one higher than the current value. In case of floodplain roughness these are 0.05 and 0.15 and for channels 0.02 and 0.08. These values are chosen such that the difference is large enough to have significant influence on the modeled hydrographs, while staying within the boundaries of the Chow classification.

Analysis of scenarios

The additional channel depth and roughness values (table 2) give a set of 4x3x3=36 scenarios, including the current standard parameterization. The routing module of the PCR-GLOBWB model is run for these scenarios for the period 1979-2010, giving daily maps with discharge and flooding data that can be analyzed. The differences between the outcomes of these scenarios are expected to deviate from one another enough to make a distinction between the effects of the parameter changes.

Parameter	**Default**	**Var 1**	**Var 2**	**Var 3**
Added channel depth (m)	0	1	2.5	5
Channel friction (-)	0.04	0.02	0.08	-
Floodplain friction (-)	0.10	0.05	0.15	-

Table 2: Overview of parameters and the values that will be used in the different scenarios.

For the analysis of discharge the bias corrected NSE (NSEb) is the leading measure. It quantifies the performance of the model with respect to the timing of discharge peaks and -to a lesser extent- the steepness of rises and falls in discharge. The NSEb is calculated for each available station over the whole period of interest and plotted per selected river basin. These plots give an overview of which scenarios do best for each station and whether there are differences between stations within the same basin. For instance, the simulation performance of one station may benefit from a high channel roughness, while a station upstream along the same river does better with low friction. Such differences would give useful insight in how the model responds to changes in parameter settings and in what way the model could be improved, also looking at parameters other than those used in this study. To further visualize the effects of the parameter changes, the NSEb values of the main (most downstream) stations of each river basin are plotted.

The effect of the different parameterizations on flooding is analyzed by means of the flooded area. For each time step the model returns the fraction flooded area for each half degree cell. To get a measure that is representative for the flooding characteristics of a basin, the flooded area summed over the basin is divided by the total area. From these values five percentiles (20th, 40th, 60th, 80th and max) are retrieved for each basin/scenario combination, reducing the amount of data while still giving sufficient information on the distribution of flood extent.

4. Results and discussion

4.1 ERA-Interim vs. MERRA

The comparison of monthly variable fields of the two reanalysis products with the CRU climate dataset results in four sets of maps, three with the bias distribution (fig. 1, 3 and 4) and one with the correlation of the precipitation data (fig. 2). The bias maps of precipitation and potential evapotranspiration are scaled inversely, such that drier and wetter conditions are indicated by the same colors.

Precipitation bias and correlation

Starting with precipitation (fig. 1), both products show several regions with substantial biases with respect to the reference climate data. The map of ERA-Interim shows large regions where average rates are significantly higher, especially in Europe, South America and the southeast of North America. Of these places Europe has the highest relative bias with on average ~0.8 mm/day compared to observed rates of around 2 mm/day. This is contradictory to the results of the study of Szczypta et al., (2010), who found a 26% underestimation. Even though the reference datasets are different one would not expect such a significant difference.

When looking at Europe and the southeast of North America in the MERRA map, precipitation rates are much closer to the CRU values. For this reanalysis product the main problems lie around the equator, especially in South America. Along the northern coast of this continent precipitation is strongly overestimated, while inland the opposite is the case. Looking ahead at the other variables a similar pattern can be seen in the MERRA data. Both positive and negative biases in the variables in this region are caused by errors in the cloud cover (Bosilovich et al., 2011). Higher cloud cover lowers the amount of incoming shortwave radiation and leads to lower temperatures and that way also less evapotranspiration. On the other hand the reduction in outgoing long wave radiation increases precipitation. The combination of these effects leads to wetter conditions along the coast of South America, while the center of the continent becomes more arid due to the opposite effect. In Central Africa the same seems to be the case, although the inland effects are less explicit, and southeast Asia only has the wetter conditions. Apart from the regions along the equator MERRA precipitation does compare well quantitatively to CRU data. The biases in the Alaskan area and Greenland also occur in ERA-Interim so these may very well be improvements relative to CRU.

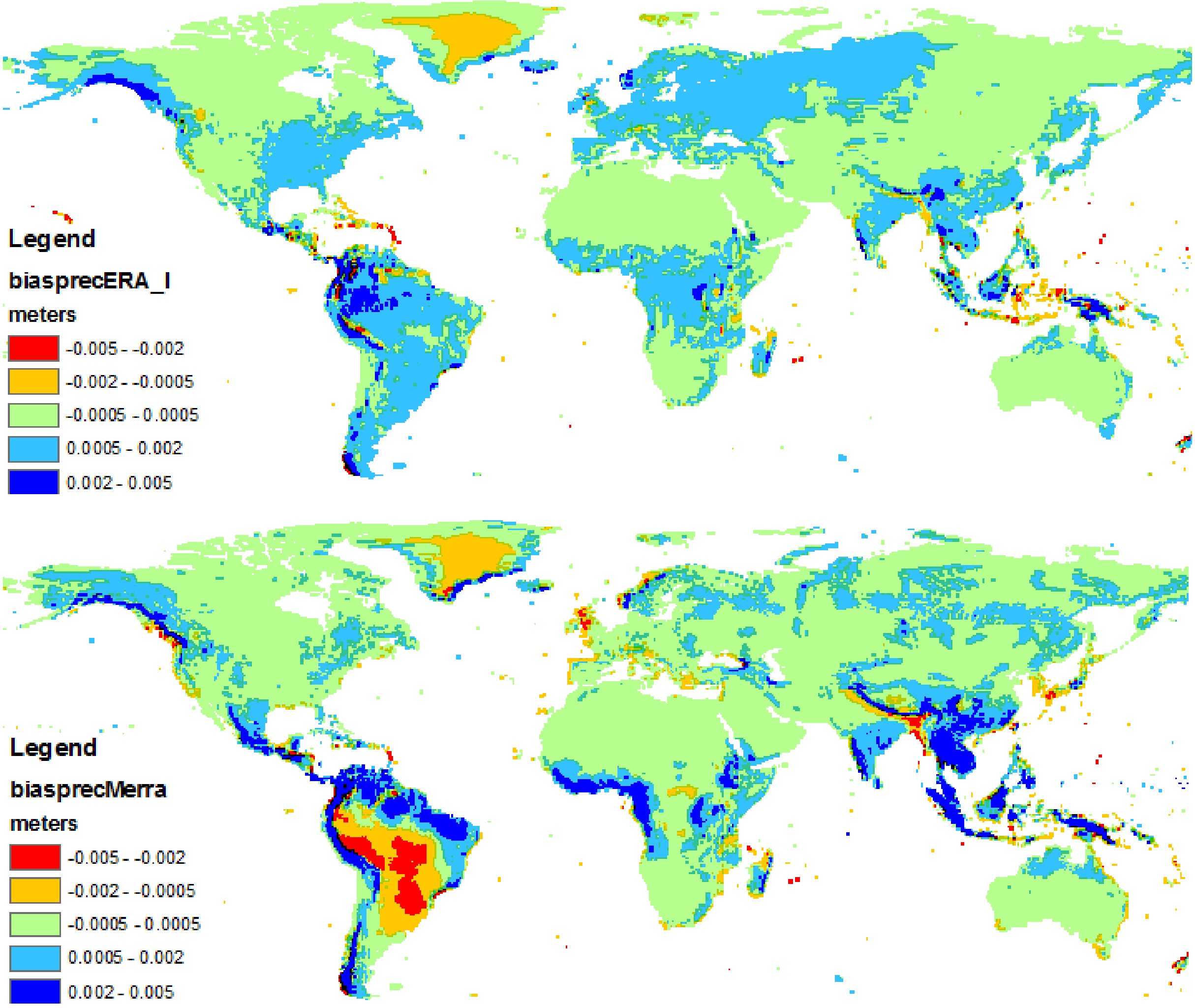


Figure 1: Precipitation biases (in m/day) of ERA-Interim (top) and MERRA (bottom) compared to CRU.

Next to the amount of produced runoff correct simulation of seasonality is essential in hydrological modeling. Therefore, the temporal signal in precipitation of the model input should be as close to reality as possible. To test this the correlation is calculated (fig. 2), again using the CRU dataset as reference. The resulting maps show broad resemblance when it comes to areas with very low correlation, which are concentrated in arid regions. The correlation in these regions responds relatively strong to deviations in precipitation, as the average value is low. Apart from these regions, ERA-Interim does better for apparently the whole Earth’s land surface, which is quite remarkable. Even in areas where

ERA-Interim precipitation is more biased, like Europe and the southeast of North America, it seems to perform better when it comes to seasonality.

