Application of global hydrological datasets for river basin modelling

Toepassing van wereldwijde hydrologische datasets voor modellering van stroomgebieden

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op woensdag 23 mei 2018 des middags te 2.30 uur

door

Patricia López López

geboren op 2 november 1987 te Granada, Spanje

Promotor:

Prof. Dr. ir. M. F. P. Bierkens

Copromotoren:

Dr. ir. G. Sterk

Dr. J. Schellekens

Application of global hydrological datasets for river basin modelling

Promotor:

Prof. Dr. ir. M. F. P. Bierkens

Copromotoren:

Dr. ir. G. Sterk

Dr. J. Schellekens

Examination committee:

Prof. Dr. A. P. J. de Roo, European Comission, Joint Research Centre

Prof. Dr. ir. Gabriëlle J. M. de Lannoy, KU Leuven

Prof. Dr. Wouter Dorigo, Vienna University of Technology

Prof. Dr. Jaap Kwadijk, Deltares

Prof. Dr. ir. Hubert Savenije, Delft University of Technology

This research was funded by a grant from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 603608, Global Earth Observation for integrated water resource assessment: eartH2Observe.

ISBN number 978-90-6266-503-7

Published by Faculty of Geosciences, Universiteit Utrecht, The Netherlands, in:

Utrecht Studies in Earth Sciences (USES) 152

Cover picture: Francisco Javier López Mendoza

Printed in the Netherlands by Ipskamp.

Utrecht Studies in Earth Sciences 152

Application of global hydrological datasets for river basin modelling

Patricia López López

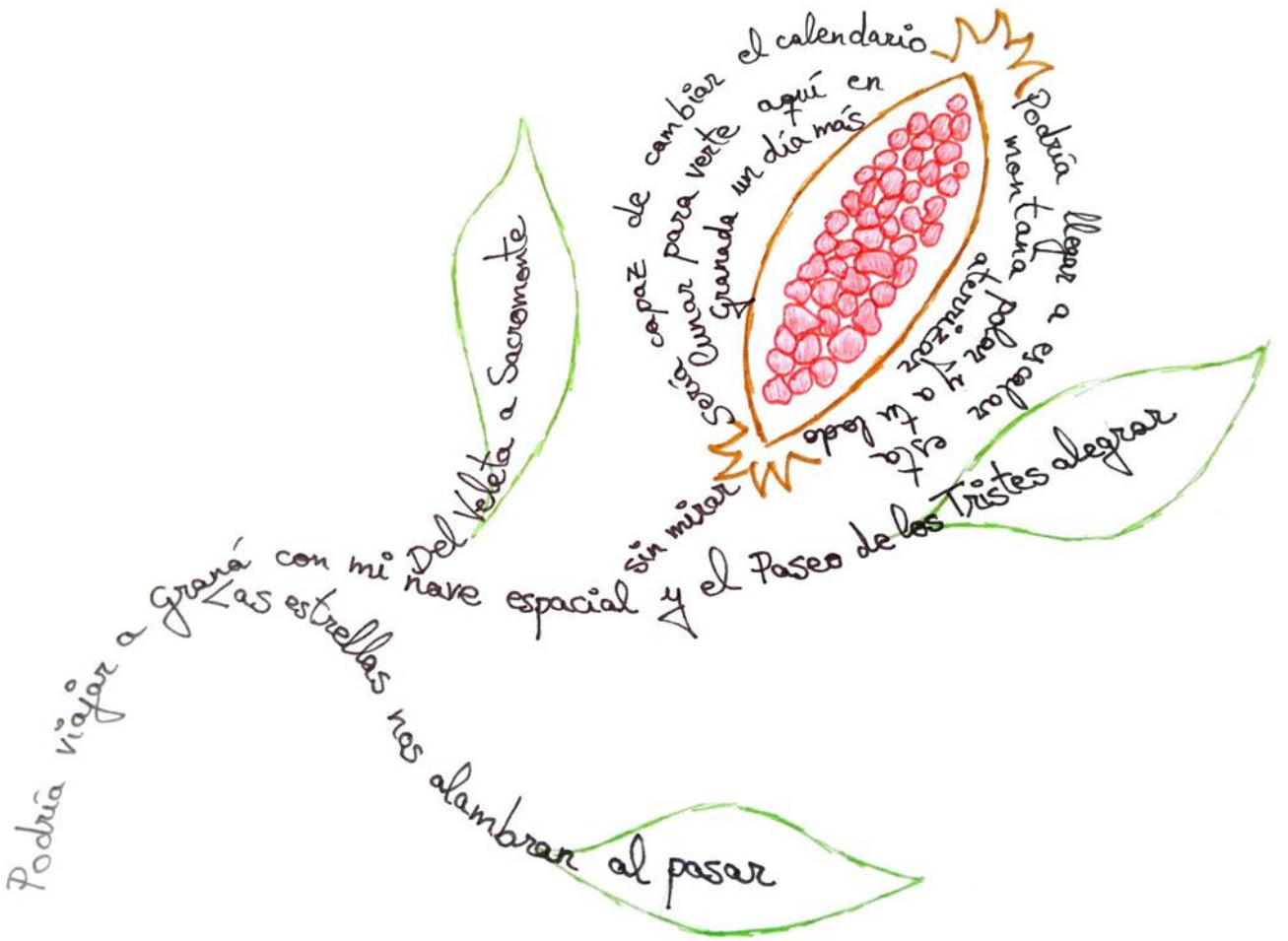
Utrecht 2018

Department Physical Geography

Faculty of Geosciences – Utrecht University

Utrecht Studies in Earth Sciences

Local editors:



EN GRANADA, SUPERSUBMARINA

Contents

1 Introduction

1.1 Background

Hydrology is the branch of science concerned with the distribution and circulation of water of the Earth and their physical and chemical properties in relation with the surrounding environment (Linsley, 1975). During the last decades, the number and diversity of water-related challenges in hydrological systems has increased. These include population growth, rapid urbanization and industrialization which result in growing water needs, as well as increased water demands by agriculture, fishery and inland navigation. Moreover, climate change intensifies these challenges impacting water resources in many river basins around the world (Devia et al., 2015). A thorough study of the different hydrological processes at global and river basin scale is needed in order to overcome these challenges. Nowadays, hydrological models have been developed to analyze, understand and explore solutions for integrated and sustainable water resources management, in order to support decision makers and operational water managers.

1.1.1 Hydrological modelling

Hydrological models have evolved from the Rational Method (Mulvaney, 1850) to recent spatially distributed models (Schellekens et al., 2016; Sutanudjaja et al., 2018). Nowadays, the choice of a hydrological model from the wide available range is complex and depends on the specific pursued task (Todini, 2007). Hydrological models can serve many purposes: (i) better understanding of the hydrological processes involved in the water cycle and how climate, human and natural variations may affect these; (ii) generation of retrospective water resources accounts for design of infrastructures and policy planning; (iii) near-real time monitoring of water availability and (iv) water resources forecasting, with special importance of extreme events, such as floods and droughts (Van Dijk and Renzullo, 2011).

Hydrological models are typically applied in a river basin for water resources management at a regional or local scale. During the last decades, a number of large-scale hydrological models (and land-surface models) have been developed to quantify the global water cycle components and to analyze the human being influence and climate change impact on water resources over continental to global domains. VIC (Liang et al., 1996), HTESSEL (Balsamo et al., 2009), LISFLOOD (Van Der Knijff et al., 2010), WATERGAP (Döll et al., 2009), PCR-GLOBWB (Sutanudjaja et al., 2018), ORCHIDEE (d'Orgeval et al., 2008), SURFEX-TRIP (Decharme et al., 2010) and W3RA (Van Dijk et al., 2014) are some examples of large-scale models for water resources applications (Bierkens et al., 2015).

However, to apply large- and local-scale hydrological models for water resources management at the river basin scale, hydro-meteorological information at fine spatial and temporal resolution is needed. Nonetheless, many river basins worldwide are poorly gauged or ungauged with limited or non-available in situ data (Hrachowitz et al., 2013). Recently, satellite-derived observations have emerged as a promising complement and/or alternative to in situ datasets.

1.1.2 Satellite-derived observations for hydrological modelling

In the last few decades, radar and satellite technologies have greatly improved and have become more broadly available, providing diverse hydro-meteorological datasets at fine spatial and temporal resolutions. Satellite-derived datasets cover large areas (many of them have a near-global coverage) over long time periods and are able to reflect the spatial patterns and temporal variability of the measured variable (Wagner et al., 2009).

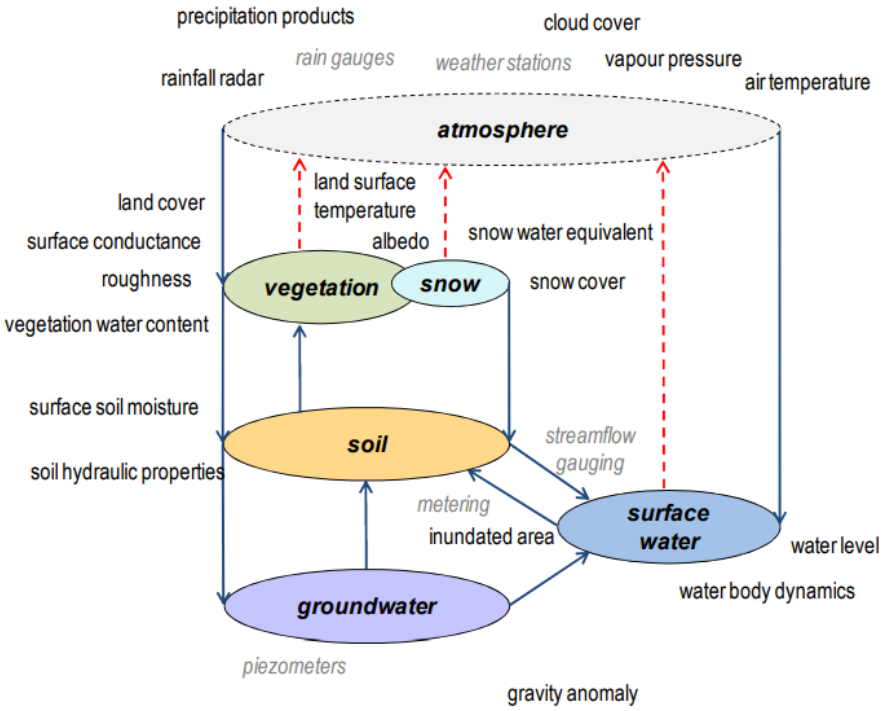


Figure 1.1 Diagram showing in which parts of the hydrological system, satellite observations could be used to support hydrological modelling. Conventionally used on-ground observations are listed in italics (from Van Dijk and Renzullo, 2011).

The integration of satellite observations into hydrological models can provide important benefits for water resources management (Schultz and Engman, 2000; Schmugge et al., 2002; Fernandez Prieto et al., 2009; Wagner et al., 2009; Escobar et al., 2017), especially in areas where in situ information is limited or not available (Sivapalan et al., 2003; Hrachowitz et al., 2013). Previous studies have proven that satellite observations can improve spatial and temporal accuracy of hydrological model estimations. Van Dijk and Renzullo (2011) provided a schematic diagram showing in which parts of the hydrological system, satellite observations could be used to support hydrological modelling (*Figure 1.1)* distinguishing five groups that could be independently or jointly applied: (i) atmosphere (to improve meteorological forcing data, mainly precipitation), (ii) vegetation and snow (to classify the landscape into land cover types, to estimate evapotranspiration, snow cover, vegetation structure and biomass), (iii) soil (to estimate surface soil moisture and soil hydraulic properties), (iv) surface water (to measure water level and water extent) and (v) groundwater (to increase the information on groundwater dynamics).

Therefore, satellite-derived observations of hydro-meteorological variables offer a wealth of opportunities to improve hydrological modelling. However, to reach the full potential of satellite-derived observations in hydrology, some challenges need to be overcome, including the continuity of satellite missions and datasets; the finding of more computationally efficient techniques for modelling at very fine spatial resolution, data assimilation, parameter optimization and state updating; and the better understanding of the relationship between satellite- and model-derived variables.

Numerous approaches can be followed to use the valuable information of satellite-derived observations in hydrology, which depend on the observations, characteristics of the basin, model structure, aimed task and in situ data availability, among others. Below, the possible methodologies to integrate satellite-derived observations into hydrological models to improve water resources estimates are discussed in three subsections (*Figure 1.2*): (i) meteorological datasets, (ii) model parameters estimation and data assimilation and (iii) model evaluation.

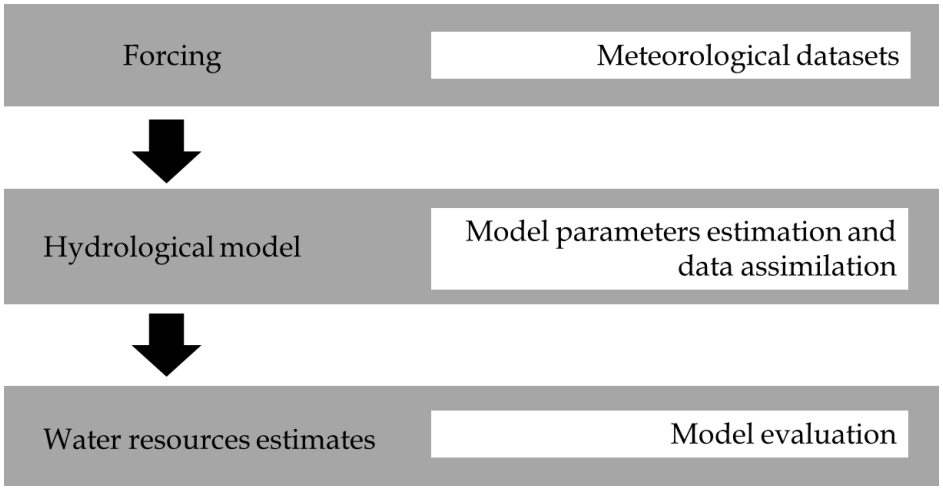


Figure 1.2 Scheme of possible approaches to use satellite-derived observations for improving hydrological modelling.

1.1.2.1 Meteorological datasets

Meteorological datasets are used to force the hydrological models and can have significant impact on model simulations (Fekete et al., 2004). Precipitation and evapotranspiration are the main drivers of the hydrological processes in the water cycle. Historically, meteorological data have been obtained from in situ weather stations (rain gauges, evapotranspiration pans, etc.). However, many river basins around the world are still poorly gauged or ungauged (Sivapalan et al., 2003). Moreover, meteorological data measured from ground stations are only representative for a limited area around the location of the instruments (Bohnenstengel et al., 2011) and, at the usual densities at which they are available, cannot effectively capture spatial variability. In recent years, global datasets of meteorological variables, such as precipitation, evapotranspiration, temperature, wind speed and solar radiation, among others, have been developed.

Scientific efforts have been made to optimally combine all data sources to generate consistent meteorological products for hydrological modelling. For example, precipitation datasets at regional and global scale can be obtained blending information from in situ, satellites and models. The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN; Hsu et al., 1997), the Climate Prediction Center morphing method to derive precipitation (CMORPH; Joyce et al., 2004), the Global Satellite Mapping of Precipitation project (GSMaP; Kubota et al., 2007), the Tropical Rainfall Measuring Mission (TRMM; Huffman et al., 2007), the Global Precipitation Measurement (GPM; Hou et al., 2014) and the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS; Funk et al., 2015) are some examples of precipitation datasets merging in situ, satellite and radar observations. Climate models are also used in combination with the previous sources to provide not only global precipitation products, but also other meteorological variables, such as temperature, humidity, etc., in the form of reanalysis: the European Centre for Medium-Range Weather Forecasts global atmospheric reanalysis data (ECMWF ERA-Interim; Dee et al., 2011), the Climatic Research Unit Time Series (CRU TS; Harris et al., 2014), the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data (WFDEI; Weedon et al., 2014) and the Multi-Source Weighted-Ensemble Precipitation (MSWEP; Beck et al., 2017a), among others.

Most of the meteorological products mentioned above are available at spatial resolutions between 0.25^{o} and 1^{o}, which limits their ability to accurately reproduce meteorological phenomena at the river basin scale. Opportunites for further development of meteorological datasets include the increase in spatial and temporal resolution using additional information from auxiliary variables, such as vegetation greenness, biomass, topography, etc. (Long et al., 2016; Ezzine et al., 2017), and the use of alternative methodologies for merging satellite, radar, in situ and model estimates (Beck et al., 2017a).

1.1.2.2 Model parameters estimation and data assimilation

An appropriate model representation of the hydrological processes in a river basin is crucial to provide accurate model prediction. Satellite-derived observations can be used for a better understanding of these processes (Seneviratne et al., 2010; Hafeez et al., 2011) leading to improved model simulations through calibration and assimilation techniques (Parajka et al., 2006; Beck et al., 2009; Thirel et al., 2013, Thiemig et al., 2013).

A large suite of calibration approaches exist, such as the Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, 1992). Traditionally, calibration techniques have used in situ streamflow observations to estimate parameters of hydrological models. However, this information is not always available and satellite-derived observations become a promising additional data source that allows calibrating for multiple variables, not only streamflow. Moreover, calibration based on satellite-derived observations partly overcomes the problem of over-parameterization that can occur when calibrating to limited hydrometric observations (Fenicia et al., 2007; Gupta et al., 2008). Various studies have shown the potential of satellite-derived observations, such as surface soil moisture, evapotranspiration or total water storage data for calibrating model

parameters (Meesters et al., 2005; Campo et al., 2006; Immerzeel and Droogers, 2008; Lo et al., 2010; Wanders et al., 2014a).

Data assimilation comprises a set of mathematical techniques that jointly use observations and model predictions to improve knowledge of the system states. These techniques are frequently classified in sequential (derived from the classic linear Kalman filter, e.g. ensemble Kalman filter; Kalman, 1960) and variational algorithms (e.g. 3DVAR; Parrish and Derber, 1992). Data assimilation techniques can be used to optimally combine observations and model simulations in hydrology. Data assimilation algorithms have been broadly applied in hydrological modelling, assimilating mainly in situ observations (Rakovec et al., 2012; Abaza et al., 2014). Global satellite-derived observations can be alternatively or additionally used fFor improving model simulations through assimilation. Previous studies have shown the benefit of assimilating satellite-derived observations, such as surface soil moisture and snow (Andreadis and Lettenmaier, 2006; Lievens et al., 2015; Tangdamrongsub et al., 2015).

In spite of the last advances in model parameters estimation and data assimilation into hydrological models using satellite-derived observations, further investigations need to be carried out to solve some problems. First, there are often conceptual differences between modelled and satellite-derived variables (e.g. surface soil moisture depth). Second, satellite-derived observations can be at coarser spatial resolutions than the hydrological model scale (e.g. total water storage data at ~0.25^{o} from the Gravity Recovery and Climate Experiment, GRACE or surface soil moisture data at ~0.5^{o} from the Advance Microwave Scanning Radiometer - EOS, AMSR-E). Third, the effectiveness of calibration and data assimilation techniques using satellite-derived observations depends on the dominant hydrological processes in the basin (e.g. surface soil moisture may improve model simulations in areas where evapotranspiration is limited by water availability and may prove of limited use under different conditions). Lastly, finding the optimal parameter set for each model unit and/or updating model states via data assimilation using satellite-derived observations require further development in computationally efficient techniques.

1.1.2.3 Model evaluation

Hydrological models need to be evaluated against observations to identify those processes or quantities that are not well represented in the model and should be modified for improving future model simulations. Moreover, model evaluation helps to understand which situations are optimal for the model to be applied and which are inadvisable.

Instead of using in situ data which provide information at selected locations, satellite-derived observations allow a spatially consistent model evaluation. For example, Van Dijk and Renzullo (2009) compared total water storage estimates from a hydrological model to satellite-derived observations from GRACE, identifying errors in the model description of soil and groundwater dynamics. However, sometimes it is difficult to find out where differences between model- and satellite-derived data come from (forcing data, model parameters, initial conditions, model structure, etc.). Moreover, satellite-derived observations can equally contain considerable errors; for example, those derived from the

scaling process when the spatial resolution of the hydrological model and the observed data for evaluation are different.

1.2 Motivation: the eartH2Observe project

This research was carried out in the framework of the eartH2Observe project (Global Earth Observations for integrated water resources assessment, http://www.earth2observe.eu/). EartH2Observe was funded by the Seventh Framework Programme for Research and Technological Development (FP7) of the European Union (EU) with a total duration of 4 years, from January 2014 until December 2017. The eartH2Observe consortium was composed of 27 partners and 4 associate partners, 23 from the EU and 8 from non-EU.

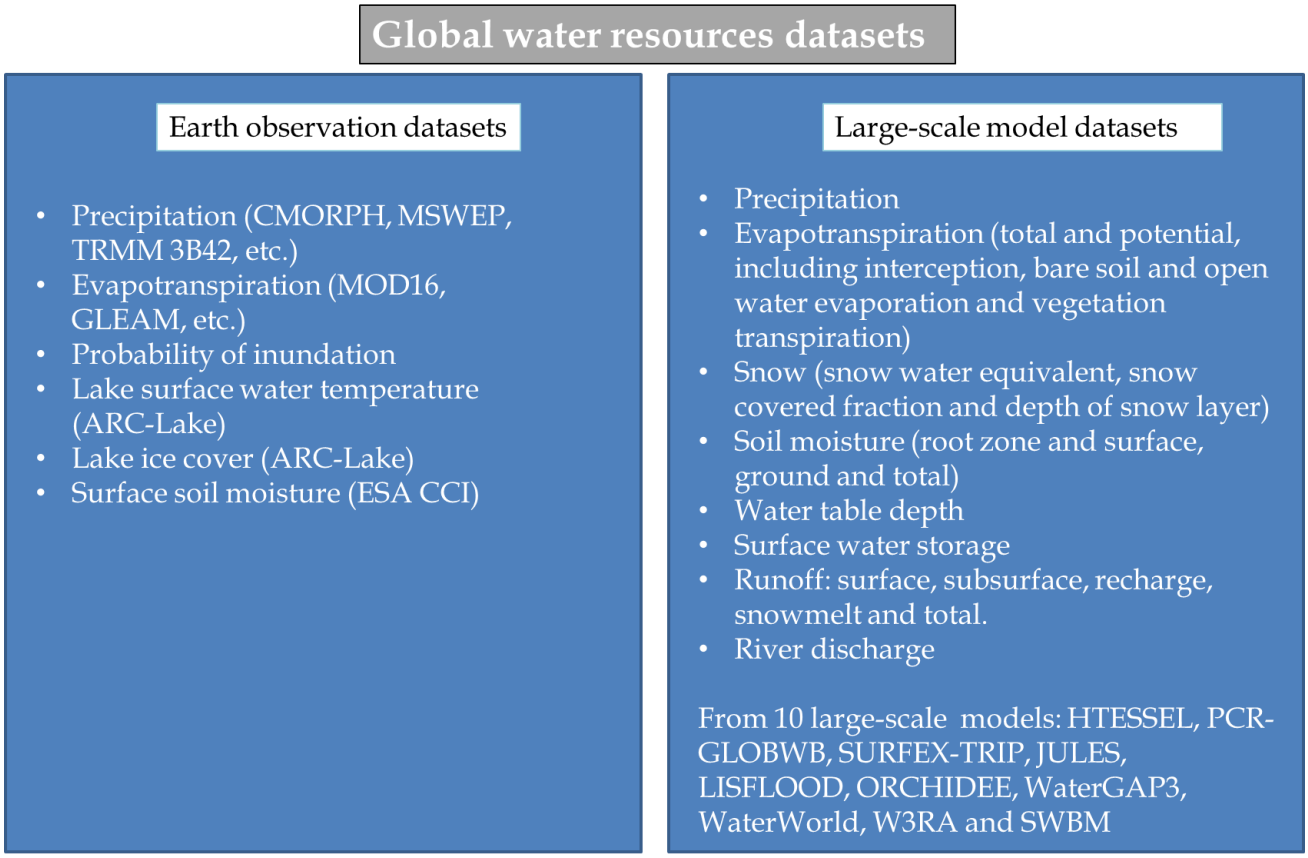


Figure 1.3 Global water resources datasets based on global Earth observation, in situ and large-scale model datasets constructed under the eartH2Observe project.

The main aim of the eartH2Observe project is to integrate available global Earth observations, in situ datasets and models to construct comprehensive and consistent global water resources datasets of several decades (at least 33 years length, from January 1979 until December 2012). The project has tested new Earth observation data sources, extended existing processing algorithms and combined data from multiple satellite missions in order to improve the overall spatial and temporal resolutions and reliability of Earth observation data included in the final datasets. The constructed datasets can particularly be of great value in data-sparse regions, where water management is constrained by a lack of in situ

information. The resulting global water resources datasets were structured in two groups: (i) Earth observation datasets and (ii) large-scale model datasets (*Figure 1.3*), which have been made available through an open Water Cycle Integrator data portal (WCI).

The usability and added value of the constructed datasets were verified and demonstrated in six case studies in eight different countries across the world aiming at improve the efficiency of regional water distribution. The case studies cover multiple continents, a variety of hydro-meteorological conditions and differ in degree of data richness (e.g. Mediterranean – Spain and Morocco –, Estonia, Ethiopia, Colombia, Bangladesh and Australia/New Zealand). For example, the Mediterranean case study focused on the usability of the datasets for assessing drought events comparatively for a data rich area (Spain) and a data poor area (Morocco), whereas the Bangladesh case study focused on flood events in a glacier-fed water system where monsoon periods govern intra-annual water availability.

This PhD research contributes to the local application and evaluation of the generated global water resources datasets in a selection of the case studies included in the eartH2Observe project (Morocco, Colombia, Bangladesh and Australia). To this end, it tests the applicability of resulting datasets in combination with large- and local-scale hydrological models using a variety of techniques, such as data assimilation, in order to improve water resources assessment in the case studies regions.

1.3 Research objective and research questions

The overall objective of this thesis is *to evaluate the applicability of global water resources datasets (including satellite-derived observations, in situ data and models) for hydrological modelling at the river basin scale*. In view of the complexity of this goal, this research focuses on applying and testing techniques that can potentially be implemented in any river basin around the world. Four case study areas were selected with different hydro-meteorological characteristics, water management conditions and in situ data richness: Brahmaputra basin in Bangladesh, Magdalena-Cauca basin in Colombia, Oum Er Rbia basin in Morocco and Murrumbidgee basin in Australia. To pursue this objective, the research presented in this thesis will attempt to answer the following specific research questions (all at the river basin scale):

1. What are the possibilities and limitations of large-scale hydrological models for water resources and flood estimation? (Chapter 2)

2. Does spatial precipitation downscaling improve the accuracy of satellite-derived precipitation datasets? (Chapter 3)

3. What is the impact of precipitation spatial resolution on streamflow model estimations? (Chapter 3)

4. Does calibration solely based on satellite-derived observations, such as soil moisture and evapotranspiration, improve large-scale model simulations? And how does this relate to calibration based on in situ discharge? (Chapter 4)

5. Is it possible to bring closer large-scale model simulations to those obtained with locally calibrated models with the assimilation of satellite-derived soil moisture observations? (Chapter 5)

6. Is it possible to improve large- and local-scale model simulations assimilating satellite-derived observations, such as soil moisture and discharge? And how does this relate to assimilating in situ discharge? (Chapters 5 and 6)

7. What is the improvement on model estimations that can be achieved assimilating satellite-derived observations in comparison to increasing the accuracy of meteorological forcing data (precipitation)? (Chapters 5 and 6)

This research contributes to a better understanding of the techniques needed to integrate global datasets into hydrological models for improving water resources estimation at the river basin scale. This research focuses on the science that is required to effectively use satellite-derived observations globally available in hydrology, in particular how downscaling, model parameters estimation and data assimilation techniques can be developed and applied. However, this research has not only scientific significance, but also societal and economic, as it contributes to ensure an adequate application of global information at a regional scale supporting water resources and hazard management (e.g. floods and droughts) over the past 35 years, which has a great added value for river basins where scarce in situ information is available.

1.4 Outline

This thesis is structured in seven chapters. In addition to the introduction (Chapter 1) and synthesis (Chapter 7) chapters, this thesis contains five other chapters corresponding to five peer reviewed publications that were produced during this study. These have already been published or are in the process of being published, contributing to the scientific understanding of the application and evaluation of global datasets in hydrological modelling. The sequence of chapters does not follow a chronological order. Instead, the publications are arranged according to the topic and following the same line exposed in *1.1 Background,* starting with those on model application (Chapters 2 and 3), moving to those on model parameters estimation and calibraton (Chapter 4) and ending with those on data assimilation (Chapters 5 and 6). In all the chapters the impact of forcing meteorological datasets on model estimates was investigated. A scheme of the hydrological models and the integration techniques applied in each chapter of this PhD thesis is shown in *Figure 1.4*.

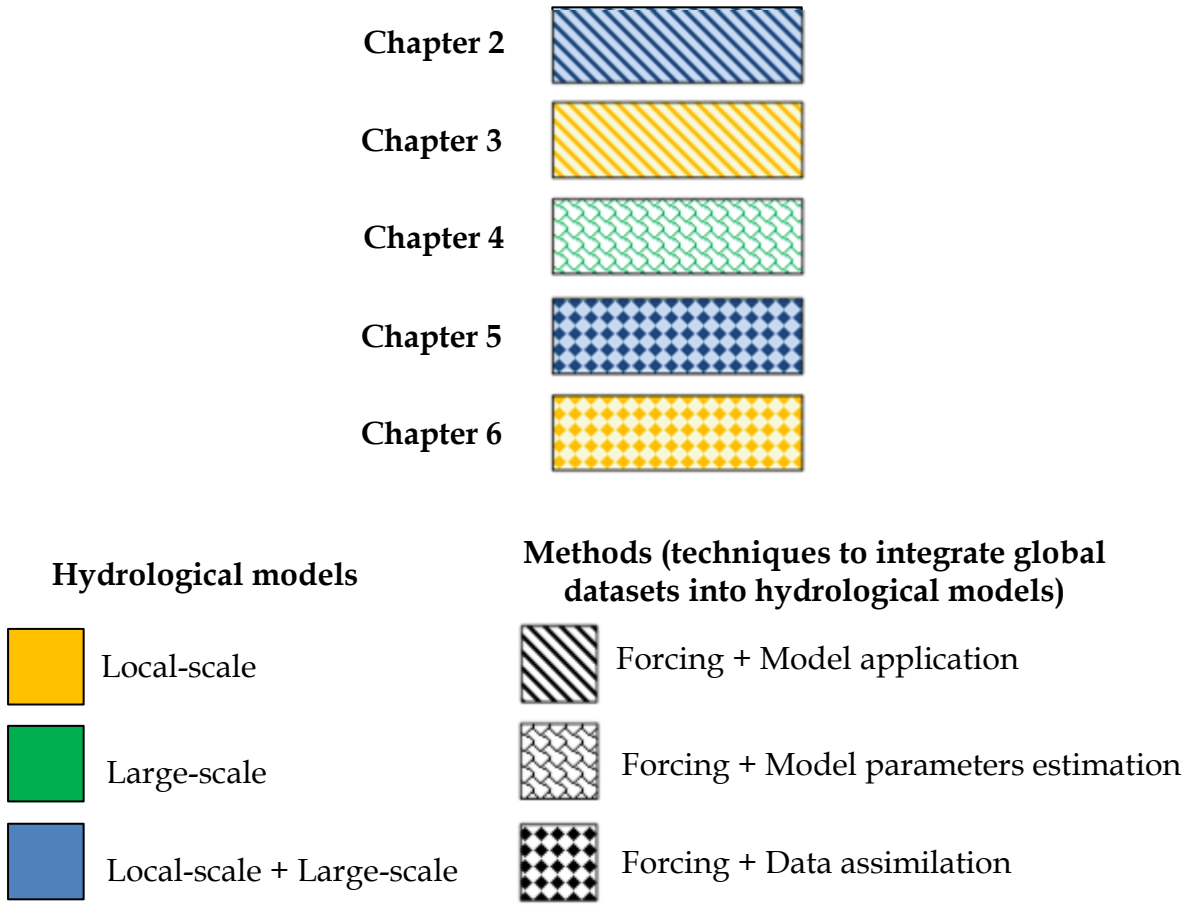


Figure 1.4 Scheme of the hydrological models and the integration techniques applied in each chapter of this PhD thesis.

Chapter 2 evaluates several large-scale hydrological models and global datasets, based on satellite-derived observations, in situ data and models, for water resources and flood estimation in the Brahmaputra River basin, compared to the available in situ data and a locally calibrated model.

Chapter 3 discusses how sensitive streamflow simulations of a locally calibrated hydrological model are to precipitation spatial resolution and data quality in the Magdalena-Cauca River basin.

Chapter 4 presents the results of calibrating a large-scale hydrological model using in situ gauged discharge and satellite-derived soil moisture and evapotranspiration observations in the Oum Er Rbia River basin. Furthermore, three different global precipitation products are compared and used as model forcing.

Chapter 5 investigates the impact of assimilating in situ gauged discharge and satellite-derived soil moisture observations on large-scale hydrological model simulations in the Murrumbidgee River basin, compared to the available in situ data and a locally calibrated model.

Chapter 6 investigates the impact of assimilating in situ gauged and satellite-derived discharge observations on local-scale hydrological model simulations in the Magdalena-Cauca River basin. Moreover, the hydrological model is driven by local and global precipitation datasets.

2 Evaluation of an ensemble of large-scale hydrological models for water resources and flood estimation

This chapter is based on:

López López P., Sultana T., Abdulla Hel Kafi Md., Shahadat Hossain Md., Saleh Khan A., and Sohel Masud Md. (in review). Evaluation of an ensemble of large-scale hydrological models for water resources and flood estimation in the Brahmaputra River basin. Water Resources Management.

Abstract

Estimation of quantum of flow in the Brahmaputra River basin is crucial for establishing effective water resources and flood risk management in Bangladesh. To this end, adequate hydro-meteorological datasets are needed. However, the area lacks of a consistent and dense network of stations to measure the required variables, such as precipitation and streamflow. In this study, an ensemble of large-scale hydrological models and datasets based on Earth observations and reanalysis were used for water resources and flood estimation in the Brahmaputra River basin for 1980 to 2012 period. Five large-scale models (HTESSEL, LISFLOOD, PCR-GLOBWB, SURFEX-TRIP and WATERGAP3) and two multi-model ensembles were evaluated at different temporal resolutions in terms of discharge and flood estimates and their performance was compared with a local-scale model (NAM+MIKE BASIN). Two global precipitation datasets (WFDEI and MSWEP) were used to drive all the hydrological models.

Results suggest that both global precipitation datasets underestimated rainfall in the pre-monsoon, monsoon and post-monsoon periods (April to October). Lower underestimation values were found with MSWEP precipitation dataset. Discharge simulations of all the large-scale models showed fairly reasonable agreement with streamflow observations, even better than those provided by the local-scale model. Flood evaluation results revealed that, in spite of magnitude differences, all the hydrological models could give acceptable estimates of peak flows, especially for the extreme flood events of 1987, 1988, 1998, 2004 and 2007. The multi-model ensembles were found to perform better than all the individual models. This study demonstrates the potential of large-scale hydrological models and datasets for water resources and flood management applications in data scarce regions.

2.1 Introduction

Bangladesh is the ultimate passage of the Ganges, Brahmaputra and Meghna (GBM) rivers. Its geographical location at the deltas of the GBM river basins and the regional meteorological conditions, convert Bangladesh into a highly flood prone area (Mirza, 2003). Historical data have shown that during the past forty years, Bangladesh has suffered extreme floods during the years of 1987, 1988, 1998, 2004 and 2007, making an average of one flood in six years (Ghatak et al., 2012). Problems derived from these extreme floods are increased due to the rapid population growth and the unplanned socio-economic development of the cities (Dewan, 2015; Haque and Jahan, 2015). According to the IPCC Special Report on Extreme Events (SREX; Murray and Ebi, 2012), more than 60% of Bangladesh was inundated during the 1987, 1988 and 1998 floods, causing thousands deaths and millions homeless.

Previous studies have pointed out the complexity of these extreme floods and have investigated possible causes (Ahmed and Mirza, 1998; Mirza, 2003). The main cause is the monsoon precipitation regime over the GBM river basins, which concentrates more than 80% of the country annual precipitation between June and September (Mirza, 2011; Bajracharya et al., 2015). Moreover, accelerated retreat of glaciers, El Nino Southern Oscillation (ENSO), backwater affects, deforestation and synchronization of flood peaks of the GBM rivers could be other factors leading to extreme flooding (Mirza, 2011).

Further research on the spatio-temporal variability of floods is needed to improve flood risk management in Bangladesh (Dewan, 2015). To this end, appropriate hydro-meteorological information and long-term flood records are crucial (Haque and Jahan, 2015). However, GBM river basins, especially the Brahmaputra River basin, lack of an adequate monitoring and observation systems of the required variables, such as precipitation and streamflow.

Recently, a suite of large-scale hydrological and land surface models have been developed, such as SURFEX-TRIP (Decharme et al., 2010), HTESSEL (Balsamo et al., 2009), LISFLOOD (Van Der Knijff et al., 2010), PCR-GLOBWB (Sutanudjaja et al., 2018) and WATERGAP3 (Döll et al., 2009); among others. Moreover, several global datasets based on Earth observations, in situ data and reanalysis have been generated, including precipitation (Beck et al., 2017a), soil moisture (Owe et al., 2008), snow (Hall et al., 2006) and total water storage (Tapley et al., 2004); among others. These global models and data may constitute an alternative or complement to in situ information in poorly or ungauged river basins, such as the Brahmaputra River basin.

In this study, an ensemble of large-scale hydrological models and datasets based on Earth observations, in situ data and reanalysis were used for monitoring water resources and flood estimation in the Brahmaputra River basin over the past 33 year-period from 1980 to 2012. Moreover, a comparison with a locally calibrated hydrological model was carried out providing insight into how large-scale hydrological models predictions perform in contrast to local-scale model estimates.

This study aims to evaluate an ensemble of large-scale hydrological models and datasets for water resources and flood estimation. Understanding the potential benefit of global hydrological models and Earth observations for water resources and flood management may have a great added value for river basins where scarce in situ hydro-meteorological information is available.

2.2 Study area

The Brahmaputra River basin (*Figure 2.1*) is one of the largest river basins in the world with an area of approx. 556,365 km^2. It includes territories from four different countries: China (50.4%), India (34.9%), Bangladesh (7.4%) and Bhutan (7.2%). The Brahmaputra River originates from the Hindu Kush Himalaya (HKH) region, which is one of the most dynamic and complex mountainous systems in the world. It flows from the great glacier mass of Chema-Yung-Dung in the Kailas range of southern Tibet at an elevation of 5,300 m, traversing China, India and Bangladesh before emptying into the Bay of Bengal, with a total length of 2,900 km (Bajracharya et al., 2015).

The Brahmaputra River drains diverse environments, including the cold dry plateau of Tibet, the rain-drenched Himalayan slopes, the landlocked alluvial plains of Assam and the vast deltaic lowlands of Bangladesh (Goswami, 2008). Three main physiographic regions can be distinguished inside the basin (Immerzeel, 2008): Tibetan Plateau (TP; > 3,500 m), Himalayan Belt (HB; 100 – 3,500 m) and Flood Plain (FP; < 100 m).

In the Brahmaputra River basin, around 80% of annual precipitation falls during the monsoon period (June to September), with mean annual precipitation of 2,300 mm. Spatial distribution of precipitation ranges from 1,200 mm in the central-eastern area to over 6,000 mm in the southern Himalayan slopes. Spatial distribution of average temperature varies from -1.5^{o}C in TP to 22^{o}C in FP. Average discharge of the Brahmaputra River is approx. 20,000 $m^3\ s^{-1}$.

Total water use estimated in the basin is around 27,457 $Mm^3\ y^{-1}$, out of which 2% is used in China, 1% in Bhutan, 43% in India and 54% in Bangladesh. Sector wise water uses in the basin are: 89% in agriculture, 9% in domestic uses and 2% in the industrial sector.

Water resources in Bangladesh almost entirely depend on the Brahmaputra River basin. Water resources management in the country is more and more problematic due to the growing demands by agriculture, fishery, inland navigation and the urban and industrial sectors. One of the main problems is that water availability is highly seasonal depending on the magnitude and duration of the monsoon period, resulting frequently in floods. These often cause widespread loss of life, crops, livelihood and properties. Several previous studies have attempted to analyze and improve flood risk management in the area (Paul, 1997; Islam and Sado, 2000; Mirza, 2001, 2003; Dewan, 2015; Haque and Jahan, 2015).

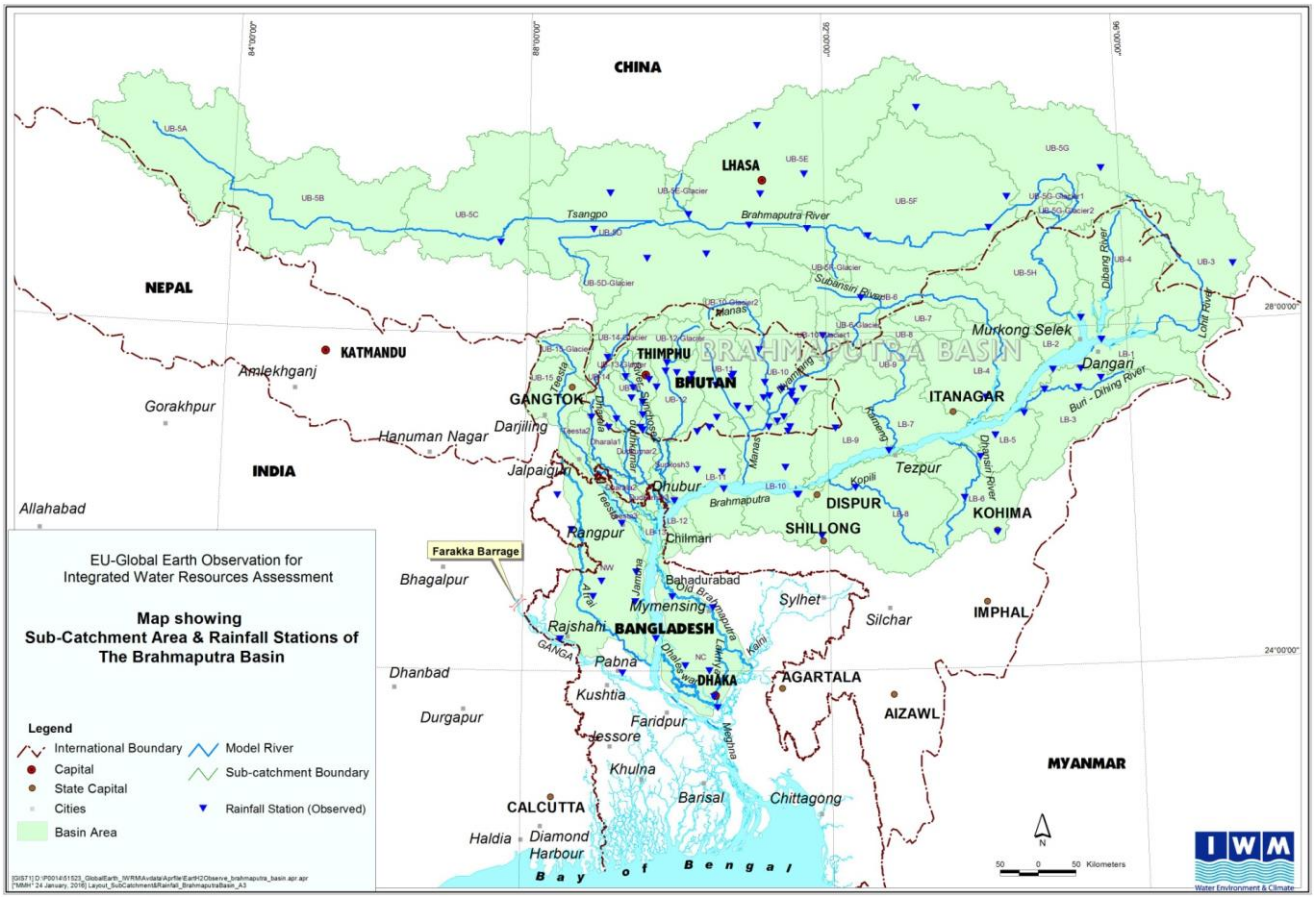


Figure 2.1 Map of the Brahmaputra River basin.