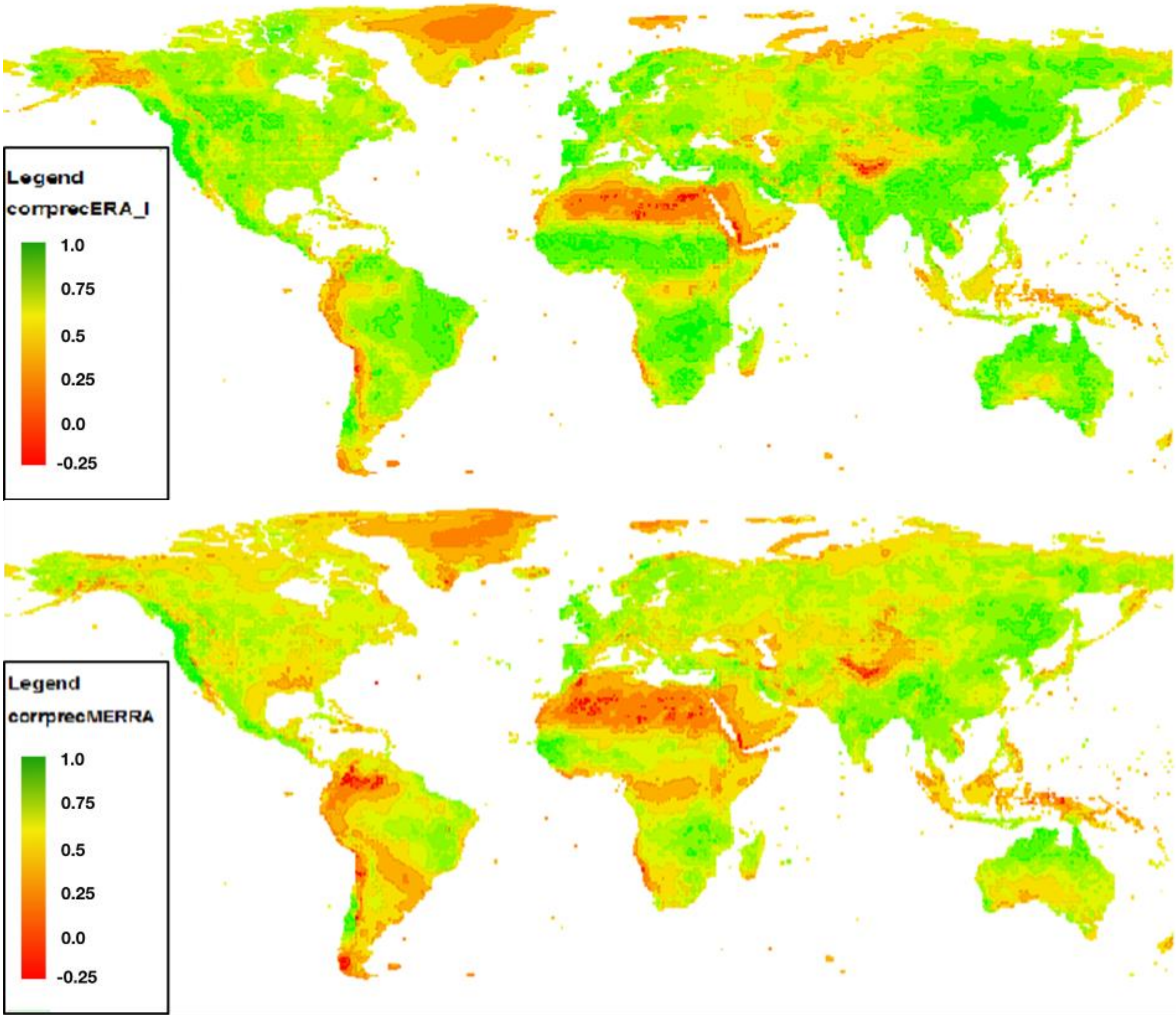


Figure 2: Correlation of precipitation between ERA-Interim (top) / MERRA (bottom) and CRU.

Potential evapotranspiration bias

Data on potential evapotranspiration is used as a limit for actual evapotranspiration, which is calculated by the hydrological model. Therefore, biases in this variable may not be directly translated to deviating modeling results for each region in the world, still for reliability it is important that values are realistic. It is also interesting to see whether effects of biases in potential evapotranspiration and precipitation either amplify or neutralize each other in relationship with the land surface water balance. For instance, the aforementioned problems of MERRA related to cloud cover lead to significant consequences for the amount of water that is available for runoff as both the input and output of the system make conditions either wetter or drier.

For both reanalyses the biases are concentrated in areas with extreme climates (fig. 3). In deserts potential evapotranspiration is often high, but as there is little water available the actual quantity is limited. This does therefore not harm the quality of the product much. Apart from the regions around the equator in the MERRA data and desert-like areas there are no spots in either map that show prominent deviation from the reference dataset. Overall, MERRA is more biased than ERA-Interim, but not such that it is decisive in the selection process.

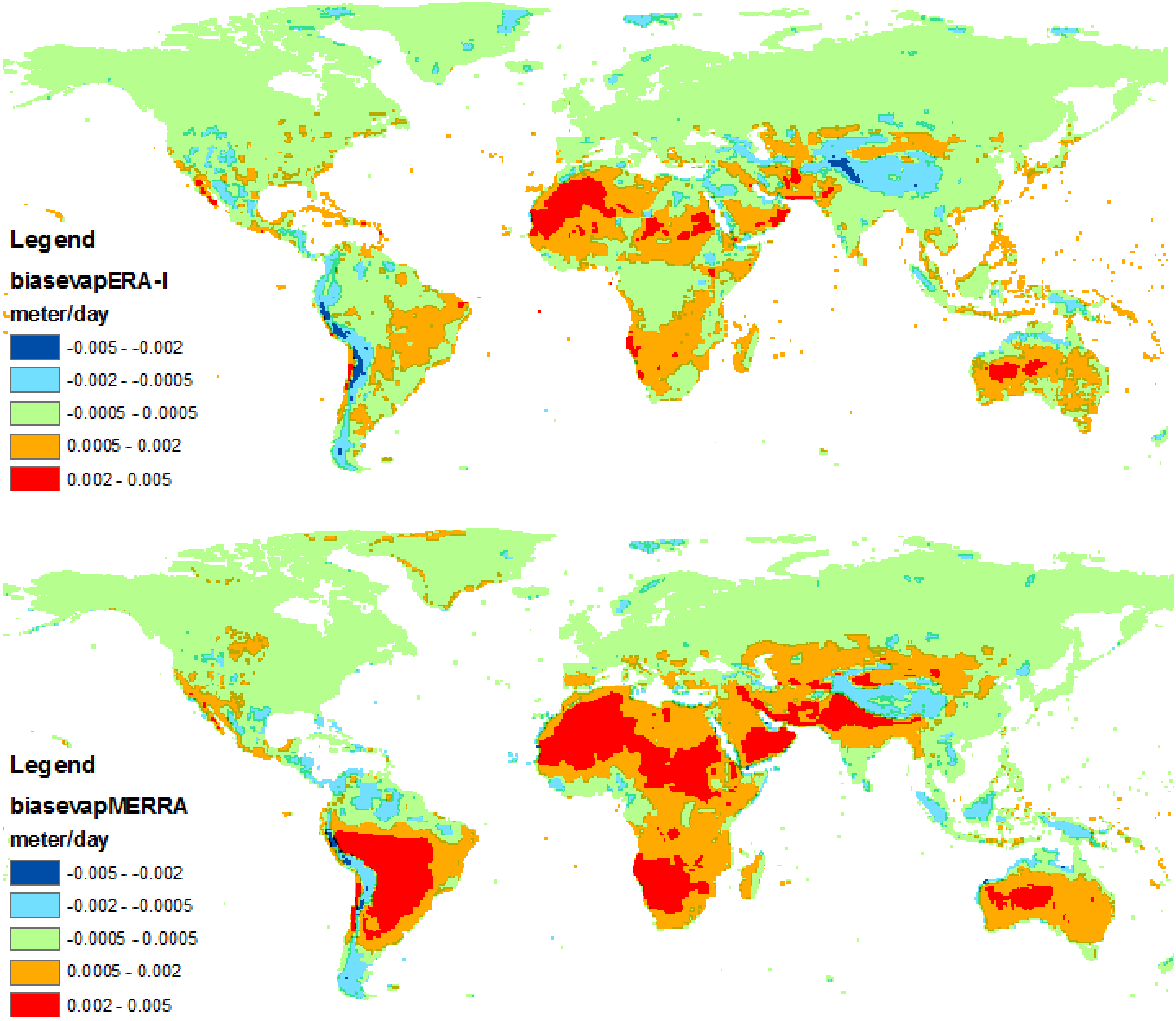


Figure 3: Biases in potential evapotranspiration (in m/day) of ERA-Interim (top) and MERRA (bottom) compared to CRU.

Temperature bias

Temperature is used to calculate actual evapotranspiration, which is limited by the potential values afterwards. In order to get correct runoff amounts it is thus important that this variable is realistic. At first sight, the maps of temperature bias of the two reanalyses are much more fragmented (fig. 4). The reason for this is that compared to the other variables temperature is much better recorded all over the world and therefore relies less on modeled fields and extrapolation of observations. Still, there are substantial temperature differences between the two reanalysis products and the climate data, while on the other hand there are quite some resemblances between the two. The major differences are located in South America and Africa, which can be related to the cloud cover problems of MERRA. Even when taking this out of consideration MERRA temperature is still more biased then that of ERA-Interim.