2.3 Hydrological models

2.3.1 Large-scale models

Five different large-scale models were used, including PCR-GLOBWB (van Beek et al., 2018), WATERGAP3 (Döll et al., 2009), SURFEX-TRIP (Decharme et al., 2010), HTESSEL (Balsamo et al., 2009) and LISFLOOD (Van Der Knijff et al., 2010). They are spatially distributed models providing a grid-based representation of the hydrological processes. Three of the models are Global Hydrological Models (GHMs): PCR-GLOBWB, WATERGAP3 and LISFLOOD and two are Land Surface Models (LSMs): HTESSEL and SURFEX-TRIP. None of the five large-scale models was locally calibrated for the Brahmaputra River basin. For a more detailed description of the models, the reader is referred to Beck et al. (2017b).

2.3.2 Local-scale model

MIKE 11 NAM (Nedbør Afrstrømnings Model) in combination with MIKE BASIN was used as local-scale hydrological model. NAM is a lumped conceptual rainfall-runoff model part of the MIKE 11 river modelling system which simulates the hydrological processes in sub-basins (Nielsen and Hansen, 1973; Havnø et al., 1995). In this study, the Brahmaputra River basin was divided into a total of 54 sub-basins (*Figure 2.1*) and MIKE 11 NAM was applied

per sub-basin to obtain runoff time series. These runoff values were used as input for MIKE BASIN. MIKE BASIN is a model to route runoff representing rivers and tributaries by a network of branches and nodes (DHI Water and Environment, 2001). From hereafter this model combination will be referred as NAM+MIKE BASIN.

NAM+MIKE BASIN, forced with in situ precipitation, was locally calibrated by the Institute of Water Modelling (IWM) in Bangladesh using in situ streamflow observations at the closest station to the outlet of the Brahmaputra River basin, Bahadurabad station (*Figure 2.1*), to obtain the optimal model parameters set. The model was calibrated for 2002-2007 period and it was validated for 2007-2014 period, due to in situ precipitation data availability.

2.4 Data

2.4.1 Meteorological data

Wind speed, air temperature, specific humidity, surface air pressure, shortwave radiation and longwave radiation datasets were based on the WATCH Forcing Data methodology applied to ERA-Interim reanalysis (WFDEI; Weedon et al., 2014). Two different precipitation datasets were used to force the models: (i) the WFDEI precipitation and (ii) the Multi-Source Weighted-Ensemble Precipitation (MSWEP; Beck et al., 2017a). MSWEP is the result of combining gauge, satellite and reanalysis data. Daily meteorological datasets were available from January 1980 to December 2012.

To evaluate WFDEI and MSWEP precipitation datasets, in situ precipitation measurements from 114 gauging stations (*Figure 2.1*) located in India (28), Bhutan (66) and China (20) were used. In situ precipitation data were provided by the Institute of Water Modelling (IWM) in Bangladesh from January 2002 to December 2012.

2.4.2 Discharge data

Daily discharge data were provided by the Institute of Water Modelling (IWM) in Bangladesh at Bahadurabad station (*Figure 2.1*) from January 1980 to December 2012.

2.5 Methodology

2.5.1 Evaluation of precipitation estimates

Initially, in situ precipitation values at the gauging stations were interpolated to obtain spatial maps covering the entire Brahmaputra River basin. The Thiessen polygon method was used for precipitation interpolation, due to its robustness and simplicity (Liu et al., 2015).

WFDEI and MSWEP precipitation datasets were evaluated by comparing them to in situ precipitation values at daily and monthly temporal resolutions for 2002-2012 period. This analysis was carried out both for the entire Brahmaputra River basin and for individual sub-basins. For the entire basin comparison, Root Mean Square Error (*RMSE*) and percent

bias between daily estimated and observed precipitation were calculated and the results were averaged per month. These metrics provide insight about how well WFDEI and MSWEP precipitation datasets reproduce the intra-annual variability of precipitation, with special attention to the pre-monsoon, monsoon and post-monsoon periods.

For the sub-basin to sub-basin comparison, in situ, WFDEI and MSWEP precipitation estimates were spatially averaged for each of the 54 sub-basins. Comparing average precipitation per sub-basin gives important information about how useful precipitation products are for hydrological modelling (Mei et al., 2014).

Several performance indicators were calculated between WFDEI and MSWEP and in situ precipitation estimates per sub-basin, including (i) continuous statistics: coefficient of determination (R^2), bias and *RMSE* and (ii) categorical statistics: Probability Of Detection (*POD*) and False Alarm Ratio (*FAR*). Continuous and categorical indicators were used to assess how WFDEI and MSWEP estimate the amount of precipitation and their rainfall detection capabilities, which are relevant for the posterior water resources and flood modelling.

2.5.2 Evaluation of discharge estimates

The five large-scale models were run globally as part of the eartH2Observe project in two phases. In phase 1, models were forced with WFDEI data at 0.5° spatial resolution. In phase 2, models were forced with MSWEP data at 0.25° spatial resolution. The most-up-to-date models with the newest components were used for each phase. For a more detailed description of phases 1 and 2, the reader is referred to Schellekens et al. (2016), Beck et al. (2017b) and Schellekens et al. (2017). The local-scale hydrological model was also run in the same two phases, although there were no modifications in the model components between phases 1 and 2. From hereafter, results derived from phase 1 and phase 2 will be referred as TIER 1 and TIER 2, respectively.

In order to analyze the integrated response of most hydrological processes and to assess surface water availability in the basin, simulated discharge of each individual model was evaluated. Various previous studies have proven that combining the outputs from multiple models usually improves model accuracy (Viney et al., 2009; Beck et al., 2017b). Therefore, in this study, not only simulated discharge time series of the individual models were evaluated, but also two combinations of the discharge estimates of the individual models into ensembles. First, the mean simulated discharge of the five large-scale models was calculated. From hereafter, the resulting discharge estimates will be referred as "MEAN". Second, simulated discharges of the individual large-scale models were bias corrected using a mean–standard deviation (μ– σ) matching. The bias corrected discharge values q_{sim}^{new} ($m^3\ s^{-1}$) were obtained as follows

$$q_{sim}^{new} = \frac{\sigma_{obs}}{\sigma_{sim}}(q_{sim} - \overline{q_{sim}}) + \overline{q_{obs}} \qquad (2.1)$$

where $\overline{q_{sim}}$ and $\overline{q_{obs}}$ are the means of the simulated and observed discharges ($m^3\ s^{-1}$) and σ_{sim} and σ_{obs} are the standard deviations of the simulated and observed discharges. The bias corrected mean simulated discharge of the five large-scale models was calculated. From hereafter, the resulting discharge estimates will be referred as "BS-MEAN". MEAN and BS-MEAN simulated discharges of the large-scale models were computed with TIER 1 and TIER 2.

Simulated discharge estimates were compared with in situ observations at Bahadurabad station. Various performance metrics were computed at daily and monthly temporal scales, including Kling-Gupta efficiency (*KGE*; Gupta et al., 2009) and Pearson's correlation coefficient (*r*). *KGE* equally measures bias and differences in amplitude and timing, whereas *r* is a measure of reproducing correct timing of high and low discharge values. Traditionally, Nash Sutcliffe efficiency (*NSE*; Nash and Sutcliffe, 1970) is used to compare modelled and observed discharge. However, in this study, *KGE* was used to avoid underestimating the variability of values.

2.5.3 Evaluation of flood estimates

Flood estimates were evaluated in terms of timing of occurrence and magnitude at Bahadurabad station for 1980-2012 period. The timing of occurrence of maximum daily peak flows was obtained using in situ discharge observations. Furthermore, critical peak flow periods were identified.

Next, the maximum of the daily discharge values was obtained per year. From hereafter, they will be referred as MAD (Maximum Annual Discharge). Simulated and observed MAD values were compared by visual inspection of the time series and the computation of four performance metrics, including *KGE, r, RMSE* and percent bias (Mediero et al., 2015).

Lastly, MAD values were ranked and fitted to four statistical distributions (*Table 2.1*), including Normal, Pearson Type III, two parameter Lognormal and Gumbel's Extreme value I. *RMSE* and *r* were used to test the fit of every distribution function to the data (Bezak et al., 2014). Simulated and observed flood frequency distributions were compared. For more detailed information on model fitting and selection, the reader is referred to Stasinopoulos and Rigby (2007) and Villarini et al. (2009).

2.5.4 Evaluation of the 1987, 1988, 1998, 2004 and 2007 extreme floods

According to previous studies (Mirza, 2003; Ghatak et al., 2012; Dewan, 2015), the five most extreme floods that have occurred in Bangladesh during the 1980-2012 period were in 1987, 1988, 1998, 2004 and 2007. These flood events devastated Bangladesh causing huge loss of life, crops and properties and thus, were further analyzed here.

Table 2.1 Summary of the probability distribution functions considered to model the Maximum Annual Discharge (MAD) values.

Identifier	Probability distribution function	Mathematical formulation
Normal	Normal	$x = \sigma Z + \mu$ where μ is the mean and σ is the standard deviation
PearsonIII	Pearson Type III	$x_T = \bar{x} + K_T S_x$ $K_T = \frac{2}{g_y}\left(1 + \frac{Z_T g_y}{6} - \frac{{g_y}^2}{36}\right)^3 - \frac{2}{g_y}$ where $K_T$is the frequency factor and Z_T is the standard normal variate
2P-Lognormal	Two parameter Lognormal	$x_T = e^{(\bar{y} + Z_T S_y)}$ $Z_T = \frac{(1-1/T)^{0.135} - (1/T)^{0.135}}{0.1975}$ where $\bar{y}$ is the mean of the y data, S_y is the standard deviation of the y data and $y = \ln(x)$
GumbelEV1	Gumbel's Extreme Value I	$x_T = u + \alpha y_T$ $u = \bar{x} - 0.5772\alpha$ $\alpha = S_x\sqrt{6}/\pi$ $y_T = -\ln(-\ln(1-1/T))$

The precipitation regimes associated to the five extreme floods were evaluated during the monsoon period (June to September). The multi-year monthly time average precipitation estimated with in situ, WFDEI and MSWEP precipitation datasets were compared with monthly precipitation values of every extreme flood over the entire basin. Absolute deviations of monthly precipitation from the long-term average were computed.

Models representation of the absolute amount, the timing and the duration of the peak flows during these extreme floods were assessed with additional performance indicators, including observed maximum discharge ($max.Qobs$) and differences between simulated and observed maximum discharge in magnitude ($dmax\ Qsim$), time (dT) and duration ($durT$).

2.6 Results

2.6.1 Evaluation of precipitation estimates

WFDEI and MSWEP precipitation estimates were compared with in situ precipitation over the entire Brahmaputra River basin and for the individual sub-basins. Results of the analysis for the whole basin are shown in *Figure 2.2*. MSWEP performed similarly to WFDEI in *RMSE* terms (*Figure 2.2* left), with values between 5.90 and 9.61 mm d^{-1}, an average of 7.72 mm d^{-1} during monsoon precipitation period (between June and September). In terms of percent bias (*Figure 2.2* right), MSWEP performed better than WFDEI. For MSWEP, percent bias suggested underestimation of precipitation by 17.53% in the pre-monsoon, monsoon and post-monsoon periods (April to October), whereas WFDEI underestimated precipitation by 24.12% during those periods.

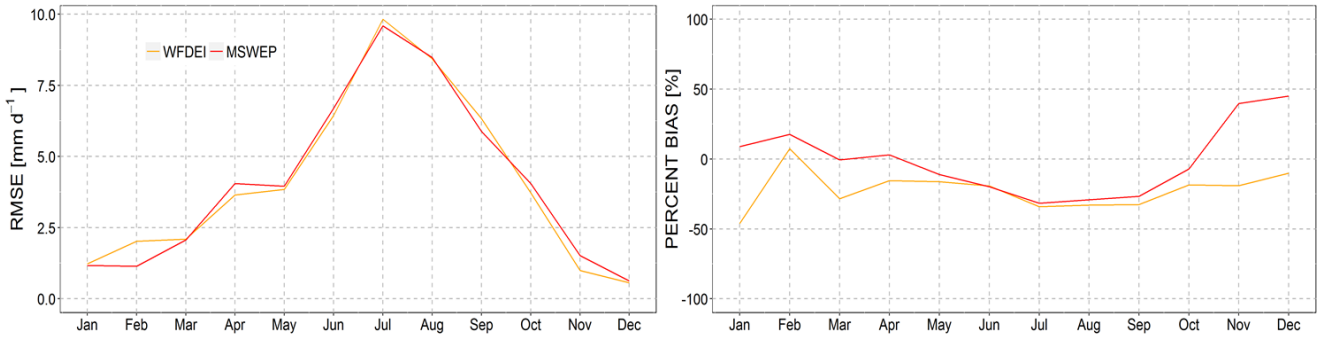


Figure 2.2 Monthly average of daily *RMSE* (left) and percent bias (right) between WFDEI (orange) and MSWEP (red) and in situ precipitation estimates for the entire Brahmaputra River basin for January 2002 – December 2012.

Results of the sub-basin to sub-basin comparison are shown in *Figure 2.3* and *Figure 2.4*. R^2, bias and *RMSE* spatial maps derived from daily precipitation estimates per sub-basin are shown in *Figure 2.3*. MSWEP showed an overall better performance than WFDEI in the entire basin, with R^2 values varying from 0.10 to 0.44. Higher R^2 values were found in the north and north western part of the basin (in the TP), possibly because of the limited number of rain gauge stations available in this region. Bias results indicated that WFDEI precipitation underestimated precipitation in most part of the basin, especially in the central and north-eastern areas (in the HB). Lower biases were found with MSWEP. *RMSE* obtained with WFDEI and MSWEP precipitation estimates were similar in spatial pattern, but of higher magnitude in the case of WFDEI.

Results of the daily categorical performance indicators are shown in *Figure 2.4* for a selection of three sub-basins with different hydro-meteorological characteristics located in the three main physiographic zones of the basin: UB-5E sub-basin in the TP, UB-07 sub-basin in the HB and LB-03 sub-basin in the FP. WFDEI and MSWEP provided relatively similar results, with overall average daily POD values of 0.97 during monsoon period. POD

values were lower during the dry season, especially for the sub-basins located in TP, such as UB-5E sub-basin (*Figure 2.4* upper left). In the TP, precipitation pattern is largely influenced by topography leading to heavy orographic rainfall events that are difficult to detect. WFDEI and MSWEP had more or less similar results for FAR estimation. FAR was lower during the monsoon months than in the rest of the year.

2.6.2 Evaluation of discharge estimates

Once WFDEI and MSWEP precipitation datasets had been evaluated, the capabilities of large- and local-scale models for monitoring water resources were assessed comparing their discharge estimates at Bahadurabad station. As explained in *2.5.2 Evaluation of discharge estimates*, TIER 1 results were derived in phase 1 of the eartH2Observe project with models forced with WFDEI data at 0.5° spatial resolution and TIER 2 results were derived in phase 2 with models forced with MSWEP data at 0.25° spatial resolution. Simulated and observed daily discharge time series from TIER 1 and TIER 2 for January 1995 – December 2005 are shown in *Figure 2.5*. On the whole, simulated discharge captured well the intra-annual variability of discharge observations. The highest discharge values during the monsoon precipitation period were underestimated with all the large- and local-scale models in TIER 1 and TIER 2 (with the exception of PCR-GLOBWB in TIER 2, which overestimated discharge). TIER 1 (orange) discharge estimates showed larger differences with in situ data, compared with TIER 2 (red).

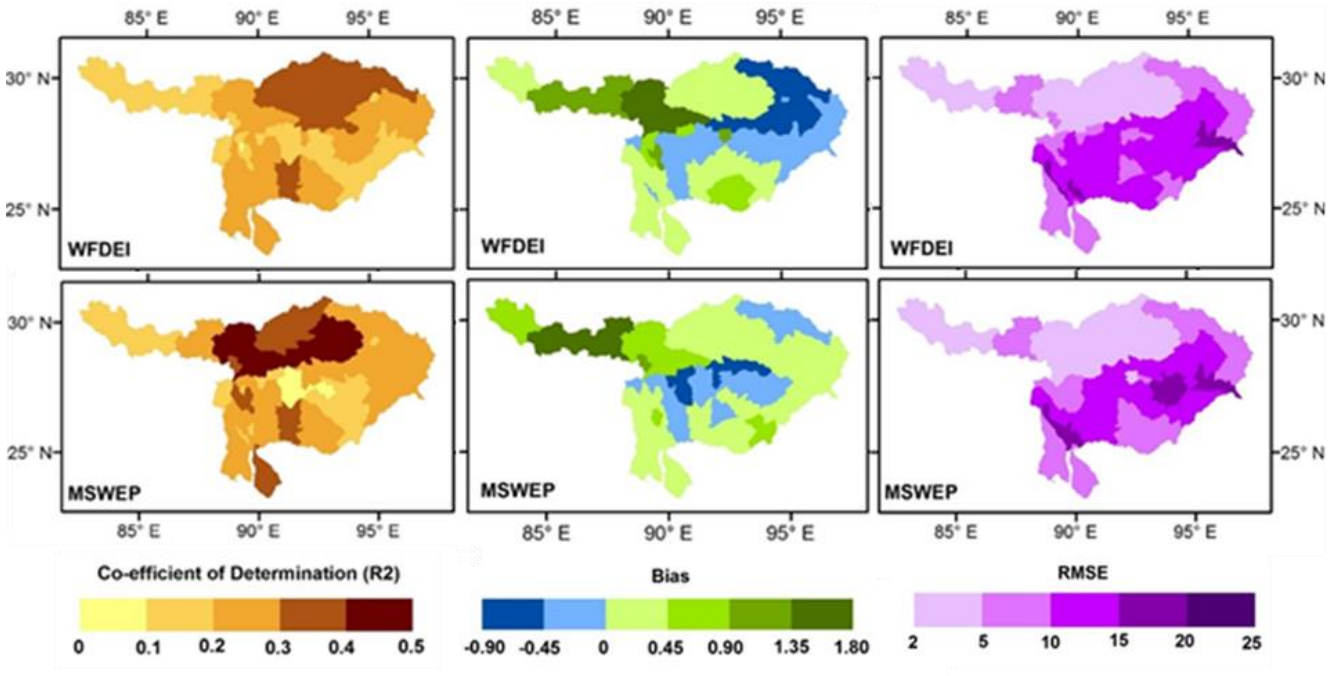


Figure 2.3 Daily R^2 [-] (left column), bias [-] (middle column) and *RMSE* [mm d^{-1}] (right column) between WFDEI (first row) and MSWEP (second row) and in situ precipitation estimates over the Brahmaputra River basin for January 2002 – December 2012.

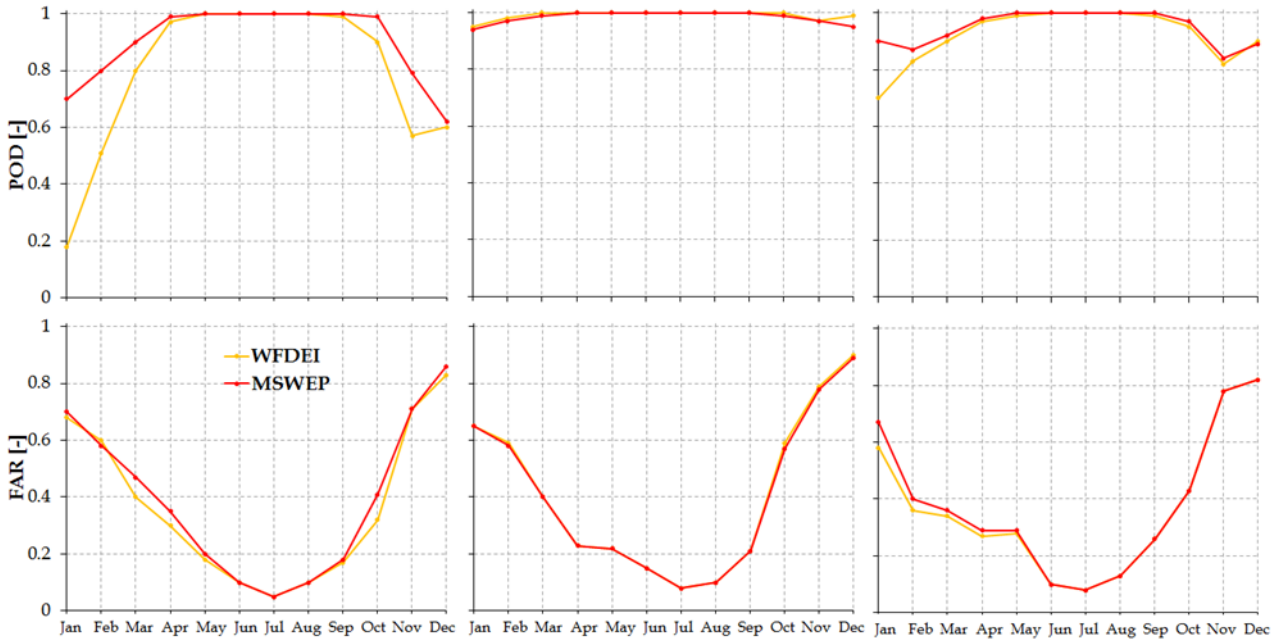


Figure 2.4 Monthly average of daily POD and FAR between WFDEI (orange) and MSWEP (red) and in situ precipitation estimates over three sub-basins UB-5E (left column), UB-07 (middle column) and LB-03 (right column) in the Brahmaputra River basin for January 2002 – December 2012.

KGE and *r* values between daily and monthly simulated and observed discharge are shown in *Table 2.2*. Simulations of the large-scale models in TIER 1 and TIER 2 showed fairly high *KGE* values, ranging from 0.42 to 0.77, and *r* values, varying from 0.84 to 0.96. An average increase of 0.16 in *KGE* was found when moving from TIER 1 to TIER 2 discharge estimates. Improvements in *r* values were of lower magnitude (average increase of 0.02). These results indicated that the change in forcing from TIER 1 to TIER 2 led to a better representation of discharge variability and reduced bias. Overall, the models performances improved when using TIER 2 compared with TIER 1.

KGE and *r* values obtained with the large-scale models simulations were similar or even higher than those obtained with the local-scale model, showing their potential for surface water estimation in the basin. The multi-model ensembles, MEAN (*Figure 2.5g*) and BS-MEAN (*Figure 2.5h*), accurately reproduced the timing of discharge estimates, with the highest daily and monthly *r* values. BS-MEAN showed the best performance in amplitude and timing, with *KGE* and *r* values exceeding 0.90.

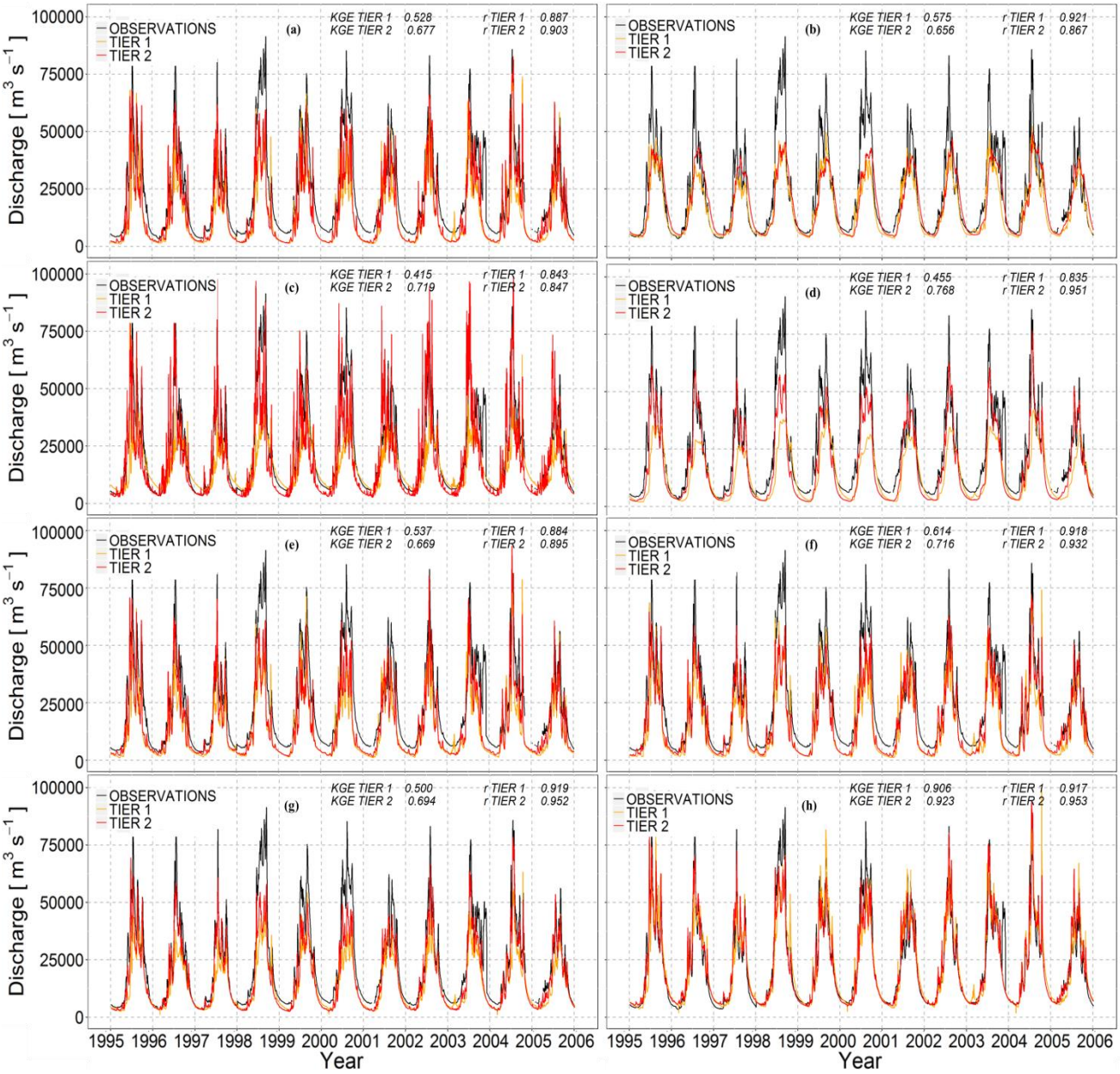


Figure 2.5 Daily observed and simulated discharge from TIER 1 (orange) and TIER 2 (red) at Bahadurabad station for January 1995 – December 2005. (a) HTESSEL, (b) LISFLOOD, (c) PCR-GLOBWB, (d) SURFEX-TRIP, (e) WATERGAP3, (f) NAM+MIKE BASIN, (g) MEAN and (h) BS-MEAN.

Table 2.2 KGE and *r* values between simulated and observed discharge from TIER 1 and TIER 2 at Bahadurabad station for January 1980 – December 2012.

	TIER 1				TIER 2			
	Daily		Monthly		Daily		Monthly	
	KGE [-]	*r* [-]	*KGE* [-]	*r* [-]	*KGE* [-]	*r* [-]	*KGE* [-]	*r* [-]
HTESSEL	0.53	0.89	0.51	0.95	0.68	0.90	0.66	0.96
LISFLOOD	0.58	0.92	0.59	0.95	0.66	0.87	0.69	0.89
PCR-GLOBWB	0.42	0.84	0.42	0.91	0.72	0.85	0.77	0.92
SURFEX-TRIP	0.46	0.84	0.48	0.87	0.77	0.95	0.78	0.96
WATERGAP3	0.54	0.88	0.52	0.94	0.67	0.90	0.66	0.95
NAM+MIKE BASIN	0.61	0.92	0.61	0.92	0.72	0.93	0.71	0.93
MEAN	0.50	0.92	0.50	0.95	0.69	0.95	0.76	0.97
BS-MEAN	0.91	0.92	0.94	0.95	0.92	0.95	0.96	0.97

2.6.3 Evaluation of flood estimates

Initially, timing of occurrence of floods and critical peak periods in the basin were investigated. The occurrence of maximum daily peak discharges was obtained using in situ discharge observations for the 1980-2012 period. Maximum daily discharge values occurred during the monsoon period, from June to September. Most of the peak flows were concentrated in July (44%) and August (34%). Maximum discharge values occurred 13% of the time in September and the remaining 9% occurred in June. Furthermore, two critical peak periods were identified per year: a first flood event that occurred between the end of July and the beginning of August and a late peak flow that arrived at the end of August. The MAD frequently occurred during the first critical peak period.

The representation of peak flows magnitude over time was analyzed. Simulated and observed MAD time series at Bahadurabad station for January 1980 – December 2012 are shown in *Figure 2.6*. The five most extreme floods occurred in 1987, 1988, 1998, 2004 and 2007 (*Figure 2.6*), with the highest peak flow in 1988 reaching a discharge value of approx. 100,000 $m^3 s^{-1}$ in only one day. Three flood-poor years were found in 1986, 1994 and 2006. Flood events of intermediate magnitude were also identified in 1984, 1991 and 1995.

The general behavior of the Brahmaputra River was poorly reproduced by all models in TIER 1. The tendency of the models to underestimate the peak flows was likely to be caused by the underestimation of precipitation of WFDEI compared to MSWEP (see *2.6.1 Evaluation of precipitation estimates*). Correspondence between observed and simulated MAD increased with the TIER 2 model runs. Exceptions like the years 1995-2000, where TIER 1 and TIER 2 simulations underestimated the flood severity, may be attributed either to errors in the hydrological models and in the precipitation data or in the analysed MAD time series.

LISFLOOD was barely able to reproduce flood time series with TIER 1 and TIER 2 (*Figure 2.6b*). Similar agreements between observed and simulated MAD series with HTESSEL (*Figure 2.6a*), SURFEX-TRIP (*Figure 2.6d*) and NAM+MIKE BASIN (*Figure 2.6f*) were found. WATERGAP3 (*Figure 2.6e*) simulated peak flows with the highest accuracy, surpassing the local-scale model performance. The highest correspondence between observed and simulated peak flows was obtained with BS-MEAN (*Figure 2.6h*).

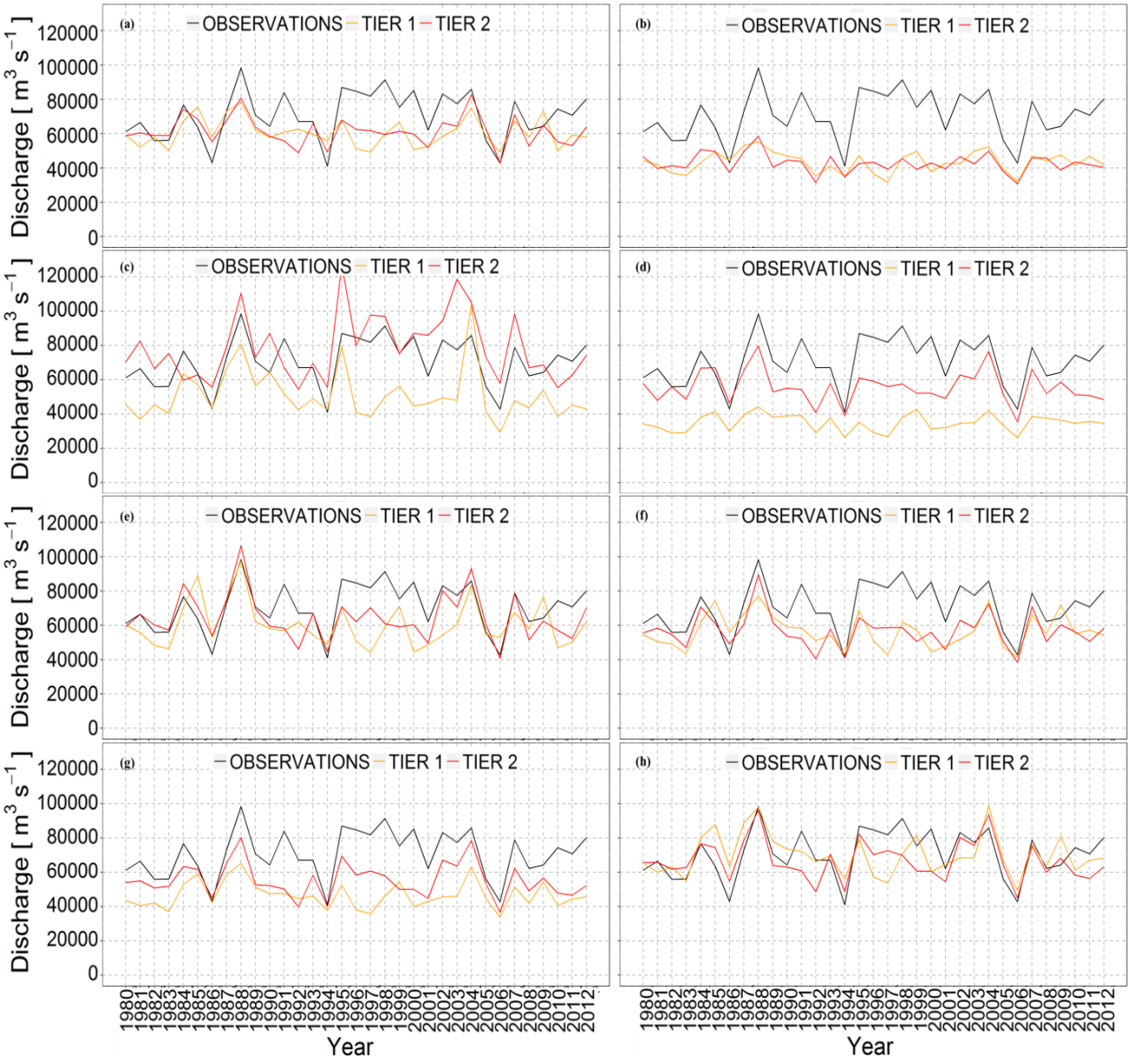


Figure 2.6 Simulated and observed MAD from TIER 1 (orange) and TIER 2 (red) at Bahadurabad station for January 1980 – December 2012. (a) HTESSEL, (b) LISFLOOD, (c) PCR-GLOBWB, (d) SURFEX-TRIP, (e) WATERGAP3, (f) NAM+MIKE BASIN, (g) MEAN and (h) BS-MEAN.

Apart from the visual inspection of peak flows in time, *KGE, r, RMSE* and percent bias between simulated and observed MAD were calculated and are shown in *Figure 2.7*. With all the models and combinations, considerably high values of *KGE* and *r*, and relatively low values of percent bias and *RMSE* were obtained (especially with TIER 2).

Discharge estimates with TIER 2 produced average increases of 48% in *KGE* and 40% in *r* and average decreases of 23% in *RMSE* and 22% in percent bias compared to TIER 1 results. The upgrading of structural components of the large-scale models and their increased spatial resolution from 0.5° to 0.25° could be the cause of these performance improvements. However, NAM+MIKE BASIN was not modified neither in spatial resolution nor model structure, and *KGE* and *r* values increased in 37% and 51%, respectively. Therefore, performance improvements may be mainly due to differences in WFDEI and MSWEP precipitation datasets (see *2.6.1 Evaluation of precipitation estimates*).

Between all the individual large- and local-scale models, WATERGAP3 showed the highest *KGE* (0.66) and *r* (0.67) values and the lowest *RMSE* (12692 m^3s^{-1}) and percent bias (-8.7%) values. The best values of all performance indicators were obtained with BS-MEAN.

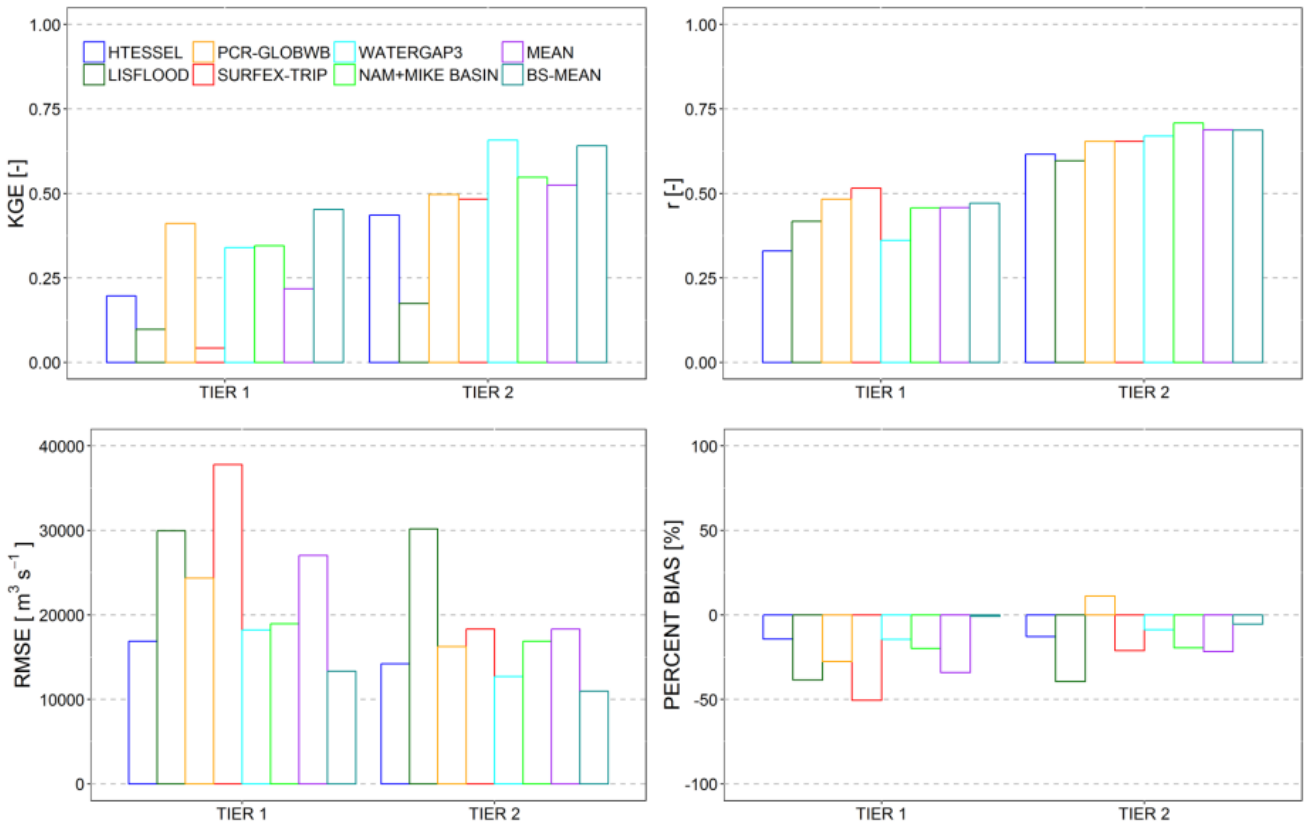


Figure 2.7 KGE, r, RMSE and percent bias between simulated and observed MAD from TIER 1 and TIER 2 at Bahadurabad station for January 1980 – December 2012.

Simulated and observed MAD time series were ranked and fitted to four probability distribution functions (*Table 2.1*) in order to derive the risk of occurrence of any flood event. In terms of *r* and *RMSE*, the best fit to data was provided by PearsonIII and 2P-Lognormal distribution functions. In *Figure 2.8*, results of simulated and observed MAD are shown for the four considered distribution functions. Individual local and large-scale models significantly underestimated the observed flood frequency (with the exception of PCR-GLOBWB TIER 2 simulations), with an average error in the 50-year flood discharge of 54%. The underestimation is reduced to 23% when using TIER 2 simulations. However, despite of this strong discrepancy in absolute values, observed and modeled frequency distributions showed a good correspondence. WATERGAP3 (between the individual models simulations) gave the closest discharge values to those obtained with discharge observations, especially for large return periods. BS-MEAN simulations reproduced the flood probability distributions with the highest accuracy.