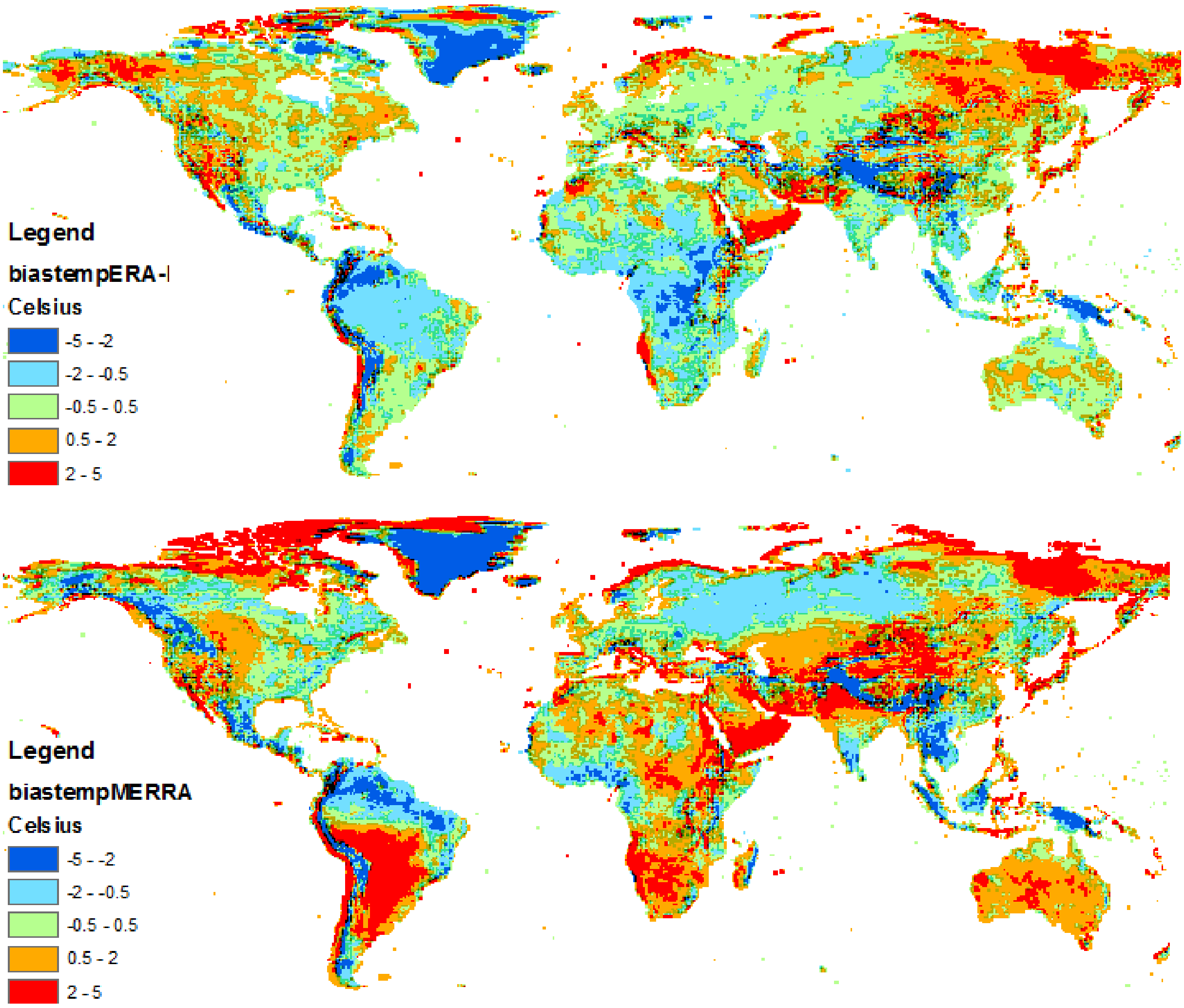


Figure 4: Temperature biases of ERA-Interim (top) and MERRA (bottom) compared to CRU.

4.2 Analysis of PCR-GLOBWB monthly discharge

From the analysis of the meteorological fields from the reanalyses it can be concluded that ERA-Interim shows more resemblances to the climate data than MERRA. This was however just a preliminary investigation that serves as material that can be used in the interpretation of the model simulation of global river discharge. Monthly values of modeled discharge are analyzed using the GRDC dataset.

Quantitative comparison of discharge

This part of the research starts by making a selection of observation stations from the GRDC dataset in order to downsize the amount of data. For each large river basin one station is picked, ending up with a total of 67 stations that give a good representation of global discharge. Figure 5 shows the locations of the selected stations and the ratio of simulated and observed discharge and in figure 6 the average values (in m^3/s) are plotted against one another. Additionally, figure 7 shows a similar plot, but here discharge is divided by the upstream catchment area. By scaling this to mm/year it can be compared to annual precipitation quantities and the data is sorted from arid to humid conditions. The combination of these figures gives an overview of how the model does quantitatively with meteorological input from the two reanalyses. The model results using CRU are also included in the plot for reference.

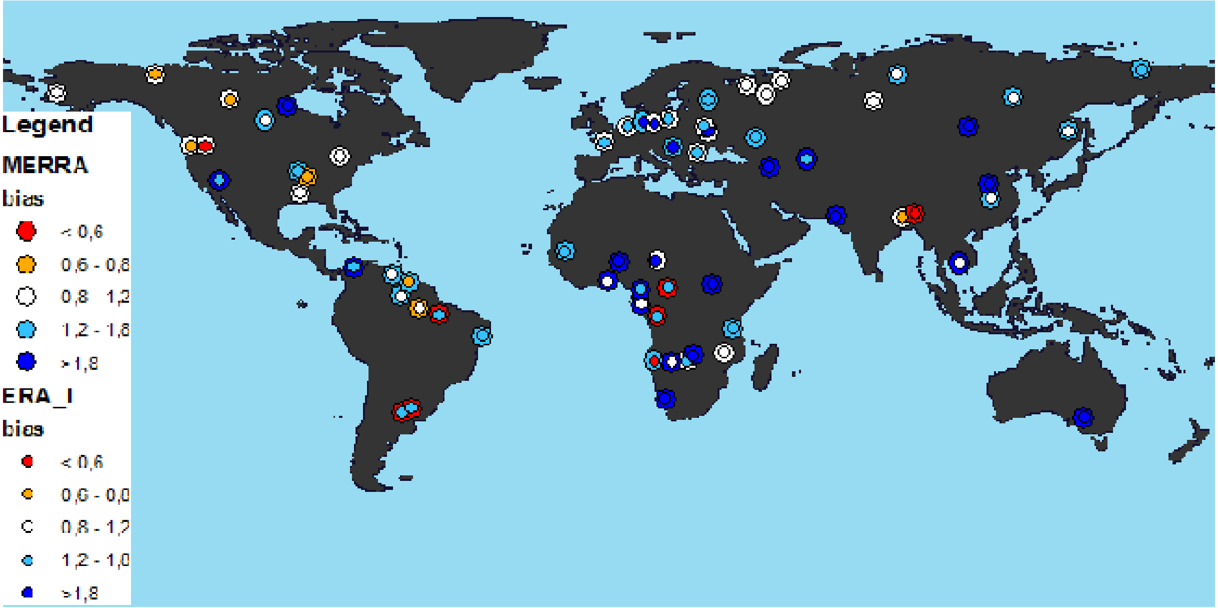


Figure 5: Ratio between simulated and observed discharge for a selection of observation stations, resulting from forcing PCR-GLOBWB with MERRA and ERA-Interim.

The simulated discharge of ERA-Interim and MERRA is overall slightly greater than that of CRU. The cumulative bias of CRU is close to zero, which shows that it was justified to use its data as reference in the previous step of the study. Comparing the results from the two products shows that MERRA has a few more stations for which the discharge deviates quite a bit from the identity line, both above and below it. Many of the stations that have significant bias are located in the problematic regions of MERRA, for instance the rivers Orinoco, Uruguay and Congo, as can be seen in the station map. The station in the Amazon basin has a relatively small bias, as it is affected by both drier and wetter conditions, respectively up- and downstream.