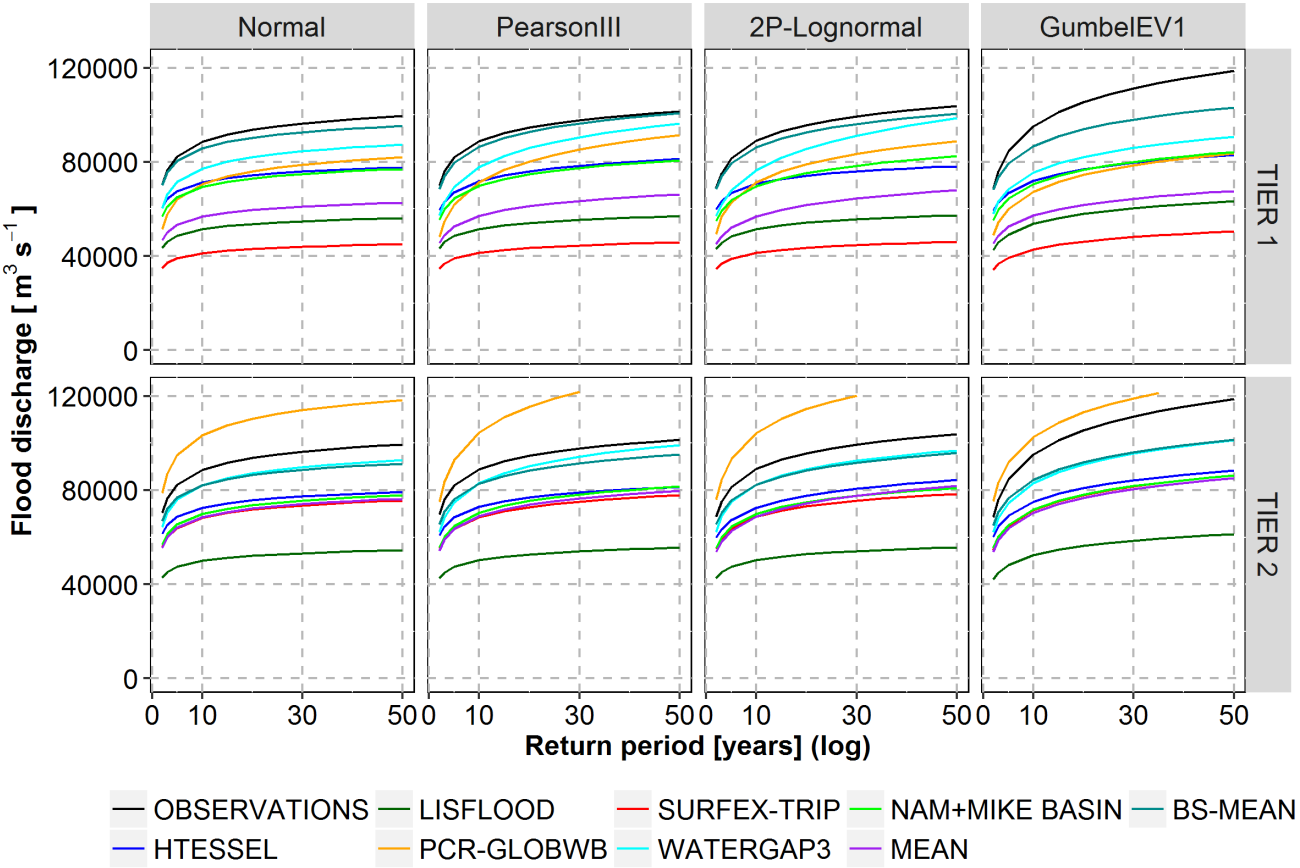


Figure 2.8 Simulated and observed flood frequency distributions derived from MAD from TIER 1 (first row) and TIER 2 (second row) at Bahadurabad station for January 1980 – December 2012.

2.6.4 Evaluation of the 1987, 1988, 1998, 2004 and 2007 extreme floods

Monthly precipitation values were obtained with in situ, WFDEI and MSWEP data and the multi-year monthly time average was calculated. Long-term average precipitation was compared with monthly precipitation of the five extreme floods during the monsoon period: June (*Figure 2.9a*), July (*Figure 2.9b*), August (*Figure 2.9c*) and September (*Figure 2.9d*).

In situ precipitation data was available for 2002-2012 period, thus only 2004 and 2007 extreme floods could be used in the comparison with WFDEI and MSWEP estimates. The WFDEI dataset notably underestimated precipitation. MSWEP precipitation estimates were closer to in situ observations, with an average bias of 7.32% (with the exception of the 2004 flood in July).

For the five extreme floods, monsoon precipitation was above the long-term average, especially during the months when most of the peak flows occurred (July and August). For the 1987 extreme flood, the highest precipitation values occurred in July (506.61 mm month^{-1}, *Figure 2.9*b), whereas in 1988 two waves of floods were identified, with higher precipitation in July (deviations of 21.71%) and August (deviations of 45.43%). During the 1998 extreme flood, precipitation peaked in July with maximum values of 418.177 mm month^{-1}. For the 2004 extreme flood, precipitation was 30.92% higher than the multi-year average, with maximum values in July (532.53 mm month^{-1}). During the 2007 extreme flood, the highest precipitation occurred also in July, with deviations from the long-term average of 11.71%.

Models representation of the absolute amount, the timing and the duration of the peak flows during the 1987, 1988, 1998, 2004 and 2007 extreme floods was further investigated. *Figure 2.10* shows the relative difference in magnitude, timing and duration between simulated and observed maximum annual discharge of the five most extreme floods during 1980-2012 period. The observed peak flows varied from 73000 m^3 s^{-1} in 1987 to 98300 m^3 s^{-1} in 1988. Maximum observed discharge in 1998 was 91404 m^3 s^{-1}. Peak flows of 85879 m^3 s^{-1} and 78877.4 m^3 s^{-1} were experienced in 2004 and 2007, respectively.

In terms of the simulated high flows, results revealed that all the models underestimated the five peaks ($dmax\,Qsim < 0$, *Figure 2.10* first row), with some exceptions. The underestimation is of lower magnitude with TIER 2 simulations. Furthermore, *Figure 2.10* (second row) shows that TIER 1 simulations of all models tended to bring forward flood peaks ($dT > 0$), with the exception of SURFEX-TRIP. TIER 2 simulations of all models estimates the timing of the peaks fairly well (with the exception of some models for the 2004 flood), with an average anticipation of less than 1 day.

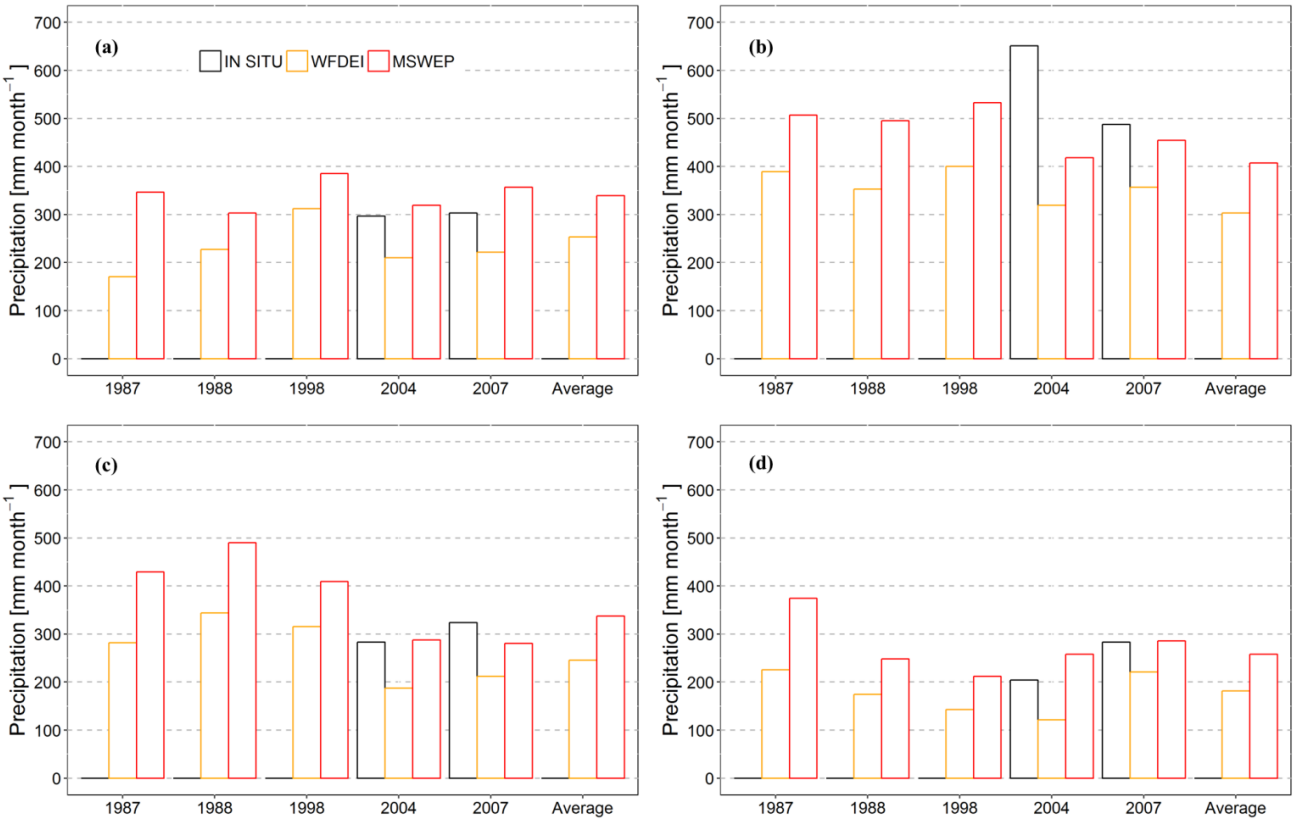


Figure 2.9 Monthly WFDEI (orange) and MSWEP (red) precipitation over the entire Brahmaputra River basin for the 1987, 1988, 1998, 2004 and 2007 extreme floods during the monsoon period: (a) June, (b) July, (c) August and (d) September.

Duration of discharge values above the daily Q_{95} (57,738.31 m^3 s^{-1}) were ultimately analyzed (*Figure 2.10* third row). The 1987, 1988, 2004 and 2007 extreme floods lasted 18, 31, 25 and 19 days, respectively, according to the observed discharge values. The long duration of the 1998 extreme flood of 82 days surpassed all past records. For most of July and from the beginning of August to the middle of September, discharge was consistently higher than the daily Q_{95} . Duration of discharge values above the daily Q_{95} obtained with TIER 1 and TIER 2 estimates were lower than those obtained with in situ data, with average differences in the order of 14 and 6 days, respectively (with the exception of the 1998 flood). All the models highly underestimated the duration of the 1998 flood.

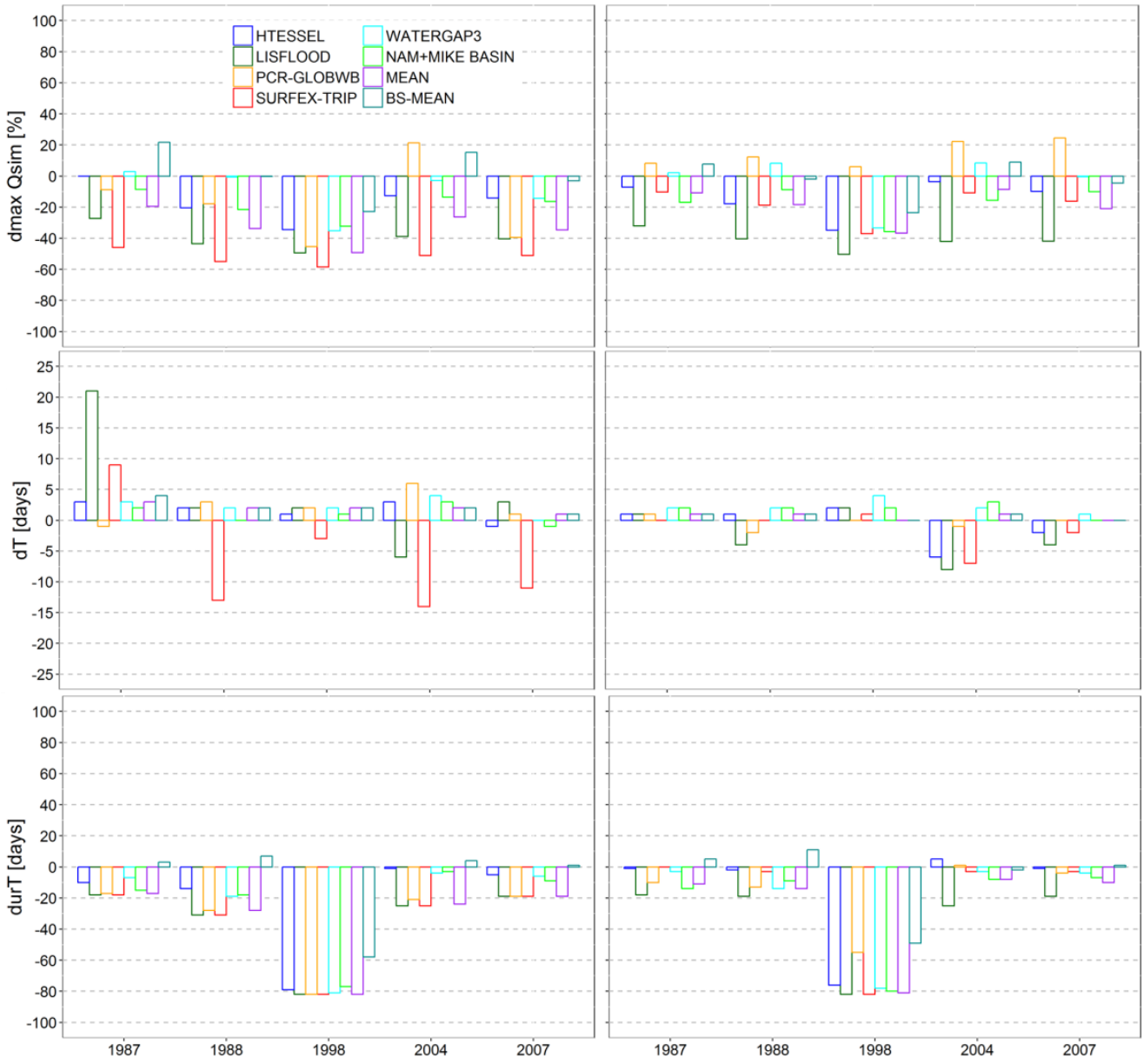


Figure 2.10 Difference between simulated and observed MAD in magnitude (*dmax Qsim*, first row), timing (*dT*, second row) and duration (*durT*, third row) at Bahadurabad station for the 1987, 1988, 1998, 2004 and 2007 extreme floods (TIER 1: left column and TIER 2: right column).

2.7 Discussion

In this study, an ensemble of large-scale hydrological models and Earth observations and reanalysis datasets in comparison with a local-scale model were used for water resources and flood estimation. Firstly, WFDEI and MSWEP precipitation datasets were evaluated. Results revealed acceptable precipitation estimates with both global products, in line with previous studies in the Brahmaputra River basin (Immerzeel et al., 2008; Bajracharya et al., 2015). Additional use of Earth observations, especially to improve the quality and

availability of precipitation datasets, may highly benefit water resources and flood monitoring (Atkinson, 2013; Beck et al., 2017a).

Secondly, five large-scale models, two multi-model combinations of them and a local-scale model were evaluated in terms of discharge and flood estimates. Simulated discharge showed a good agreement with in situ observations. Flood events were well captured with most of the models, but a consistent underestimation was found. The individual large-scale models simulated peak flows similarly or even better than the local-scale model. The multi-model ensembles provided the closest discharge and flood estimates to the observations (especially BS-MEAN). These results are in agreement with previous studies emphasizing the benefits of multi-model combinations (Viney et al., 2009; Yang et al., 2015). Therefore, it would be worthwhile to stimulate the development of global models and Earth observations and to further explore more sophisticated multi-model combination techniques, e.g. different bias correction methods (Bohn et al., 2010; Bierkens, 2015).

Discharge and flood evaluation results showed models performance improvements when moving from TIER 1 to TIER 2 simulations. Possible reasons behind this improvement could be the upgrading of the large-scale model components, the increase in the spatial resolution from 0.5^{o} to 0.25^{o} and the use of different precipitation datasets (WFDEI and MSWEP). When increasing the spatial resolution, a better delineation of the basin and its sub-basins, a better characterization of the drainage network and a more accurate parameterization of soil and land cover are achieved improving models performance. This improvement was found not only in the simulations of the large-scale models, but also in those from the local-scale model, which was not modified neither in its structural definition nor the spatial resolution. Therefore, precipitation may be considered the main driving force behind the better models performance. Alternative approaches, such as running the large-scale models at different spatial resolutions with the same precipitation dataset or vice versa, could allow to better quantify the impact of every modelling aspect (Lehner et al., 2006; Te Linde et al., 2008).

Previous studies on analyzing and estimating floods have pointed out the limitations of using the Maximum Annual Discharge (MAD) as the unique hydrological variable to detect extreme flood events. Some relatively large flood events could be not identified with MAD, as only one flood per year is considered. Moreover, the MAD series results in usually rather small samples of flood events (Bezak et al., 2014). The Peaks Over Threshold (POT) appears as an alternative variable to identify flood events. In the POT series, all the flood events that exceed a given threshold or base level are included, regardless of the time that they occurred (Lang et al., 1999). This selection criterion entails to consider a wider range of events and it gives the possibility to control the number of flood events to be included in the analysis (Mediero et al., 2014). POT and seasonal maximum discharge series could be used as ancillary data for flood detection (Mediero et al., 2015).

Furthermore, MAD series were fitted to four probability distribution functions to obtain the risk of occurrence of a flood event. Other probability functions, such as the Generalized

Pareto distribution or the Extreme Value Type I distribution, could be tested and compared (Halbert et al., 2016; Nagy et al., 2017).

Special attention was paid to the 1987, 1988, 1998, 2004 and 2007 extreme floods. Results are in line with previous studies on the impact of historical floods in Bangladesh, with the most devastating events in 1988, due to the occurrence of two precipitation peaks in July and August, and 2004, due to the high precipitation values occurred in July (Dewan, 2015). Peak flow during the 1998 flood was considerably high, even though precipitation within the Brahamputra River basin was not extreme. Mirza (2003) explained that the high flows of 1998 flood were due to the times of occurrence of the peak flows in the Ganges, Brahmaputra and Meghna Rivers, which occurred in the same month (September) and only two days apart. This situation was unusual, as only 13% of the maximum discharge values occurred in September. Similarly, during the 1988 flood event, the Ganges and Brahmaputra Rivers peaked at the same time, originating extremely high flows. Therefore, when evaluating extreme floods in Bangladesh, the interaction between the Ganges, Brahmaputra and Meghna Rivers should be taken into account.

Due to the limited availability of in situ precipitation data, different time periods were used for precipitation (10-year period from 2002 to 2012) and discharge and flood evaluation (33-year period from 1980 to 2012). Moreover, estimates of discharge and peak flows were solely evaluated at the location of Bahadurabad station. More studies using a wider spectrum of sub-basins within the Brahmaputra River basin and for longer time periods may provide additional information about hydro-meteorological temporal and spatial variability in the basin for a more exhaustive evaluation of models. However, this research serves as an example of the added value of global hydrological models and Earth observations for estimating water resources and flood events within a poorly gauged river basin.

2.8 Conclusions

This study investigated how an ensemble of large-scale hydrological models driven with global remotely sensed and reanalysis datasets can estimate water resources and floods in the Brahmaputra River basin, compared with a locally calibrated model. The conclusions of the study are as follows:

(i) MSWEP and WFDEI precipitation datasets underestimated precipitation in the pre-monsoon, monsoon and post-monsoon periods (April to October) over the entire basin, especially in the central and north-eastern areas. Lower underestimation values in the order of 17% were found with MSWEP precipitation dataset.

(ii) TIER 1 and TIER 2 discharge model simulations showed reasonable correspondence with observations at monthly and daily temporal scale. An average increase of 0.16 in *KGE* was found when moving from TIER 1 to TIER 2 discharge estimates. These results suggest an important reduction of the bias and a better representation of the discharge variability with TIER 2 simulations, leading to an overall upgraded performance.

(iii) TIER 1 simulations poorly reproduced peak flows, showing a tendency to underestimate extreme floods. In spite of the strong discrepancy in flood magnitude, the ensemble of large-scale models and the local-scale model could estimate peak flows quite well. Flood estimates were highly improved with TIER 2 simulations, with average increases of 48% in *KGE* and 40% in *r* and average decreases of 23% in *RMSE* and 22% in percent bias compared to TIER 1 simulations.

(iv) Individual large- and local-scale models were able to estimate the extreme floods of 1987, 1988, 1998, 2004 and 2007. Magnitude of the peak flows was underestimated, whereas their timing was fairly well captured, with an average anticipation of less than 1 day.

(v) Flood evaluation results revealed that, despite of magnitude differences, models could provide acceptable estimates of flood probability distributions, especially with PearsonIII and 2P-Lognormal distributions.

(vi) The multi-model ensemble of the large-scale models, BS-MEAN, was found to perform better than all the individual models for discharge and flood estimation.

3 Impact of high spatial resolution precipitation on streamflow simulations

This chapter is based on:

López López P., Immerzeel W. W., Rodríguez Sandoval E. A., Sterk G. and Schellekens J. (in review). Impact of high spatial resolution precipitation on streamflow simulations. Frontiers in Earth Sciences.

Abstract

Precipitation is one of the most important components of the water cycle and its accurate spatial and temporal representation is fundamental for hydrological modelling. In the present study, we investigated the impact of spatial resolution of various precipitation datasets on discharge estimates. First, a new precipitation spatial downscaling procedure was developed and applied to four gridded global precipitation datasets based on (i) solely satellite observations: CMORPH and PERSIANN, (ii) satellite and in situ observations: TRMM and (iii) satellite and in situ observations and reanalysis data: MSWEP. The here presented downscaling methodology blended global precipitation datasets with data on vegetation and topography to improve the representation of precipitation spatial variability. Second, interpolated in situ, non-downscaled (25 km) and downscaled (1 km) precipitation data were used to force a grid-distributed version of the HBV-96 rainfall-runoff model for the Magdalena River basin in Colombia.

Results showed that MSWEP and TRMM outperformed CMORPH and PERSIANN precipitation datasets. The downscaling procedure resulted in considerable improvements in coefficient of determination, root mean square error and bias in comparison with in situ precipitation observations. Discharge model estimates were also in better agreement with the observations when the model was forced with the downscaled precipitation. Model performance was improved with Kling Gupta efficiency increases in the order of 0.1 to 0.5. Moreover, better discharge simulations were obtained using downscaled precipitation compared to using only in situ precipitation data when using less than 100 stations.

3.1 Introduction

Precipitation is a key component of the water cycle, playing a crucial role in hydro-meteorological and environmental processes (Goovaerts et al., 2000; Schuurmans and Bierkens, 2007; Langella et al., 2010). An accurate knowledge of precipitation is essential for water resources management and to predict extreme weather events, such as floods and droughts (Arnaud et al., 2002; Vischel and Lebel, 2007; Tramblay et al., 2011).

Precipitation datasets can be obtained from in situ weather stations, but many river basins around the world are still poorly gauged (Loukas and Vasiliades, 2014) or ungauged (Sivapalan et al., 2003). Moreover, measurements from in situ gauges are only representative over a limited distance around the location of the instruments (Collischonn et al., 2008; Bohnenstengel et al., 2011) and their location is often biased towards accessible lower lying areas. These conventional ground observations cannot effectively capture the spatial variability of precipitation and therefore, may be insufficient/non reliable for hydrological modelling (Javanmard et al., 2010).

Satellite-based precipitation observations may be an attractive alternative to in situ measurements. Remote Sensing (RS)-derived precipitation datasets cover large areas (many of them have a near-global coverage) over long time periods, reflecting the spatial patterns and temporal variability of precipitation (Adler et al., 2001). A series of gridded global precipitation datasets, including Earth observations, in situ datasets and models, have been recently developed at a regional and global scale: the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN; Hsu et al., 1997), the Climate Prediction Center morphing method to derive precipitation (CMORPH ; Joyce et al., 2004), the Global Satellite Mapping of Precipitation project (GSMaP ; Kubota et al., 2007), the Tropical Rainfall Measuring Mission (TRMM; Huffman et al., 2007), the European Centre for Medium-Range Weather Forecasts global atmospheric reanalysis data (ECMWF ERA-Interim ; Dee et al., 2011), the NASA Global Precipitation measurements mission (GPM; Hou et al., 2014), the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data (WFDEI; Weedon et al., 2014), the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS; Funk et al., 2015) and the Multi-Source Weighted-Ensemble Precipitation (MSWEP; Beck et al., 2017a), among others.

However, the spatial resolution of these precipitation datasets may still be too coarse for hydrological applications at the smaller river basin scale. Several previous studies have developed downscaling, interpolation and aggregation methodologies to increase the spatial resolution of RS-derived precipitation, often in combination with in situ observations (see overview in Table 1). The resulting precipitation datasets may provide a better representation of the spatial variability of precipitation to be used for hydrological applications.

Table 3.1 Summary of relevant studies where downscaling methodologies are developed to increase the spatial resolution of RS-derived precipitation datasets.

Study	Study area	Study period – time scale	Precipitation dataset	Model	Predictors	In situ observations	Residuals interpolation
Immerzeel et al. (2009)	Iberian Peninsula (Spain and Portugal)	2001-2007 – yearly	TRMM 3B42 and TRMM 3B43	Simple exponential regression	NDVI (SPOT)	–	Spline tension interpolator
Quiroz et al. (2011)	Altiplano centred around Lake Titicaca (Peru and Bolivia)	1999-2003 – daily	In situ	Wavelet analysis	NDVI (SPOT)	–	–
Jia et al. (2011)	Qaidam basin (China)	1999-2009 – yearly	TRMM 3B43	Multiple linear regression	DEM (SRTM) and NDVI (SPOT)	–	Spline tension interpolator
Duan and Bastiaanssen (2013)	Lake Tana basin (Ethiopia) and Caspian Sea region (Iran)	1998-2004 and 1999-2003 – yearly with monthly disaggregation	TRMM 3B43	Simple non-linear regression (second order polynomial and power functions)	NDVI (SPOT)	Geographical Differential Analysis and Geographical Ratio Analysis calibration methods	Spline tension interpolator
Fang et al. (2013)	Xiao River basin (China)	Six rainstorm events between 2001 and 2010 – daily	TRMM 3B42	Multiple linear regression	DEM (SRTM), pre-storm maximum temperature, humidity, slope, aspect	–	Spline tension interpolator

					and topographical roughness		
Park (2013)	South Korea	October 2005 – monthly	TRMM 3B43	Multiple linear regression	DEM and NDVI	–	Area to point simple kriging
Hunink et al. (2014)	Tungurahua province as part of the Patate River basin (Ecuador)	1998-2011 – weekly	TRMM 3B42	Multiple linear regression	DEM (SRTM), NDVI (SPOT) and TRMM 2B31	Spatial correction with in situ observations	Spline tension interpolator
Chen et al. (2014)	North China	2008-2011 – yearly	TRMM 3B43	Geographically weighted regression	DEM (SRTM) and NDVI (SPOT)	–	Spline tension interpolator
Teng et al. (2014)	Zheijiang Province (China)	2011-2012 – daily	TRMM 3B42	Regression kriging	DEM (ASTER GDEM) and rain gauge data	Regression of rain gauge data	Regression kriging
Shi and Song (2015)	Tibetan Plateau (China, Nepal, Bhutan and India)	2001-2012 – yearly with monthly disaggregation	TRMM 3B43	Random forest algorithm (non-parametric statistical regression)	DEM (SRTM), EVI (MOD13A3), aspect, slope, latitude and longitude	Geographical Differential Analysis and Geographical Ratio Analysis calibration methods	Spline tension interpolator
Ceccherini et al. (2015)	South America and West Africa	1951-2015 (various) – yearly	GPCC, TRMM 3B43, PERSIANN CDR, CMOPRH, CHIRPS, RFE	Geographically weighted regression	DEM (SRTM) and EVI (MOD13A3)	–	Cubic spline interpolation

			and TAMSAT				
Xu et al. (2015)	South China	1998-2013 – monthly	TRMM 3B43	Combination of multi-fractal analysis and stepwise and ANN regression models	DEM (SRTM), latitude and longitude	–	–
Long et al. (2016)	Qinghai Lake basin (China)	1998-2016 – daily	TRMM 3B42	Shuffled Complex Evolution (SCE)	DEM (SRTM), topographical ruggedness index, slope-aspect-wind direction angle, maximum air temperature and average humidity	Double kernel smoothing model	Spline tension interpolator
Ezzine et al. (2017)	Morocco	1998-2012 – yearly	TRMM 3B43	Stepwise multiple regression	DEM (SRTM), NDWI (SPOT) and distance from sea	–	Spline tension interpolator

Previous studies have demonstrated that spatial distribution of precipitation is one of the main sources of uncertainty in hydrological modelling (Berne et al., 2004; Sangati and Borga, 2009). The impact of spatial representation of precipitation on hydrological model estimates is complex and it depends on the type of precipitation (Bell and Moore, 2000), the hydrological characteristics of the basin (soils, geology, river morphology, vegetation cover, etc.), the hydrological model structure (Koren et al., 1999) and the considered spatial and time scales (Segond et al., 2007). Some studies have stated that the impact of precipitation spatial resolution on streamflow simulation was not significant (Gascon et al., 2015; Niaka et al., 2017). However, most of the studies (Andréassian et al., 2001; Smith et al., 2004; Schuurmans and Bierkens, 2007; Wagener et al., 2007; Fu et al., 2011; Zocatelli et al., 2011; Arnaud et al., 2011; Emmanuel et al., 2012; Zhao et al., 2013; Lobligeois et al., 2014) have shown that better model performances were obtained when representation of the spatial variability of precipitation was improved. Hence, accurate precipitation datasets at relative fine spatial resolution are needed to obtain reliable model estimates.

The present study aimed to analyze how sensitive streamflow simulations of a distributed hydrological model are to precipitation spatial resolution and data quality in the Magdalena River basin in Colombia. To this end, two main steps were taken: (i) estimating high spatial resolution precipitation datasets using a combination of downscaling frameworks and (ii) forcing a distributed hydrological model with coarse and high spatial resolution precipitation datasets and assessing its performance.

Four different daily RS-derived precipitation datasets at a spatial resolution of 25 km, including CMORPH, MSWEP, PERSIANN and TRMM, were downscaled to 1 km and combined with ground observations from March 2000 to December 2012. The downscaling methodology was based on earlier work of Duan and Bastiaanssen (2013), Hunink et al. (2014) and Ceccherini et al. (2015) and auxiliary information from vegetation response, elevation, slope and aspect at a fine spatial scale was used. A Geographically Weighted Regression (GWR) algorithm (Chen et al., 2014; Ceccherini et al., 2015), in which regression parameters varied with location, was applied. Furthermore, high resolution RS-derived precipitation was corrected with in situ observations to improve the accuracy of precipitation estimates (Cheema and Bastiaanssen, 2012; Long et al., 2016).

To evaluate the impact of precipitation spatial resolution on streamflow simulations, interpolated in situ, non-downscaled and downscaled RS-derived precipitation datasets were used to force the distributed hydrological model OpenStreams wflow-hbv. Lastly, the effect on streamflow of a decrease in the number of weather stations used for precipitation downscaling was investigated. More than 40 hydrological simulations were carried out with different network densities of precipitation data (Vischel and Lebel, 2007; Bardossy and Das, 2008).

The new contributions of this work include the comparison of four recently developed precipitation datasets, and the exploration of the impact of their spatial variability in a distributed hydrological model used to estimate streamflow in the Magdalena River basin. Furthermore, understanding the influence of the number of ground measurements used for

precipitation downscaling on streamflow estimations may have broader implications for similar but data-poor river basins globally.

3.2 Study area

The study area is the Magdalena River basin (*Figure 3.1*), which is the largest fluvial system in Colombia, draining an area of approx. 257,000 km^2 (about 24% of the total territory of the country). The Magdalena River originates from headwaters in the Colombian Andes at an elevation of 3,700 m and it runs for 1,612 km into the Western Caribbean, in the Atlantic Ocean (Restrepo and Kjerfve, 2000). The main tributary of the Magdalena River is the Cauca River on the Western part of the watershed.

Average annual precipitation for the Magdalena River basin is approx. 2,150 mm year^{-1}, with large inter-annual variability, especially due to the effect of the El Niño-Southern Oscillation (ENSO) phenomenon (Hoyos et al., 2013). Precipitation ranges from 1,000 mm year^{-1} in the eastern mountains to more than 5,000 mm year^{-1} in the western region of the basin. The climate in the basin is characterized by two wet periods (March-May and October-November) and two dry periods (December-March and June-September). Average annual air temperature is approx. 28^{o}C and average annual evapotranspiration is approx. 1,630 mm year^{-1}. Average annual discharge at the outlet of the basin is approx. 7,200 m^3 s^{-1}, varying from 4,050 m^3 s^{-1} in March to 10,200 m^3 s^{-1} in November (Camacho et al., 2008).

The main cities of Colombia, including Cali, Bogotá, Medellín and Barranquilla are situated in the Magdalena River basin and almost 80% of Colombia's population lives within the basin. During the last decades, the basin has witnessed considerable changes in land use, water, soil losses and a rapid increase of natural resources exploitation due to the economic development in the area (Restrepo and Syvitski, 2006). This recent situation has increased the pressure on the Magdalena River basin, which is the main source for human water consumption, agriculture, hydropower generation, industrial activities and ecosystems support.

3.3 Data

3.3.1 Precipitation data

3.3.1.1 Satellite-based precipitation data

Four satellite-based precipitation datasets were used in this study:

- Climate Prediction Center MORPHing technique - CMORPH:

CMORPH precipitation dataset is derived from passive microwave observations from low-Earth orbiting satellites exclusively (such as AMSR-E and TMI aboard NASA's Aqua and TRMM spacecraft), and whose features are transported via spatial propagation information entirely obtained from geostationary satellite infrared data. The technique applied to derive CMORPH precipitation data is not a precipitation estimation algorithm, but a means to combine precipitation estimates from existing microwave precipitation algorithms. Infrared

data are used to transport the microwave-derived precipitation features during periods when microwave data are not available at a location. (Joyce et al., 2004).

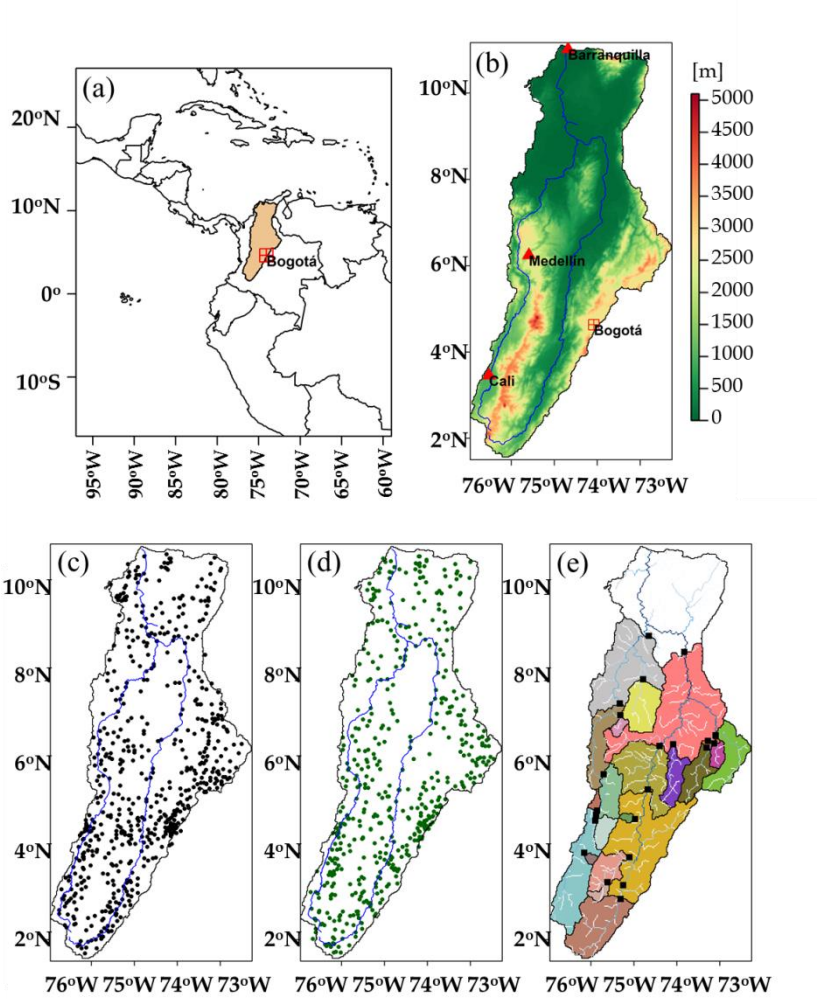


Figure 3.1 (a) Location of the Magdalena River basin in Colombia; (b) topography, main river network and urban areas within the basin; (c) weather stations used for deriving the downscaled precipitation datasets; (d) weather stations used for evaluation and (e) discharge stations and drainage sub-basins.

- Multi-Source Weighted-Ensemble Precipitation – MSWEP:

MSWEP precipitation dataset combines a wide range of data sources, including gauges, satellites and atmospheric reanalysis models. The long-term mean of MSWEP is based on the elevation-corrected CHPclim dataset but replaced with more accurate regional datasets where available. A correction for gauge under-catch and orographic effects is introduced by inferring catchment-average P from discharge observations at 13,762 stations across the globe. The temporal variability of MSWEP precipitation was determined by weighted averaging precipitation anomalies from seven precipitation datasets: two based solely on interpolation of in situ observations (CPC Unified and GPCC), three on satellite remote sensing (CMORPH, GSMaP-MVK and TMPA 3B42RT) and two on atmospheric model

reanalysis (ERA-Interim and JRA-55). For each grid cell, the weight assigned to the gauge-based estimates was calculated from the gauge network density, while the weights assigned to the satellite- and reanalysis-based estimates were calculated from their comparative performance at the surrounding gauges (Beck et al., 2017).

- Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – PERSIANN:

PERSIANN precipitation dataset are derived using artificial neural network function classification/approximation procedures based on both infrared and daytime visible imagery by geostationary satellites (such as GOES-8 and GMS-5). Model parameters of PERSIANN precipitation algorithm are updated from passive microwave observations from low-Earth orbiting satellites (Hsu et al., 1997).

- Tropical Rainfall Measuring Mission – Multi-satellite Precipitation Analysis 3B42 - TRMM:

TRMM precipitation dataset combines remote observations such as precipitation radar, passive microwave and infrared from multiple low-Earth orbiting and geostationary satellites and ground observations. Version 7 TRMM precipitation estimated by the 3B42 algorithm was used in this study. TRMM precipitation estimates are produced in four stages: (i) the passive microwave precipitation estimates are calibrated and combined, (ii) the infrared precipitation estimates are created using the calibrated microwave precipitation, (iii) the microwave and infrared estimates are combined, and (iv) rescaling to monthly data is applied (Huffman et al. 2007).

Satellite-based precipitation data were provided at 25 km spatial resolution by the Consiglio Nazionale delle Richerche (CNR) in Italy, as part of the FP7 European project eartH2Observe (Levizzanni and Dorigo, 2017).

3.3.1.2 In situ precipitation data

Precipitation data from 1118 weather stations within the Magdalena River basin were used. Daily rainfall data was provided by the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) of Colombia, covering the period from March 2000 to December 2012. The locations of the weather stations are shown in *Figure 3.1*.

3.3.2 Air temperature and evapotranspiration data

Daily air temperature and evapotranspiration data were obtained from the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data (WFDEI) at a spatial resolution of approx. 50 km (Weedon et al., 2014). The FAO Penman-Monteith equation was used to derive reference potential evapotranspiration. Air temperature and reference potential evapotranspiration were downscaled from 50 km to 1 km resolution using the e2o-downscaling-tools (Weiland et al., 2015; Schellekens and Weiland, 2017).

3.3.3 Vegetation response data: Enhanced Vegetation Index (EVI)

The Enhanced Vegetation Index (EVI) is an indicator of plant greenness or photosynthetic activity based on how different surfaces reflect different wavelengths of light. The EVI was developed as an alternative vegetation index to overcome some limitations of the Normalized Difference Vegetation Index (NDVI), such as saturation of the signal in densely vegetated and humid areas. The EVI was developed to maintain high sensitivity to changes in areas with dense biomass, to reduce the influence of the atmospheric conditions in the index value and to minimize canopy background variations. Some previous studies proved successful EVI applications in areas having high biomass, including the Amazon forest (Huete et al., 2002; Bradley et al., 2011). In this study, EVI data (MOD13A2) derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Terra satellite (from hereafter referred as EVI) were used. EVI data are provided every 16 days at 1 km spatial resolution. The post-processing steps include a re-projection and masking of water bodies. Water bodies may present negative EVI values leading to problems in the subsequent regression. Therefore, water bodies were masked and corresponding EVI values were removed. Subsequently, EVI values were interpolated for those areas. The water body mask was obtained from the NASA Shuttle Radar Topography Mission (SRTM) Water Body Data dataset developed by the National Geospatial-Intelligence Agency (SWBD, 2017).

3.3.4 Elevation, slope and aspect data

Digital Elevation Model (DEM) data from the NASA Shuttle Radar Topographic Mission (SRTM) distributed by the United States Geological Survey (USGS) was used. The vertical error of the DEM is less than 16 m (Sun et al., 2003). The DEM at a spatial resolution of 3 arc-second (approx. 90 m) was resampled (averaged) to 1 km resolution. Slope and aspect were extracted from the DEM using QGIS (QGIS, 2017).

3.3.5 Discharge data

Daily discharge data from 22 gauging stations along the Magdalena River and its tributaries were used for hydrological model evaluation and provided by the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) of Colombia. *Figure 3.1* shows the location of discharge stations and *Table 3.2* includes the location, river and sub-basin area for each discharge station.

3.4 Methodology

3.4.1 Hydrological model

The distributed hydrological model OpenStreams wflow-hbv (Schellekens, 2014) was used. The OpenStreams wflow-hbv model is based on the HBV-96 model (Saelthun, 1996) and it is programmed in the PCRaster-Python environment (Karssenberg et al., 2010). It is freely available through the OpenStreams project (Schellekens, 2016).

Table 3.2 Discharge stations used for validation.

Station name	Station ID IDEAM	Latitude	Longitude	River	Area [km²]
La Nueva	21257100	4.80	-74.97	Recio	610
Mateguadua	26107130	4.02	-76.16	Tuluá	664
Puente Carretera	21137030	3.26	-75.25	Aipe	746
Puente Negro	26147140	4.99	-75.86	Risaralda	1055
El Cóndor	22027010	3.33	-75.62	Ata	1058
Puente Anori	27027090	7.20	-75.32	Anori	1310
San Gil	24027010	6.55	-73.13	Fonce	1849
Cartago	26127040	4.76	-75.90	La Vieja	2736
Puerto Araujo-Automat.	23127020	6.53	-74.09	Carare	5300
La Ceiba	24017640	6.45	-73.31	Suárez	6831
Piedras Cobre-Automat.	22057010	3.91	-75.11	Saldaña	7009
Remolino	24047020	6.61	-73.28	Suárez	9312
El Jordán	24037360	6.73	-73.10	Chicamocha	10197
La Esperanza	27037010	8.03	-74.79	Nechí	13508
La Virginia-Automat.	26177030	4.89	-75.88	Cauca	22814
Puente Iglesias	26207030	5.83	-75.71	Cauca	29022
Apavi	26247030	7.47	-75.33	Cauca	38807
Puente Santander-Automat.	21097070	2.94	-75.31	Magdalena	15705
Puerto Salgar-Automat.	23037010	5.48	-74.68	Magdalena	56905
Puerto Berrío-Automat.	23097030	6.49	-74.40	Magdalena	74410
Regidor	25027410	8.67	-73.83	Magdalena	137636
Barbosa	25027530	9.04	-74.65	Magdalena	228192

OpenStreams wflow-hbv is applied on a cell-by-cell basis and for each grid cell it determines the water balance considering the following three components: precipitation-snow routine (including interception), soil moisture routine and runoff response routine. To simulate the different runoff processes, the soil is divided into two layers: the upper and lower zone. Daily total runoff of every grid cell results from adding direct runoff, interflow from the upper soil zone and baseflow from the lower soil zone. The total runoff is accumulated from all grid cells and routed using a kinematic wave function to obtain river discharge. A schematic representation of the hydrological model is given in *Figure 3.2*.

Land cover information was obtained from the global land cover map GlobCover-2009 derived from observations of MERIS sensor on board the ENVISAT satellite mission (Arino et al., 2010). Soil information was obtained from the Food and Agriculture Organization (FAO) Digital Soil Map of the World (DSMW, 2007). The model version used in this study does not include reservoirs. Daily simulations at 1 km spatial resolution were carried out for the time period March 2000 to December 2012.

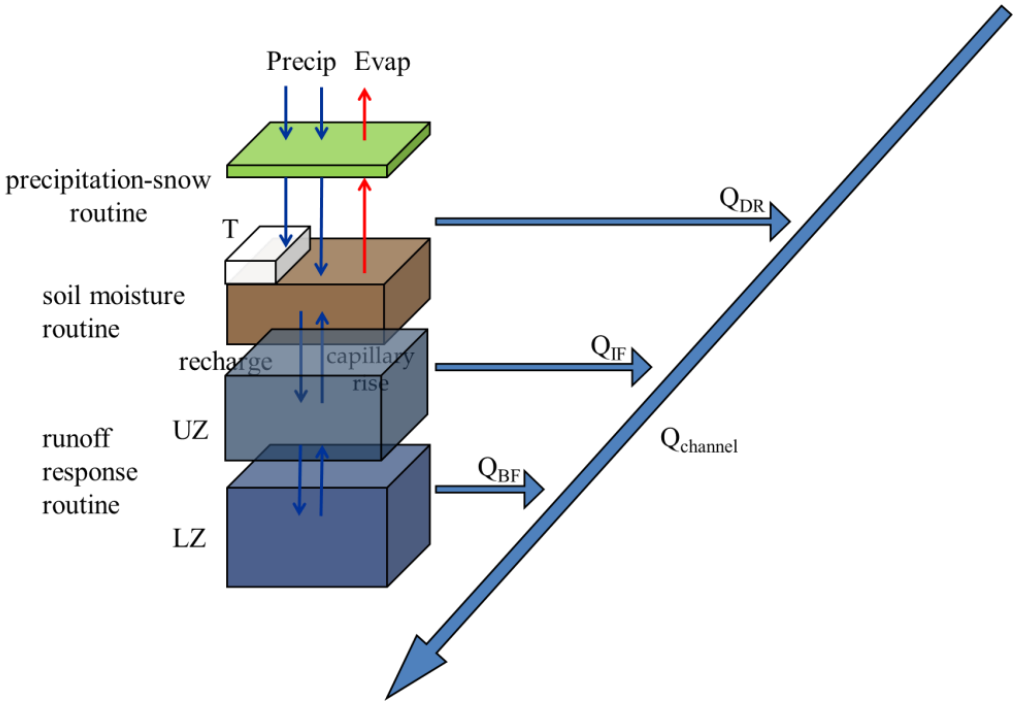


Figure 3.2 OpenStreams wflow-hbv model structure.

3.4.2 Downscaling precipitation

Additional information from four environmental variables was used as a proxy to downscale precipitation, including the vegetation response (EVI), elevation (DEM), slope and aspect. The same downscaling procedure was followed for the four global precipitation datasets.

Initially, the relation between EVI and precipitation was analysed. The sensitivity of the vegetation state to precipitation is cumulative (Immerzeel et al., 2005; Gessner et al., 2013), which means that vegetation indexes, such as NDVI and EVI, in response to precipitation have a lag time. Quiroz et al. (2011) estimated the lag time in 1-3 months and Hunink et al. (2014) considered a lag time of 1 week in their weekly regression models between precipitation, elevation, climatology and NDVI. In this study, EVI was used to avoid limitations of NDVI (see 3.3.3 *Vegetation response data: Enhanced* Vegetation Index (EVI)). An analysis was carried out to estimate the lag time between precipitation and EVI. Basin average values of precipitation and EVI were calculated for the entire time period (March 2000-December 2012). Correlation values between precipitation and EVI were obtained considering lag times between 1 week and 3 months (*Table A. 1*). The highest correlation values were obtained with a lag time between precipitation and EVI of 1 month.

The precipitation downscaling procedure consisted of three main steps (Step 1 and Step 2 on the spatial downscaling at monthly temporal scale and Step 3 on the disaggregation from monthly to daily temporal scale):

- Step 1: modelling and prediction of monthly precipitation at fine scale using Geographically Weighted Regression (GWR).
- Step 2: spatial correction with in situ observations to include the variability not explained by the models.
- Step 3: disaggregating monthly values to daily precipitation estimates.

A more detailed description of the steps follows:

Step 1: Geographically Weighted Regression modelling and prediction

(i) Daily precipitation values were summed to obtain monthly precipitation estimates (P^{25km}).

(ii) Post-processed EVI values were averaged at monthly scale (EVI^{1km}).

(iii) DEM, slope, aspect and monthly EVI were aggregated by pixel averaging from 1 km spatial resolution to the spatial resolution of the global precipitation datasets, 25 km: DEM^{25km}, $SLOPE^{25km}$, $ASPECT^{25km}$ and EVI^{25km}.

(iv) A functional relationship was established between monthly P^{25km} and DEM^{25km}, $SLOPE^{25km}$, $ASPECT^{25km}$ and EVI^{25km} using a Geographically Weighted Regression (GWR). The GWR is a local form of multiple linear regression and can be written as follows:

$$P_{estim,j}^{25km} = a_j^{25km}(u) + b_j^{25km}(u) * DEM^{25km} + c_j^{25km}(u) * SLOPE^{25km} + d_j^{25km}(u) * ASPECT^{25km} + e_j^{25km}(u) * EVI_{j+1}^{25km} \quad (3.1)$$

where $a_j^{25km}(u)$, $b_j^{25km}(u)$, $c_j^{25km}(u)$, $d_j^{25km}(u)$ and $e_j^{25km}(u)$ were the intercept and the slope parameters, varying with location (u) and $P_{estim,j}^{25km}$ is the estimated precipitation at 25 km resolution at month j. The GWR supplied the coefficient of determination of the regression R^2 per pixel.

(v) GWR was repeated excluding alternatively $SLOPE^{25km}$, $ASPECT^{25km}$ and EVI^{25km}; and $P_{estim,j}^{25km}$, GWR parameters and R^2 were re-computed:

$$P_{estim,j}^{25km} = a_j^{25km}(u) + b_j^{25km}(u) * DEM^{25km}$$

$$P_{estim,j}^{25km} = a_j^{25km}(u) + b_j^{25km}(u) * DEM^{25km} + c_j^{25km}(u) * SLOPE^{25km} + d_j^{25km}(u) * ASPECT^{25km} \quad (3.2)$$

$$P_{estim,j}^{25km} = a_j^{25km}(u) + b_j^{25km}(u) * DEM^{25km} + e_j^{25km}(u) * EVI_{j+1}^{25km}$$

Step (v) was performed because in some cases the relationship between precipitation and auxiliary variables was stronger excluding some of them. For example, the relationship

between precipitation and EVI may be weak in areas where the land use was fragmented, hence excluding EVI improved the model regression performance. All possible linear combinations of variables (separately) were analysed and the previous ones resulted in higher R^2.

(vi) For each grid cell, the $P_{estim,j}^{25km}$ and the parameters of the GWR associated with the highest coefficient of determination R^2 were selected and used for further downscaling. Residuals were computed as the difference between $P_{estim,j}^{25km}$ and the initial precipitation dataset, P_j^{25km}, at 25 km resolution. Assumptions of normality, non-autocorrelation and homocedasticity of the residuals were checked. Normality of the residuals was verified using histograms of the residuals and the Shapiro-Wilk test (Shapiro and Wilk, 1965). Homocedasticity was checked with the Breusch-Pagan test (Breusch and Pagan, 1979) and autocorrelation was verified with the Durbin-Watson test (Durbin and Watson, 1951).

(vii) The GWR parameters were downscaled to 1 km using a cubic spline tension interpolator (Ceccherini et al., 2015), obtaining $a_j^{1km}(u)$, $b_j^{1km}(u)$, $c_j^{1km}(u)$, $d_j^{1km}(u)$ and $e_j^{1km}(u)$.

(viii) Monthly precipitation at 1 km resolution, $P_{estim,j}^{1km}$, was estimated applying the obtained GWR model with DEM, slope, aspect and EVI at 1 km:

$$P_{estim,j}^{1km} = a_j^{1km}(u) + b_j^{1km}(u) * DEM^{1km} + c_j^{1km}(u) * SLOPE^{1km} + d_j^{1km}(u) * ASPECT^{1km} + e_j^{1km}(u) * EVI_{j+1}^{1km} \quad (3.3)$$

Step 2: Spatial correction with in situ observations

In a second step a spatial correction using in situ observations of precipitation from weather stations was carried out:

(ix) Monthly precipitation values estimated at 1 km resolution were extracted for the location of weather stations: $P_{estim,j}^{station}$.

(x) The difference between the in situ precipitation for each station ($P_j^{station}$) and $P_{estim,j}^{station}$ was calculated:

$$\Delta P_j^{station} = P_j^{station} - P_{estim,j}^{station} \quad (3.4)$$

(xi) $\Delta P_j^{station}$ were spatially interpolated using a cubic spline tension interpolator to a 1 km resolution, obtaining ΔP_j^{1km}. Inverse distance weighting, nearest neighbor algorithm and cubic spline interpolation methods were tested and cubic spline outperformed the other methods.

(xii) By adding the $P_{estim,j}^{1km}$ to the residuals obtained in the previous step, ΔP_j^{1km}, the high resolution monthly precipitation at 1 km was obtained ($P_{downs,j}^{1km}$):

$$P_{downs,j}^{1km} = P_{estim,j}^{1km} + \Delta P_j^{1km} \quad (3.5)$$

Step 3: Disaggregating from monthly precipitation into daily values

For hydrological modelling and prediction of daily discharge, daily precipitation estimates are needed. However, the procedure described until here estimated precipitation at monthly temporal resolution. In the third step, the fraction of precipitation per day derived from precipitation datasets at 25 km resolution was used to disaggregate the high-resolution monthly precipitation at 1 km resolution into maps at daily time scale. Duan and Bastiaanssen (2013) applied a similar methodology to disaggregate annual to monthly precipitation. The steps are described as follows:

(xiii) Precipitation at 25 km resolution that occurs during the ith day was divided between the monthly total precipitation to obtain the fraction of precipitation per day i:

$$F_i^{25km} = \frac{P_i^{25km}}{\sum_{i=1}^{n} P_i^{25km}} \quad (3.6)$$

where n was the number of days of each month.

(xiv) The fractions of precipitation per day, F_i^{25km}, were interpolated using a cubic spline tension interpolator into a spatial resolution of 1 km, obtaining F_i^{1km}.

(xv) Monthly precipitation estimates at 1 km resolution were multiplied times the corresponding fractions to obtain daily precipitation:

$$P_{downs,i}^{1km} = F_i^{1km} \, x \, P_{downs,j}^{1km} \quad (3.7)$$

3.4.3 Precipitation evaluation

To ensure an independent evaluation of precipitation before and after downscaling, a split sample approach of the 1118 weather stations was used. The precipitation of 616 stations was used for deriving the downscaled precipitation datasets and the remaining 502 stations were used for evaluation. These two station groups were randomly selected such that they were equally distributed in space over the catchment.

The accuracy of the different precipitation datasets at 25 km and 1 km resolutions was assessed by a number of commonly used performance indicators (Immerzeel et al., 2008; Ceccherini et al. 2015): coefficient of determination (R^2), difference between RS-derived and in situ precipitation values and Root Mean Square Error (*RMSE*).

3.4.4 Hydrological modelling and evaluation

The daily precipitation datasets at 25 km and 1 km resolutions were used to force the OpenStreams wflow-hbv model. Firstly, in situ daily precipitation values at weather locations were interpolated using the inverse distance weighting algorithm to create spatial maps at 1 km resolution (various interpolation techniques were tested, including splines, inverse distance weighting and kriging with external drift, and inverse distance weighting outperformed the other methods). Then, OpenStreams wflow-hbv was calibrated and validated using interpolated in situ precipitation during 2000 - 2012. The year 2000 was used to spin up the model until reaching a dynamical steady state. The time periods 2001-2004 and 2005-2012 were used for calibration and validation, respectively. Kling-Gupta efficiency (*KGE*; Gupta et al., 2009) was selected as the optimization criterion, to avoid problems that could occur when Nash-Sutcliffe efficiency (*NSE*; Nash and Sutcliffe, 1970) is used for model calibration (e.g. high sensitivity to extreme values). Model parameters were calibrated to optimize *KGE* values at 22 discharge stations.

Secondly, the impact on discharge simulations of a decrease in the number of weather stations used in the DMD model was analysed. Fourteen weather station networks composed of 0, 4, 8, 10, 20, 30, 40, 50, 60, 80, 100, 200, 400 and 616 stations were selected from the 616 rain gauges used for deriving the downscaled precipitation datasets. A stratified sampling technique was used to build the weather station networks. This technique avoids problematic networks that may result from simple randomly sampling, aiming uniform station networks homogenously distributed over the basin. The specific steps to build each station network are described as follows:

(i) The area of the basin was divided into different grid cells using a spatial resolution of ca. 50 km.

(ii) A random sample of size n_i from each grid cell was extracted aiming for $n_1 + n_2 + \cdots + n_k = n$, where n is the total number of weather stations of the network (4, 8, 10, 20, 30, 40, 50, 60, 80, 100, 200, 400). Minimum distance between stations was considered to avoid taking stations that were very close to each other.

(iii) Step (ii) was repeated 10 times to generate multiple realizations of each subset of stations per grid cell of size n_i, which results in 10 different configurations of the entire station network. This reduces the effect of poorly distributed networks and the influence of undetected inhomogeneous station records.

(iv) Precipitation values at weather stations were interpolated using the inverse distance weighting algorithm to create spatial maps at 1 km resolution for each of the 10 station network configurations.

(v) Precipitation derived from each of the 10 station network configurations was evaluated by calculating r^2 and $RMSE$ and the spatial configuration with the highest performance was selected.

This technique was applied repeatedly to obtain networks of *n* number of rain gauge stations. Similar sampling approaches have been successfully used in previous studies to analyze the effect of sample size on precipitation and hydrological modelling (Janis et al., 2004; Bardossy and Das, 2008; Xu et al., 2013).

Every generated precipitation dataset was used to force the OpenStreams wflow-hbv and the effect on simulated streamflow was evaluated. Various statistical indicators were used to evaluate model performance: *KGE*, Pearson's correlation coefficient (*r*) and *RMSE*. *KGE* equally measures bias and differences in timing and amplitude, whereas *r* measures mainly differences in timing of high and low discharge and *RMSE* differences in magnitude. Additional performance indicators could be calculated to further analyze the impact of rain intensity, location and time errors, among others (Thiemig et al., 2012).

3.5 Results and discussion

3.5.1 Precipitation evaluation

3.5.1.1 Assessing the performance of precipitation datasets at 25 km

In order to evaluate how well the original RS-derived precipitation datasets at 25 km resolution perform, *Figure 3.3a*, *Figure 3.3b* and *Figure 3.3c* show the boxplots of three performance indicators, including the difference (*Figure 3.3a*), *RMSE* (*Figure 3.3b*) and r^2 (*Figure 3.3c*) between monthly RS-derived and in situ observed precipitation. From *Figure 3.3a*, MSWEP and TRMM precipitations show the largest agreement with in situ data (differences close to 0). CMORPH slightly underestimates precipitation, while PERSIANN largely overestimates it. From *Figure 3.3b*, PERSIANN exhibits the largest *RMSE* values, followed by CMORPH. Low values of *RMSE* come from MSWEP and TRMM. From *Figure 3.3c*, MSWEP provides the highest r^2 values. High r^2 values are also displayed with TRMM, while CMORPH and PERSIANN show lower values. To complete the evaluation of the RS-derived precipitation datasets at 25 km resolution, *Figure 3.3d* shows the boxplot of climatology of monthly RS-derived and in situ observed precipitation. MSWEP and TRMM capture the intra-annual variability of precipitation during the year well, whereas CMORPH and PERSIANN show larger differences with ground data. PERSIANN displays a good agreement during the dry period from December to March, but highly overestimates precipitation during the rest of the year, especially from May to September. CMORPH is able to reproduce precipitation variability, but with a consistent underestimation.

As expected, MSWEP was the precipitation dataset with the best performance since this dataset is the result of combining several precipitation data sources, including gauge, satellite and reanalysis data and therefore, it takes full advantage of the complementary nature of the different sources (Schellekens et al., 2017). Although CMORPH performance could be considered acceptable in terms of low *RMSE* and relatively high r^2 values, it tends to systematically underestimate precipitation for each month. Dinku et al. (2009) showed similar CMORPH results for Colombia and they partly attributed them to the orographic warm rain processes. PERSIANN shows the worst performance, which is consistent with previous studies in the area (de Goncalves et al., 2006; Ceccherini et al., 2015). Performance

differences could be also due to the fact that CMORPH and PERSIANN do not include a posterior global gauge correction, whereas MSWEP and TRMM do. Accompanying the RS-derived precipitation by an error product could benefit their use (Zeweldi and Gebremichael, 2009).

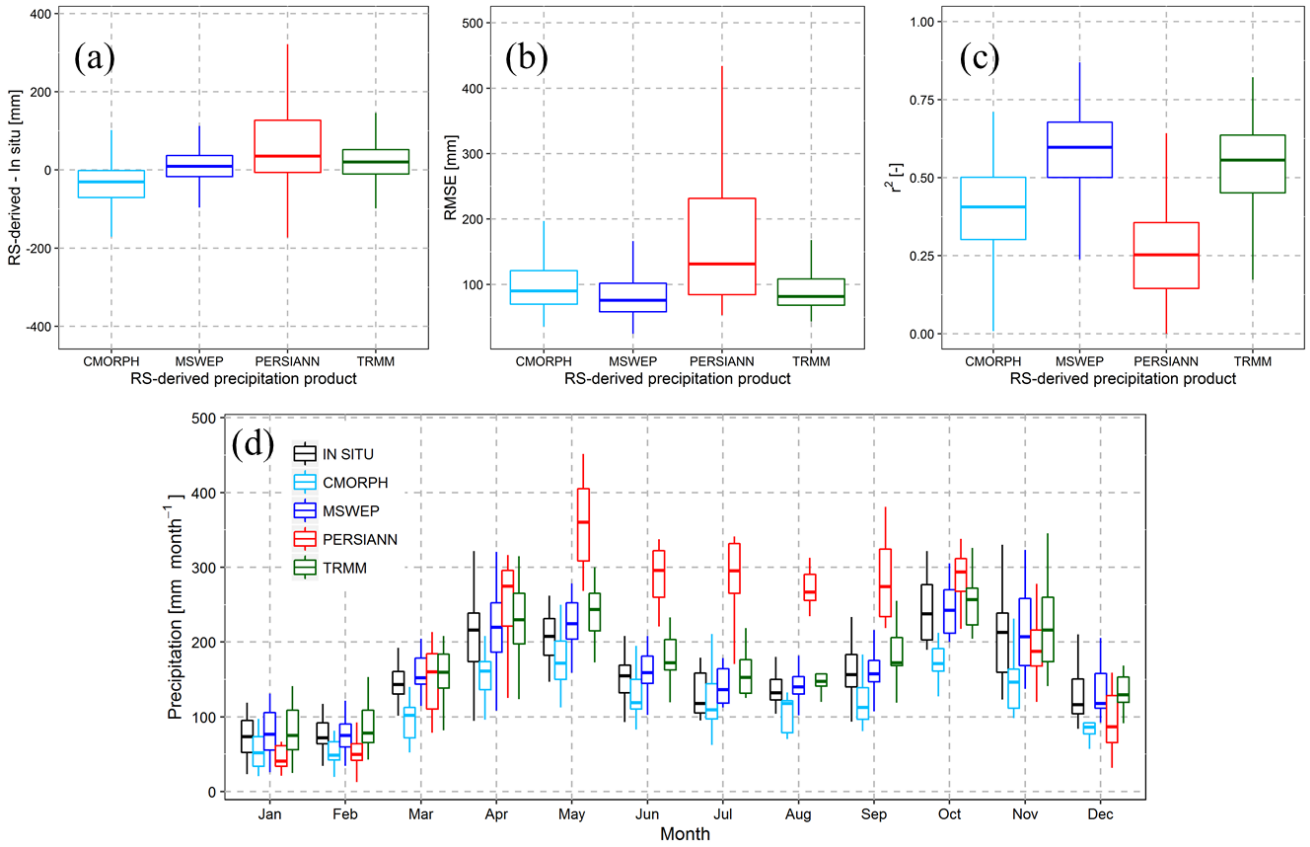


Figure 3.3 Boxplots of (a) the difference, (b) *RMSE* and (c) r^2 between monthly RS- derived precipitation at 25 km resolution and observed precipitation for the 502 validation weather stations. (d) Boxplot of climatology of monthly RS-derived and observed precipitation at 25 km resolution (average of 502 validation weather stations).

3.5.1.2 Precipitation downscaling analysis

Figure 3.4 summarizes the results derived from the GWR analysis (Step 1) at 25 km resolution. GWR model parameters and variables to estimate MSWEP precipitation are shown for May 2003. Although in the GWR model, slope and aspect were considered as predictors, results indicated that their contribution is minimal. Therefore and for practical reasons, *Figure 3.4* shows only EVI and DEM predictors and the associated model parameters. Model parameters show different spatial patterns. Slope parameter for DEM is positively related to precipitation in the northern and southern parts of the basin, but a negative relationship exists in the central region (high precipitation and low elevation). Slope parameter for EVI shows positive values for the relationship between EVI and precipitation in most part of the basin, with higher values in the central-eastern region.

A lag time of 1 month between precipitation and EVI was determined using average values over the entire basin (*Table A. 1*). A spatial analysis of the correlation between precipitation and EVI with different lag times by regions may be done in the future, limiting the use of EVI predictor to those areas with higher correlations. Other regression methodologies, such as multiple linear regression or exponential regression (see overview in *Table 3.1*), could be used. However, the GWR model can capture the spatial variability in the relationship between precipitation, DEM, slope, aspect and EVI, which would not be noticed in the other regression models. Additional variables, such as humidity, wind speed or topographical roughness, may be included as model predictors to further improve precipitation estimates.

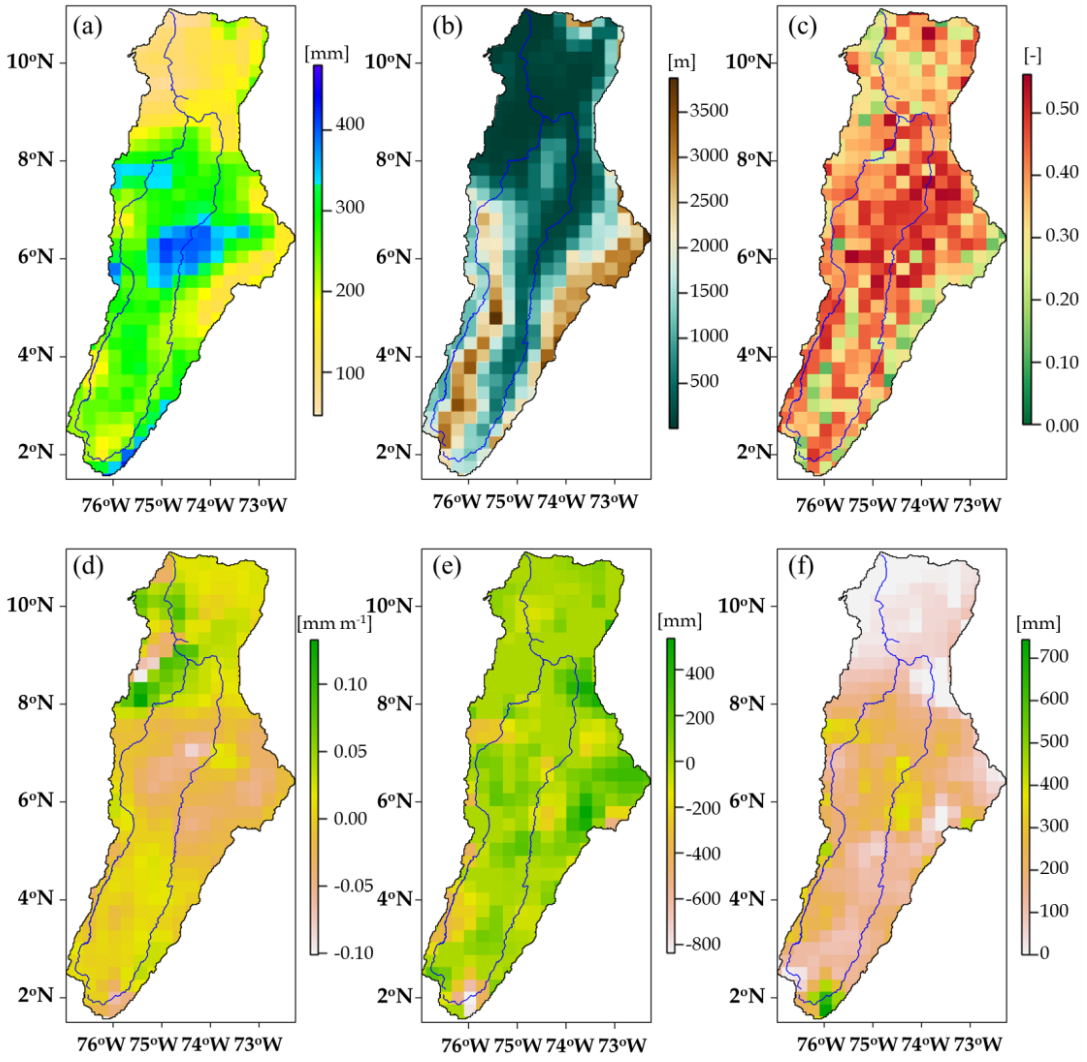


Figure 3.4 Spatial distribution of (a) average monthly MSWEP precipitation, (b) DEM, (c) average monthly EVI, (d) slope parameter for DEM, (e) slope parameter for EVI and (f) intercept in May 2003 at 25 km resolution.

Once the GWR analysis was completed, a spatial correction using the in situ residuals was carried out to estimate precipitation at 1 km resolution. *Figure 3.5* shows precipitation datasets at 25 km (first row), at 1 km (second row) and the residuals at 1 km (third row) for April 2000. *Figure 3.5* reveals that precipitation in the central part of the basin is indeed much higher (~ 400 mm month^{-1}) than in northern and southern regions (~ 100 mm month^{-1}). The range of values is consistent with the monthly precipitation analysis done by IDEAM in Colombia (IDEAM, 2017). Spatial patterns of CMORPH, MSWEP and TRMM precipitation datasets are similar, except for some areas in the south-eastern part of the basin (CMORPH in the first row of *Figure 3.5*). PERSIANN differs in the spatial variability of precipitation, providing higher estimates in most of the basin than the remaining precipitation datasets.

The general precipitation patterns are well captured by the P_{estim}^{1km} (second row), with the central part wetter, corresponding well with the original precipitation datasets. Residuals (ΔP^{1km}, third row) show the amount of precipitation that cannot be explained by the GWR model compared to in situ measurements. Negative residuals (yellow) indicate regions where precipitation is overestimated. Positive residuals (blue) depict areas where precipitation is less than expected according to in situ information. Residuals of CMORPH, MSWEP and TRMM are close to zero (green) in the majority of the basin, whereas residuals of PERSIANN are around -150 mm month^{-1} with values closer to zero in the southern part. In all precipitation datasets, some isolated small areas with higher residuals can be observed, which could be related to the complex orography and the strong spatial variability of precipitation.

3.5.1.3 Assessing the performance of downscaled precipitation

To assess the performance of downscaled precipitation datasets R^2, the difference between RS-derived and in situ precipitations and *RMSE* were calculated. Each indicator was obtained for each weather station and was plotted and spatially interpolated using cubic splines. *Figure 3.6* shows the difference between precipitation datasets at 25 km resolution (first row) and 1 km resolution (second row) and in situ observations, giving an indication of the accuracy of the final and original precipitation datasets over the entire basin.

At 25 km resolution (first row), MSWEP and TRMM show similar spatial variabilities with the lowest absolute differences, negative in the northern part (Sierra Nevada) and positive in the southern part. CMORPH shows negative differences for the whole basin, except for some spots, indicating an underestimation of precipitation. On the other hand, PERSIANN overestimates precipitation in the northern and central regions and differences are reduced in the upstream part of the river.

At 1 km resolution (second row), differences are highly reduced and their spatial patterns are similar for the four datasets. These similarities in the precipitation estimates are due to the high density of weather stations. The downscaled PERSIANN dataset is an exception in the north-western region, which is due to the high bias of this product at 25 km resolution.

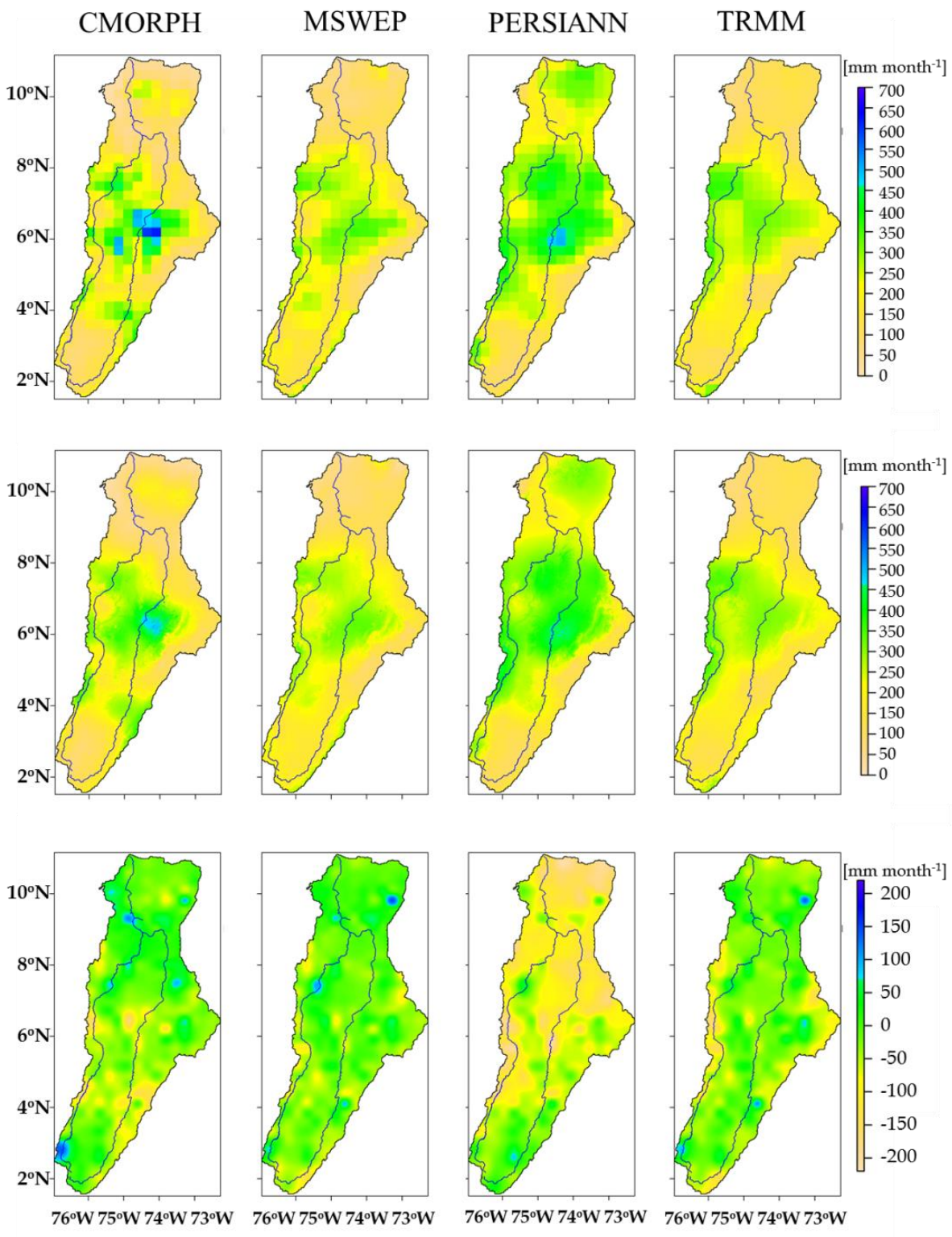


Figure 3.5 Spatial distribution of CMORPH, MSWEP, PERSIANN and TRMM precipitation datasets at 25 km (P^{25km}, first row) and 1 km (P^{1km}_{estim}, second row) and spatial distribution of residuals at 1 km ($\Delta\ P^{1km}$, third row) for April 2000.

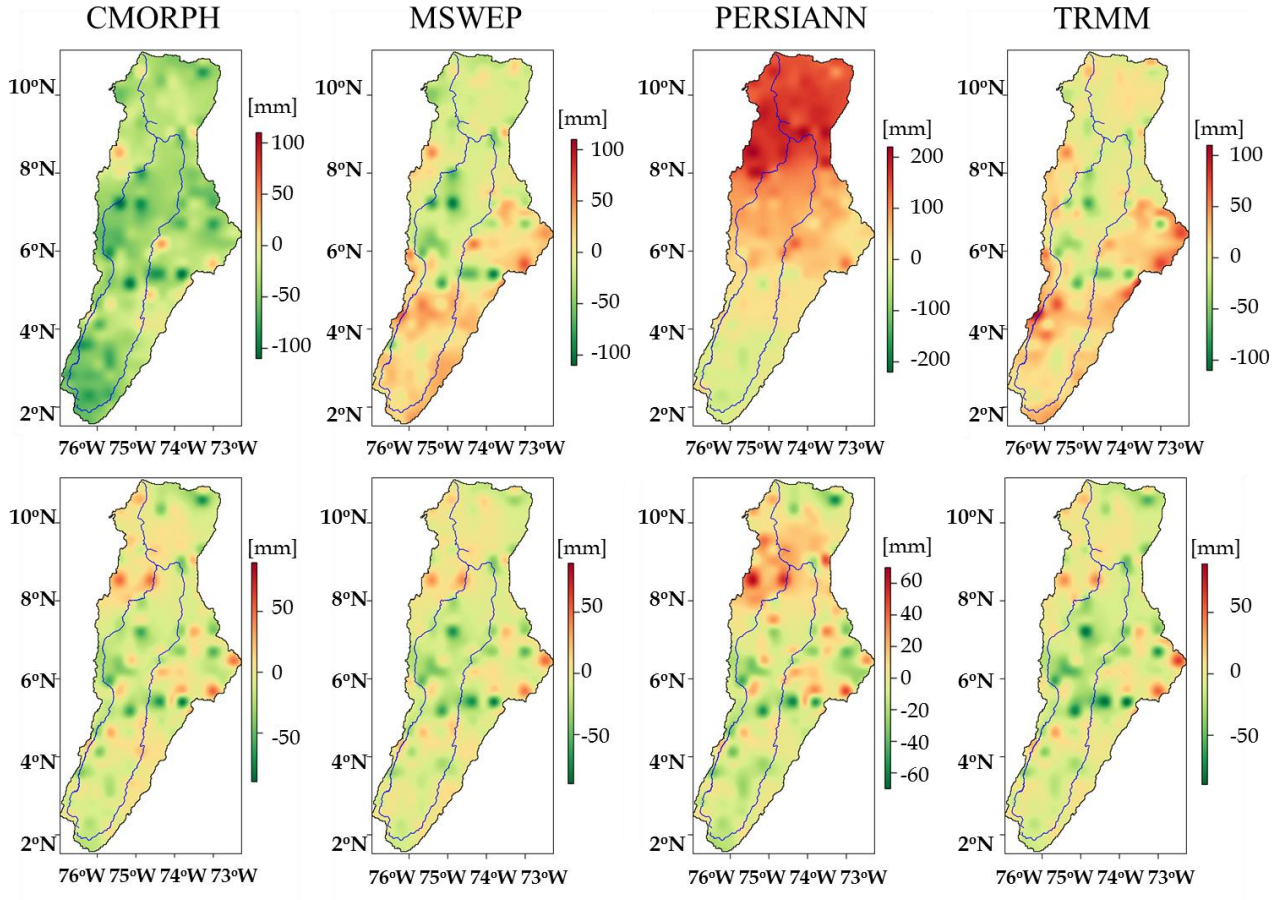


Figure 3.6 Interpolated difference between CMORPH, MSWEP, PERSIANN and TRMM precipitation datasets at 25 km (P^{25km}, first row) and 1 km (P^{1km}_{downs}, second row) and in situ observations for the 2000 - 2012 period.

Table 3.3 summarizes the statistical indicators of P^{25km}, P^{1km}_{estim} and P^{1km}_{downs}. These indicators were calculated by averaging the individual indicators of all validation weather stations. *Table 3.3* shows that the GWR analysis alone already improves the accuracy, with decreased *RMSE* and difference between RS-derived and in situ precipitation and increased R^2 (except for PERSIANN due to the very poor performance of the original dataset). This shows that the use of the GWR model is a necessary step to improve the accuracy of RS-derived precipitation estimates. The spatial correction with in situ observations further improves the accuracy increasing R^2 and reducing *RMSE* and differences with ground data. Duan and Bastiaanssen (2013) also highlighted the significance of the downscaling procedure previously to the calibration with in situ information. The high density of weather stations in the basin makes that the improvement due to the spatial correction with ground precipitation data (~ 83%) is higher than the one due to the GWR analysis alone (~ 17%). Lower density networks may limit the spatial correction impact, emphasizing the importance of the GWR analysis in data-poor river basins.

Table 3.3 Statistical indicators of CMORPH, MSWEP, PERSIANN and TRMM precipitation datasets versus in situ data (average of 502 validation weather stations).

	RS-derived – in situ [mm]			***RMSE* [mm]**			**R^2 [-]**		
	P^{25km}	P^{1km}_{estim}	P^{1km}_{downs}	P^{25km}	P^{1km}_{estim}	P^{1km}_{downs}	P^{25km}	P^{1km}_{estim}	P^{1km}_{downs}
CMORPH	-19.93	-13.38	2.48	102.74	98.36	76.23	0.40	0.42	0.61
MSWEP	14.11	13.14	3.16	87.25	86.26	75.35	0.58	0.59	0.62
PERSIANN	45.30	46.89	2.25	163.88	164.74	77.76	0.26	0.25	0.59
TRMM	25.25	23.31	3.42	95.02	93.77	75.69	0.53	0.54	0.62

3.5.2 Hydrological modelling and evaluation

The daily precipitation datasets at 25 km and 1 km resolutions with and without spatial correction with ground data (P^{25km}, P^{1km}_{estim} and P^{1km}_{downs}) were used to force the OpenStreams wflow-hbv model. Previously, the model was calibrated and validated using in situ discharge data (*Table A. 2*). Further model parameters calibration could have been done, but the idea was to fairly represent the main hydrological processes to assess the impact of various precipitation datasets and their spatial variability on simulated discharge.

Figure 3.7 shows the simulated and observed discharge at La Esperanza station. As expected, discharge estimates when the model was driven with MSWEP and TRMM at 25 km resolution exhibit the highest agreement with observations. The lowest performance is provided when the model was forced with PERSIANN at 25 km, with a significant discharge overestimation from April to September, in line with precipitation evaluation results. CMORPH results emphasize the limitations of this dataset for precipitation estimation in the basin, which tends to slightly underestimate discharge.

Driving the model with the downscaled precipitation datasets improves discharge simulation performance increasing *KGE* values. For MSWEP and TRMM (*Figure 3.7b* and *Figure 3.7d*), the application of the downscaling procedure improves discharge estimates to a lesser extent (ΔKGE ~ 0.02) than when CMORPH (ΔKGE ~ 0.10) and PERSIANN (ΔKGE ~ 1.61) are used (*Figure 3.7a* and *Figure 3.7c*). Due to the good model performance when the original MSWEP and TRMM datasets were used (*KGE ~ 0.76*), there are no significant discharge differences between using downscaled precipitation spatially corrected (P^{1km}_{downs}) and not spatially corrected with in situ data (P^{1km}_{estim}). In this study, the spatial resolution of the hydrological model, 1 km, was finer compared to that of precipitation, 25 km. For hydrological models at 25 km grid or coarser, it is expected that the precipitation downscaling procedure does not lead to further improvements on discharge estimates.

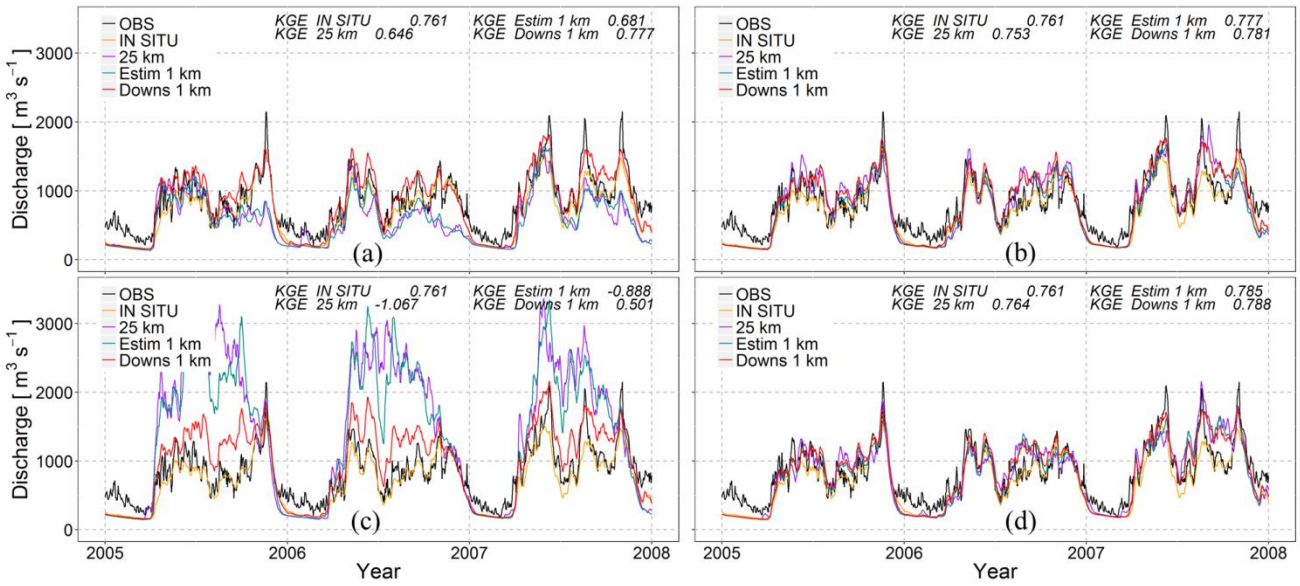


Figure 3.7 Daily observed discharge (black) and estimated discharge (orange, purple, blue and red) time series at La Esperanza station (27037010) when the model was forced with (a) CMORPH, (b) MSWEP, (c) PERSIANN and (d) TRMM for the 2005 - 2007 period. The orange lines represent estimates when in situ precipitation data was used. The purple lines represent estimates when precipitation datasets at 25 km, P^{25km}, were used. The blue and red lines represent estimates when precipitation datasets at 1 km without spatial correction, P^{1km}_{estim}, and with spatial correction, P^{1km}_{downs}, were respectively used.