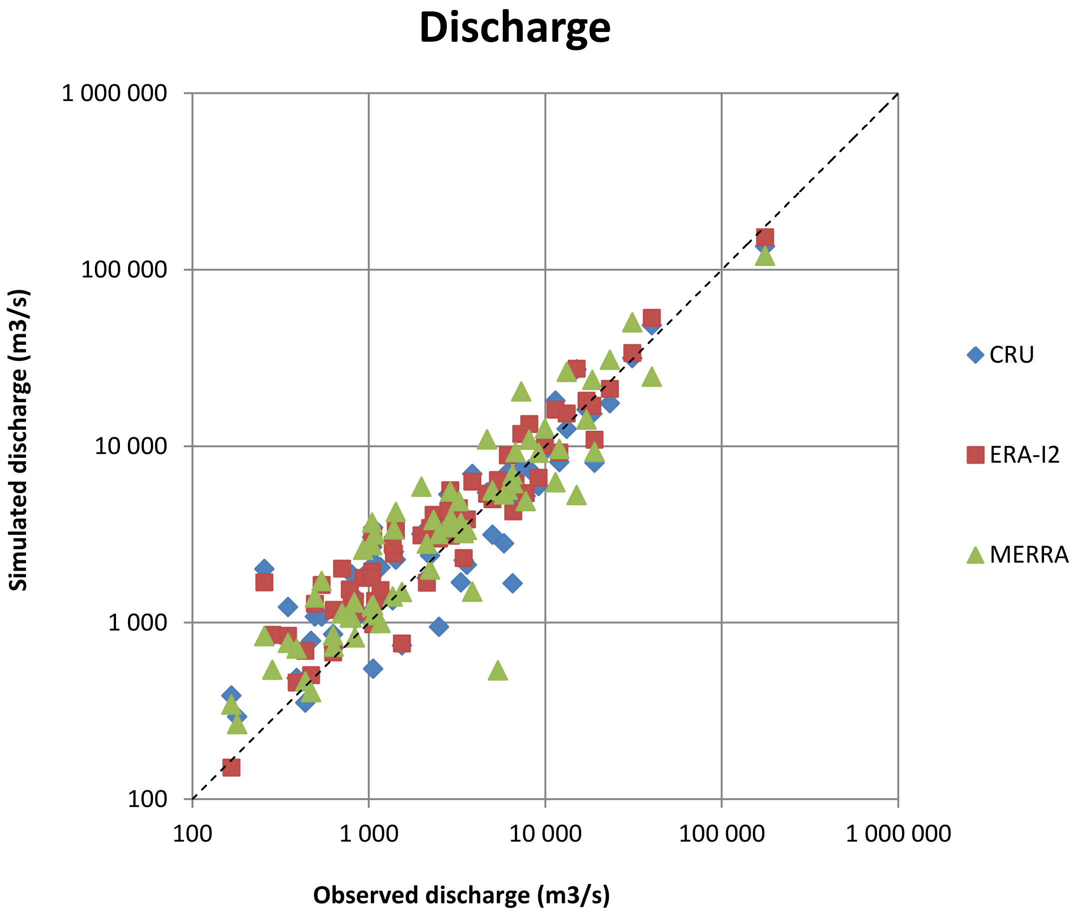


Figure 6: Scatterplot of simulated versus observed river discharge for the selection of observation stations.

Furthermore, the station map shows that discharge of European rivers for ERA-Interim is significantly higher than that of the MERRA run and biased compared to observations. This is another feature that can be linked to the findings in the previous section. One other striking difference in the station map is the difference between MERRA and ERA-Interim for the northwest of North America. ERA-Interim produces less discharge than MERRA for most of the stations located in this area. Referring back to the analysis of the meteorological fields the only explanation can be found in the difference in temperature. ERA-Interim is on average slightly warmer than CRU, where MERRA is mainly colder. The resulting difference in actual evapotranspiration apparently has significant consequences for the runoff that is generated.

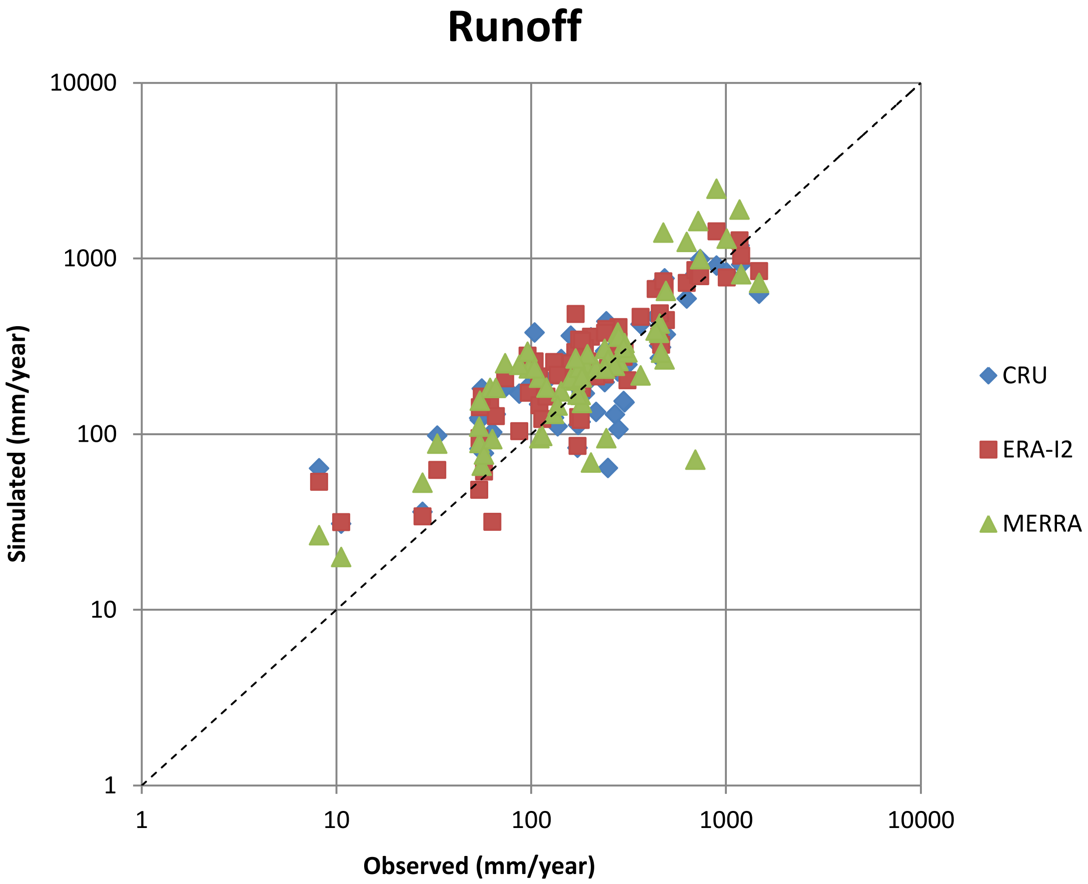


Figure 7: Scatterplot of simulated versus observed effective runoff for the selection of observation stations. Derived from discharge in figure 6 and basin area.

The plot of generated runoff shows that ERA-Interim is more consistent for different degrees of humidity. Here it is clear that MERRA is most problematic in tropical regions, specifically those of South America and Africa. In arid to moderate climates ERA-Interim runoff deviates slightly more, which can be explained for a great deal by the bias in precipitation over Europe.

Correlation of monthly discharge

From the analysis of the seasonality of precipitation, ERA-Interim came out better than MERRA. Figure 8 shows whether the same counts for the resulting discharge with another reference dataset. Looking at South America and Africa it is clear that the problems of MERRA are continued in the seasonality of the discharge. The correlation between MERRA and the GRDC discharge is worse than that of ERA-Interim for most of the stations in these regions. For example the river Congo scores 0.143 with MERRA and 0.543 with ERA-Interim. On the other hand the bias in precipitation over Europe and North America does not have major consequences for the correlation of monthly discharge. The average correlation with ERA-Interim is 0.05 higher (0.85 against 0.8), which is not very significant.

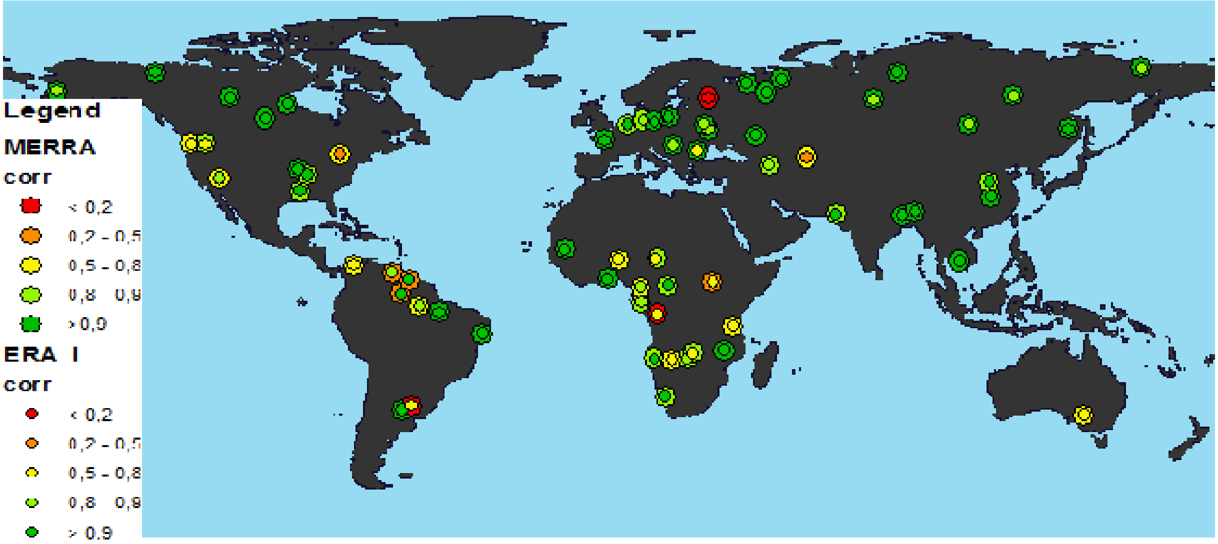


Figure 8: Correlation between observed discharge and simulations of PCR-GLOBWB using MERRA and ERA-Interim.