Figure 3.8 summarizes discharge simulation performances for the 22 gauging stations through boxplots of their *KGE* and *r* values when the model is forced with precipitation datasets at 25 km and 1 km resolutions (P^{25km}, P^{1km}_{estim} and P^{1km}_{downs}). *KGE* and *r* values obtained when the model was forced with the precipitation datasets at 25 km resolution (purple) are lower than when in situ interpolated data was used (orange). Possible reasons could be the lower quality of RS-derived precipitation datasets compared to in situ data and that model parameters were calibrated with ground precipitation data. Some previous studies calibrated the hydrological model for each precipitation dataset to analyze the sensitivity of model parameters to precipitation (Andreassian et al., 2001; Nkiaka et al., 2017). In this study, model parameters are not optimized for each forcing aiming to avoid correcting precipitation errors through fine-tuning the hydrological processes representation in the model.

In agreement with precipitation results in *Table 3.3*, *KGE* and *r* increase when the downscaled precipitation without the spatial correction was used (blue). This improvement is higher when the model was forced with the spatially corrected downscaled precipitation (red). When analyzing the discharge performance obtained with P^{1km}_{downs}, the downscaling procedure manages to reduce the initial differences between the original MSWEP, TRMM, CMORPH and PERSIANN precipitation datasets providing comparable averaged *KGE* (~ 0.57) and *r* values (~ 0.75). These values are similar to those obtained when the model was forced with in situ data, *KGE = 0.57* and *r = 0.74* (except for PERSIANN).

Due to the application of the precipitation downscaling procedure, *KGE* increases were in the order of ~ 0.10-0.50, indicating a considerable improvement in model performance in terms of its bias and its ability to capture the overall flow variability between extreme events (Lobligeois et al., 2014). Correlation increases were in the order of ~ 0.02-0.40, meaning that forcing the model with downscaled precipitation also improves the timing and the shape of the hydrographs.

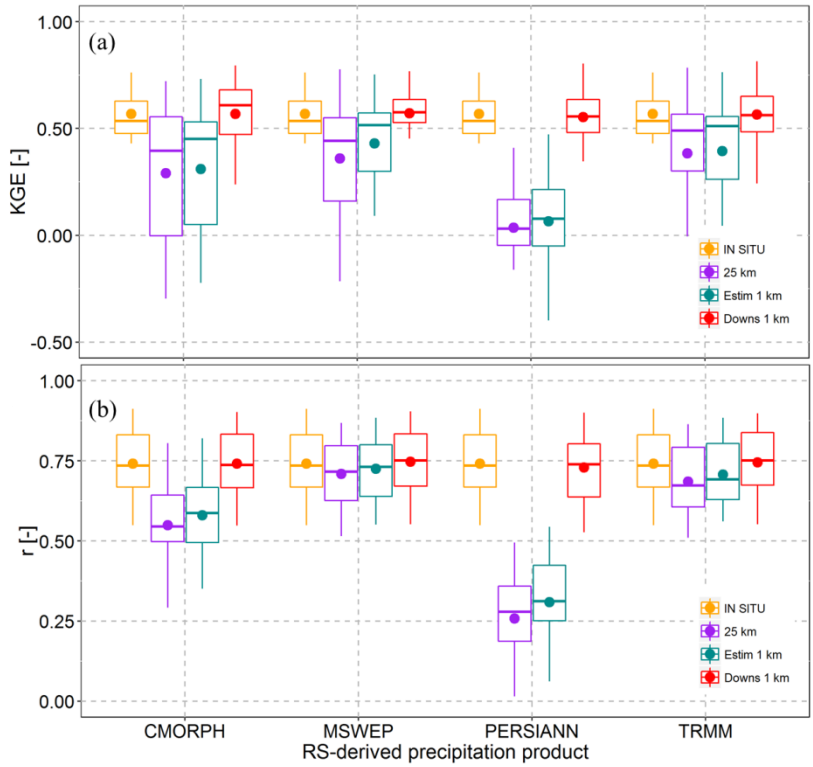


Figure 3.8 (a) *KGE* and (b) *r* values between daily simulated and observed discharge at the 22 discharge stations when the model was forced with in situ precipitation data (orange) and precipitation datasets at 25 km (P^{25km}, purple) and at 1 km without spatial correction (P^{1km}_{estim}, blue) and with spatial correction (P^{1km}_{downs}, red) for the 2000 - 2012 period. Numbers are the average *KGE* and *r* values for the 22 discharge stations.

Once the impact of the downscaling procedure on discharge estimates was analyzed, the influence of the number of stations used in that procedure was further investigated. *Figure 3.9* shows that when increasing the number of weather stations in the downscaling procedure, the performance of the hydrological model driven by the precipitation datasets improves. CMORPH precipitation is increased from approx. 110 mm month^{-1} to 200 mmm month^{-1} and therefore, discharge is also increased from approx. 3000 m^3 s^{-1} to 4000 m^3 s^{-1}. PERSIANN precipitation is reduced from approx. 260 mm month^{-1} to 180 mmm month^{-1} and therefore, discharge also decreases from approx. 7,000 m^3 s^{-1} to 4,500 m^3 s^{-1}. In general, no further variations occur when increasing the number of weather stations above 100. Contrary to CMORPH and PERSIANN, MSWEP and TRMM do not show a monotonic tendency to increase or decrease when the number of stations increases. Precipitation and

discharge vary similarly until they reach constant values of approx. 190 mm month^{-1} and 4,600 m^3 s^{-1}, respectively (100 stations).

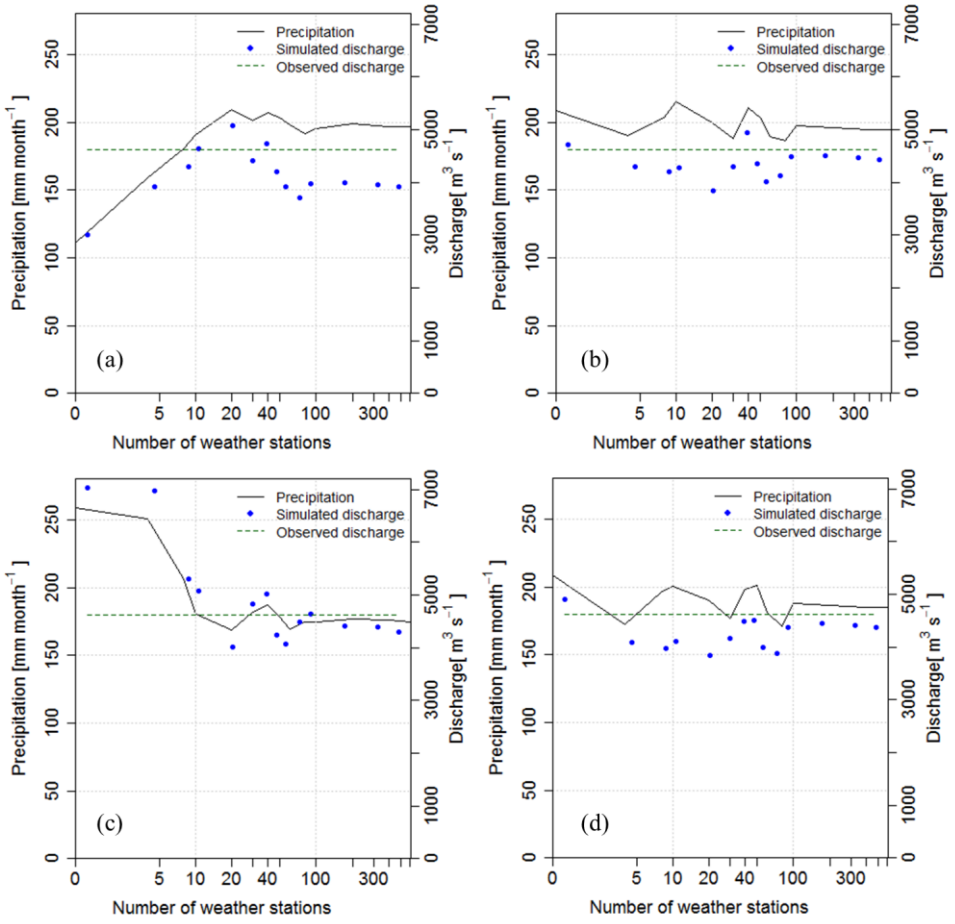


Figure 3.9 Monthly areal average precipitation (left vertical axis) and simulated (blue dots) and observed (green dashed line) average discharge (right vertical axis) at Barbosa station (25027530) at 1 km resolution for the 2000 - 2012 period, derived with different number of weather stations (0, 4, 8, 10, 20, 30, 40, 50, 60, 80, 100, 200, 400 and 616). (a) CMORPH, (b) MSWEP, (c) PERSIANN and (d) TRMM. Logarithmic scale was used for the horizontal axis.

Figure 3.10 and *Figure 3.11* summarize *KGE* and *r* results for the 22 discharge stations when changing the number of weather stations in the precipitation downscaling procedure. Average values were calculated from *KGE* and *r* values obtained at all discharge stations.

As expected, the lowest model performance was observed when downscaled precipitation was not spatially corrected with in situ data (with some exceptions for TRMM and MSWEP). In general, *KGE* and *r* values increase with increasing the number of weather stations in a monotonic tendency. *KGE* and *r* variation ranges differ between precipitation datasets. Average *KGE* values vary from 0.31 to 0.57 for CMORPH (*Figure 3.10a*) and from 0.07 to 0.55 for PERSIANN (*Figure 3.10c*). *KGE* and *r* values obtained when using MSWEP and TRMM show lower improvements when the numbers of stations increases, varying from 0.43 to 0.57 for MSWEP (*Figure 3.10b*) and from 0.39 to 0.57 for TRMM (*Figure 3.10d*).

In spite of this general trend, sometimes the increase of the number of stations does not imply an increase in *KGE* and *r* values, or even causes a decrease. A possible reason behind this may be related to the inclusion of weather stations with lower quality data which would deteriorate the RS-derived precipitation datasets. Previous studies (Bardossy and Das, 2008; Xu et al., 2013) have found similar results and attributed them to the contribution of individual weather stations and their location within the basin. Xu et al. (2013) explained that when a low number of stations is used, their location plays an important role. Although 10 stations cannot reproduce the spatial variability of precipitation in comparison with all 616 stations, if they would be more evenly spatially distributed within the basin that would significantly impact model results. A possibility for future research would be to further investigate the weather station networks that lead to a decrease in *KGE* and *r* values by regions, especially in mountainous areas where heavy orographic precipitation events occur, which highly contribute to discharge. Moreover, alternative sampling techniques could be used.

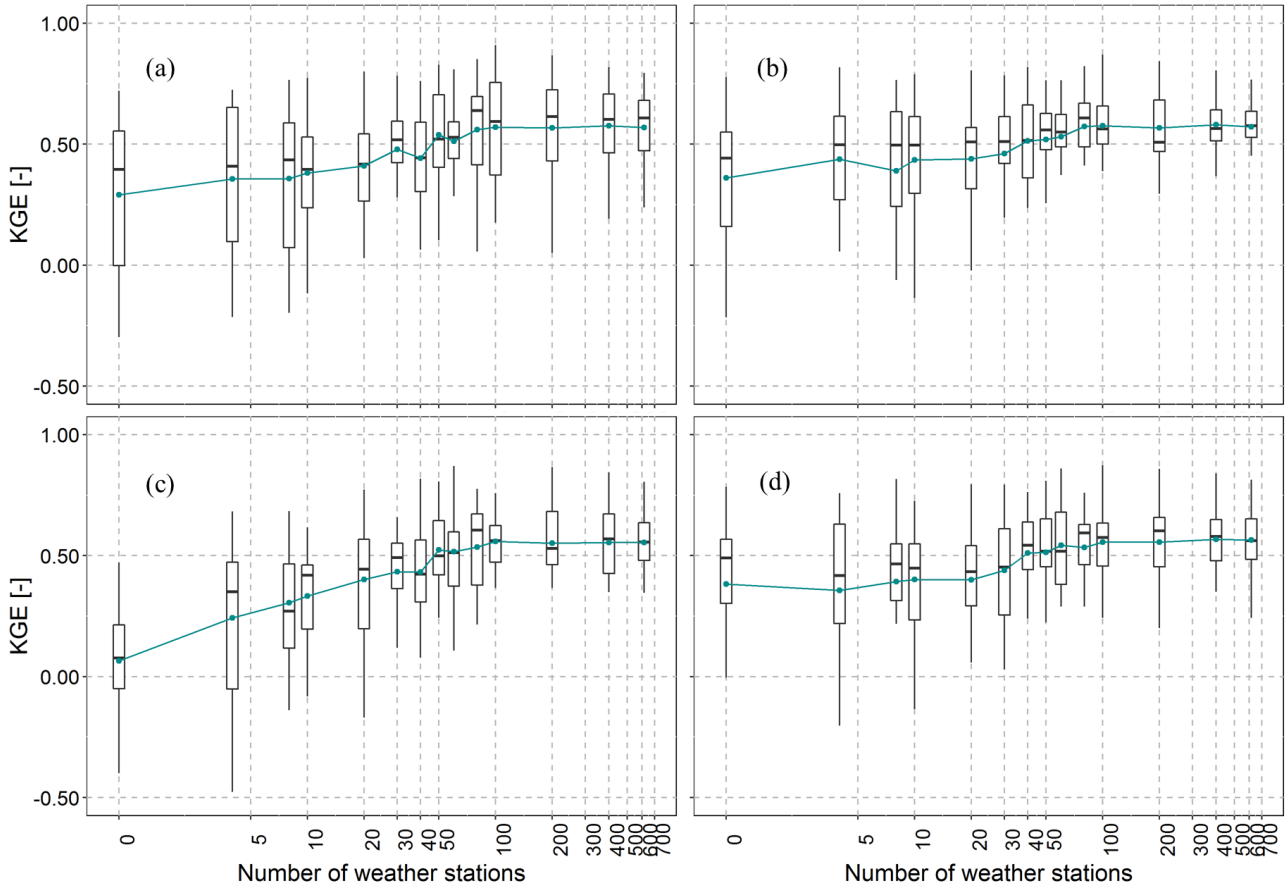


Figure 3.10 KGE values (y-axis) between daily simulated and observed discharge when the model is forced with (a) CMORPH, (b) MSWEP, (c) PERSIANN and (d) TRMM precipitation datasets at 1 km resolution for the 2000 - 2012 period, derived with different number of weather stations (x-axis; 0, 4, 8, 10, 20, 30, 40, 50, 60, 80, 100, 200, 400 and 616). Logarithmic scale was used for the horizontal axis.

Increasing the number of stations above 100 did not further increase *KGE* or *r* values (the increase in average and median *KGE* and *r* values becomes relatively small). These results are specific for this basin, hydrological model, precipitation datasets and resolutions. The impact of the number of weather stations and the spatial variability of precipitation on discharge estimates might differs depending on the topography, soil type, basin hydrological characteristics and the type of precipitation (Janis et al., 2002; Bardossy and Das, 2008).

In this study, precipitation at two spatial resolutions, 25 km and 1 km, was used. The impact of intermediate spatial resolutions, such as 20, 15, 10, 5 or 2 km, on discharge estimates could be further investigated (Gascon et al. 2015). Moreover, this study could be complemented with the analysis of the influence of temporal variability of precipitation on discharge estimates.

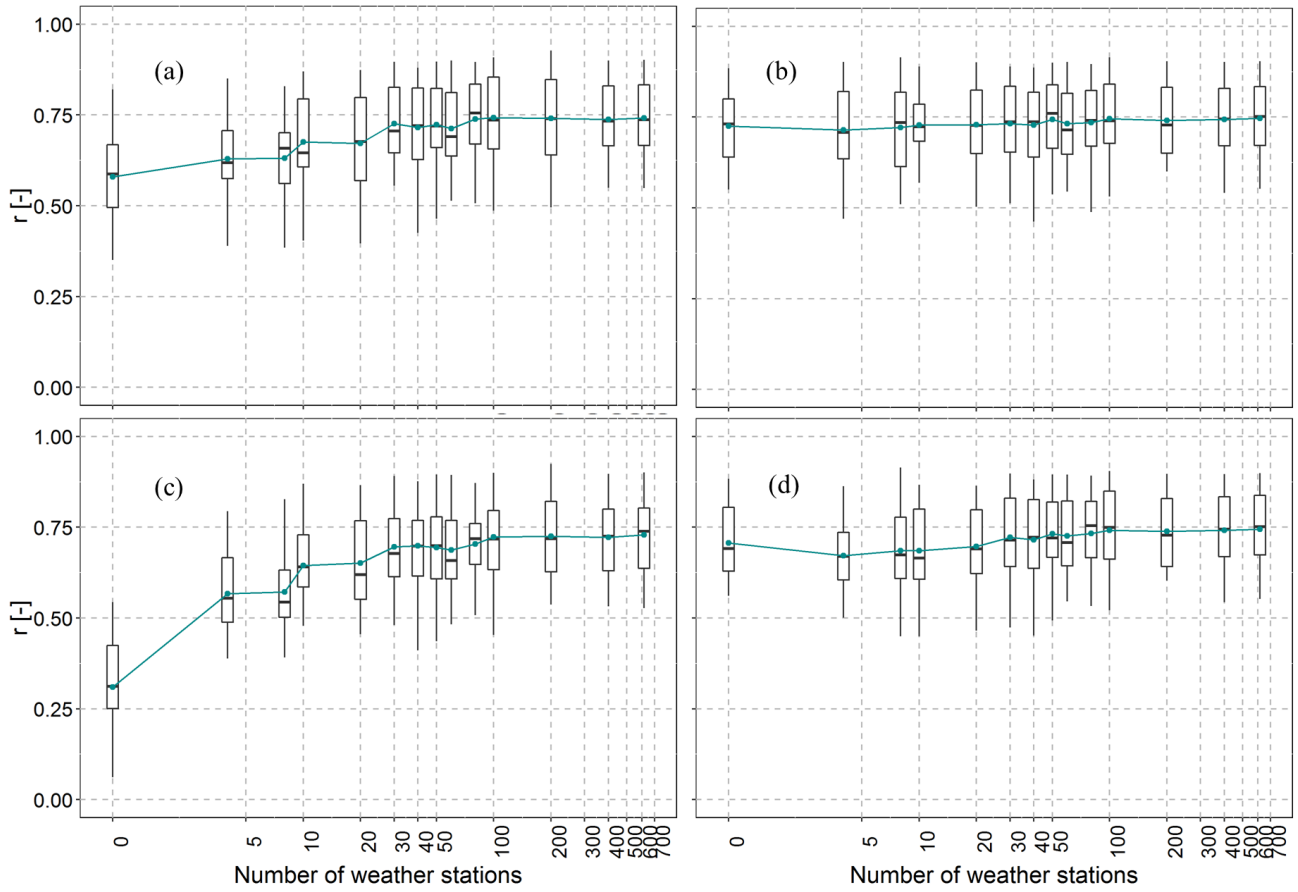


Figure 3.11 r values (y-axis) between daily simulated and observed discharge when the model is forced with (a) CMORPH, (b) MSWEP, (c) PERSIANN and (d) TRMM precipitation datasets at 1 km resolution for the 2000 - 2012 period, derived with different number of weather stations (x-axis; 0, 4, 8, 10, 20, 30, 40, 50, 60, 80, 100, 200, 400 and 616). Logarithmic scale was used for the horizontal axis.

Once the impact of increasing the number of weather stations for precipitation downscaling was analyzed, in situ precipitation values at the same samples of weather stations were interpolated (using the inverse distance weighting algorithm) to obtain different spatial maps at 1 km resolution. These maps, based only on in situ data, were used to force the hydrological model and discharge simulation performance results are shown in *Figure 3.12*.

As expected, increasing the number of weather stations used for ground precipitation interpolation improves *KGE* (*Figure 3.12a*) and *r* (*Figure 3.12b*) values. Discharge performance when the model was forced with downscaled precipitation is better than when the model was forced with in situ interpolated precipitation, considering the same weather station network. For example, for a network of 20 weather stations, higher *KGE* and *r* values were obtained when driving the model with downscaled MSWEP precipitation (*KGE = 0.44* and *r = 0.73*) compared to those obtained when in situ interpolated precipitation was used (*KGE = 0.36* and *r = 0.66*). This may be due to the downscaled datasets that better capture precipitation spatial variability combining different sources of information (satellite and ground precipitation, vegetation, elevation, slope and aspect). These results show the potential of the described downscaling methodology for ungauged river basins or with limited number of weather stations.

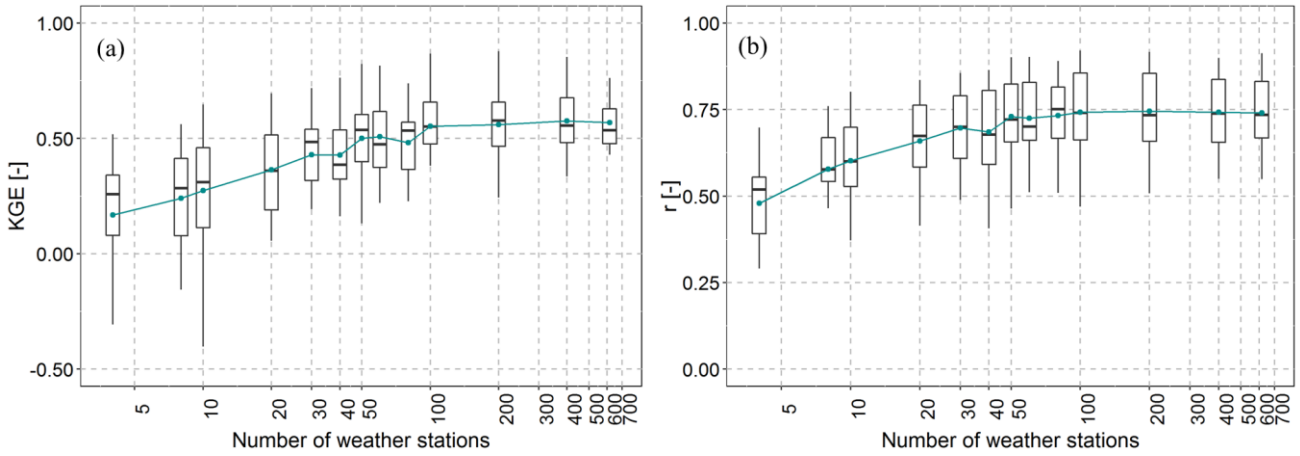


Figure 3.12 (a) *KGE* and (b) *r* values (y-axis) between daily simulated and observed discharge when the model is forced with ground interpolated precipitation at 1 km resolution for the 2000 - 2012 period, derived with different number of weather stations (x-axis; 4, 8, 10, 20, 30, 40, 50, 60, 80, 100, 200, 400 and 616). Logarithmic scale was used for the horizontal axis.

3.6 Conclusions

This study investigated how discharge estimates of a distributed hydrological model, OpenStreams wflow-hbv, can be influenced by precipitation quality and spatial variability in the Magdalena River basin in Colombia. The effect of precipitation on streamflow simulations was assessed by running the model with in situ and RS-derived precipitation datasets at non-downscaled and downscaled spatial resolutions. The conclusions of the study are as follows:

(i) At 25 km resolution, MSWEP and TRMM outperformed CMORPH and PERSIANN precipitation datasets.

(ii) The downscaling procedure resulted in highly improved accuracy of the four precipitation datasets. Vegetation, elevation, slope and aspect were successfully used in a GWR model as proxies to precipitation leading to improvements of approx. 17% in *RMSE*, R^2 and bias compared to in situ observations. The spatial correction with ground precipitation data further improved them, due to the high density of weather stations in the basin. Lower density networks may limit the spatial correction impact, highlighting the significance of the GWR analysis in data-poor river basins.

(iii) Discharge model estimates were in better agreement with observations when MSWEP and TRMM were used to force the model compared to CMORPH and PERSIANN precipitation datasets at 25 km resolution.

(iv) Forcing the model with the downscaled precipitation datasets considerably improved discharge model estimates, with *KGE* increases in the order of 0.10 to 0.50.

(v) A higher number of weather stations was necessary in the spatial correction with in situ data of CMORPH and PERSIANN to achieve similar discharge performances than those obtained with MSWEP and TRMM. However, increasing the number of stations above 100 did not further improve model estimates with any precipitation dataset.

(vi) The downscaling of RS-derived precipitation datasets resulted in better discharge estimates compared to using only in situ precipitation data when using less than 100 weather stations.

Although results depend on the specifics of each basin, the present study showed that an accurate representation of precipitation spatial variability may help to improve streamflow simulations. Downscaling procedures, such as the one used in this study, make globally available RS-derived precipitation datasets an interesting alternative/complement to ground data for hydrological modelling in poor-gauged or ungauged river basins.

4 Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products

This chapter is based on:

López López P., Sutanudjaja E. H., Schellekens J., Sterk G. and Bierkens M. F. P.(2017). Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. Hydrology and Earth System Sciences, 21, 3125–3144, https://doi.org/10.5194/hess-21-3125-2017.

Abstract

A considerable number of river basins around the world lack sufficient ground observations of hydro-meteorological data for effective water resources assessment and management. Several approaches can be developed to increase the quality and availability of data in these poorly gauged or ungauged river basins, and among those, the use of Earth observations products has recently become promising. Earth observations of various environmental variables can be used potentially to increase the knowledge about the hydrological processes in the basin and to improve streamflow model estimates, via assimilation or calibration. The present study aims to calibrate the large-scale hydrological model PCR-GLOBWB using satellite-based products of evapotranspiration and soil moisture for the Moroccan Oum Er Rbia basin. Daily simulations at a spatial resolution of 5 arc min are performed with varying parameters values for the 32-year period 1979-2010. Five different calibration scenarios are inter-compared: (i) reference scenario using the hydrological model with the standard parameterization, (ii) calibration using in situ observed discharge time series, (iii) calibration using GLEAM actual evapotranspiration time series, (iv) calibration using ESA CCI surface soil moisture time series and (v) step-wise calibration using GLEAM actual evapotranspiration and ESA CCI surface soil moisture time series. The impact on discharge estimates of precipitation in comparison with model parameters calibration is investigated using three global precipitation products, including EI, WFDEI and MSWEP.

Results show that GLEAM evapotranspiration and ESA CCI soil moisture may be used for model calibration resulting in reasonable discharge estimates (NSE values from 0.5 to 0.75), although better model performance is achieved when the model is calibrated with in situ streamflow observations. Independent calibration based on only evapotranspiration or soil moisture observations improves model predictions to a lesser extent. Precipitation input affects to discharge estimates more than calibrating model parameters. The use of WFDEI precipitation leads to the lowest model performances. Apart from the in situ discharge

calibration scenario, the highest discharge improvement is obtained when EI and MSWEP precipitation products are used in combination with a step-wise calibration approach based on evapotranspiration and soil moisture observations. This study opens up the possibility to use globally available Earth observations and reanalysis products of precipitation, evapotranspiration and soil moisture into large-scale hydrological models to estimate discharge at the river basin scale.

4.1 Introduction

To assess and manage the available water resources within a river basin, good estimates of hydro-meteorological data, such as precipitation, temperature and streamflow, are required. Yet many river basins around the world still have a limited number of in situ observations, being either ungauged (Sivapalan et al., 2003) or poorly gauged (Loukas and Vasiliades, 2014). Ungauged or poorly gauged river basins also include those basins where data are inaccurate, scarce, intermittent or collected at different temporal resolutions, leading to the problem that it is not clear how to integrate these data consistently into hydrological models (Winsemius et al., 2009). As a result, the limited availability and poor quality of data induces large uncertainty in model outputs from these river basins (Seibert and Beven, 2009). Developing novel strategies to enhance available datasets and hydrological models is one of the key strategies when working in ungauged basins (Hrachowitz et al., 2013).

To overcome the lack of hydro-meteorological data, a promising approach is the use of the recently developed Earth observations and reanalysis products to supplement the available data. In the last decades, radar and satellite technologies have improved and have become more broadly available providing diverse hydro-meteorological datasets at finer spatial and temporal resolutions: precipitation (Joyce et al., 2004; Huffman et al., 2007), soil moisture (Njoku et al., 2003; Dorigo et al., 2015), total water storage (Tapley et al., 2004), evapotranspiration (Bastiaanssen et al., 1998; Nishida, 2003; Miralles et al., 2011b), etc.

Previous studies have demonstrated the possibility of using these global datasets to better understand the hydrological processes in a river catchment (Kite and Droogers, 2000; Vereecken et al., 2008; Seneviratne et al., 2010; Hafeez et al., 2011) and to improve streamflow model estimates through assimilation (Parajka et al., 2006; Roy et al., 2010; Brocca et al., 2012; Thirel et al., 2013; López López et al., 2016) and/or calibration techniques or a priori determination of model parameters (Jacobs et al., 2003; Beck et al., 2009). Calibration approaches based on multiple remotely sensed variables have some advantages in comparison with traditional calibration approaches using only observed and modelled hydrographs in a limited number of locations. Fenicia et al. (2007) and Gupta et al. (2008) recognized that traditional calibration may lead to over-parameterization, i. e. similar model results are obtained with more than one parameters combination, whereas calibrating to multiple variables - step-wise calibration - may partly resolve the problem of non-uniqueness and it helps to a better understanding of the processes happening within the catchment.

Several studies have investigated calibration approaches based on variables different to streamflow. Campo et al. (2006) used soil moisture information from radar images from ERS-2 sensors to parameterize the hydrological model MOBIDIC. Immerzeel and Droogers (2008) calibrated the hydrological model SWAT based on satellite evapotranspiration from MODIS satellite images. Lo et al. (2010) improved the parameter estimation of the Community Land Model 3.0. using GRACE total water storage data while Isenstein et al. (2015) calibrated the VIC hydrological model using snow covered area from MODIS satellite data. Others have combined remotely sensed variables with in situ streamflow

observations for calibration. In Rientjes et al. (2013), the HBV model was calibrated on satellite based evapotranspiration from MODIS and streamflow. Wanders et al. (2014) calibrated model parameters of LISFLOOD based on discharge and soil moisture observations acquired by AMSR-E, SMOS and ASCAT while Sutanudjaja et al. (2014) calibrated the large-scale model PCR-GLOBWB using streamflow and soil water index information derived from the ERS scatterometers. At a global scale, Beck et al. (2016) used parameter regionalization to calibrate a HBV model. However, the simultaneous use of more than one environmental variable different to streamflow for calibration is rare. A calibration approach using various variables, independently and in combination with streamflow observations, may further improve model performance and contribute to a better understanding of hydrological processes. In the present study, this is tested by comparing multiple calibration scenarios based on evapotranspiration, soil moisture and discharge data.

The previously mentioned calibration experiments were performed for well studied river basins, such as the Rhine-Meuse river basin, with a good coverage of in situ hydro-meteorological data. In the present study area, the Oum Er Rbia River basin located in Morocco, ground observations are spatially sparse and limited in number classifying it as a poorly-gauged river basin. The region frequently suffers from water scarcity and droughts and water availability is the main factor influencing socio-economic development, mostly driven by agriculture (Houdret, 2008). The studies of Tramblay et al. (2012), Tramblay et al. (2016) and Ouatiki et al. (2017) are testimony to the relevance of this area. Therefore, developing new strategies to model this watershed is highly relevant to improve water management and assessment of the water availability within the basin.

This study aims to calibrate a large-scale hydrological model, PCR-GLOBWB 2.0 (Sutanudjaja et al., 2016, 2017), using soil moisture and evapotranspiration observations alone and to compare its discharge estimates to those obtained when the model is traditionally calibrated to streamflow data. We use the evapotranspiration product generated by an enhanced version of the GLEAM model (GLEAM v3.0; Martens et al., 2016b) in combination with the surface soil moisture product from ESA CCI (Dorigo et al., 2015). Both products are derived from satellite data. Furthermore, the influence of precipitation forcing is considered and three different global precipitation products are used and inter-compared: ERA-Interim reanalysis data (EI; Dee et al., 2011), WATCH Forcing Data methodology applied to ERA-Interim reanalysis data (WFDEI; Weedon et al., 2014) and Multi-Source Weighted-Ensemble Precipitation data by merging gauge, satellite and reanalysis data (MSWEP; Beck et al., 2017a).

Five different calibration approaches are performed by using five calibration scenarios that include streamflow, soil moisture and evapotranspiration: (i) reference scenario using the hydrological model with the standard parameterization, (ii) calibration using in situ observed discharge time series, (iii) calibration using GLEAM actual evapotranspiration time series, (iv) calibration using ESA CCI surface soil moisture time series and (v) step-wise calibration using GLEAM actual evapotranspiration and ESA CCI surface soil moisture time series. The above is repeated for each of the selected global precipitation

product. A priori, it is expected that calibrating to streamflow observations yields the best discharge estimates, and that the step-wise calibration using soil moisture and evapotranspiration provides better results than the calibration scenarios based only on soil moisture or evapotranspiration.

The novel aspects and new contributions of this work include the use and comparison of three different and recently generated global precipitation products, the exploration of calibration techniques based on Earth observations of soil moisture and evapotranspiration and their application into a large-scale hydrological model to provide streamflow estimates in the ungauged river basin of Oum Er Rbia in Morocco. Furthermore, understanding the potential gain of calibrating large-scale models to remotely sensed observations may have benefits for water resources management in data-poor river basins globally.

4.2 Study area

The study area is the Oum Er Rbia River basin, which is situated in the central-west region of Morocco between the Atlas Mountains to the south and the Mesetian area to the north flowing into the Atlantic Ocean (*Figure 4.1*). The basin's topography ranges from 2,800 m in the southern upstream zone to 150 m in the northern downstream zone. The Oum Er Rbia is the second largest river in Morocco with a total length of 550 km and it drains an area of approximately 38,025 km^2.

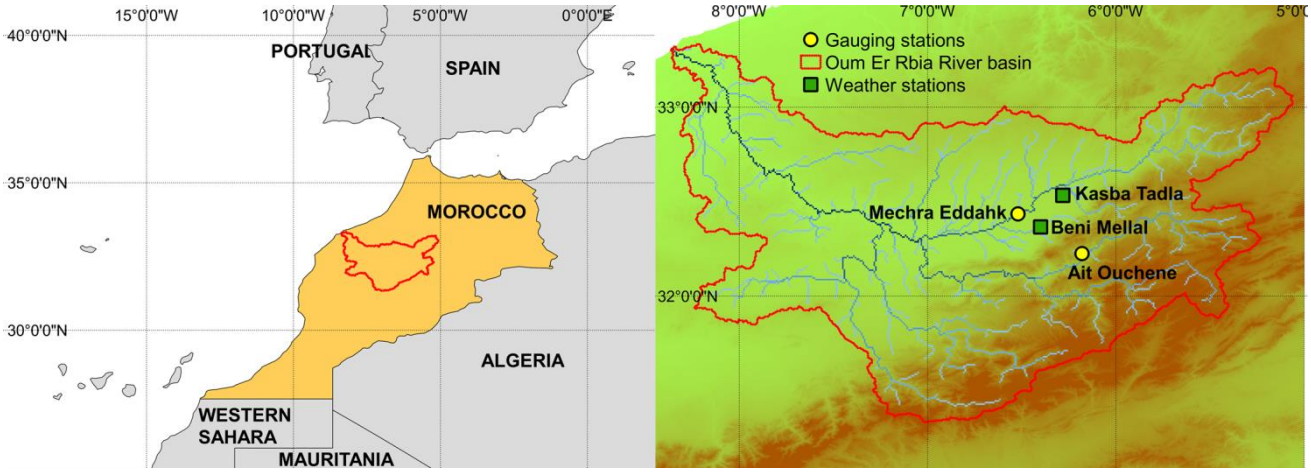


Figure 4.1 Oum er Rbia River basin and its location in Morocco (the delineation of the catchment is physically based). Yellow points represent the gauging stations and green squares represent the weather stations.

The lowlands of the basin are mainly covered with rain-fed and irrigated agriculture fields and the upstream regions are a combination of Mediterranean forests, woodlands and scrubs. The geology of the area is mostly composed of limestone, marls and sandstone with a karst aquifer in the Atlas Mountains and a multi-layered system of superficial and deep aquifers in the western plains (Bouchaou et al., 2009).

The climate in the coastal and mountainous areas is Mediterranean, characterized with high temperatures in the summer and warm autumn and winter months with rainfall, and semi-arid in the central plain (Jones et al., 2013). Precipitation increases from downstream to upstream areas in the mountains. The mean annual precipitation and temperature are 400 mm and 18°C, respectively. Approximately 70% to 80% of the annual rainfall is concentrated in the period from October to May.

4.3 Hydrological model

The large-scale hydrological model PCR-GLOBWB 2.0 (Sutanudjaja et al., 2016, 2017) was used at a spatial resolution of 5 arc min (approximately 10 km at the equator) and at a daily temporal resolution. PCR-GLOBWB is a leaky-bucket type of model applied on a cell-by-cell basis. Figure 2 illustrates a schematic representation of the structure of PCR-GLOBWB model. For each grid cell and time step, the model determines the water balance considering the following water storage components: soil moisture, groundwater, surface water, interception storage and snow. The soil is divided into three vertical layers representing the top 5 cm of soil (depth $Z_1 \leq 5$ cm), the following 25 cm of soil (depth $Z_2 \leq 30$ cm) and the remaining 120 cm of soil (depth $Z_3 \leq 150$ cm), in which the storages are symbolized as S_1, S_2 and S_3, respectively. The underlying groundwater store (S_4) consists of two layers: an active or renewable layer and a non-active or non-renewable layer of fossil water, in which the storages are symbolized as S_{4act} and S_{4fos}, respectively. The model also includes the water exchange processes between the top layer and the atmosphere (precipitation, evapotranspiration and snowmelt), between the soil layers (percolation and capillary rise) and between the soil layers and the active layer of the groundwater store (groundwater recharge, discharge to baseflow and capillary rise). Each grid cell is divided into sub-grids considering variations of elevation, vegetation, soil and land cover. Five land cover types are distinguished: irrigated paddy field, irrigated non-paddy field, grassland (short natural vegetation), forest (tall natural vegetation) and open water. To compute the total runoff of every grid cell, the model includes direct runoff (Q_{DR}), shallow sub-surface flow from the third soil layer (Q_{SF}), and baseflow from the active groundwater layer (Q_{BF}). The total runoff is accumulated from all grid cells and routed along the drainage network to obtain the river discharge ($Q_{channel}$). The PCR-GLOBWB model version used here (Sutanudjaja et al., 2016) simulates water availability and water abstraction, including reservoirs and domestic, industrial, livestock and irrigational water demands. The following subsections briefly describe the model components and the parameters relevant for the present calibration study. The reader is referred to Sutanudjaja et al. (2011) and Sutanudjaja et al. (2014) for a more detailed explanation.

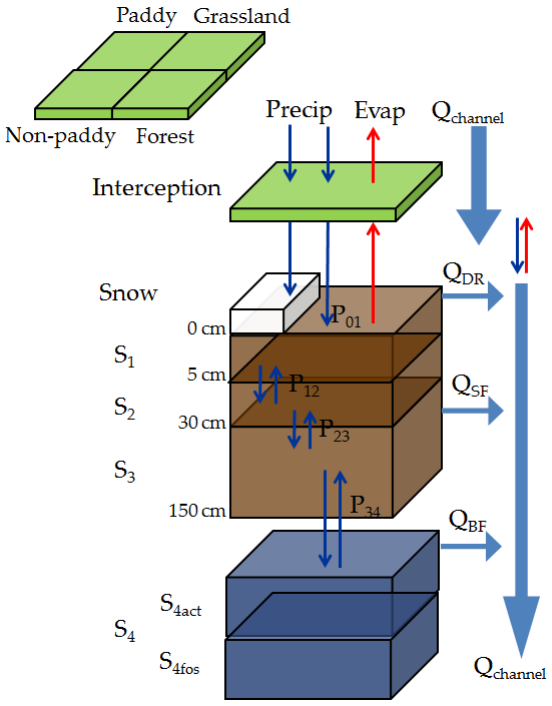


Figure 4.2 PCR-GLOBWB model structure, adapted from van Beek et al. (2011).

4.3.1 Direct or surface runoff

The amount of water that goes into the soil is the net precipitation (P_n) resulting from the surplus of precipitation above the interception capacity and the excess melt water from the snow pack. P_n is partitioned into direct runoff (Q_{DR}) and net infiltration to the first soil layer (P_{01}). The partitioning is done using the Improved Arno Scheme (Hagemann and Gates, 2003), in which the fraction of saturated soil of a cell is estimated based on the cell-minimum capacity (W_{min}), the cell-average actual storage ($W_{act} = S_1 + S_2 + S_3$) and the water capacity for the entire soil profile ($W_{max} = SC_1 + SC_2 + SC_3$, SC_n: soil water capacity for layer n). If $W_{min} = 0$, direct runoff always occurs for a rainfall event. If $W_{min} > 0$, an event P_n only generates runoff Q_{DR} if $W_{act} > W_{min}$. W_{min} is therefore an important parameter which governs runoff generation response time.

4.3.2 Vertical water exchanges between soil and groundwater stores and shallow sub-surface flow

Net infiltration water into the first soil layer (P_{01}) is transferred through the remaining soil layers. Vertical water exchanges occur between the first and the second layers (P_{12}), between the second and the third layers (P_{23}) and between the third soil layer and the active layer of the groundwater store (P_{34}). P_{12}, P_{23} and P_{34} consist of downward percolation and upward capillary rise, which depend on the degree of saturation ($s_1 = {S_1}/{SC_1}$, $s_2 = {S_2}/{SC_2}$ and $s_3 = {S_3}/{SC_3}$) and the unsaturated hydraulic conductivity of each soil layer (K_{sat1}, K_{sat2} and K_{sat3}). If $s_1 > s_2$, percolation is equal to K_{sat1}; whereas if $s_2 > s_1$, capillary rise is equal to $K_{sat2}\ x\ (1 - s_1)$, being $(1 - s_1)$ the moisture deficit in the first soil

layer. K_{sat1}, K_{sat2} and K_{sat3} control the vertical fluxes between the soil layers and the groundwater store which affect significantly to the ground water recharge. Moreover, K_{sat3} influences the shallow sub-surface flow from the third soil layer (Q_{IF}).

4.3.3 Baseflow

The last component that contributes to the total runoff for each grid cell is the baseflow from the active groundwater layer (Q_{BF}). Q_{BF} is calculated as $Q_{BF} = S_{4act}\, x\, J$, where J is the baseflow recession coefficient and depends on the aquifer transmissivity and the aquifer specific yield. Therefore, J controls the direct contribution of groundwater store to the total runoff and hence, to the river discharge.

4.3.4 Evapotranspiration

Actual evapotranspiration consist of transpiration (E_t), bare soil evaporation from the top soil layer (E_b), open-water evaporation (E_w), interception loss (E_i) and evaporation from the melt water store in the snow pack (E_s). Each evapotranspiration component is calculated using the reference potential evapotranspiration ($E_{p,0}$) as basis and the corresponding factor coefficients related with vegetation cover fraction, crop and land cover type, surface water bodies, water stress and the interception flux.

4.4 Data

4.4.1 Meteorological data

The meteorological data required to force PCR-GLOBWB are air temperature, precipitation and reference potential evapotranspiration. Air temperature and precipitation were obtained from the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data (WFDEI) at an original spatial resolution of 0.5° (Weedon et al., 2014). Reference potential evapotranspiration was obtained through the FAO Penman-Monteith equation. Precipitation, air temperature and reference potential evapotranspiration were downscaled from the original spatial resolution to a 0.08° grid. Precipitation and air temperature were downscaled using precipitation and temperature lapse rates derived from the 10' CRU-CL2.0 data (New et al., 2002) through a linear regression analysis (Sutanudjaja et al., 2011). Reference potential evapotranspiration was downscaled using the e2o-downscaling-tools (Weiland et al., 2015; Schellekens and Weiland, 2017).

To test model sensitivity to precipitation, air temperature and reference potential evapotranspiration were fixed and two additional global precipitation products were used: (i) ERA-Interim reanalysis data (EI) from the European Centre for Medium-range Weather Forecasts (ECMWF) at 0.5° resolution (Dee et al., 2011) and (ii) Multi-Source Weighted-Ensemble Precipitation data (MSWEP) by merging gauge, satellite and reanalysis data at 0.25° resolution (Beck et al., 2017a).

The three global precipitation products were inter-compared and interpolated to the two weather station locations found inside the Oum Er Rbia basin (http://www.wmo.int/pages/themes/climate/), Beni Mellal and Kasba Tadla (*Figure 4.1*).

Kling-Gupta efficiency (*KGE*), Nash-Sutcliffe efficiency (*NSE*), Pearson's correlation coefficient (*r*) and Percent Bias (*PBias*) between the interpolated and in situ ground daily data were calculated. A description of the performance metrics with their mathematical formulation is included in *4.5.2 Performance* metrics. These metrics were selected to have detailed information about differences between precipitation products.

4.4.2 Discharge data

Daily river gauge data were obtained from the Oum Er Rbia Hydraulic Agency (ABHOER). In situ data from two gauges in the western region of the basin were used in this study (*Figure 4.1*): Ait Ouchene and Mechra Eddahk. *Table 4.1* summarizes some key hydrological data.

Table 4.1 Hydrological and geographical information of the analyzed catchments at the Oum er Rbia basin.

Station name	River	Upstream basin area [km²]	Outlet location		Elevation [m]
			Longitude	Latitude	
Ait Ouchene	El Abid	2350	6°10′48′′ W	32°13′30′′ N	1070
Mechra Eddahk	Oum er Rbia	6555	6°31′12′′ W	32°26′6′′ N	406

4.4.3 Evapotranspiration data

The GLEAM (Global Land Evaporation Amsterdam Model - http://www.gleam.eu/ -) evapotranspiration product version 3.0a (GLEAM_v3.0a), generated by VU Amsterdam in collaboration with Ghent University (Miralles et al., 2011a,b; Martens et al., 2016b), was used to calibrate PCR-GLOBWB. The product consists of a global dataset based on reanalysis net radiation and air temperature, satellite and gauged-based precipitation, Vegetation Optical Depth (VOD) and snow water equivalents spanning the 35-year period 1980-2014. To generate the GLEAM evapotranspiration product, the GLEAM model separately estimates the different components of terrestrial evaporation, including transpiration, interception loss, bare-soil evaporation, snow sublimation and open-water evaporation. To this end, it consists of four modules: evaporation, stress, soil-water balance and rainfall interception (Martens et al., 2016a). GLEAM (0.25° resolution) was interpolated with distance-weighted average remapping to a 0.08° grid for the period 1980-2010. GLEAM actual evapotranspiration thus obtained was subsequently compared to simulated actual evapotranspiration by PCR-GLOBWB.

4.4.4 Soil moisture data

The ESA CCI surface soil moisture combined product version 2.2 (ESA CCI SM v02.2 CP) was generated as part of the European Space Agency (ESA) soil moisture Climate Change Initiative (CCI) project by the Vienna University of Technology (http://www.esasoilmoisture-cci.org/). A dataset for the 35-year period 1980-2014 of

surface soil moisture was produced using C-band scatterometer data (ERS-1/2 AMI scatterometer, MetOp Advanced Scatterometer -ASCAT-) and multi-frequency radiometer data (SMMR, SSM/I, TMI, AMSR-E, Windsat and AMSR2). Soil moisture retrieved using satellite active microwave data and satellite microwave radiometry were merged to make best use of soil moisture data from the different available satellites and sensors (Liu et al., 2011, 2012; Dorigo et al., 2015). ESA CCI surface soil moisture combined product represents approximately a top soil layer depth of 0.5 - 2 cm. Similarly to GLEAM evapotranspiration, ESA CCI SM product at 0.25° resolution was interpolated with distance-weighted average remapping to a 0.08° grid for the period 1980-2010.

ESA CCI surface soil moisture observations were compared to simulated soil moisture of the first of the three vertical soil layers in PCR-GLOBWB (top 5 cm of soil). Due to differences in layer depth and/or data characteristics, systematic biases between modelled and observed soil moisture may exist (Reichle and Koster, 2004). To overcome this expected discrepancy and match the remotely sensed observations to the statistics of corresponding hydrological model simulations, a mean-standard deviation ($\mu - \sigma$) matching (Draper et al., 2009) was used. This technique was implemented to rescale simulated soil moisture against ESA CCI surface soil moisture time series to have the same mean and variance.

The adjusted simulated surface soil moisture values θ'_{sim} were calculated as

$$\theta'_{sim} = \frac{\sigma_{\theta_{obs}}}{\sigma_{\theta_{sim}}} x (\theta_{sim} - \overline{\theta_{sim}}) + \overline{\theta_{obs}} \qquad (4.1)$$

where θ_{sim} are the simulated soil moisture values, θ_{obs} is the ESA CCI soil moisture observations, $\sigma_{\theta_{sim}}$ and $\sigma_{\theta_{obs}}$ are the standard deviations of the simulated and observed soil moisture values and $\overline{\theta_{sim}}$ and $\overline{\theta_{obs}}$ are the means of the simulated and observed soil moisture values.

When comparing the original and the rescaled soil moisture, it is observed that the mean-standard deviation technique effectively removes the biases between the simulated and observed soil moisture time series (*Figure B. 1*).

4.5 Methodology

4.5.1 Calibration and validation strategy

Alternative single objective calibration approaches based on discharge, actual evapotranspiration and surface soil moisture and a multiobjective calibration approach based on actual evapotranspiration and surface soil moisture were inter-compared. Five different calibration scenarios were carried out. Calibration scenario S0 represents the reference calibration scenario, which was not locally calibrated for the Oum Er Rbia basin, but uses a-priori model parameters derived from vegetation, soil properties and geological information at a global scale (latest model version of PCR-GLOBWB). Calibration scenario S1 aims to calibrate the hydrological model using in situ discharge observations, following

the traditional calibration approach. Calibration scenarios S2 and S3 use GLEAM actual evapotranspiration and ESA CCI surface soil moisture time series for calibration, respectively. Calibration scenario S4 represents the multiobjective calibration approach and it consists of a step-wise calibration scheme that attempts to combine the strengths of calibration scenarios S2 and S3. Step one is simply scenario S2, where all the model parameters are allowed to be adjusted based on GLEAM actual evapotranspiration. In step two, those parameters that are clearly identified by calibration scenario S2 are held constant and the remaining parameters are allowed to be adjusted according to ESA CCI surface soil moisture, calibration scenario S3.

The five calibration scenarios were analyzed for each of the three global precipitation products to study their impact on model parameters calibration and model performance. The calibration scenarios are described in *Table 4.2*, including the scenario identifier.

Table 4.2 Calibration scenarios

Scenario identifier	**Description**
S0	Reference scenario
S1	Calibration using in situ-observed discharge time series
S2	Calibration using GLEAM actual evapotranspiration time series
S3	Calibration using ESA CCI surface soil moisture time series
S4	Step-wise calibration: using GLEAM actual evapotranspiration and ESA CCI surface soil moisture time series

For the calibration using in situ observed discharge time series (S1), two river gauge observation time series were used (*4.4.2 Discharge* data). The objective function to maximize for the calibration scenarios was Kling-Gupta efficiency (*KGE*), instead of the traditional Mean Squared Error (*MSE*) or Nash Sutcliffe efficiency (*NSE*) to avoid underestimating the variability of values (Gupta et al., 2009). The mathematical formulation and description of the used objective function are included in *4.5.2 Performance* metrics.

To calibrate PCR-GLOBWB for each of the three precipitation products, 81 runs with different parameter values were simulated: minimum soil water capacity (W_{min}), soil saturated hydraulic conductivites (K_{sat1}, K_{sat2} and K_{sat3}) and baseflow recession coefficient (J). These model parameters, which vary spatially over the basin, influence different model parts of the model behaviour, as it was explained in *4.3 Hydrological model*. For the variation of the parameter values, spatially uniform prefactors were used: f_w, f_K and f_j (*Table 4.3*). The remaining model parameters were kept fixed.

The prefactors to vary model parameter values were referred to the parameters of the S0 calibration scenario. The spatial distribution of the parameters W_{min}, K_{sat} and J used in S0 scenario can be found in *Figure 4.3*.

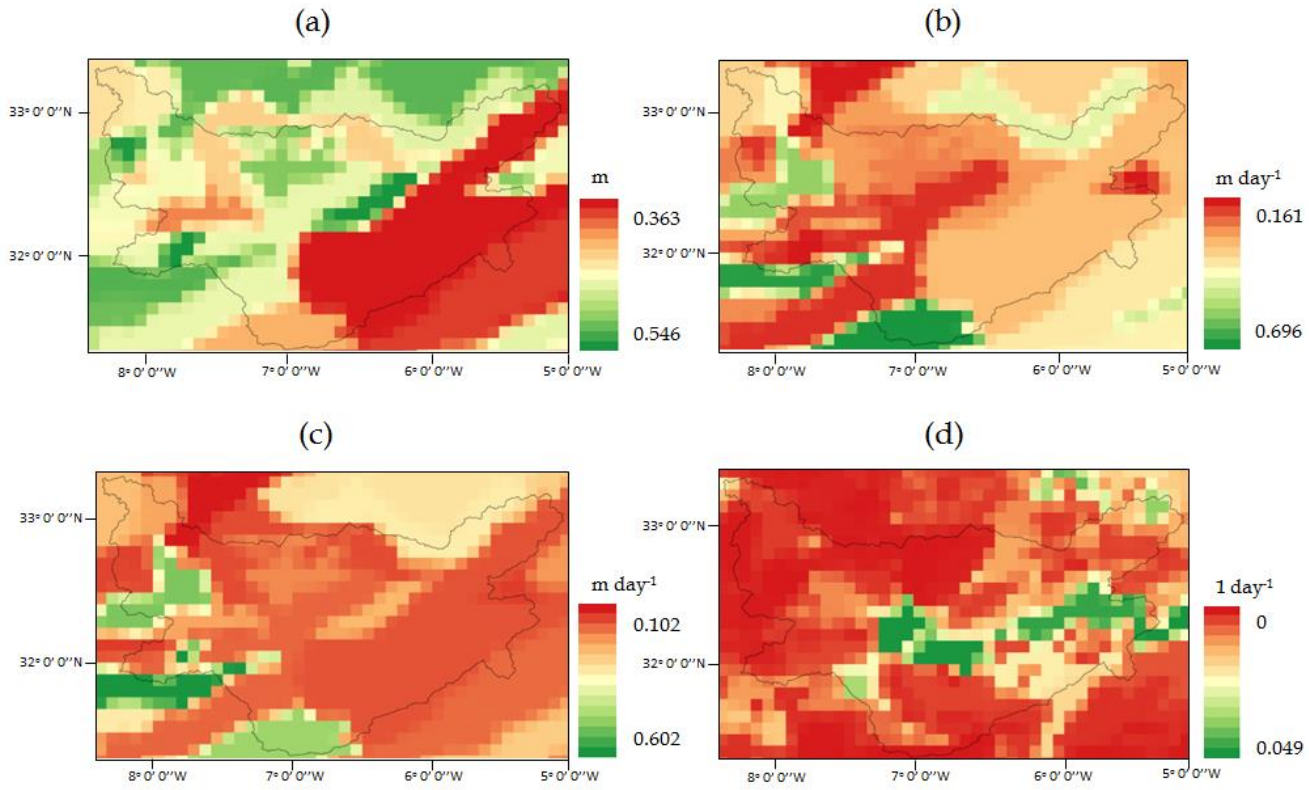


Figure 4.3 Initial model parameter values for the S0 calibration scenario (reference): (a) total soil water storage capacity ($W_{max} = SC_1 + SC_2 + SC_3$), (b) saturated hydraulic conductivity of the first and second soil layers (K_{sat1} and K_{sat2}), (c) saturated hydraulic conductivity of third soil layer (K_{sat3}) and (d) baseflow recession coefficient (J).

Furthermore, the uncertainty of reference potential evapotranspiration ($E_{p,0_{ref}}$) was also investigated using a correction prefactor, f_e, to this model variable. Considered values for f_e prefactor are included with the previously mentioned ones in *Table 4.3*.

As reference calibration scenario, S0 prefactors are: $f_w = 1$, $f_k = 0$, $f_j = 1$ and $f_e = 1$. The model performances of all the simulations were evaluated for each of the five calibration scenarios to identify the best prefactor sets as the calibrated prefactor sets.

All the simulations were performed at a daily temporal resolution for the 32-year period 1979-2010. The 2-year period 1979-1980 was used to spin up the hydrological model until reaching a dynamically steady state. The model was calibrated based on monthly values of discharge, actual evapotranspiration and surface soil moisture. Validation was also carried out at a monthly temporal resolution but exclusively for streamflow, aiming to analyze if similar discharge estimates may be obtained with a calibrated model based on remotely sensed observations (S2, S3 and S4), in comparison with a model traditionally calibrated to in situ discharge data (S1). The 13-year period 1981-1993 was used for model calibration and during the 17-year period 1994-2010, the model was validated.

Table 4.3 Parameter values used in the calibration processes.

Parameters ID	Description	Pre-factors	Parameter values
W_{min}	Minimum soil water capacity	$f_w \in \{0.75, 1, 1.25\}$	$W_{min} = f_w \cdot W_{max}$
K_{sat1}	Saturated hydraulic conductivity of first soil layer	$f_k \in \{-0.25, 0, 0.25\}$	$\log(K_{sat1}) = f_k + \log(K_{sat1_{ref}})$
K_{sat2}	Saturated hydraulic conductivity of second soil layer	$f_k \in \{-0.25, 0, 0.25\}$	$\log(K_{sat2}) = f_k + \log(K_{sat2_{ref}})$
K_{sat3}	Saturated hydraulic conductivity of third soil layer	$f_k \in \{-0.25, 0, 0.25\}$	$\log(K_{sat3}) = f_k + \log(K_{sat3_{ref}})$
J	Baseflow recession coefficient	$f_j \in \{-0.5, 0, 0.5\}$	$\log(J) = f_j + \log(J_{ref})$
$E_{p,0}$	Reference potential evapotranspiration	$f_e \in \{0.75, 1, 1.25\}$	$E_{p,0} = f_e \cdot E_{p,0_{ref}}$

4.5.2 Performance metrics

To inter-compare the three global precipitation products six metrics were used: Nash-Sutcliffe efficiency (*NSE*), Kling-Gupta efficiency (*KGE*), Pearson's correlation coefficient (*r*) and Percent Bias (*PBias*). Moreover, one of those metrics, *KGE*, was chosen as objective function to calibrate and validate model performance for each calibration scenario. *NSE*, *PBias* and *r* were also used as additional assessment measurements in the validation procedure.

Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970), *NSE*, is defined as

$$NSE = 1 - \frac{\sum_{t=1}^{n}[x(t) - y(t)]^2}{\sum_{t=1}^{n}[y(t) - \bar{y}]^2} \quad (4.2)$$

where $x(t)$ and $y(t)$ are the modeled and observed variable at t time step (months), $\bar{y}$ is the mean of observed data and n is the total number of observations. *NSE* is widely used for calibrating and validating hydrological models in terms of discharge. *NSE* varies from $-\infty$ to 1. If $NSE = 0$, modeled values perform as well as the mean of the observations. If $NSE < 0$, modeled values perform worse than the mean of the observations.

Gupta et al. (2009) analyzed various decompositions of *NSE* and proposed an alternative model performance criteria, Kling-Gupta efficiency (*KGE*), to avoid the problems that can be derived of using the *NSE* criterion (e.g. high sensitivity to extreme values and bias). *KGE* is given as

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \quad (4.3)$$

where r represents the Pearson's correlation coefficient, α is the ratio between the variance of the modeled variable and the variance of the observed variable and β is the ratio between the mean of the modeled variable and the mean of the observed variable, i.e. βrepresents the bias. Analogous to *NSE, KGE* ranges from $-\infty$ to 1 with an ideal value of 1. *KGE* measures simultaneously bias, variability and correlation.

Pearson's correlation coefficient (Pearson, 1986), r, measures the degree of linear association between modeled and observed values and it is defined as

$$r = \frac{\sum_{t=1}^{n}(x(t)-\bar{x})(y(t)-\bar{y})}{\sqrt{\sum_{t=1}^{n}(x(t)-\bar{x})^2}\ \sqrt{\sum_{t=1}^{n}(y(t)-\bar{y})^2}} \quad (4.4)$$

where $x(t)$ and $y(t)$ are the modeled and observed variable at t time step (months), $\bar{y}$ is the mean of observed data, $\bar{x}$ is the mean of modeled data and n is the total number of observations. r varies within the interval [-1,1]. r is mainly used in hydrological modelling to evaluate the timing of modeled to observed time series.

Percent Bias indicates the average tendency of the modeled values to over- or underestimate the observations. *PBias*, is calculated in percentage terms as

$$PBias = 100 \times \frac{\sum_{t=1}^{n}(x(t)-y(t))}{\sum_{t=1}^{n} y(t)} \quad (4.5)$$

The optimal value of *PBias* is 0.

When the performance metrics were calculated between simulated and observed soil moisture estimates, the subscript $_{sm}$ was added to the metric, i.e. NSE_{sm}, KGE_{sm}, r_{sm} and $PBias_{sm}$. Similarly, when comparing actual evapotranspiration estimates, precipitation and discharge, the added subscripts were $_{evap}$, $_{precip}$ and $_{q}$, respectively.

4.6 Results

4.6.1 Inter-comparison of precipitation products

To inter-compare precipitation products, the annual mean precipitation for the study time period (1979-2010) for each forcing dataset was calculated (*Figure 4.4a, Figure 4.4b* and *Figure 4.4c*). In addition to the spatial resolution difference, MSWEP was able to capture the rainfall pattern over the Atlas Mountains rather well, which was only roughly distinguished by WFDEI and unrecognized by EI. The finer spatial resolution and the combination of reanalysis, satellite and in situ data are probably the reasons for its more plausible spatial pattern. Furthermore, climatology of precipitation products was analyzed (*Figure 4.4d*). WFDEI ranged from 4.50 mm in July to 57 mm in February, whereas EI and MSWEP showed a lesser variability with precipitation values from 10.50 mm in July to 42.60 mm in November. Smaller differences between WFDEI and EI and MSWEP were observed

during the summer months. EI and MSWEP showed similar temporal precipitation patterns. Annual mean precipitation over the entire basin obtained with MSWEP (355.15 mm) was approximately 80 mm higher than with EI (276.67 mm). Similar annual median values were obtained with the three global precipitation products, although the distribution of WFDEI highly differed from the other two products.

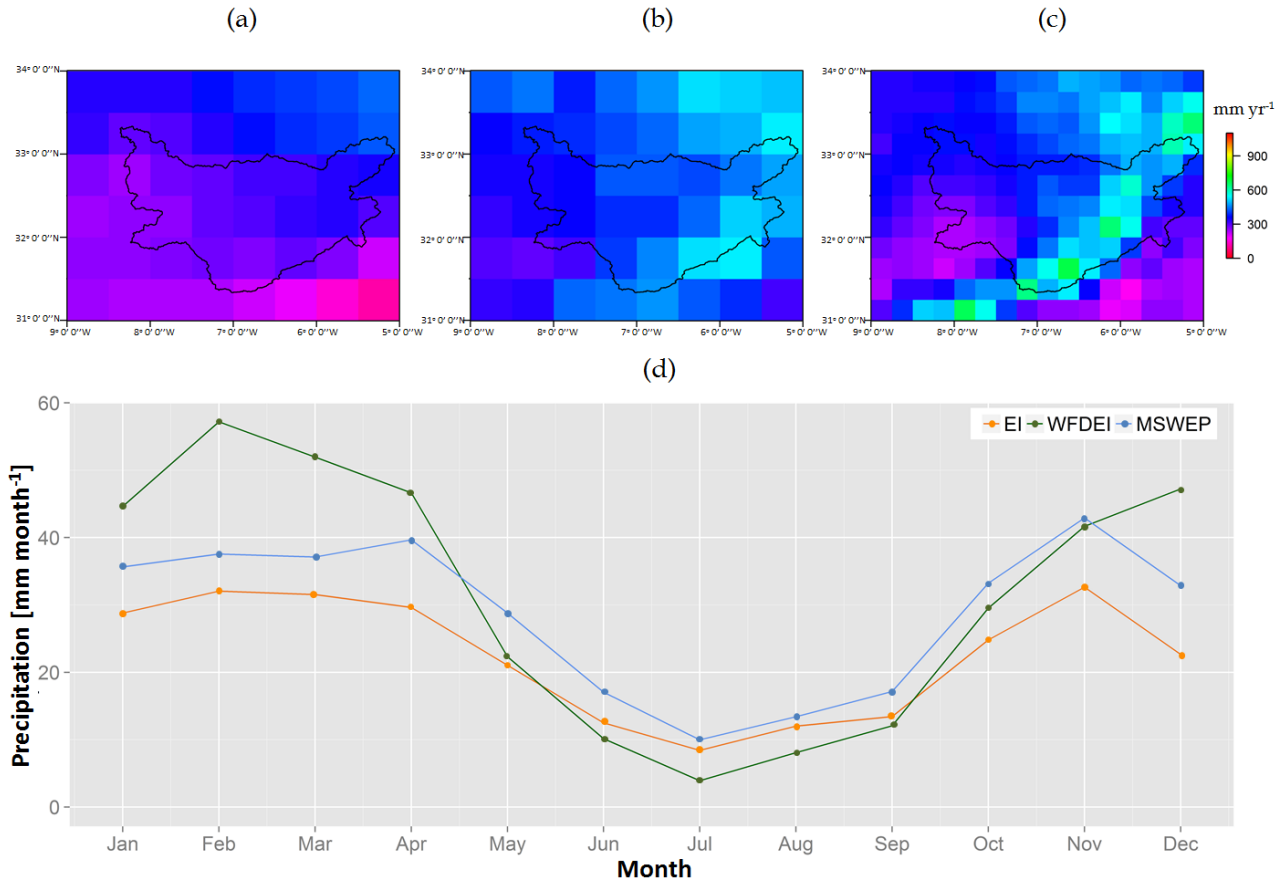


Figure 4.4 (a) EI annual mean precipitation, (b) WFDEI annual mean precipitation and (c) MSWEP annual mean precipitation for 1979–2010 time period and (d) climatology of EI, WFDEI and MSWEP precipitation products.

Moreover, various performance metrics between the interpolated and in situ ground data were calculated and shown in *Figure 4.5*. Overall, EI and MSWEP provided a better fit to the station data compared to WFDEI, with higher KGE_{precip}, NSE_{precip} and r_{precip} than WFDEI. When comparing EI with MSWEP, similar values of KGE_{precip} and NSE_{precip} were found, whereas higher differences existed in r_{precip} and $PBias_{precip}$. In terms of correlation MSWEP showed the best performance, but EI showed the lowest *PBias* at both weather stations, with a value of less than 10%. Only two weather stations were found within the basin for the previous analysis. These measurements were considered too scarce to cover the basin and to discard the precipitation product with the lowest performance (WFDEI). Therefore, the three global precipitation products were used to calibrate PCR-GLOBWB under the five calibration scenarios.

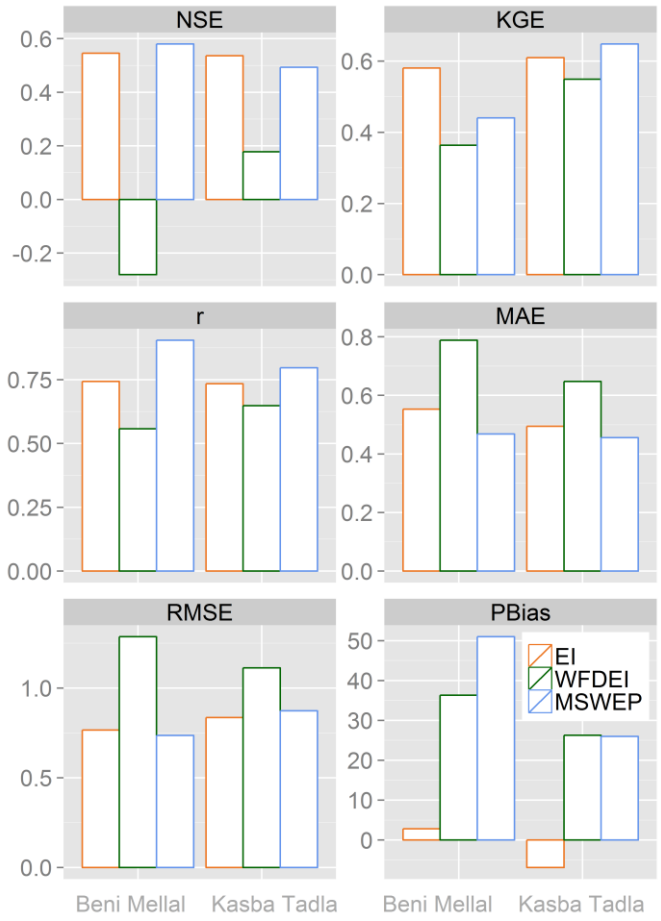


Figure 4.5 Performance metrics of daily EI, WFDEI and MSWEP precipitation products at Beni Mellal and Kasba Tadla weather stations, including *KGE, NSE, r* and *PBias*.

4.6.2 Calibration results

Model parameters were calibrated using discharge, evapotranspiration and soil moisture observations through five different calibration scenarios for the time period 1981-1993. *Figure 4.6* shows results of all runs produced in this study for different calibration scenarios based on: in situ discharge observations (S1) at Ait Ouchene (*Figure 4.6*a) and Mechra Eddahk (*Figure 4.6*b), GLEAM actual evapotranspiration (S2, *Figure 4.6*c) and ESA CCI surface soil moisture (S3, *Figure 4.6*d). For each sub-figure in *Figure 4.6*, *KGE* results (y-axis) of using the three precipitation products were plotted in different rows (top: EI, middle: WFDEI and bottom: MSWEP) and prefactor values were plotted in different columns (x-axis, first column: f_e, second column: f_j, third column: f_k and fourth column: f_w). Each scatterplot contains 81 dots representing each run with a different combination of parameter values. This means that the *KGE* values are the same in the four scatterplots of a row (y- axis), but in each of these scatterplots, they were plotted against a different prefactor (x-axis). With *Figure 4.6*, prefactor, and therefore parameter, ranges leading to better and worse performances could be distinguished, as well as their global optimal values. If no optimal value could be inferred, prefactors from the calibration scenario S0 were maintained ($f_e = 1$, $f_j = 0$, $f_k = 0$ and $f_w = 1$).

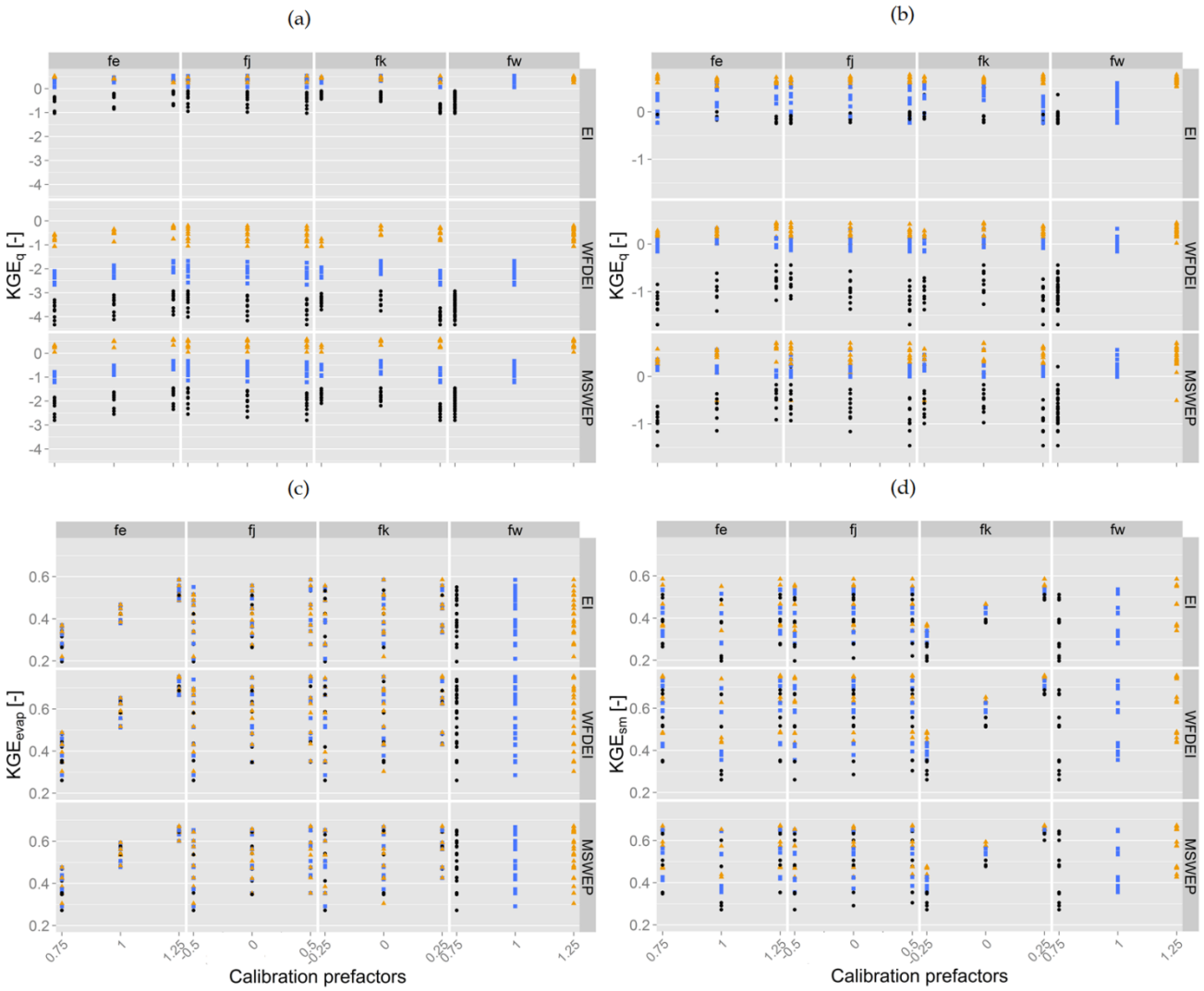


Figure 4.6 Scatter plots of discharge performance indicator *KGE* based on the monthly observations versus pre-factors f_e, f_j, f_k and f_w for the calibration scenarios S1 ((a) Ait Ouchene (b) Mechra Eddahk), S2 (c) and S3 (d). In each panel, columns indicate the different calibrated pre-factors and rows indicate the three global precipitation products used as model forcing. Different colours and dot shapes indicate different f_w values.

Once the best runs for each calibration scenario were identified, their discharge performance was checked at the two gauging stations: Mechra Eddahk, in *Figure 4.7*, and Ait Ouchene, in *Figure B. 2*. Observed discharge (y-axis) and estimated discharge (x-axis) were plotted in *Figure 4.7* for the five calibration scenarios. Different rows in *Figure 4.7* indicate the three global precipitation products (top: EI, middle: WFDEI and bottom: MSWEP) and different columns indicate the five calibration scenarios (first column: S0, second column: S1, third column: S2, fourth column: S3 and fifth column: S4). The performance indicators *NSE* and *KGE* for discharge were included in every scatterplot in *Figure 4.7* (NSE_q and KGE_q).

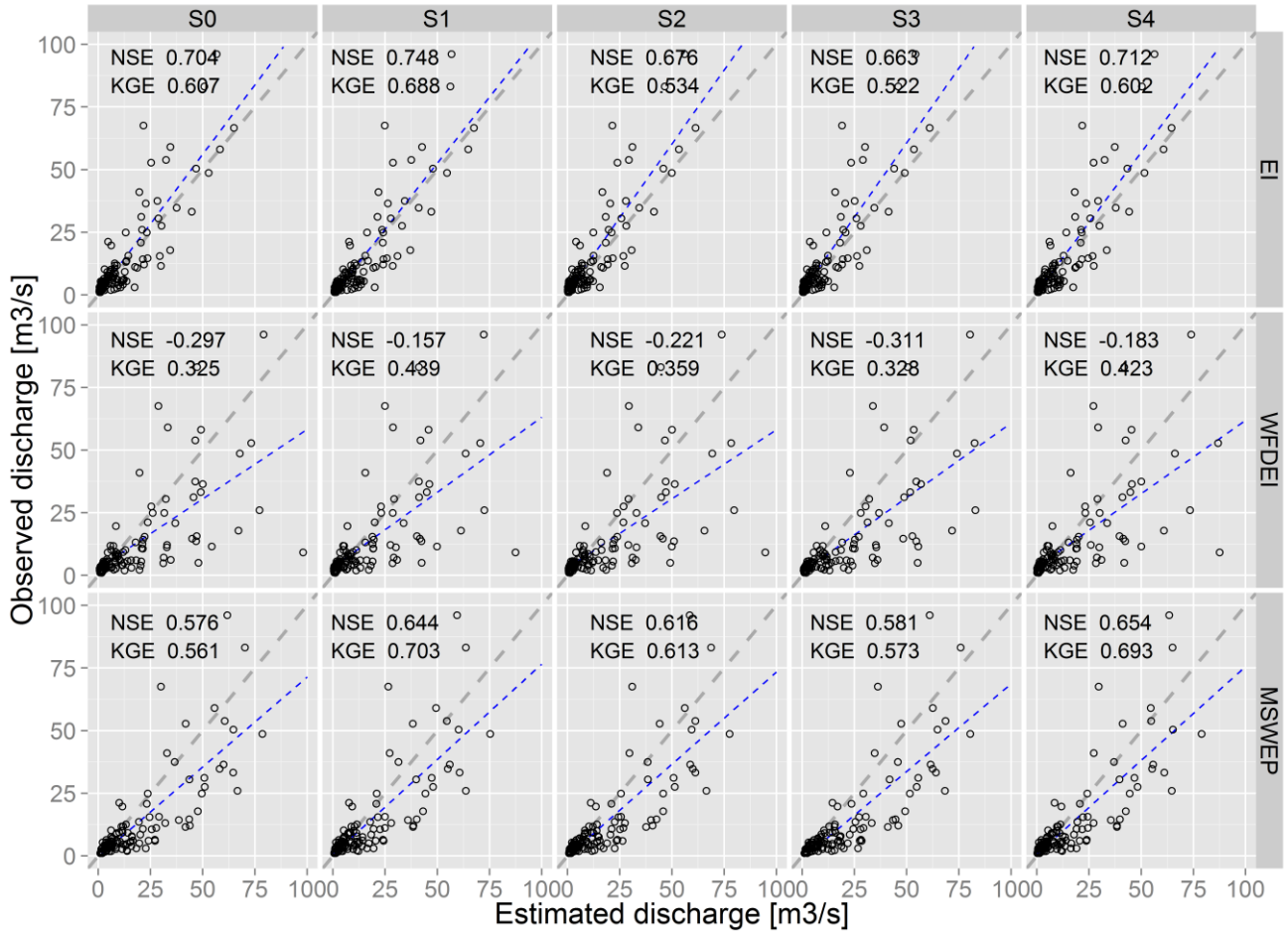


Figure 4.7 Scatter plots of monthly estimated discharge (x-axis) and observed discharge (y-axis) at Mechra Eddahk. Rows indicate the three global precipitation products and columns indicate the five calibration scenarios.

To summarize results shown in *Figure 4.6* and *Figure 4.7*, *Table 4.4* includes for each calibration scenario the identified optimal parameters values and the KGE_q performance values at Ait Ouchene and Mechra Eddahk.

4.6.2.1 Calibration using in situ observed discharge time series (S1)

Figure 4.6a and *Figure 4.6b* (calibration scenario S1) are similar. From these figures, f_e (first column) and f_w (fourth column) were well identified by discharge calibration at both gauging stations when forced with any of the three precipitation products. $f_e = 1.25$ and $f_w = 1.25$ led to the highest KGE_q values. However, it was not possible to identify the best prefactors of f_j (second column) and f_k (third column). There were no clear and distinct maximum values in the scatterplots of these figures, hence $f_j = 0$ and $f_k = 0$ were used.

From *Figure 4.7* (second column), the highest discharge performance was obtained when the model was calibrated with in situ discharge observations (S1).

Table 4.4 Parameter identifiability values and optimal values for each calibration scenario.

Calibration scenario	f_e	f_j	f_k	f_w	*KGE* (Ait Ouchene)	*KGE* (Mechra Eddahk)
S0	1	0	0	1	0.470/ − 1.906/ − 0.542[2]	0.607/0.325/0.561
S1	1.25	NI[1]	NI	1.25	0.510/ − 0.494/0.520	0.688/0.439/0.703
S2	1.25	NI	NI	NI	0.508/ − 0.580/0.342	0.602/0.423/0.693
S3	NI	NI	0.25	1.25	0.487/ − 0.607/0.331	0.634/0.369/0.613
S4	1.25	NI	0.25	1.25	0.478/ − 0.768/0.271	0.522/0.328/0.573

[1] NI indicates that the parameter was not identifiable. [2] *KGE* values are obtained from observed and simulated discharge when PCR-GLOBWB is forced with EI/WFDEI/MSWEP

For all the calibration scenarios, a few general observations could be made. Scatterplots (*Figure 4.7*) highlighted an overall better agreement and a lower bias between discharge observations and estimates for the Ait Ouchene (*Figure B. 2*) than for Mechra Eddahk station. KGE_q values at Ait Ouchene station for calibration scenario S0 were lower than for Mechra Eddahk station. This may be due to their different locations within the basin, the former one being situated in the Atlas Mountains, where precipitation estimates can be less accurate, and in a tributary of the Oum Er Rbia River, whose representation in PCR-GLOBWB can be limited by the model spatial resolution.

Scatterplots (*Figure 4.7*) also showed that estimated discharges were closer to observed discharges at both gauging stations when PCR-GLOBWB was forced with EI precipitation. Moreover, scatterplots indicated a worse agreement and a tendency to overestimate discharge when WFDEI and MSWEP were used. KGE_q values for the reference calibration scenario S0 at Mechra Eddahk were 0.607, 0.325 and 0.561 when EI, WFDEI and MSWEP were used as forcing data respectively. These performance discrepancies were related with the differences between EI, WFDEI and MSWEP precipitation products discussed in *4.6.1 Inter-comparison of precipitation products*. The lower quality of WFDEI in this region compared with the other precipitation datasets may be a possible reason of the lower discharge performance. When MSWEP was compared with in situ precipitation data, performance in terms of correlation was higher than EI. However, EI showed less bias. The higher performance of discharge estimates when PCR-GLOBWB was forced with EI may be due to this bias difference and that the validation was carried out at a monthly temporal resolution, reducing the impact of correlation.

4.6.2.2 Calibration using GLEAM actual evapotranspiration (S2)

Figure 4.6c (calibration scenario S2) indicated that only prefactor f_e (first column) could be clearly identified (the highest KGE_{evap} values were obtained with $f_e = 1.25$), whereas the remainder of the prefactors (f_j, f_w and f_k) were non identifiable, suggesting that evapotranspiration-based calibration may be unreliable in their identification. Therefore, model run with prefactors $f_e = 1.25$, $f_j = 0$, $f_k = 0$ and $f_w = 1$ was considered as the calibrated run based on the evapotranspiration performance.

From *Figure 4.7* (third column), results indicated an increase of KGE_q and NSE_q values when GLEAM evapotranspiration was used for model calibration compared to the reference scenario (S0, first column of *Figure 4.7*). However, higher model performance values were obtained when calibrating based on in situ discharge observations (S1, second column of *Figure 4.7*).

4.6.2.3 Calibration using ESA CCI surface soil moisture time series (S3)

Figure 4.6d (calibration scenario S3) indicated that prefactors f_k (third column) and f_w (fourth column) could be identified, $f_w = 1.25$ and $f_k = 0.25$. There was a clear maximum value of KGE_{sm} with these prefactors values. Prefactors f_e (first column) and f_j (second column) were not identifiable when soil moisture was used for calibration. Therefrom, the calibrated run based on soil moisture performance was the model run with prefactors $f_e = 1$, $f_j = 0$, $f_k = 0.25$ and $f_w = 1.25$. This implies that ESA CCI soil moisture may be used to indirectly tune groundwater recharge by calibrating the upper soil saturated hydraulic conductivities, K_{sat}.