4.3 Conclusion on reanalysis comparison

The combined analysis of monthly meteorological variable fields and the resulting discharge from the PCR-GLOBWB model gives enough information to decide which of the reanalysis products is best to use in the remainder of this study and in future research. ERA-Interim turns out to be the more reliable product as it is both less biased and more accurate when it comes to seasonality in comparison with MERRA. The cloud cover problem of MERRA is a major flaw, leading to incorrect discharge quantities in regions around the equator, especially for river basins in South America. Also, the temporal correlation between MERRA discharge and observations is consequently lower than that of ERA-Interim.

Considering the remaining part of this study, this analysis points out some things to keep in mind, concerning weaknesses in the meteorological forcing of ERA-Interim. Particularly the bias in precipitation over Europe and in temperature over the northwest of North America have consequences for discharge amounts in these areas. On the other hand, the temporal correlation of monthly discharge shows only few stations with critically low rating, which corresponds to the good correlation between ERA-Interim and CRU precipitation.

4.4 Free flooding vs. no flooding

The next step in this study starts with running PCR-GLOBWB with a temporal resolution of 1 day using ERA-Interim variables for meteorological forcing. For the routing part of the model two scenarios are applied, respectively with and without flooding. In the first scenario rivers can spread out laterally when river stages are higher than a certain critical value, while in the alternative case river water remains within the channel under any condition. The model results are used in a preliminary analysis of how simulated discharge responds to different model conditions. Also, a new selection of river basins and GRDC stations needs to be made, based on this analysis, because of the changed temporal resolution.

Analysis of discharge

Figure 9 shows a plot of the bias corrected Nash-Sutcliffe model efficiency coefficient (NSEb) of the two model runs containing data for the 67 selected GRDC stations with daily discharge observations. To see how NSEb is influenced by the intensified peaks the same plot is shown for 5-day averages.

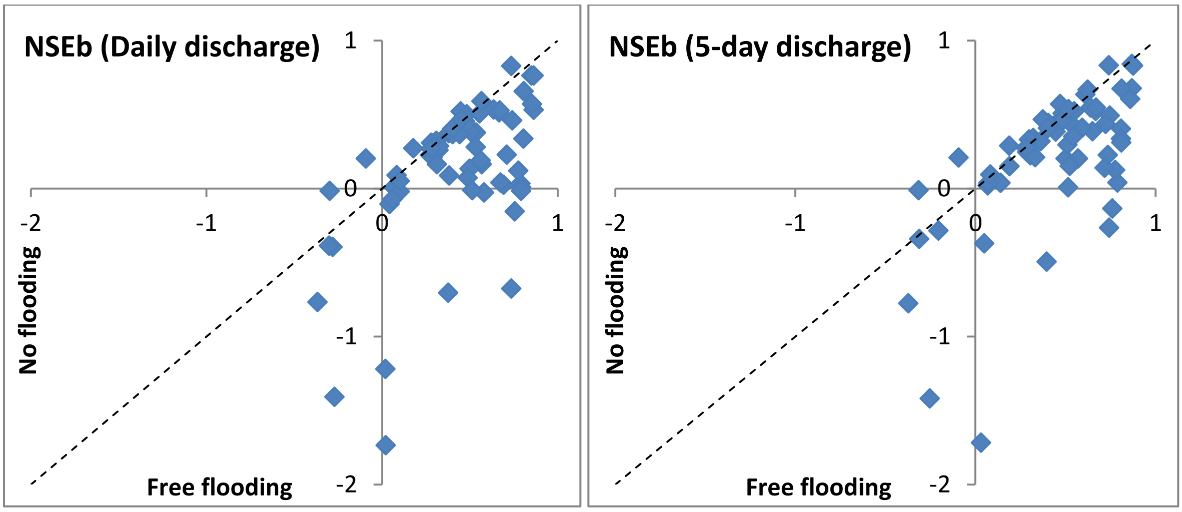


Figure 9: Scatterplots of Nash-Sutclliffe model efficiency for scenarios with and without flooding.

It is clear that for most of the stations the free flooding scenario produces discharge signals that are more similar to observations, even using 5-day averages. There are a few stations that score better without flooding but the difference between the values is in no case significant. The only significant difference between the two plots is the improved NSEb of the stations along the

Rhine and the Maranho, which values go from 0 to around 0.35, but it does not change which scenario does best.

Basin selection

A selection of the available daily discharge data is made based on the analysis of the results of the model run with free flooding. The stations in table 3 are chosen based on the criteria that were described earlier. Some basic statistics are included in the table. As can be seen there is quite some variation in the number of days that the discharge logs cover. Especially two of the stations in the Yangtze basin have a low coverage. This should be considered when looking at the outcomes of the analysis.

id	***River***	***Cor***	***NSE***	***NSEb***	***bias***	***nrDays***
1147010	*Congo*	0.395	-2.021	0.068	0.343	11683
2181900	*Yangtze*	0.863	0.663	0.736	-0.138	366
2903420	*Lena*	0.527	0.274	0.273	0.041	9131
3649950	*Tocantins*	0.925	0.309	0.855	0.469	10945
4127503	*Mississippi*	0.843	0.578	0.710	0.209	11572
4207900	*Fraser*	0.308	-0.346	-0.281	0.052	10958
4208025	*Mackenzie*	0.756	0.219	0.475	-0.289	11172
6435060	*Rhine*	0.891	-0.165	0.794	0.460	11688
6742900	*Danube*	0.835	-0.532	0.674	0.388	10586

Table 3: Selection of GRDC stations that are used in the remaining parts of the study.

4.5 Parameter variations

The final part of this study consists of a sensitivity analysis of the GHM PCR-GLOBWB to changes in parameters related to river regulation. Both discharge and flooding are looked at to see the consequences both in- and outside channels. The parameters that are variable are channel depth and floodplain and channel surface roughness, leading to a set of 36 scenarios.

Model efficiency in simulating discharge

In the analysis of discharge the NSEb is again the leading statistic in determining the similarity between modeled and observed time series. For the interpretation of the results, the NSEb values of the 36 different model runs are drawn in boxplot graphs (fig. 10). For the main station per basin the range of the resulting NSEb is plotted, each time keeping one parameter fixed.

Looking at the plots there are groups of rivers that have similar modeling performance and response to changes in the parameterization. The first group contains the high latitude rivers Lena, Mackenzie and Fraser. The overall NSEb of these rivers is low and for Fraser even none of the scenarios has a positive score. These rivers are relatively independent of the floodplain surface roughness and channel depth, which indicates that flooding does not occur much in these areas. For most of the time discharge is low in arctic rivers except from spring time, when melting of snow and ice produces a large amount of runoff in a short period of time. The score on modeling efficiency is mainly determined by the timing of this flood peak. The highest NSEb is achieved with a high value for channel roughness, which shows that with the current parameterization the modeled flood peak arrives downstream too soon. There are several possible explanations for this, of which premature snow and ice melting is the most obvious. Another factor that could play a role is river ice, which increases friction and stores river water that is released later. River ice is included in PCR-GLOBWB, but this analysis may indicate that the influence is underestimated.