From *Figure 4.7* (fourth column), scatterplots indicated an improvement in the correspondence between observed and estimated discharge compared to the non-calibrated scenario (S0, first column of *Figure 4.7*). Similarly to calibration scenario S2 (third column of *Figure 4.7*), this improvement was lower than when the model was calibrated based on ground discharge observations (S1, second column of *Figure 4.7*).

The calibrated runs based on evapotranspiration (S2, third column of *Figure 4.7*) and soil moisture (S3, fourth column of *Figure 4.7*) resulted in lower discharge performances compared to the reference scenario (S0) at some cases, e.g. when EI precipitation was used at Mechra Eddahk location, KGE_q *(S0) = 0.607,* KGE_q *(S2) = 0.534* and KGE_q *(S3) = 0.522*.

4.6.2.4 Step-wise calibration using GLEAM actual evapotranspiration and ESA CCI surface soil moisture time series (S4)

Calibration scenario S4 attempted to combine the strengths of scenarios S2 and S3. In the first step, the model was calibrated using GLEAM evapotranspiration (S2, *Figure 4.6c*). From *Figure 4.6c*, only f_e prefactor was well identified (the highest KGE_{evap} value was obtained with $f_e = 1.25$). In the second step, f_e prefactor that had been identified was held constant and the remaining three prefactors were allowed to be calibrated according to ESA CCI soil moisture (S3, *Figure 4.6d*). From *Figure 4.6d*, f_w and f_k were identifiable (the highest KGE_{sm} values were obtained with $f_w = 1.25$ and $f_k = 0.25$). As a result, for calibration scenario S4, the prefactors identified during the evapotranspiration calibration (S2): $f_e = 1.25$ and during the soil moisture calibration (S3): $f_w = 1.25$ and $f_k = 0.25$ were adopted. This step-wise calibration approach using multiple system variables allowed to identify more parameters than when those variables were separately used. Nonetheless, neither of the steps in calibration scenario S4 allowed the clear identification of f_j, so its value for the calibration scenario S0 was used, $f_j = 0$.

From *Figure 4.7* (fifth column), calibration using GLEAM evapotranspiration and ESA CCI soil moisture led to further improvements than when these observations were separately

used. For example, when MSWEP precipitation was used to model discharge at Mechra Eddahk station, KGE_q varied between 0.703, 0.693, 0.613 and 0.573 for calibration scenarios S1, S4, S2 and S3, respectively (KGE_q = *0.561* for the reference scenario S0). At Ait Ouchene station (*Figure B. 2*), KGE_q varied between 0.520, 0.342, 0.331 and 0.271 for calibration scenarios S1, S4, S2 and S3, respectively (KGE_q = *-0.542* for the reference scenario S0).

4.6.3 Validation results

Once the model had been calibrated for each calibration scenario and each precipitation product, comparisons between estimates (before and after the calibration) and observations of actual evapotranspiration, surface soil moisture and discharge were carried out for the validation time period (1994-2011). To perform the analysis of these results, time series plots were included in *Figure 4.8* and *Figure 4.9*.

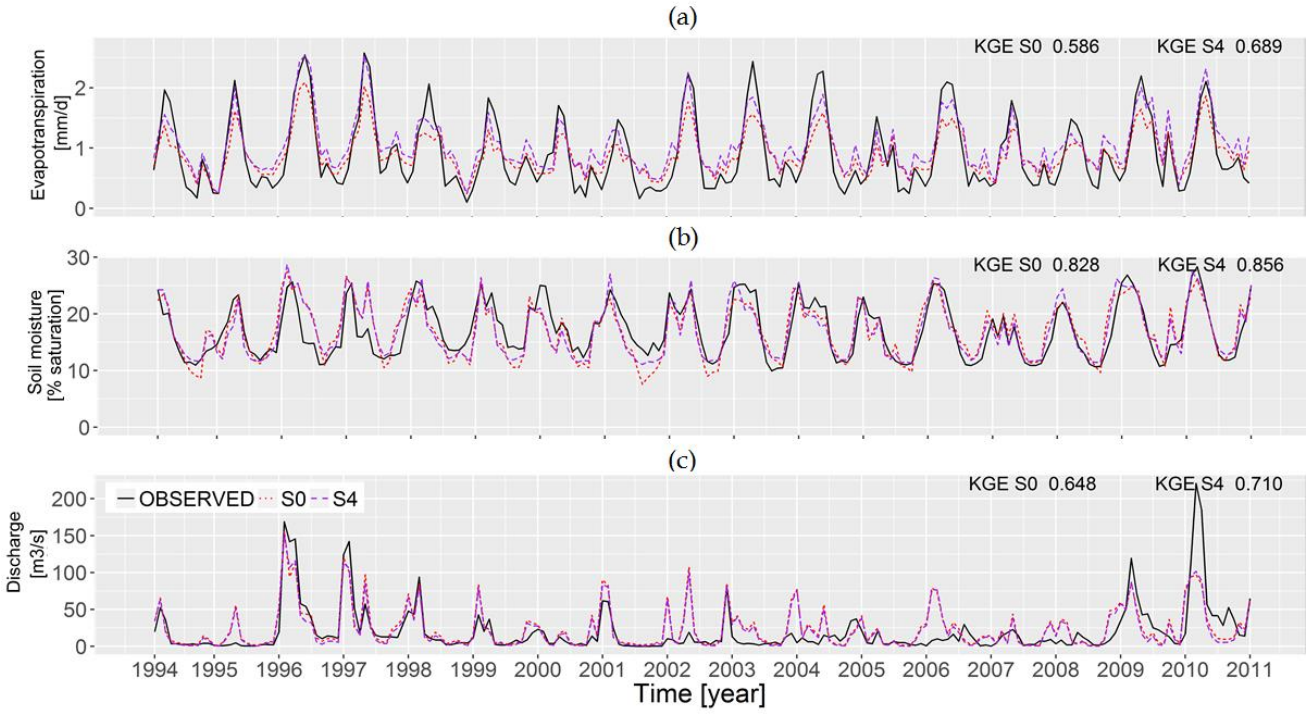


Figure 4.8 (a) Monthly GLEAM actual evapotranspiration (black) and estimated actual evapotranspiration (red and purple) time series over the entire Oum er Rbia basin. (b) Monthly ESA CCI soil moisture (black) and estimated soil moisture (red and purple) time series over the entire Oum er Rbia basin. (c) Monthly observed discharge (black) and estimated discharge (red and purple) time series at Mechra Eddahk. The red dashed lines represent estimates from calibration scenario S0 (reference scenario). The purple dashed lines represent the calibrated time series from calibration scenario S4 which are taken from the runs that yield the best simulations. Estimated time series over the entire Oum er Rbia basin for the validation time period obtained with MSWEP precipitation are shown.

In *Figure 4.8a*, simulated actual evapotranspiration time series of the reference run (S0, red dashed line) and the step-wise calibrated run (S4, purple dashed line) were plotted against GLEAM actual evapotranspiration observations (black line). Similarly as *Figure 4.8a*, *Figure 4.8b* shows simulated surface soil moisture of the reference run (S0, red dashed line) and the step-wise calibrated run (S4, purple dashed line) plotted against ESA CCI surface soil moisture time series (black line). The rescaled soil moisture time series (after mean-standard deviation matching technique applied, see *4.4.4 Soil moisture data*) are shown. In *Figure 4.8c*, estimated discharge of the reference run (S0, red dashed line) and the step-wise calibrated run (S4, purple dashed line) were plotted against discharge observations (black line) at Mechra Eddahk. *KGE* values for actual evapotranspiration, surface soil moisture and discharge were included in *Figure 4.8a*, *Figure 4.8b* and *Figure 4.8c*. For the sake of simplicity, only results when the model was forced with MSWEP precipitation are shown.

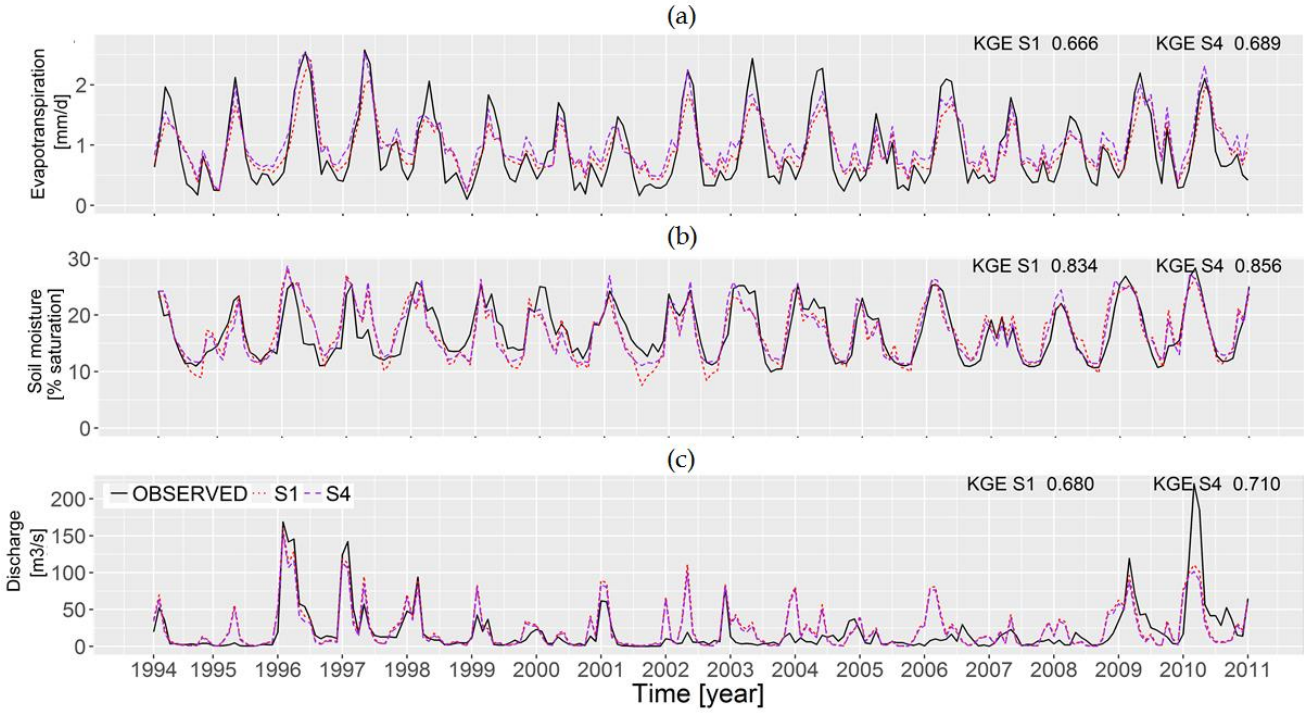


Figure 4.9 (a) Monthly GLEAM actual evapotranspiration (black) and estimated actual evapotranspiration (red and purple) time series over the entire Oum er Rbia basin. (b) Monthly ESA CCI soil moisture (black) and estimated soil moisture (red and purple) time series over the entire Oum er Rbia basin. (c) Monthly observed discharge (black) and estimated discharge (red and purple) time series at Mechra Eddahk. The red dashed lines represent estimates from calibration scenario S1. The purple dashed lines represent the calibrated time series from calibration scenario S4 which are taken from the runs that yield the best simulations. Estimated time series over the entire Oum er Rbia basin for the validation time period obtained with MSWEP precipitation are shown.

Similarly to *Figure 4.8*, *Figure 4.9* shows simulated evapotranspiration (*Figure 4.9a*), surface soil moisture (*Figure 4.9b*) and discharge (*Figure 4.9c*) against observations. However, in *Figure 4.9*, estimates of the discharge-calibrated run (S1, red dashed line) and the step-wise calibrated run (S4, purple dashed line) were plotted against observations (black line). With *Figure 4.9*, the impact on calibration of using in situ discharge and remotely sensed evapotranspiration and soil moisture observations could be compared.

Reference run (S0) provided evapotranspiration (*Figure 4.8a*) and soil moisture (*Figure 4.8b*) estimates fairly close to observations with $KGE_{evap} = 0.586$ and $KGE_{sm} = 0.828$. Discharge estimates of calibration scenario S0 performed well at Mechra Eddahk (*Figure 4.8c*). Discharge performance of the reference run (S0) was lower at Ait Ouchene. From *Figure 4.8a*, the calibration procedure based on GLEAM evapotranspiration and ESA CCI soil moisture (S4) produced an increase of 18% in KGE_{evap} compared to the reference run (S0). From *Figure 4.8b*, estimated and observed surface soil moisture time series showed a good correspondence. KGE_{sm} difference of 0.028 was found between the reference (S0) and the step-wise (S4) calibration scenarios. From *Figure 4.8c*, the step-wise calibrated run (S4) reproduced the monthly observed discharge well, except some simulated extreme peaks which were not observed, e.g. January and June in 2002 and some which were not simulated properly, e. g. January and May in 1996 and 1997. This lack of fit may be due to errors in the precipitation data, because higher discharge differences were shown when WFDEI and MSWEP products were used in comparison to EI product. Other possible reasons may be related with model structural deficiencies. When comparing discharge time series (*Figure 4.8c*), calibration scenario S4 slightly improved KGE_q compared to calibration scenario S0, with KGE_q values varying from 0.648 to 0.710.

From *Figure 4.9a*, an increase of 14% in KGE_{evap} was produced when the discharge calibration scenario (S1) was used compared to the reference run (S0). This improvement in evapotranspiration estimates was higher when calibrating the model using GLEAM evapotranspiration and ESA CCI soil moisture (S4, $KGE_{evap} = 0.689$) than when calibrating it only for ground discharge (S1, $KGE_{evap} = 0.666$). Similarly to *Figure 4.9a* showing the evapotranspiration comparison, *Figure 4.9b* indicated that a higher KGE_{sm} value was obtained when using GLEAM and ESA CCI observations for calibration (S4, $KGE_{sm} = 0.856$) than when calibration was based on in situ discharge observations (S1, $KGE_{sm} = 0.834$). From *Figure 4.9c* comparing discharge time series at Mechra Eddahk, step-wise calibration scenario S4 led to an increase of 10% in KGE_q, compared to the increase of 5% obtained when discharge observations were used for calibration (S1). However, at Ait Ouchene the discharge performance improvement was lower when evapotranspiration and soil moisture observations (S4) were used for calibration than when in situ discharge measurements were used (S1). This may be due to the lower performance of the reference run (S0) at Ait Ouchene.

To further understand the added value of using GLEAM evapotranspiration and ESA CCI soil moisture data for model calibration and inter-compare the impact of the calibration scenarios, the variations of KGE_q, NSE_q, r_q and $PBias_q$ between each calibration scenario (S1, S2, S3 and S4) and the reference calibration scenario (S0) were calculated and plotted for the

validation time period in *Figure 4.10*. Rows indicate the three global precipitation products and columns indicate the performance indicators. The variations of the performance metrics are shown with barplots for the two gauging stations, Ait Ouchene and Mechra Eddahk. At each location, a positive value of KGE_q, NSE_q, $PBias_q$ and r_q means that either S1, S2, S3 or S4 scenario obtained a higher skill score than S0, whereas a negative value means that those scores decreased after calibration.

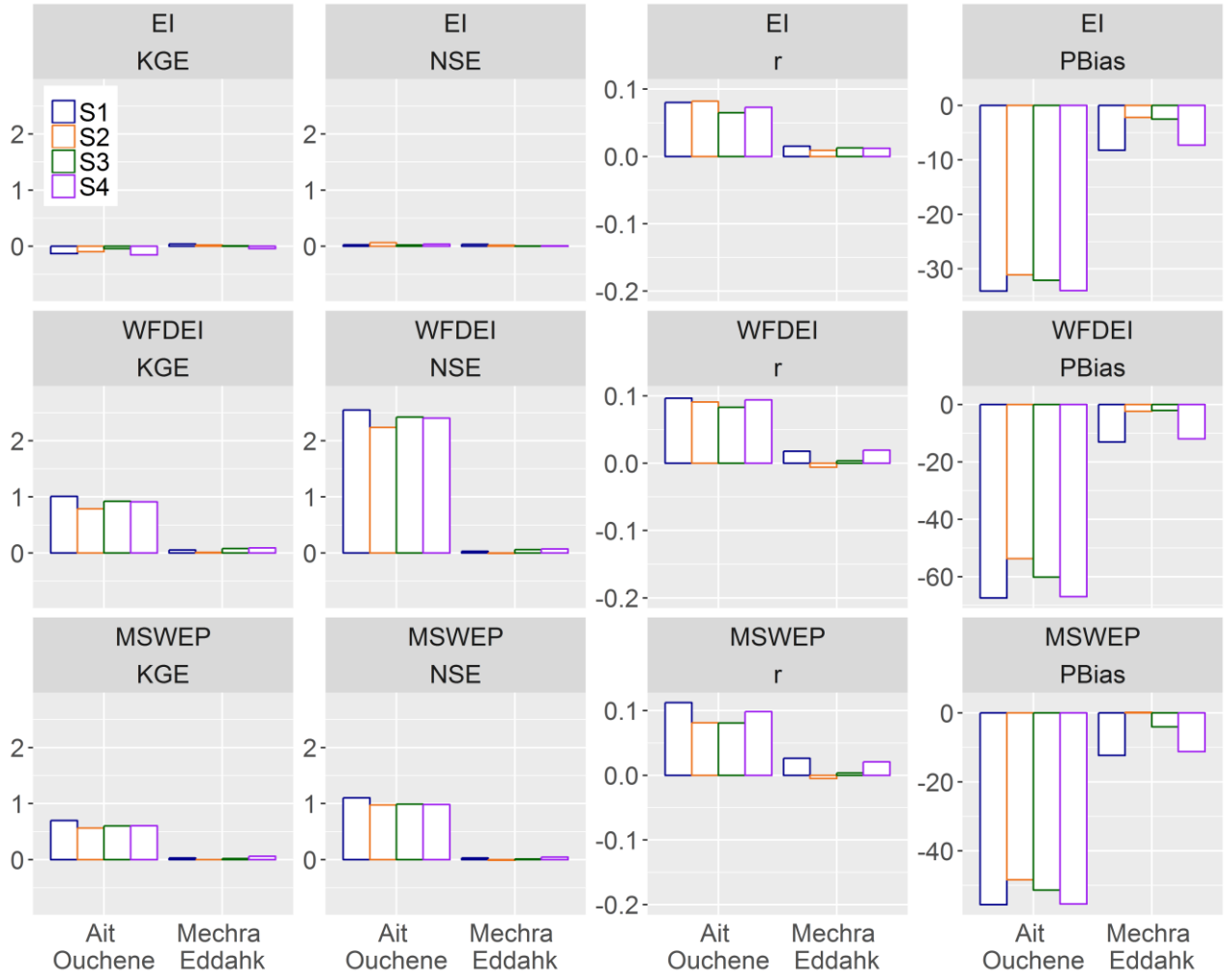


Figure 4.10 KGE, NSE, r and *PBias* variations comparing monthly discharge estimates of calibration scenarios S1, S2, S3 and S4 with S0. Rows indicate the three global precipitation products and columns indicate the performance metrics.

Figure 4.10 showed that variations of the performance indicators were lower when EI precipitation product was used. The highest differences between the calibration scenarios were obtained when the model was forced with WFDEI precipitation. In the inter-comparison of the calibration scenarios, calibration scenario using in situ observed discharge data (S1) obtained overall the highest increase of KGE_q, NSE_q and r_q and the highest reduction of $PBias_q$ when any of the precipitation products were used, as it was expected. Similar KGE_q and NSE_q increases and $PBias_q$ decreases were obtained when the

model was calibrated using only soil moisture (S3) and using the combination of evapotranspiration and soil moisture (S4), but larger improvements in r_q were obtained with the step-wise calibration scenario (S4). KGE_q, NSE_q and r_q gains when comparing calibration scenarios S2 and S0 were positive, but of a lower magnitude than when model was calibrated in scenarios S3 and S4. The higher performance of scenario S4 may be due to the fact that this calibration approach used multiple system variables providing more hydrological information and allowing the identification of more parameters than when those variables were separately used.

In each barplot of *Figure 4.10*, metrics improvements were larger at Ait Ouchene station than at Mechra Eddahk station. This is due to the lower discharge performance for the reference calibration scenario S0 at the former gauging location. Note that in some cases where the change in KGE_q was negative (e.g. when EI precipitation was used at Ait Ouchene station), this is because although there was an improvement in the KGE_q performance indicator during the calibration time period, when calculating it for the validation time period, it is possible that the metric slightly worsens. Note that some variations in NSE_q, $PBias_q$ and r_q were small or close to 0, because its calibration was optimised for KGE_q and not for those particular metrics in terms of discharge. A possible route to overcome this problem may be to use various performance indicators (for example, *KGE, NSE, PBias* and *r*) as objective functions to optimize in each calibration scenario, instead of using a single one. This multiobjective calibration approach may further improve discharge model estimates.

4.7 Discussion and conclusions

This study investigated alternative routes to calibrate the large-scale hydrological model PCR-GLOBWB using Earth observations globally available for the data-poor river basin of Oum Er Rbia in Morocco. Three global precipitation products, EI, WFDEI and MSWEP, were inter-compared and applied to force PCR-GLOBWB. Five different calibration scenarios were followed where GLEAM actual evapotranspiration and ESA CCI surface soil moisture data were used to identify model parameters with the aim to improve discharge estimates. In situ discharge observations were also used for calibration, as they are traditionally used to calibrate hydrological models.

Results showed that GLEAM actual evapotranspiration and ESA CCI soil moisture observations may be used to calibrate determined PCR-GLOBWB model parameters. GLEAM actual evapotranspiration was used to calibrate the reference potential evapotranspiration (f_e) as expected, affecting the water exchange between the top soil layer and the atmosphere and hence the soil water balance. ESA CCI soil moisture data was used to calibrate the minimum soil water capacity (f_w) and the saturated hydraulic conductivities of the soil layers (f_k), determining the surface runoff generation response, the shallow sub-surface flow and the groundwater recharge. However, calibration using only GLEAM evapotranspiration data or only ESA CCI soil moisture can result in more than one parameters combination to be optimal in terms of discharge (overparametrization or equifinality problem). To overcome this problem, a step-wise calibration scenario based on

both observations, evapotranspiration and soil moisture, was necessary to identify the optimal values of reference potential evapotranspiration (f_e), runoff-infiltration partitioning parameters (f_w) and the soil saturated hydraulic conductivity (f_k). Nonetheless, neither of these observations was used to calibrate the baseflow from the active groundwater layer (f_j). To identify baseflow recession coefficient parameter (f_j) a multiobjective calibration approach to streamflow observations could be followed. Similarly to Fenicia et al. (2007), multiple objective functions may be optimized in sequential steps for high flows, low flows and timing.

Spatially uniform prefactors for the entire Oum Er Rbia basin were used for the variation of the parameter values in this study. Developing novel calibration strategies where prefactors and so, model parameters vary with soil type, land use, elevation and/or other characteristics within the basin would be a promising research route to investigate. Furthermore, the present work inter-compares five calibration scenarios using a brute force method, where several combinations of parameters values were tested and the best performing was selected. For these combinations, and due to computational limitations, only four prefactors were considered leading to 81 model runs per precipitation product. Using more prefactor values and therefore, more runs may improve the estimation of the optimal parameters set for each calibration scenario. A suggestion for future studies may be to use an Ensemble Kalman Filter to calibrate the hydrological model, as previously presented in literature (Moradkhani et al., 2005; Wanders et al., 2014a). Furthermore, the validation of this study was carried out exclusively on streamflow. Other validation approaches, including the empirical orthogonal functions, wavelet analysis or their combination, may be another promising way towards a more in-depth validation of distributed hydrological models (Mascaro et al., 2015; Koch et al., 2015; Fang et al., 2015).

A step-wise calibration approach based on GLEAM actual evapotranspiration and ESA CCI soil moisture resulted in discharge estimates of acceptable accuracy (Moriasi et al., 2007), compared to discharge estimates derived from a model that was calibrated to in situ discharge measurements. Traditional calibration to in situ discharge measurements resulted in the highest model performance. A model calibrated only on evapotranspiration or soil moisture observations achieved a lower discharge performance than when they were used together.

In the inter-comparison between the three global precipitation products, WFDEI showed the lowest performance, whereas EI and MSWEP performed quite well. Apart from the in situ discharge calibration scenario, the highest discharge improvement was obtained when the two latter forcing data were used in combination with a step-wise calibration approach based on evapotranspiration and soil moisture observations.

Results indicated that precipitation impact on streamflow estimates was more significant than the one derived from calibrating model parameters, thus the lower quality of WFDEI compared to EI and MSWEP, decreased model performance and calibration was biased in order to compensate precipitation errors. Further investigation of the effect of precipitation

errors on model efficiency, but also on model parameters estimation may be an interesting route for hydrological research (Andréassian et al., 2004; Looper et al., 2012).

Although there is still room for further research, this study showed that globally available Earth observations, such as evapotranspiration or soil moisture, can be used to further parameterize large-scale hydrological models providing reasonable discharge estimates at regional or basin scale. In principle, these calibration approaches can be applied and investigated in other basins without or with limited in situ ground hydro-meteorological data (ungauged basins), not only to estimate discharge, but also to improve the understanding of the hydrological processes in the basin. Results suggested the potential of using other satellite products for hydrological modelling studies, including soil moisture products such as AMSR-E (Njoku et al., 2003) and SMOS (Kerr et al., 2001), evapotranspiration products such as SEBAL (Bastiaanssen et al., 1998) and MOD16 (Nishida, 2003), total water storage products such as GRACE (Tapley et al., 2004), etc. The spatial information of these satellite-based products could be used in a different way than the one explained in this study. For example, a calibration scenario based on a pixel by pixel, instead of basin average, comparison of surface soil moisture and actual evapotranspiration model estimates and observations could further improve discharge estimates. This calibration approach would have to take into account the spatial variability of the variables over the basin. Previous studies investigated how to incorporate spatial information into hydrological models using innovative spatial performance metrics to analyze the spatial sensitivity of simulated land-surface patterns (Koch et al., 2017).

Future studies may investigate step-wise calibration approaches using the combined information from multiple hydrological system variables. By incorporating several data products, different parts or components of the model can be optimized to increase the overall model performance. Another approach could be to calibrate the model to different variables with multiple objective functions - multiobjective calibration - (Gupta et al., 1998; Khu and Madsen, 2005; Fenicia et al., 2007). Alternatively, these hydro-meteorological data which are globally available may be used to identify and develop relationships between different basins using similarities, classification and scaling frameworks, as presented in previous studies (Samaniego et al., 2010; Kumar et al., 2013).

5 Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations

This chapter is based on:

López López P., Wanders N., Schellekens J., Renzullo L. J., Sutanudjaja E. H. and Bierkens M. F. P. (2016). Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. Hydrology and Earth System Sciences, 20, 3059–3076, https://doi.org/10.5194/hess-20-3059-2016.

Abstract

The coarse spatial resolution of global hydrological models (typically > 0.25◦) limits their ability to resolve key water balance processes for many river basins and thus compromises their suitability for water resources management, especially when compared to locally tuned river models. A possible solution to the problem may be to drive the coarse-resolution models with locally available high spatial resolution meteorological data as well as to assimilate in situ and remotely sensed observations of key water cycle variables. While this would improve the resolution of the global model, the impact of prediction accuracy remains largely an open question. In this study, we investigate the impact of assimilating streamflow and satellite soil moisture observations on the accuracy of global hydrological model estimations, when driven by either coarse or high resolution meteorological observations in the Murrumbidgee River basin in Australia.

To this end, a 0.08◦ resolution version of the PCRGLOBWB global hydrological model is forced with downscaled global meteorological data (downscaled from 0.5◦ to 0.08◦ resolution) obtained from the WATCH Forcing Data methodology applied to ERA-Interim (WFDEI) and a local high resolution, gauging-station-based gridded data set (0.05◦). Downscaled satellite-derived soil moisture (downscaled from ∼ 0.5◦ to 0.08◦ resolution) from AMSR-E and streamflow observations collected from 23 gauging stations are assimilated using an ensemble Kalman filter. Several scenarios are analyzed to explore the added value of data assimilation considering both local and global meteorological data.

Results show that the assimilation of soil moisture observations results in the largest improvement of the streamflow model estimates. The joint assimilation of both stream- flow and downscaled soil moisture observations leads to further improvement in streamflow simulations (20% reduction in *RMSE*).

Furthermore, results show that the added contribution of data assimilation, for both soil moisture and streamflow, is more pronounced when the global meteorological data are used to force the models. This is caused by the higher uncertainty and coarser resolution of the global forcing.

We conclude that it is possible to improve PCR-GLOBWB simulations forced by coarse resolution meteorological data with assimilation of downscaled spaceborne soil moisture and streamflow observations. These improved model results are close to the ones from a local model forced with local meteorological data. These findings are important in light of the efforts that are currently made to move to global hyper-resolution modelling and can help to advance this research.

5.1 Introduction

In recent decades, a number of large-scale hydrological and land-surface models have been developed to quantify the global water cycle components, to analyze the human influence on the global water balance, to study climate change impact on water resources and to assess global hydrological extremes, such as drought and flood risk. VIC (Liang et al., 1994, 1996), WaterGAP (Alcamo et al., 2003), ORCHIDEE (d'Orgeval et al., 2008), HTESSEL (Balsamo et al., 2009), JULES (Best et al., 2011), PCR-GLOBWB (Van Beek et al., 2011), SURFEX-TRIP (Decharme et al., 2010, 2013) and W3RA (Van Dijk, 2010; Van Dijk et al., 2014) are some examples of large-scale hydrological models recently applied to provide water resources assessment over continental to global domains, but their coarse spatial resolution, typically between 0.5º and 1º, limits their ability to resolve key water balance processes for many river basins (Lanza et al., 1997; Wu and Li, 2009) and thus compromises their suitability for water resources management, especially when compared to locally-tuned hydrological models.

A possible solution to the problem may be to drive these coarse resolution models with high resolution meteorological data. Several meteorological forcing datasets at a global scale are available, including the European Centre for Medium-Range Weather Forecasts - EMCWF ERA-Interim - global atmospheric reanalysis data (Dee et al., 2011), the Climatic Research Unit Time Series - CRU TS - (Mitchell and Jones, 2005), the NASA reanalysis Modern-Era Retrospective Analysis for Research and Applications - MERRA - (Rienecker et al., 2011) and the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data - WFDEI - (Weedon et al., 2014). They are the result of integrating Bayesian merging of the available Earth observations, in situ datasets and models to construct consistent large-scale meteorological time series. Some recent scientific efforts are conducted to improve the quality and availability of these datasets, for example increasing their spatial and temporal resolution (Cannon, 2011; Ebtehaj and Foufoula-Georgoiu, 2013; Atkinson, 2013). The use of high spatial resolution meteorological data would indirectly improve the resolution of the large-scale model, producing higher accuracy discharge estimates. However, when models that are designed for coarse spatial resolution are used at smaller spatial scale issues may arise with the representation of field scale processes. One of the major issues in this respect is the neglect of lateral flow, misleading the representation of the complex interactions between river water and groundwater (surface runoff, subsurface runoff, soil moisture state, etc.). At the moment, more research is required to understand the gain that can be obtained using this higher spatial resolution forcing data for uncalibrated global hydrological models at finer spatial resolutions.

Another approach to bridge the gap between the different spatial scales is to assimilate in situ and remotely-sensed observations of key water cycle variables. Higher resolution satellite data contain information at finer spatial resolution and could be used to correct for sub-optimal model performance at these finer resolutions. Multiple studies have used data assimilation techniques to obtain the best possible estimate of the hydrological system status, merging the strengths of hydrological modelling and observations and mitigating their respective weaknesses (Moradkhani, 2008; Clark et al., 2008; Van Dijk et al., 2014).

Among the sequential and variational data assimilation methods, the ensemble Kalman filter (Evensen, 2003) has arguably emerged as the most popular choice for assimilation into land surface and hydrological models. The various individual components of the water cycle, such as surface water (Vrugt et al., 2006; Rakovec et al., 2012), soil moisture (Van Dijk et al., 2014; Wanders et al., 2014a), snow water (Sun et al., 2004; Moradkhani, 2008) and groundwater (Zaitchik et al., 2008; Tangdamrongsub et al., 2015), which influence the hydrological system in different ways, can be assimilated into the model.

Soil moisture assimilation has been considered to improve model estimates, due to its key role in the terrestrial water cycle and its responsibility for the partitioning of precipitation between surface water and storage through infiltration. Several studies have assimilated soil moisture data (Draper et al., 2011; Chen et al., 2011; Wanders et al., 2014b; Massari et al., 2015; Alvarez-Garreton et al., 2015; Lievens et al., 2015) both based on in situ measurements and satellite-based soil moisture products from remote observation systems, such as ASCAT (Naeimi et al., 2009), SMOS (Kerr et al., 2012) and AMSR-E (Owe et al., 2008). On the other hand, surface water information has often been used for data assimilation frameworks (Vrugt et al., 2006; Rakovec et al., 2012) because discharge provides integrated information of all hydrological states, which is often very effective in improving model simulations. However, the risk of an integrated observation is that in some scenarios accurate simulations could be obtained by adjusting the wrong states variables.

An additional improvement could be made by the assimilation of downscaled or disaggregated satellite soil moisture observations into a particular land surface model (Merlin et al., 2006; Sahoo et al., 2013). Recently, new soil moisture products of higher spatial resolution have been released (Gevaert et al., 2015) but their impact on hydrological model predictions has not yet been explored. For example, they could be used to correct for incorrectly observed (or interpolated) precipitation patterns, which directly affect the input uncertainty into the model.

Moreover, improved results can be obtained by assimilation of multiple observational datasets of different parts of the hydrological cycle into the hydrological model (Barrett and Renzullo, 2009; Reichle et al., 2014). For example, the joint assimilation of discharge and soil moisture could result in an improved understanding of the runoff generation mechanisms and increase the quality and quantity of information incorporated to the model system. However, the added value of this type of joint assimilation procedures is largely unknown and should be further investigated (Aubert et al., 2003; Lee et al., 2011).

Many data assimilation experiments have been set up in conjunction with local-scale hydrological models and the benefit of data assimilation for large-scale models remains largely an open question. In this context, it is interesting to analyze whether the accuracy of large-scale hydrological models can be improved and become more comparable with locally calibrated model estimates if satellite observations are assimilated. Understanding the potential gain of assimilating satellite observations into large-scale models is a relevant research opportunity and may have potential benefits for water resources management (Van Dijk and Renzullo, 2011). For example, in regions without or low quality

meteorological observations the use of large-scale models in combination with satellite data assimilation could improve our understanding of the available water resources.The primary goal of the present study is to investigate the impact of assimilating streamflow and satellite soil moisture observations on the accuracy of global hydrological model estimations, when driven by either coarse or highresolution meteorological observations. The Murrumbidgee River basin in the southeast of Australia was chosen as a case study for the investigation because of the variety of land uses in the area, the high level of monitoring available for a large number of relatively unimpaired catchments, and the extensive body of previous studies observing and describing the hydrologic patterns across the basin (Renzullo et al., 2014; Van Dijk and Renzullo, 2011). Eight data assimilation scenarios were considered in which discharge and soil moisture observations were either independently or jointly used, and the forcing data were obtained from either local or global datasets. In this context, comparison of the eight scenarios with the locally calibrated model estimates provides insight into how the estimations of global hydrological models driven by global forcing data can come closer to local-scale model predictions.

5.2 Study area

The selected study area was the Murrumbidgee River basin (84,000 km^2) located in south-east Australia, specifically in the south west of New South Wales (*Figure 5.1*). The Murrumbidgee River is the second largest river in the Murray--Darling system, flowing for a distance of approximately 1,600 km. Elevations range from over 1,900 m in the East to less than 50 m on the Western plain. Forest and woodland coverage dominate in the East, with pasture and cropping in the central region, and increasing grassland to the West. The Western plain is dominated by clay-loam soils and with decreasing clay content in the middle and eastern region (Peischl et al., 2012). The climate in the catchment is one of the most diverse in New South Wales, with an average annual rainfall that ranges from 1,700 mm yr^{-1} in the higher elevations of the Snowy Mountains in the east, to less than 350 mm yr^{-1} on the Western plain. Average reference evapotranspiration varies from less than 1,000 mm yr^{-1} in the south-east, to over 1,800 mm yr^{-1} in the West. Mean annual flow increases from less than 45 m^3 s^{-1} in the upstream tributaries to approximately 125 m^3 s^{-1} in the mid-Murrumbidgee (Green et al., 2011).

5.3 Hydrological models

The simulations of two distributed hydrological models, i.e. the local OpenStreams wflow-sbm and the global PCR-GLOBWB, were performed for the period 2007-2010. The period 2000-2007 was used to "spin up" the models. The local and large-scale models are described in detail in the following two sub-sections.

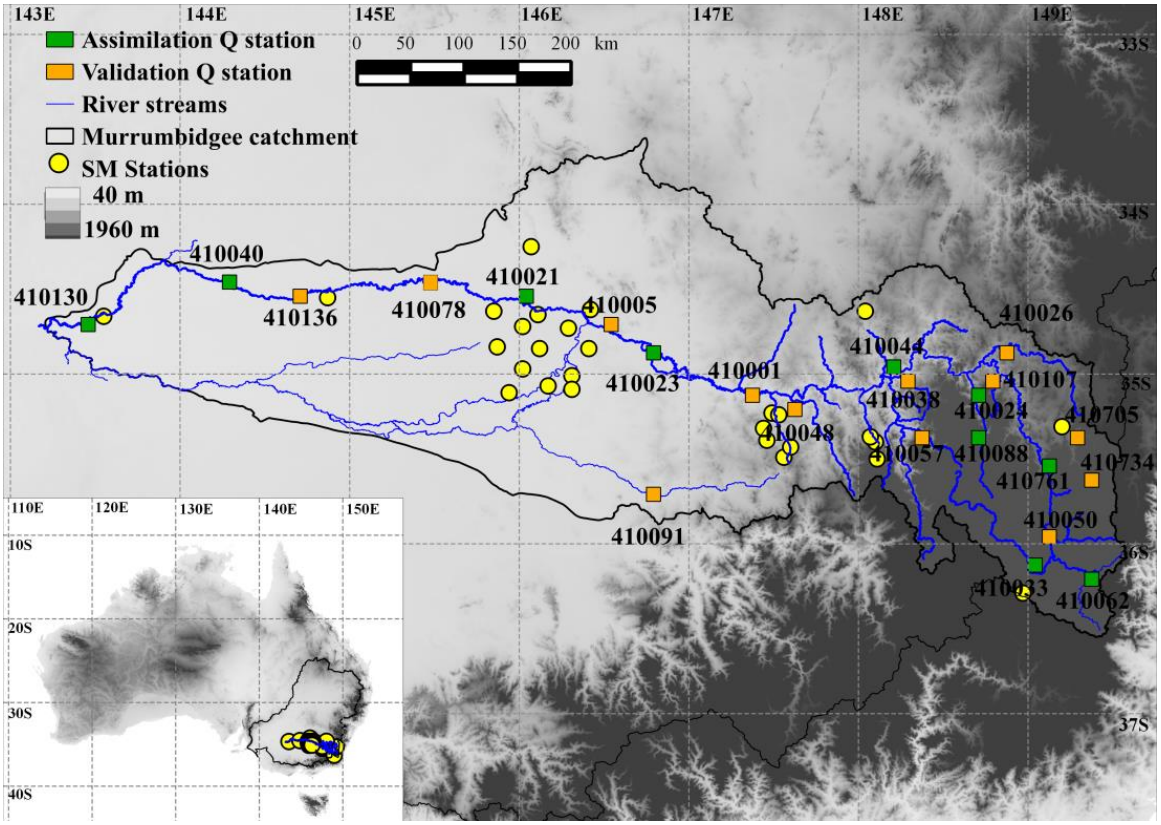


Figure 5.1 Map of the Murrumbidgee River basin and its location in Australia as part of the Murray–Darling system. Green squares indicate locations for assimilation of streamflow observations and orange squares indicate locations for evaluation of streamflow observations. Each streamflow location is identified with a gauging station number according to BoM (2015). Yellow points indicate locations of field-measured soil moisture observations.

5.3.1 Local-scale model: OpenStreams wflow-sbm

The local-scale hydrological model employed in this study was the OpenStreams wflow-sbm model (Schellekens, 2014). This is a distributed model derived from topog-sbm simple bucket model developed by Vertessy and Elsenbeer (1999). The OpenStreams wflow-sbm model (OSWS) is programmed in the PCRaster-Python environment (Wesseling et al., 1996; Karssenberg et al., 2010) and it is publicly available through the OpenStreams project (https://github.com/openstreams/wflow). The defined spatial resolution used in this study was 0.01° (approx. 1 km at the equator) and the temporal resolution was daily. A schematic representation of OSWS is given in *Figure 5.2*.

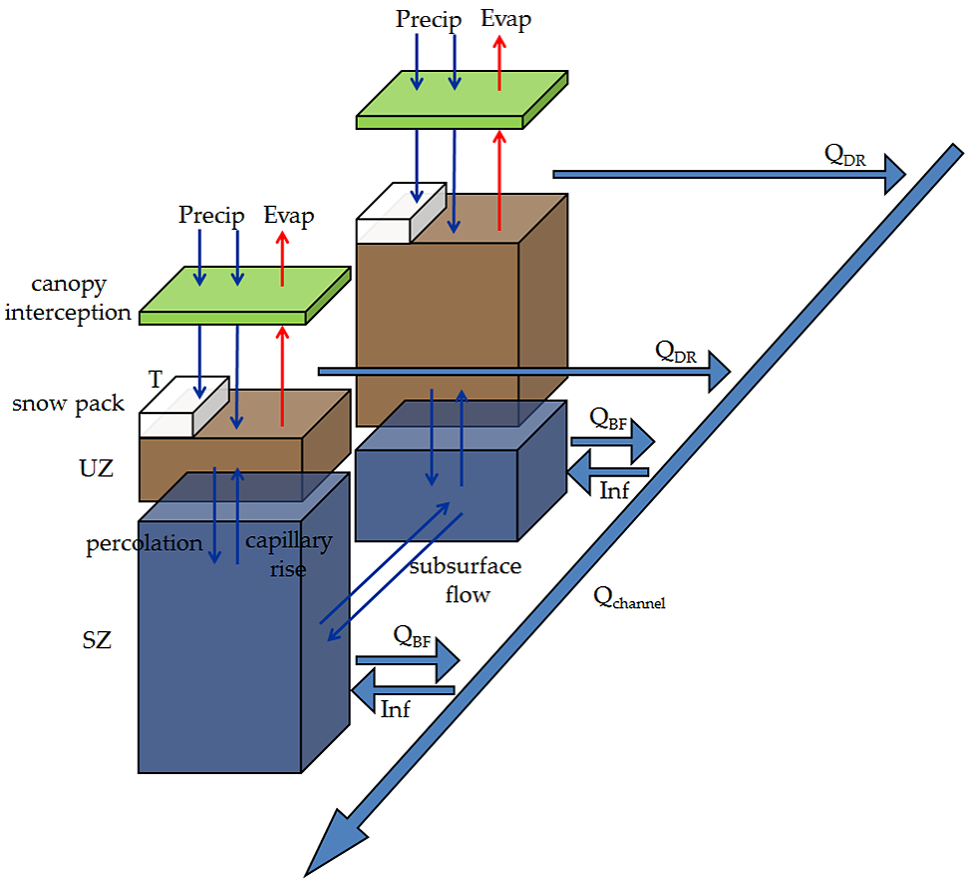


Figure 5.2 OpenStreams wflow-sbm model structure, adapted from Vertessy and Elsenbeer (1999) and Schellekens (2014). Symbols' definitions are as follows: Precip, precipitation; Evap, evaporation; T , temperature; UZ, unsaturated zone; SZ, saturated zone; $Q_{channel}$, total runoff; Q_{DR}, direct runoff; Q_{BF}, baseflow; and Inf, water flow from the river channel to the saturated zone.