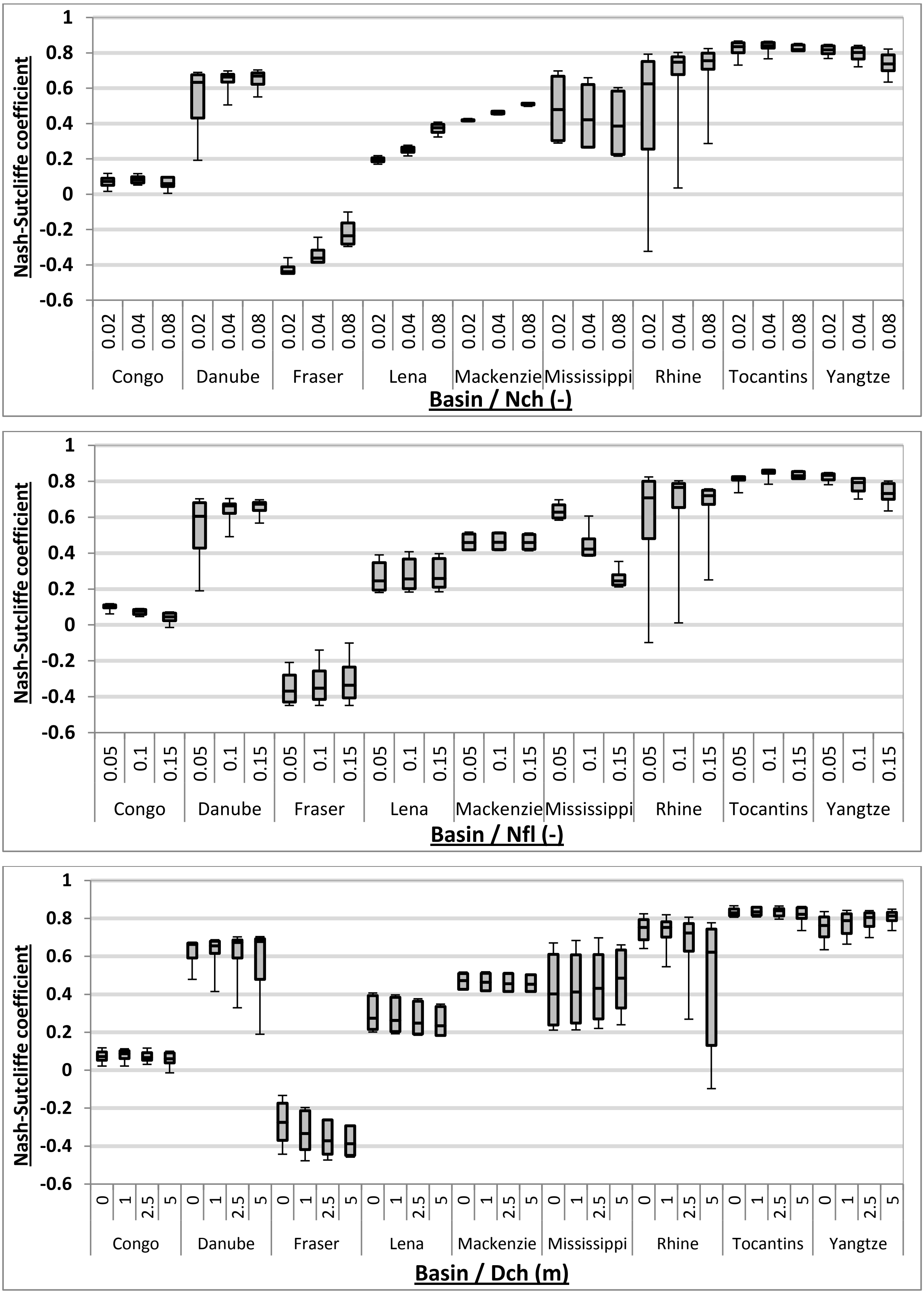


Figure 10: Boxplots representing the variability of model performance in discharge simulation. For each boxplot one parameter is kept fixed.

The second group contains the rivers Rhine, Danube and Tocantins. The link between these rivers is that their catchment size is relatively small, which makes that hydrographs rely for a great deal on the occurrence of separate rainfall events. Also the amount of precipitation that is converted to discharge is in the same order of magnitude, laying between 250 and 500 mm/year. The pattern of the boxplots shows that the model performance is influenced relatively little by changes in the parameters. The only scenarios that score significantly lower are those with both low channel and floodplain friction, especially in combination with an increased channel depth. This indicates that for these rivers modeled discharge peaks need to be decelerated to some extent for correct timing.

The rivers that are left, i.e. Congo, Mississippi and Yangtze, are relatively insensitive to changes in channel depth and roughness. The length of the flow path of these rivers is considerable, so the seasonality of the meteorological forcing is more important for the model performance than the timing and location of individual rainfall events, as is the case for the previous group. The overall score for the river Congo is not good. This is due to the fact that the intra-annual variability of the discharge is low, which makes that any deviation between modeled and observed discharge has a great impact on the NSE score.

The Mississippi has the largest variance in its model performance, with NSE scores going from 0.2 up to 0.7. The best runs are those with low floodplain friction and only with 5 meters additional channel depth, channel roughness becomes slightly more important. This indicates that with the current standard model settings extensive retention takes place under flooded conditions. This effect is reduced with a lower floodplain roughness, or with less flooding and lower channel friction.

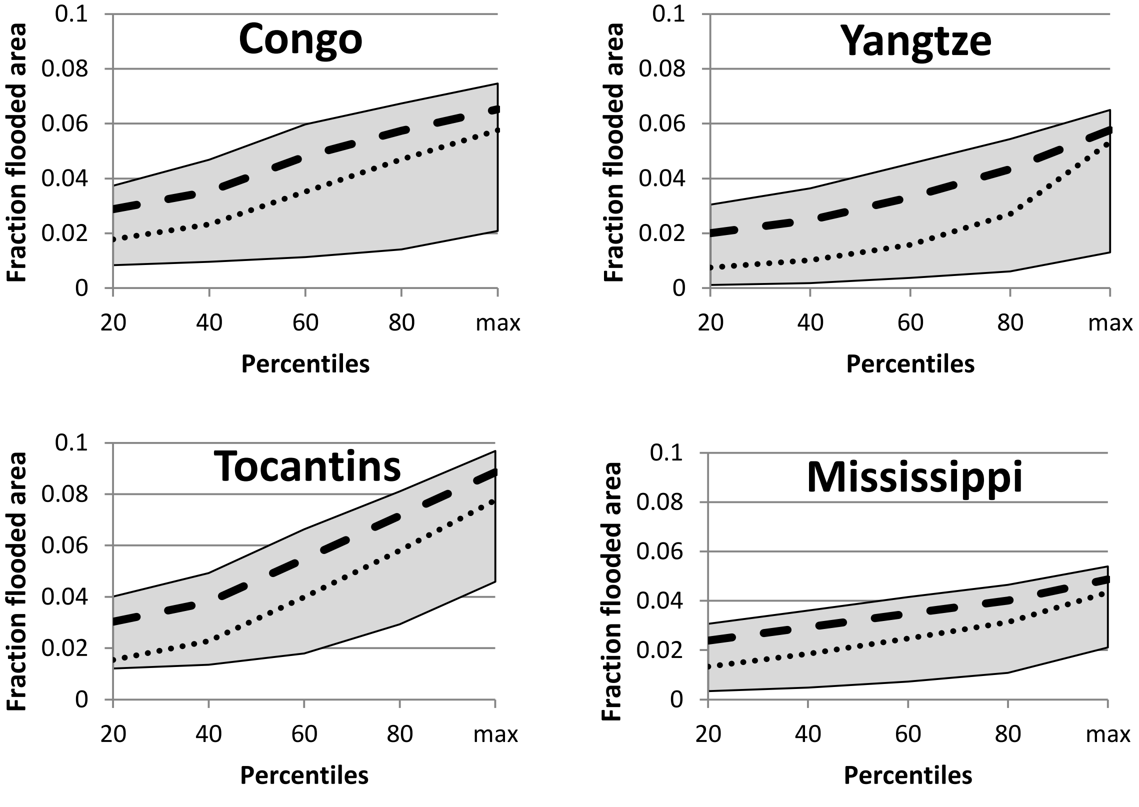
The pattern of the Yangtze river plots is the opposite of those of the rivers in the second group. For most scenarios the model performance is good, except for those with combinations of high roughness coefficients and little or no added channel depth, showing that discharge peaks should not be slowed down too much for the model to produce a hydrograph that is close to the observed situation.

Comparing the overall performance of the 36 scenarios shows that the scenario with low floodplain friction and high channel friction, in combination with none or little additional channel depth, is most stable. The only station for which this combination of parameters gives a relative low performance is that of the Tocantins basin. However, as the range in which the scores of this river fall is narrow, the negative effect is negligible.

With the Manning's values that were used here, the most stable scenario has a channel friction that is higher than the floodplain friction. This is a combination that is not expected to occur very often under natural conditions. This outcome shows that the retaining effect of flooding on discharge peaks is overestimated in the current model, while under non-flooding conditions the flow velocity should be tempered for better discharge simulation. This seems to be a contradiction, but it may be explained by the way floodwater is distributed over floodplains. It indicates that there is probably too much spreading, which makes the flood area, and along with that the average friction, larger than desired. This is a feature of the land surface elevation in the sub-grid variability and the hydrodynamic calculations that are done with it by PCR-GLOBWB.

Analysis of flooded area

The characteristics of the boxplots in figure 10 already give an indication about the role that flooding plays in correct discharge simulation and how it responds to parameter changes. For a further investigation of the flooding extent the fraction flooded area is calculated, which is simply the flooded area divided by the total area of a watershed. The focus is mainly on the variation of flood extent caused by the different scenarios. For each river and per scenario the distribution of fraction flooded area is calculated by means of the 20th, 40th, 60th, 80th percentiles and maximum of the series covering the period 1979-2010. For the selected rivers the bandwidth between minimum and maximum fraction is plotted (fig 11). In addition the lines of the current standard parameterization (dashed) and the scenario that scored best overall in the discharge reproduction (0.08/0.05/1; dotted) are included for reference.



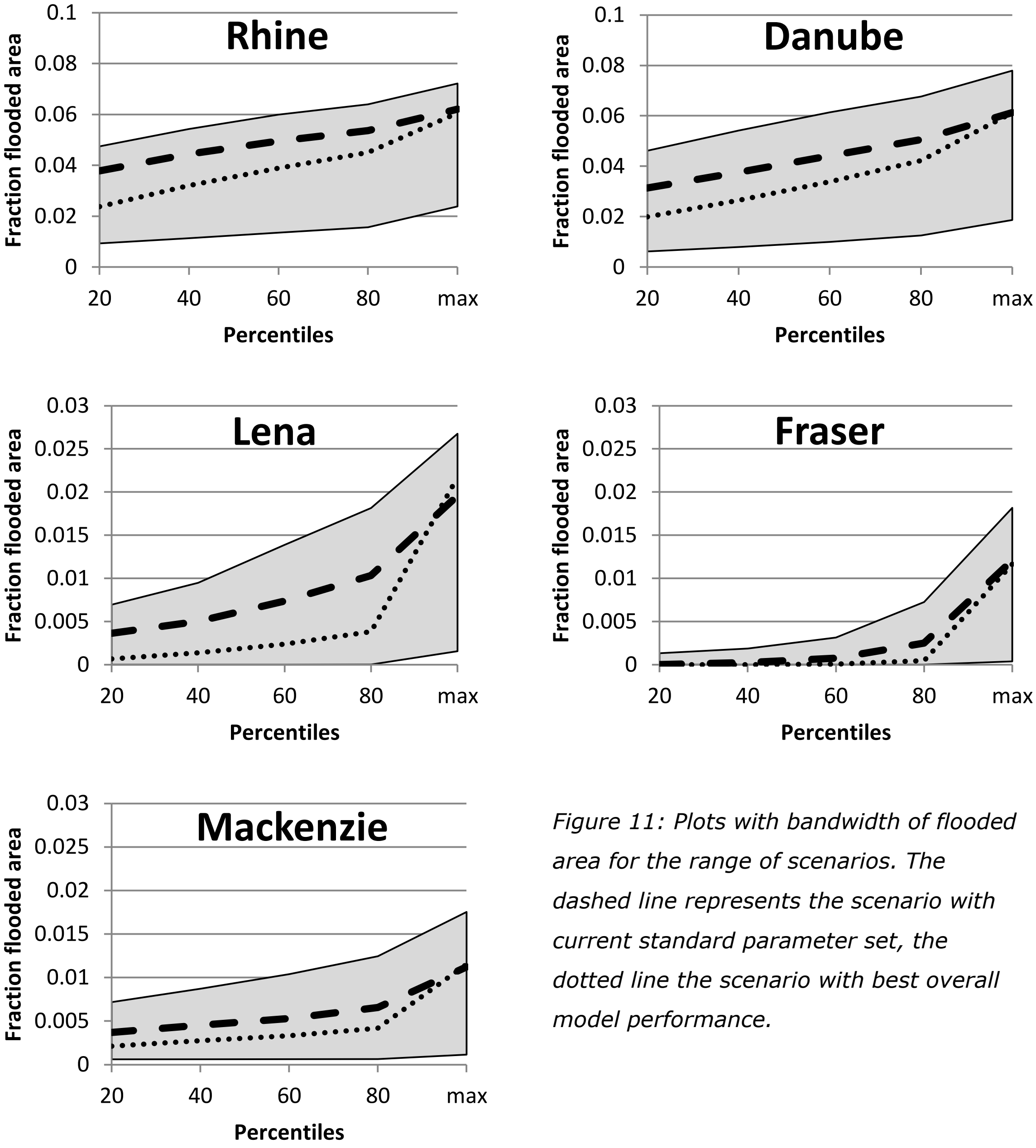


Figure 11: Plots with bandwidth of flooded area for the range of scenarios. The dashed line represents the scenario with current standard parameter set, the dotted line the scenario with best overall model performance.

Plotting the same groups of rivers together as in the previous section, shows that especially the arctic rivers have a shape deviating from the others. Because of the sharp hydrograph these rivers produce major flooding happens only during a short period of the year and even in those cases it is a smaller fraction than the other rivers. As explained previously the channel dimensions of arctic rivers are large, created by the high discharges when temperatures are rising and snow and ice starts melting. These peaks also cause the plotted lines to be more concave in comparison to those of other river types.

The plots of the six other basins are similar to one another, both in the range in which fractions fall and the stratification of the lines. The maximum extent of the larger basins, i.e. Congo, Mississippi and Yangtze, is slightly lower than that of the river Rhine, Danube and Tocantins. This can be explained by the fact that flood peaks occur less simultaneously over the basin area, as travel times of discharge between up- and downstream areas are longer.

Looking specifically at the lines within the bandwidth it is clear that the overall flood extent is lower for the alternative parameterization. The effect of a deeper channel and lower floodplain friction is apparently stronger than the increased stemming caused by a higher channel friction, although the difference becomes smaller with larger flood extent.

4.6 Combination of results and discussion

Thus far the first two research goals have been discussed, leaving the one concerning the combination of the two. Comparing the results as they have been presented up to this point should give insight in the question where the focus should be on when it comes to improving global-scale hydrologic modeling. This will lead to a discussion of the results and topics for potential future research.

Quantitative effects

The amount of discharge that is produced by a GHM is strongly related to the meteorological forcing that is applied. Compared to this the influence of parameterization on for instance discharge amounts is negligible. It is therefore important which choices are made in the selection of meteorological input. As mentioned before the choice may be different for different modeling purposes. It is also an option to enhance a dataset by combining it with another, like is done with ERA-40 and CRU TS2.1. Considering the major differences between ERA-Interim and MERRA it might be beneficiary to investigate the development of new methods for data combination.

Qualitative effects

In theory, the interaction between forcing and parameterization is more significant on a qualitative level. For instance timing and intensity of discharge peaks are determined by changes in meteorology, while channel and floodplain characteristics influence the flow velocities and the rise and fall of the water level. However, when we compare table 3 and figure 10, it is clear that the variability caused by the meteorological forcing for most basins is greater than that of the flow characteristics. Seasonal patterns, either caused by influences of rain seasons or peaks in melt water production, are hardly changed because of different friction values or additional channel depth. The European rivers are an exception, as their hydrographs contain peaks that related to rainfall events.

For a good calibration of the parameterization of the model, ideally the initial error in simulated discharge that comes from the forcing should be reduced. That way only the influence of parameter variation is left. The question is whether this is technically doable.

5. Conclusion

In this study two important components of global hydrological modeling are investigated. In the first place meteorological input variables determine the quantity and temporal variability of water availability for river discharge. On the other hand channel and floodplain properties influence the characteristics of hydrographs and river flooding. In this section the conclusions of the study are covered.

The comparison of the reanalysis products MERRA and ERA-Interim turned out to be strongly in favor of the second one. ERA-Interim performs better in all the aspects that have been analyzed, i.e. seasonal patterns in precipitation, overall quantitative resemblance in temperature, potential evapotranspiration and precipitation and simulated monthly river discharge resulting from running the PCR-GLOBWB global hydrological model with input from the two reanalyses.

The main flaw in MERRA is the cloud cover problem over (sub)tropical regions, leading to strong biases in all the meteorological variables that are of interest to this study. The biases show mainly in South America and Central Africa and lead to relatively dry inland conditions, while along the coast the opposite is the case. This results in deviating patterns and amounts of discharge for rivers in these regions, especially when their basin is mainly located either inland or close to the coast.

Although ERA-Interim comes out as the most reliable reanalysis product, it has some features that need to be considered when working with it. The most significant one is an overall positive bias in precipitation, with extremes over the western half of the Eurasian continent and the southeastern part of North America. The discharge that results from PCR-GLOBWB with ERA-Interim input is biased by up to 40% for rivers in these regions, like the rivers Rhine and Danube.

Running PCR-GLOBWB with conditions that either enable or impede flooding shows that most rivers produce more realistic discharge patterns with free flooding. Taking care of the additional spikiness of the no flooding hydrograph, by averaging the discharge over periods of 5 days, does not change that. The difference between the performances of the two scenarios gives a good indication of the possible response of simulated discharge and flooding to the parameter variations.

The analysis of discharge and flooding within PCR-GLOBWB, for the 36 scenarios, shows that there is a distinction between groups of rivers. Especially the arctic rivers respond differently to the parameter variations, as they deal with a strong discharge peak during melting season. The NSEb scores for these rivers are generally low, as any shift in the timing of the modeled discharge peak makes that high discharge values in the simulated hydrograph correspond to low values in the observed hydrograph, and vice versa. The timing of the discharge peak of arctic rivers depends on a combination of factors, which makes it hard to determine whether the analyzed parameters should be changed to improve the modeling performance. When it comes to flooding the arctic rivers are relatively insensitive to the parameter changes, as flooding only occurs during a short period of the year.

The differences between the other rivers are significantly smaller, especially when it comes to the way flooding responds to the changing parameterization. In comparison, the influence of parameter variation on NSEb scores differs much more per river, although there are groups that shows a similar pattern. The rivers Rhine and Danube are stable for most of the scenarios, only scoring lower with deep channels and low channel friction, which apparently leads to flow velocities that are too high. On the other hand, the rivers Mississippi and Yangtze are primarily sensitive to changes in floodplain friction, performing best with a low Mannings n value. This indicates, especially looking at the plot of the Mississippi, that flooding occurs anyway, but the decelerating effect should not be overestimated.

Finally, the strong influence of meteorological forcing on model results makes that it is best to look at possible improvements in that area before any fine-tuning is done on hydrological processes and the models parameterization.

6. References

Arnell, N. W. (1999), A simple water balance model for the simulation of streamflow over a large geographic domain. J. Hydrol. 217, 314–335.