OSWS model structure consists of three main routines: (i) rainfall interception (schematized by the Gash model - Gash, 1979), (ii) soil processes (schematized by the topog-sbm model) and (iii) river drainage and overland flow (modelled by the kinematic wave routing over a drainage network). The water enters each model cell from precipitation to the canopy interception storage or snow storage. The remaining liquid water infiltrates into the soil. At the same time, water is taken from the soil through evapotranspiration (based on soil water content and vegetation type). The water exchange into the soil considers two vertical soil layers, the unsaturated zone (UZ) and the saturated zone (SZ), based on topog-sbm structure. Total runoff is the sum of the direct runoff, the melt water that does not infiltrate into the soil and the baseflow (lateral subsurface flow from the saturated zone). This total runoff is conducted along the river network as discharge with kinematic wave routing.

The OSWS model was calibrated for the Murrumbidgee River basin using observations from in situ streamflow gauges (BoM, 2015) for the time period 1990-2010. These gauges are different from those considered in all the data assimilation scenarios to ensure an independent verification.

5.3.2 Large-scale model: PCR-GLOBWB

The large-scale hydrological model employed in this study was PCR-GLOBWB (Van Beek and Bierkens, 2009; Van Beek et al., 2011). Similar to OpenStreams wflow-sbm, PCR-GLOBWB is essentially a leaky-bucket type of model applied on a cell-by-cell basis. PCR-GLOBWB is coded in the PCRaster-Python environment. A spatial resolution of 0.08° (approx. 10 km at the equator) and a daily temporal resolution were used in this study. A schematic representation of PCR-GLOBWB is given in *Figure 5.3*.

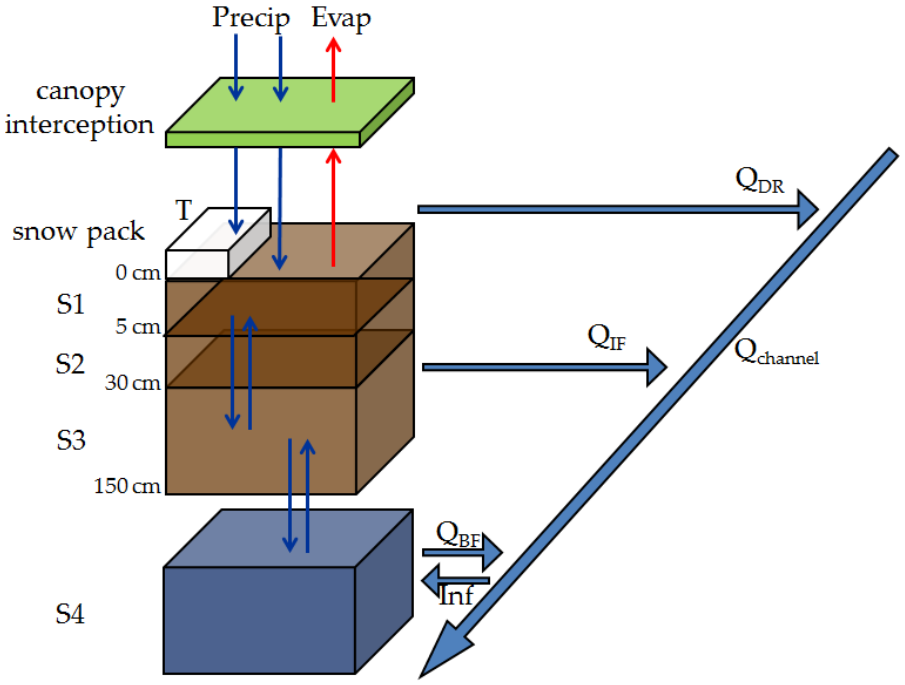


Figure 5.3 PCR-GLOBWB model structure, adapted from Van Beek et al. (2011). Symbols' definitions are as follows: Precip, precipitation; Evap, evaporation; T , temperature; S1, first soil layer; S2, second soil layer; S3, third soil layer; S4, groundwater reservoir; $Q_{channel}$, total runoff; Q_{DR}, direct runoff; Q_{IF}, intermediate flow; Q_{BF}, baseflow; and Inf, water flow from the river channel to the groundwater reservoir.

For each time step and cell, PCR-GLOBWB calculates the water balance components, including the water storage in three vertical soil layers (0-5, 5-30 and 30-150 cm) and one underlying groundwater reservoir, as well as the water exchange between the layers (percolation, capillary rise) and between the top layer and the atmosphere (rainfall, evapotranspiration and snowmelt). Sub-grid variability is taken into account considering the variations of elevation, land cover, vegetation and soil. The total runoff of a cell consists of direct runoff (saturation excess surface runoff), non-infiltrating melt water, interflow (lateral drainage from the soil profile) and baseflow (groundwater runoff from the lowest linear reservoir). The simulated runoff is routed along the river network based on the Simulated Topological Networks (STN30; Vörösmarty et al., 2000). Water abstraction and consumptive water use (domestic, industrial, livestock, irrigation) and reservoir management are included.

In contrast to the local-scale model, PCR-GLOBWB was not calibrated for the study basin. Hydrological model parameters were derived from vegetation, soil properties or geological information and estimated at a global scale.

5.4 Data

5.4.1 Meteorological forcing data

The forcing data required to drive both hydrological models are precipitation and air temperature. Two types of forcing data were used in this study: local forcing data, representing the best available data and global forcing data representing a lower spatial resolution dataset but one which is available globally.

Local precipitation and air temperature data were obtained from the gridded datasets generated by the Australian Bureau of Meteorology under the Australian Water Availability Project (AWAP; Jones et al., 2009). The data are derived from station-level daily precipitation and air temperature, and interpolated on a 0.05^o x 0.05^o grid covering the Australian continent. These data represent high resolution meteorology in this study which, we argue, will provide the modelling benchmark results.

Global precipitation and air temperature data were obtained from the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data - WFDEI - (Weedon et al., 2014). The daily global precipitation and air temperature data were provided at a spatial resolution of 0.5^o. To obtain finer spatial resolution climate maps, from 0.5^o grid to 0.08 grid, a downscaling procedure was applied based on a linear regression analysis (Sutanudjaja et al., 2011). It makes use of precipitation and temperature lapse rates derived from the 10' CRU-CL2.0 climatology dataset (New et al., 2002).

Local and global reference evapotranspiration (ET) were obtained through Hamon method (Allen et al., 1998; Lu et al., 2005).

Figure 5.4 shows the daily mean precipitation, temperature and reference evapotranspiration for the study time period (2007-2010) for both forcing datasets. Aside from the resolution difference, precipitation ranges from higher values in the mountainous regions of the catchment, with increasing variance in elevation and rainfall, to lower values in the western plain. On the contrary, temperature, and subsequently reference evapotranspiration decrease from West to East.

Local and global forcing data show some differences in their spatial distribution and magnitude. Each climate variable shows similar spatial distribution across the various resolutions (rows in *Figure 5.4*), with larger variations in the high elevation areas. However, local temperature magnitude differs in 3–4^oC with the global temperature, which is also reflected in the reference evapotranspiration. The downscaled global forcing data show a similar pattern to the global data with increasing resolution.

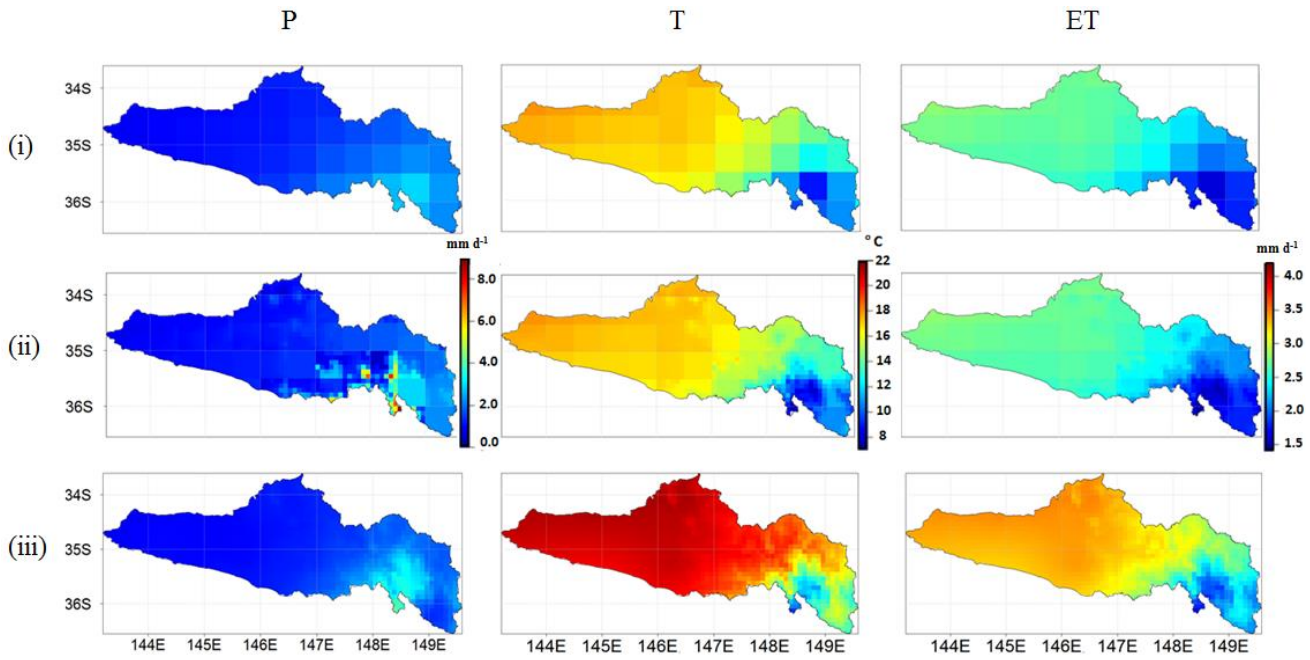


Figure 5.4 Daily mean precipitation, temperature and reference evapotranspiration for the time period 2007–2010 from the (i) global (0.5◦), (ii) downscaled global (0.08◦) and (iii) local (0.05◦) forcing datasets.

Table 5.1 shows the catchment daily mean values of the climate forcing variables for each year individually. Global temperature deviates from local approximately 3–4°C, which is also reflected in reference evapotranspiration, but less pronounced. Local and global precipitation, by contrast, is very similar for each year.

Table 5.1 Catchment daily mean (μ)/standard deviation (σ) of precipitation (P), temperature (T) and reference evapotranspiration (ET) in the Murrumbidgee River basin for 2007–2010.

	P [mm day^{-1}]		**T [°C]**		**ET [mm day^{-1}]**	
	Local μ/σ	**Global** μ/σ	**Local** μ/σ	**Global** μ/σ	**Local** μ/σ	**Global** μ/σ
2007	1.38/3.22	1.36/3.19	19.73/7.05	16.02/6.66	3.98/2.13	3.18/1.62
2008	1.24/3.19	1.26/3.32	18.89/6.65	15.18/6.24	3.88/2.06	3.00/1.51
2009	1.16/2.48	1.15/2.33	19.82/7.46	16.07/6.76	4.08/2.33	3.24/1.83
2010	2.35/5.46	2.45/5.19	18.50/6.91	15.03/6.46	3.69/2.04	2.95/1.56

5.4.2 Soil moisture data

Soil moisture observations retrieved from AMSR-E (Advanced Microwave Scanning Radiometer - EOS) brightness temperatures were provided by the Vrije Universiteit Amsterdam (VUA) in collaboration with NASA. AMSR-E is a multi-frequency passive microwave radiometer that uses C- (6.9 GHz) and X-band (10.65 and 18.7 GHz) radiance observations to derive near-surface soil moisture via the LPRM radiative transfer model

(Owe et al., 2008; De Jeu et al., 2008). In the present study, C-band AMSR-E data reported on a regular ~50 km global grid with an observation depth of 2 cm and a revisit time of 1-3 days was considered (Owe et al., 2008). Brightness temperatures from C-band AMSR-E were downscaled using the smoothing filter-based modulation technique. In this technique, brightness temperatures from the C-band (approx. 50 km) are adjusted based on data from the Ka-band (approx. 10 km). From the downscaled C-band AMSR-E brightness temperatures, soil moisture on a 0.08° spatial resolution is estimated. To improve the quality of the final soil moisture products a precipitation mask is applied (Gevaert et al., 2015).

Soil moisture observations from AMSR-E were compared to the unsaturated zone layer of OSWS and the first of the three vertical layers constituting the soil profile in each grid cell of PCR-GLOBWB. To match the remotely sensed soil moisture observations to the statistics of corresponding hydrological model states for soil water, different strategies can be followed, such as linear or minimum--maximum (MM) matching (Brocca et al., 2011), mean-standard deviation ($\mu - \sigma$) matching (Draper et al., 2009) and cumulative distribution function (CDF) matching (Reichle and Koster, 2004). In this study, a linear rescaling method was used. The converted satellite soil moisture values θ_{new}[m^3 m^{-3}] used for assimilation were calculated as

$$\theta_{new} = I_{min} + \left(\frac{I_{max} - I_{min}}{\theta_{max} - \theta_{min}}\right)(\theta - \theta_{min}) \qquad (5.1)$$

where I_{max} and I_{min} are the field capacity and the wilting point of the modelled soil moisture values at each grid cell [m^3 m^{-3}] and θ_{max} and θ_{min} are the maximum and minimumof AMSR-E satellite soil moisture values at the respective grid location [-].

In situ soil moisture observations were obtained from the Australian moisture monitoring network, OzNet (www.oznet.org.au; Smith et al., 2012). A total of 28 soil moisture monitoring stations with daily observations was used in this study for the period January 2007 to December 2010 (*Table 5.2*). Soil moisture monitoring sites were distributed evenly across 10 different study areas around and in the Murrumbidgee River basin, including the northern and eastern fringe of the catchment and those associated with the Yanco, Kyeamba Creek and Adelong Creek sites. The instrumentation at the sites measures moisture content in soil layers from either 0–8 cm or 0–5 cm depth.

5.4.3 Discharge data

Discharge observations were provided by the Bureau of Meteorology (BoM) and the Commonweath Scientific and Industrial Research Organisation (CSIRO), under the Water Information R+D Alliance (WIRADA). A total of 23 discharge monitoring stations with daily observations in the Murrumbidgee River and its main tributaries was available for the period January 2007 to December 2010. To ensure an independent evaluation of model simulations after the assimilation, a split sample approach of streamflow stations was used (e.g. Lee et al., 2012; Rakovec et al., 2012; Wanders et al., 2014b).

Table 5.2 Soil moisture monitoring sites' information.

Monitoring site name	Location		Elevation [m]
	Longitude	Latitude	
Adelong Creek 1	148.11	-35.50	772
Adelong Creek 3	148.10	-35.40	472
Adelong Creek 4	148.07	-35.37	457
Kyeamba Creek 1	147.56	-35.49	437
Kyeamba Creek 4	147.60	-35.43	296
Kyeamba Creek 6	147.46	-35.39	317
Kyeamba Creek 9	147.44	-35.32	241
Kyeamba Creek 12	147.49	-35.23	220
Kyeamba Creek 13	147.53	-35.24	261
Murrumbidgee catchment 1	148.97	-36.29	937
Murrumbidgee catchment 2	149.20	-35.31	639
Murrumbidgee catchment 3	148.04	-34.63	333
Murrumbidgee catchment 5	143.55	-34.66	62
Murrumbidgee catchment 6	144.87	-34.55	90
Murrumbidgee catchment 7	146.07	-34.25	137
Yanco 1	145.85	-34.63	120
Yanco 2	146.11	-34.65	130
Yanco 3	146.42	-34.62	144
Yanco 4	146.02	-34.72	130
Yanco 5	146.29	-34.73	136
Yanco 6	145.87	-34.84	121
Yanco 7	146.12	-34.85	128
Yanco 8	146.41	-34.85	149
Yanco 9	146.02	-34.97	122
Yanco 10	146.31	-35.01	119
Yanco 11	145.94	-35.11	113
Yanco 12	146.17	-35.07	120
Yanco 13	146.31	-35.09	121

The discharge of 10 stations was used for assimilation into the large-scale hydrological model, the remaining 13 stations were used for evaluation. Assimilation and evaluation stations were selected such they are equally distributed over the catchment and are situated both in small tributaries and the main Murrumbidgee River. *Figure 5.1* shows a map with the discharge locations. *Table 5.3* summarizes some key hydrological data.

Table 5.3 Hydrometeorological and geographical information of analyzed catchments at the Murrumbidgee River basin.

Station	Station name	Basin area [km²]	Outlet location		Mean flow [m³ s⁻¹]
			Long	Lat	
	Assimilation				
410088	Goodradigbee Riv. at Brindabella	419.66	148.73	-35.42	3.72
410062	Numeralla Riv. at Numeralla School	691.38	149.35	-35.18	1.16
410024	Goodradigbee Riv. at Wee Jasper	1050.60	148.69	-35.17	5.81
410044	Muttama Creek at Coolac	1058.49	148.16	-34.93	1.07
410033	Murrumb. Riv. at Mittagang Crossing	1809.84	149.09	-36.16	2.23
410761	Murrumb. Riv. below Lobbs Hole Creek	9332.28	149.10	-35.54	5.46
410130	Murrumb. Riv. at D/S Balranald weir	28651.21	143.49	-34.67	10.11
410023	Murrumb. Riv. at D/S Berembed weir	34133.07	146.84	-34.88	38.45
410021	Murrumb. Riv. at Darlington Point	37804.78	146.00	-34.57	29.90
410040	Murrumb. Riv. at D/S Maude weir	43110.97	144.30	-34.48	16.89
	Evaluation				
410107	Mountain Creek at Mountain Creek	140.54	148.84	-35.03	1.08
410705	Molonglo Riv. at Burbong	350.22	149.31	-35.34	0.59
410048	Kyeamba Creek at Ladysmith	350.30	147.53	-35.20	2.41
410038	Adjungbilly Creek at Darbalara	390.89	148.25	-35.02	1.49
410734	Queanbeyan Riv. at Tinderry	557.73	149.35	-35.61	0.89
410057	Goobarrag. Riv. at Lacmalac	559.77	148.35	-35.33	4.96
410026	Yass Riv. at Yass	1226.98	148.91	-34.84	1.20
410091	Billabong Creek at Walbundrie	2859.77	146.72	-35.69	2.11
410050	Murrumb. Riv. at Billilingra	3353.91	149.13	-35.98	3.84
410001	Murrumb. Riv. at Wagga Wagga	39856.21	147.37	-35.10	59.64
410005	Murrumb. Riv. at Narrandera	45321.40	146.55	-34.76	48.07
410078	Murrumb. Riv. at Carrathool	69854.30	145.42	-34.35	27.06
410136	Murrumb. Riv. at D/S Hay weir	73241.50	144.71	-34.52	25.24

5.5 Methodology

5.5.1 Data assimilation

5.5.1.1 Ensemble Kalman filter

The Ensemble Kalman filter (EnKF) is a sequential data assimilation method evolved from the standard Kalman filter (Evensen, 1994) that has been used previously for assimilation of observations into land surface and hydrological models (Chen et al., 2011; Draper et al., 2011; Wanders et al., 2014a; Tangdamrongsub et al., 2015). It is a Monte Carlo based approach that integrates an ensemble of model states forward in time to represent the error statistics of the model estimate when observations are assimilated (Burgers et al., 1998; Evensen, 2003). The state equation in a discrete form is given as

$$x_{t+1} = f(x_t, F_{t+1}, p, \varepsilon_t) \quad (5.2)$$

where f are the dynamical model equations that represents the hydrological processes in the system, x_t is the model state at time t, F_{t+1} is the forcing at time t (e.g. precipitation and temperature), p are the model parameters and ε_t is the model error.

To assimilate observations into the hydrological model, the already mentioned observations, downscaled remotely sensed AMSR-E soil moisture and discharge, can be linearly described as

$$y_t = H_t x_t + \in \quad (5.3)$$

where y_t is the observations vector, H is the observation model or operator that relates the model states x_t to the observations y, and $\in$ is the random noise with a zero mean and an error given by the observations error covariance matrix R_t.

The EnKF calculates the analysis at each time t, x_t^a, of the model forecast, x_t^f, as

$$x_t^a = x_t^f + K_t\left[y_t - H_t\left(x_t^f\right)\right] \quad (5.4)$$

where K_t is defined as the Kalman gain

$$K_t = P_t H_t^T (R_t + H_t P_t H_t^T)^{-1} \quad (5.5)$$

with H_t^T the transpose matrix of the observation model at time t (which is equal to the identity matrix after linear rescaling) and P_t the state error covariance matrix of the model prediction calculated from the spread between the different ensemble members given as

$$P_t = \frac{\sum_{n=1}^{N}\left(x_n^f - \overline{x^f}\right)\left(x_n^f - \overline{x^f}\right)^T}{N-1} \quad (5.6)$$

where $\overline{x^f}$ the ensemble average of model simulations and N the number of ensemble members considered.

5.5.1.2 Assimilating soil moisture and discharge observations

In this study, the EnKF was applied to update state variables of the large-scale hydrological model, PCR-GLOBWB, on each daily time step using downscaled remotely sensed AMSR-E soil moisture and discharge observations. We used 100 ensemble members and all the observations were calculated and assimilated at each 0.08° model grid cell for each day that observations are available. The EnKF has been implemented in the PCRaster modelling environment (Karssenberg et al., 2010).

Eight different data assimilation scenarios with PCR-GLOBWB were inter-compared and compared to the OSWS estimates without any data assimilation. The data assimilation (DA) scenarios are described in *Table 5.4*, indicating the meteorological forcing and the

observations used in each scenario. Simulations forced with local meteorological data are denoted with LOCAL and simulations forced with global meteorological data are denoted with GLOBAL. Independent assimilation of discharge (GLOBWB_Q) and soil moisture (GLOBWB_SM) were investigated, as well as the joint assimilation of both observation types (GLOBWB_SM+Q).

Table 5.4 Data assimilation scenarios including abbreviations, forcing data, hydrological model and assimilated observations.

Identifier	**DA scenarios**		
	Forcing data	**Hydrological model**	**Assimilated observations**
LOCAL GLOBWB_OL	Local (AWAP)	PCR-GLOBWB	Open-loop (none)
LOCAL GLOBWB_Q			Discharge stations
LOCAL GLOBWB_SM			AMSR-E soil moisture
LOCAL GLOBWB_SM+Q			Discharge stations and AMSR-E soil moisture
LOCAL OSWS		OpenStreams wflow-sbm (OSWS)	None
GLOBAL GLOBWB_OL	Global (WFDEI)	PCR-GLOBWB	Open-loop (none)
GLOBAL GLOBWB_Q			Discharge stations
GLOBAL GLOBWB_SM			AMSR-E soil moisture
GLOBAL_SM+Q			Discharge stations and AMSR-E soil moisture
GLOBAL OSWS		OpenStreams wflow-sbm (OSWS)	None

In the EnKF, to account for model and observations uncertainty, stochastic noise can be introduced in model forcing data, parameters, soil moisture and discharge observations. For the local and global meteorological forcing, the precipitation was perturbed with additive Gaussian white noise with standard deviation of 10% of the nominal value (Adam and Lettenmaier, 2003; Hijmans et al., 2005). The errors were assumed to be spatially uncorrelated. For the assimilation of the satellite soil moisture data, spatial information on the measurement error covariance R (*Equation 5.5*) was required. The structure of R was determined from estimates of Wanders et al. (2012) over Spain, obtained by using high resolution modelling of the unsaturated zone. The average standard error of AMSR-E is 0.049 m^3 m^{-3}. The error covariance between the discharge observations was set to zero while the standard error for the discharge observations was assumed to be 10% of the discharge. It was additionally assumed that the covariance between the satellite soil moisture observations and discharge observations equals zero. Some of the assumptions described in

this section for data assimilation were investigated through a preliminary sensitivity analysis including the number of ensemble members and the standard errors of precipitation and discharge (*Figure C. 1, Figure C. 2, Figure C. 3, Figure C. 4, Figure C. 5 and Figure C. 6*).

5.5.2 Evaluation

The impact of assimilating discharge and soil moisture observations into the large-scale hydrological model PCR-GLOBWB compared with the locally calibrated model OSWS, is separately analyzed: firstly, on soil moisture estimates and secondly, on discharge estimates. A common regular 0.08^{o} grid (approx. 10 km) was adopted for the inter-comparison of the two different resolution hydrological models estimates. For this purpose, OSWS estimates were upscaled with a linear resampling from 0.01^{o} (approx. 1 km) to 0.08^{o} (~ 10 km).

Results were produced for each of the 23 locations listed in *Table 5.3*. For practical reasons, the following section includes results for a limited number of evaluation locations only, both in the Murrumbidgee River and its tributaries. This combination thus comprises stations with varying sizes of contributing area.

To understand and inter-compare the performance of the different data assimilation scenarios described in *Table 5.3*, an extensive evaluation was carried out, including the calculation of various evaluation metrics, such as Root Mean Squared Error ($RMSE$), Mean Absolute Error (MAE), Pearson's correlation coefficient (r) and Nash Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970).

5.6 Results

5.6.1 Impact of assimilation on soil moisture estimates

The time series of simulated soil moisture for 0-5 cm for each data assimilation scenario and downscaled AMSR-E observations for the time period January 2008–May 2009 at 410057 gauging station (mountainous region) are shown in *Figure 5.5*.

The use of global forcing data produces a different dynamic response of soil moisture estimates compared to the local forcing data. This fact is due to the discrepancies between both meteorological datasets, which govern the water exchange processes between the top layer and the atmosphere (precipitation, evapotranspiration and temperature).

Even though global precipitation quite accurately depicts the overall character of the precipitation (daily mean values of the local and global precipitations show similar spatial distributions and magnitudes; see *5.4.1 Meteorological forcing data*), the global precipitation misses specific rainfall events at particular days and locations due to its lower resolution. This is especially important for warm season precipitation and regions in mountainous terrain (e.g. 410057 gauging station in *Figure 5.5*), which are dominated by convective storms. The differences in precipitation are reflected in soil moisture estimates of both hydrological models and their impact is higher in PCR-GLOBWB estimates. When PCR-GLOBWB is forced with local data, soil moisture estimates produce patterns with a more

accurate description of the small-scale variability of the observations in time. Whereas, when global forcing is used, soil moisture results in a smoother estimation of the observations.

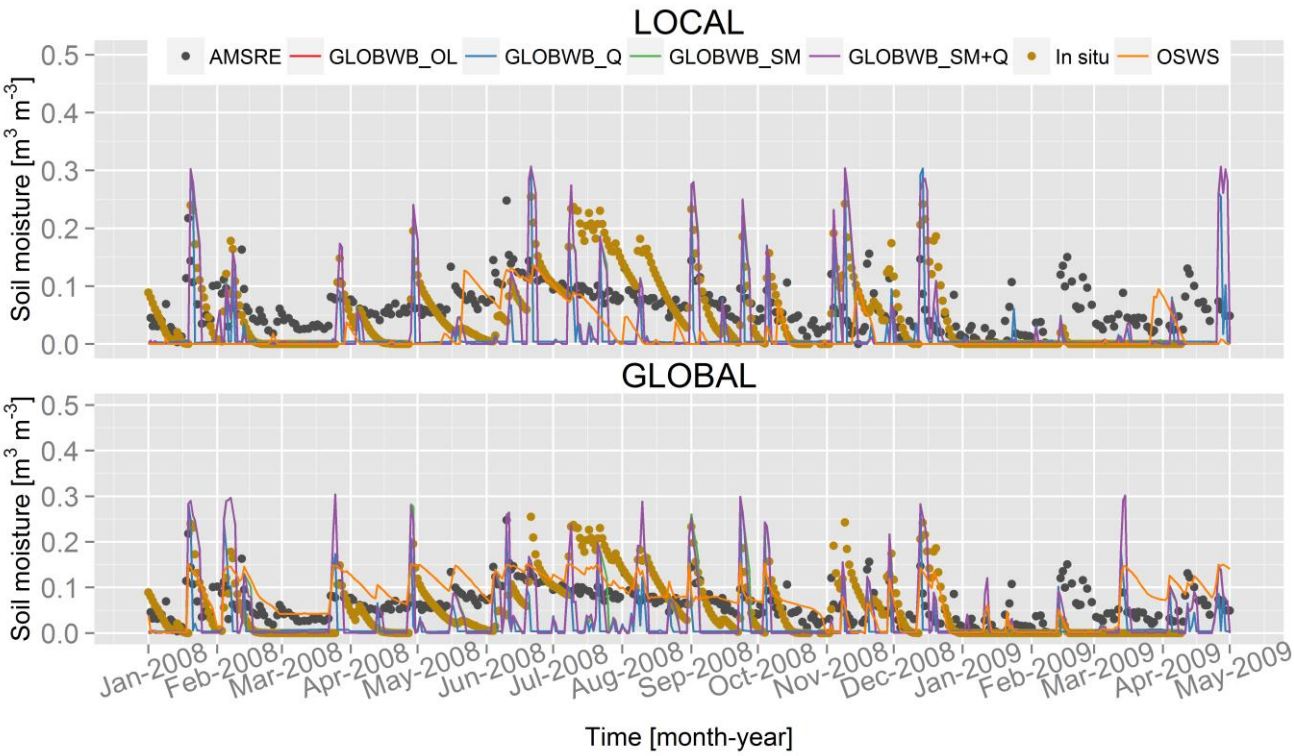


Figure 5.5 Simulated and observed soil moisture estimates at 410057 gauging station in a tributary of the Murrumbidgee River for the time period January 2008–May 2009. The upper panel shows soil moisture time series when local data are used as model forcing. Soil moisture time series obtained with the global forced models are shown in the lower panel. Each panel contains results for each data assimilation scenario plotted with different coloured lines (OSWS–orange, GLOBWB_Q–blue, GLOBWB_SM–green and GLOBWB_SM+Q–purple), downscaled AMSR-E observations with dark grey points and in situ soil moisture observations with dark yellow points.

A total of eight different data assimilation scenarios were investigated, with global and local forcing data (*Table 5.3*). In the OL scenario, no data is assimilated into the model to correct intermediate model states. In this first scenario, soil moisture observations are underestimated when local forcing is used. The assimilation of discharge observations, Q scenario, results in similar soil moisture estimates compared to the OL scenario, as expected. This similarity is caused by the assimilation procedure which constrains the model to follow the discharge observations directly affecting on groundwater and routing processes, which are a poor reflection of the upper soil moisture content from 0 to 5 cm. In contrast, the assimilation of soil moisture observations in SM and SM_Q scenarios produces a reduction of the negative bias, improving the soil moisture estimates, especially when local forcing is used.

Figure 5.6 shows the impact of each data assimilation scenario on the considered evaluation metrics (*RMSE*, *MAE* and *r*). Results of the catchment daily mean values are shown. Each histogram shows the evaluation metric on the vertical axis vs. the data assimilation scenarios on the horizontal axis. *Figure 5.6* consists of a matrix of multiple panels, with rows showing the three considered evaluation metrics and columns showing local and global forcing data.

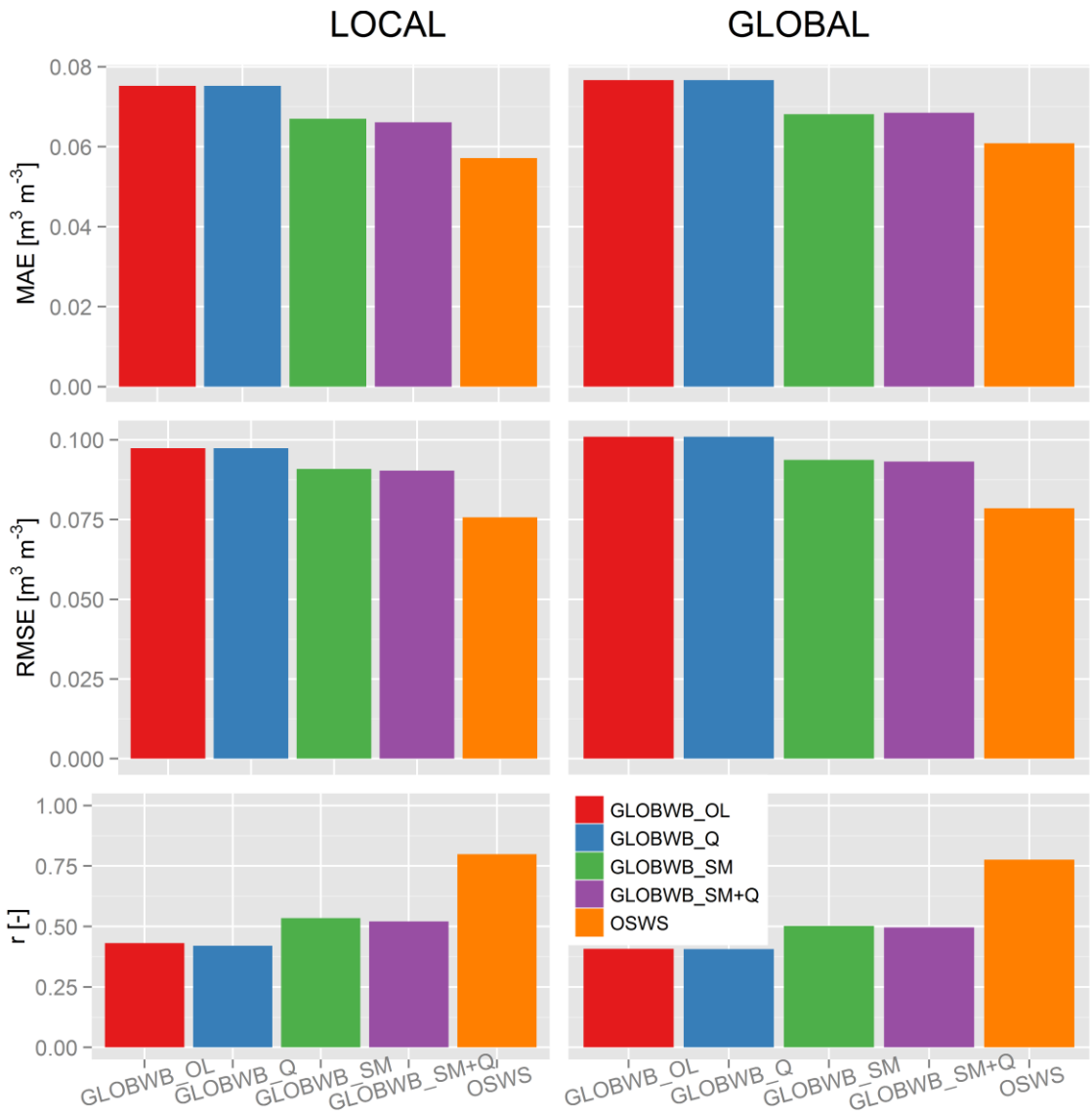


Figure 5.6 Evaluation results of the catchment daily means of soil moisture in the Murrumbidgee River basin. In the rows, three different evaluation metrics are shown, from top to bottom these are: *MAE*, *RMSE* and *r*. Columns show various forcing data: local and global. (For clarity, the exact values are included in *Table C. 1*)

From this figure, some general observations can be made. Evaluation results show differences between the results from local and global forcing of the models. For example, the use of local instead of global forcing produces a decrease of 4% and 2% in $RMSE$ and MAE, respectively and an increase of 6% in r, when the OL scenario in PCR-GLOBWB is considered.

The large-scale model, PCR-GLOBWB, without data assimilation shows a poorer performance than the locally calibrated model OSWS on soil moisture predictions. Assimilating discharge observations (Q scenario) does not lead to an improvement on soil moisture estimates, whereas downscaled AMSR-E soil moisture observations assimilation (SM scenario) results in an increase of r and a decrease of $RMSE$ and MAE. Therefore, after assimilating soil moisture observations, evaluation results of PCR-GLOBWB and OSWS are closer to each other. For example, percent differences in $RMSE$ between both models are reduced from 22% (LOCAL GLOBWB_OL and LOCAL OSWS) to 16% (LOCAL GLOBWB_SM+Q and LOCAL OSWS) and MAE from 14% (LOCAL GLOBWB_OL and LOCAL OSWS) to 16% (LOCAL GLOBWB_SM+Q and LOCAL OSWS). However, the reduction in the differences between PCR-GLOBWB and OSWS performances due to the assimilation of soil moisture and streamflow observations does not mean that both models perform similarly. Maximum r values obtained with PCR-GLOBWB are ~ 0.5, whereas for OSWS maximum r values of 0.7 are reached.

Additionally, boxplots of the catchment daily mean values are included in *Figure 5.7* considering local (upper panel) and global (lower panel) forcing. The assimilation of soil moisture observations leads to an evident improvement in the statistical distribution of PCR-GLOBWB soil moisture estimates, reducing the differences in dispersion with the observations.

Assimilating soil moisture observations and forcing the model with high spatial resolution meteorological datasets impacts the quality of soil moisture estimates with PCR-GLOBWB to a similar extent. Results indicate that the highest improved performance is achieved when their combination occurs, i.e. soil moisture observations are assimilated into a model driven by local forcing data (LOCAL GLOBWB_SM and LOCAL GLOBWB_SM+Q scenarios).

5.6.2 Impact of assimilation on streamflow estimates

The simulated and observed streamflow estimates at 410088 gauging station are shown in *Figure 5.8*. From this figure, it is clear that the peaks in streamflow are poorly estimated by PCR-GLOBWB, whereas OSWS is able to capture them with higher accuracy, independently of the forcing data used. This is most probably explained by the higher resolution and the calibration of model parameters for the study.

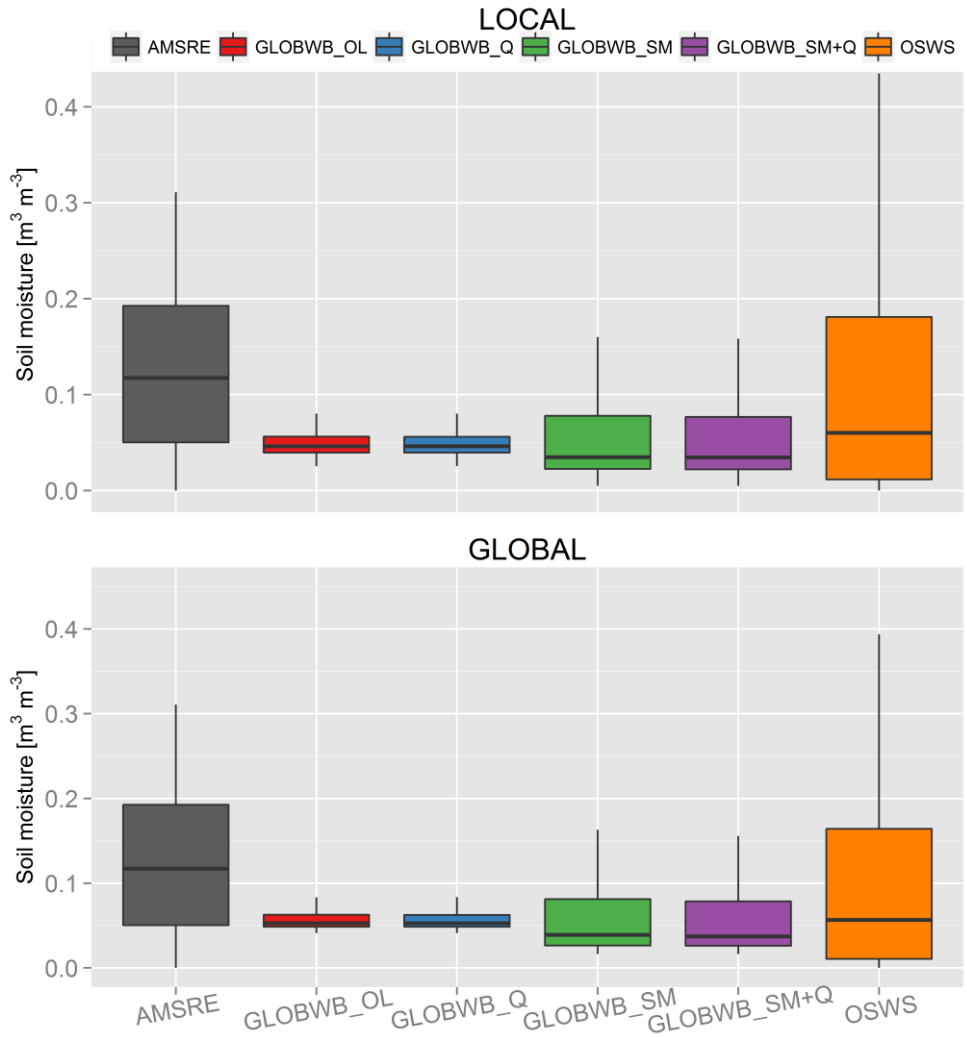


Figure 5.7 Boxplots of the catchment daily means of soil moisture in the Murrumbidgee River basin. The upper panel shows soil moisture when local data are used as model forcing. Soil moisture obtained with the global forced models is shown in the lower panel. Boxplots of each panel illustrate the first and third quantile ranges (box), the median (dark line) and the maximum–minimum range (whiskers) of soil moisture estimates.

Differences between local and global forcing data (see *5.4.1 Meteorological forcing data*) are reflected in differences in streamflow estimates from both models. When global data is used, evapotranspiration is lower; hence a higher amount of water is introduced into the models, resulting in higher streamflow estimates. By assimilating discharge and soil moisture observations, intermediate hydrological processes, including groundwater state, percolation and surface runoff among others, are corrected and errors in forcing data are reduced to improve streamflow predictions (e.g. November 2008). The best performance is achieved when soil moisture and discharge data are jointly assimilated.