Balsamo, G., S. Boussetta, P. Lopez, L. Ferranti, 2010: Evaluation of ERA-Interim and ERA-Interim-GPCP-rescaled precipitation over the U.S.A., ERA Report Series 5, 1-10.

Betts, A.K., Köhler, M., Zhang, Y., (2009) Comparison of river basin hydrometeorology in ERA-Interim and ERA-40 reanalyses with observations. J. of Geophysical research 114, D02101.

Bordi, I., Fraedrich, K., Petitta, M., Sutera, A., (2006) Large-scale assessment of drought variability based on NCEP/NCAR and ERA-40 Re-analyses. Water resources management 20, 899-915.

Bosilovich, M.G., Robertson, F.R., Chen, J., (2011) Global energy and water budget in MERRA. Journal of climate 24, 5721-5740.

Candogan Yossef, N., van Beek, L. P. H., Kwadijk, J. C. J., Bierkens, M. F. P. (2011) *Skill assessment of a global hydrological model in reproducing flow extremes.* Hydrol. Earth Syst. Sci. Discuss., 8, 3469–3505.

Challinor, A.J., Wheeler, T.R., Slingo, J.M., Craufurd, P.Q., Grimes, D.I.F., (2005) Simulation of crop yield using ERA-40: Limits to skill and nonstationarity in weather-yield relationships. J. Applied Meteorology 44, 516-532.

Chow, V.T., (1959), *Open channel hydraulics*, McGraw-Hill, New York.

Dee, D. P. and co-authors (2011) *The ERA-Interim reanalysis: configuration and performance of the data assimilation system.* Q. J. R. Meteorological Soc. 137, 553-597.

Durand, Y., Brun E., Merindol, L., Guyomarc'h, G., Lesaffre B., and Martin, E. (1993) A meteorological estimation of relevant parameters for snow models. Ann. Glaciol., 18, 65–71.

Global Runoff Data Centre (2011*) Long-Term Mean Monthly Discharges and Annual Characteristics of GRDC Station* / Global Runoff Data Centre. Koblenz, Germany: Federal Institute of Hydrology (BfG), 2011.

Kalnay, E., and co-authors, (1996), The NCEP/NCAR 40-year reanalysis project. Bul. of the Am. Met. Soc. 77, 3, 437-471.

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

Mitchell, T.D., Jones, P.D., (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. of Climatol. 25, 693-712.*

New, M., Hulme, M. and Jones, P. (1999) Representing twentieth-century space-time climate variability. Part I: Development of a 1961-1990 mean monthly terrestrial climatology*. Journal of Climate 12, 829-856.*

New, M., Hulme, M. and Jones, P. (2000) Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate*. Journal of Climate 13, 2217-2238.*

New, M., D. Lister, M. Hulme and I. Makin (2002) A high-resolution data set of surface climate over global land areas. *Climate Research 21, 1-25.*

Nijssen, B., G.M. O'Donnel, D.P., Lettenmaier, D. Lohmann and E.F.Wood (2001) Predicting the discharge of global rivers, *Journal of Climate 14: 3307-3323.*

Onogi, K., and co-authors (2007) The JRA-25 reanalysis. *J. Meteo. Soc. Japan 85, 3, 369-432.*

Reichle, R. H., Koster, R. D., De Lannoy, G. J. M., Forman, B. A., Ling, Q., Mahanama, S. P. P., Touré, A. (2011) Assessment and Enhancement of MERRA land surface hydrology estimates. *Journal of Climate, 24, 6322-6338*.

Sperna Weiland, F. C., (2011) Hydrological impacts of climate change: interpretation of uncertainties introduced by global models of climate and hydrology, PhD Thesis, *Igitur*, Utrecht University

Szczypta, C., Calvet, J.-C., Albergel, C., Balsamo, G., Boussetta, S., Carrer, D., Lafont, S., Meurey, C., (2010) Verification of the new ECMWF ERA-Interim reanalysis over France. *Hydrol. Earth Syst. Sci. Discuss. 7, 7151-7190.*

Uppala, S.M. et al. (2005), The ERA-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society 131, 2961–3012.*

Van Beek, L.P.H., Bierkens, M.F.P. (2008) The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification. *Report Department of Physical*

Geography, Utrecht University, Utrecht, The Netherlands, http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf

Vörösmarty, C.J., Sahagian, D., (2000) Anthropogenic Disturbance of the Terrestrial Water Cycle. *BioScience 50(9), 753-765.*

Wada, Y., van Beek, L.P.H., Viviroli, D., Dürr, H.H., Weingartner, R., Bierkens, M.F.P., (2011) Global monthly water stress: 2. Water demand and severity of water stress. *Water resources research 47, W07518.*

Wang, G., Dolman, A.J., Blender, R., Fraedrich, K., (2010) Fluctuation regimes of soil moisture in ERA-40 re-analysis data. *Theor. Appl. Climatol. 99, 1-8*.

Widén-Nilsson, E., S. Halldin and C. Xu (2007), Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation. *Journal of Hydrology 340, 105-118.*

Appendix A – Selection of GRDC stations

ID	**River**	**Station name**
3629000	AMAZONAS	OBIDOS
1147010	CONGO	KINSHASA
4127800	MISSISSIPPI RIVER	VICKSBURG, MS
2909150	YENISEY	IGARKA
3265601	PARANA	TIMBUES
2906900	AMUR	KOMSOMOLSK
4208025	MACKENZIE RIVER	ARCTIC RED RIVER
2181800	YANGTZE RIVER (CHANG JIANG)	HANKOU
6977100	VOLGA	VOLGOGRAD POWER PLANT
4122900	MISSOURI RIVER	HERMANN, MO.
1734500	NIGER	MALANVILLE
5204268	MURRAY	LOCK 9 UPSTREAM (764.8 KM)
1159100	ORANGE	VIOOLSDRIF
3206720	ORINOCO	PUENTE ANGOSTURA
2846800	GANGES	FARAKKA
2335950	INDUS	KOTRI
4103200	YUKON RIVER	PILOT STATION, AK
2903428	LENA	SOLYANKA
3649950	TOCANTINS	TUCURUI
2910600	OB	PROKHORKINO
2180800	HUANG HE (YELLOW RIVER)	HUAYUANKOU
6842900	DANUBE RIVER	SILISTRA
4236010	NIAGARA RIVER	QUEENSTON
4115201	COLUMBIA RIVER	BEAVER ARMY TERMINAL NEAR QUINCY, OR
2569002	MEKONG	PHNOM PENH (CHROUI CHANGVAR)
3651900	SAO FRANCISCO	TRAIPU
4208400	SLAVE RIVER	FITZGERALD (ALBERTA)
1537100	CHARI	NDJAMENA(FORT LAMY)
2998510	KOLYMA	KOLYMSKAYA
4123050	OHIO RIVER	METROPOLIS, ILL.
1749100	UBANGI	BANGUI
1673900	WHITE NILE	MONGALLA
2917100	AMU DARYA	CHATLY
2907401	SELENGA	KABANSK
4152103	COLORADO RIVER (PACIFIC OCEAN)	BELOW HOOVER DAM, AZ-NV
2851300	BRAHMAPUTRA	PANDU
1531700	VOLTA	SENCHI(HALCROW)
6970250	SEVERNAYA DVINA (NORTHERN DVINA)	UST-PINEGA
4213550	SASKATCHEWAN RIVER	THE PAS
1291100	ZAMBEZI	KATIMA MULILO
6980801	DNEPR	KIEV (KYYIV)
4214270	CHURCHILL RIVER	ABOVE RED HEAD RAPIDS
4116182	SNAKE RIVER	BELOW ICE HARBOR DAM, WA
6972430	NEVA	NOVOSARATOVKA
3103300	MAGDALENA	CALAMAR
6970701	PECHORA	UST-TSILMA
3469050	URUGUAY	SALTO
1812500	SENEGAL	BAKEL
1643100	OGOOUE	LAMBARENE

Continued		
6458010	VISTULA (WISLA)	TCZEW
6990700	KURA	SURRA
6435060	RHINE RIVER	LOBITH
1286900	RUFIJI	STIGLER
6544100	TISZA	SENTA
6340110	ELBE RIVER	NEU-DARCHAU
1338050	SANAGA	EDEA
1992700	SHIRE	LIWONDE
3618500	RIO BRANCO	CARACARAI
6970680	VYCHEGDA	FEDYAKOVO
6123100	LOIRE	MONTJEAN
6357010	ODER RIVER	HOHENSAATEN-FINOW AP
1591403	KAFUE	ITEZHI-TEZHI
6979500	PRYPYAT	MAZYR (MOZYR)
1257100	OKAWANGO	RUNDU
1255100	KUNENE RIVER	RUACANA
3308600	ESSEQUIBO	PLANTAIN ISLAND

Statement of originality of the MSc thesis

I declare that:

1. this is an original report, which is entirely my own work,

2. where I have made use of the ideas of other writers, I have acknowledged the source in all instances,

3. where I have used any diagram or visuals I have acknowledged the source in all instances,

4. this report has not and will not be submitted elsewhere for academic assessment in any other academic course.

Student data:

Name: Lex van Dolderen

UU registration number: 3021297

Date: 21-08-2014

Signature: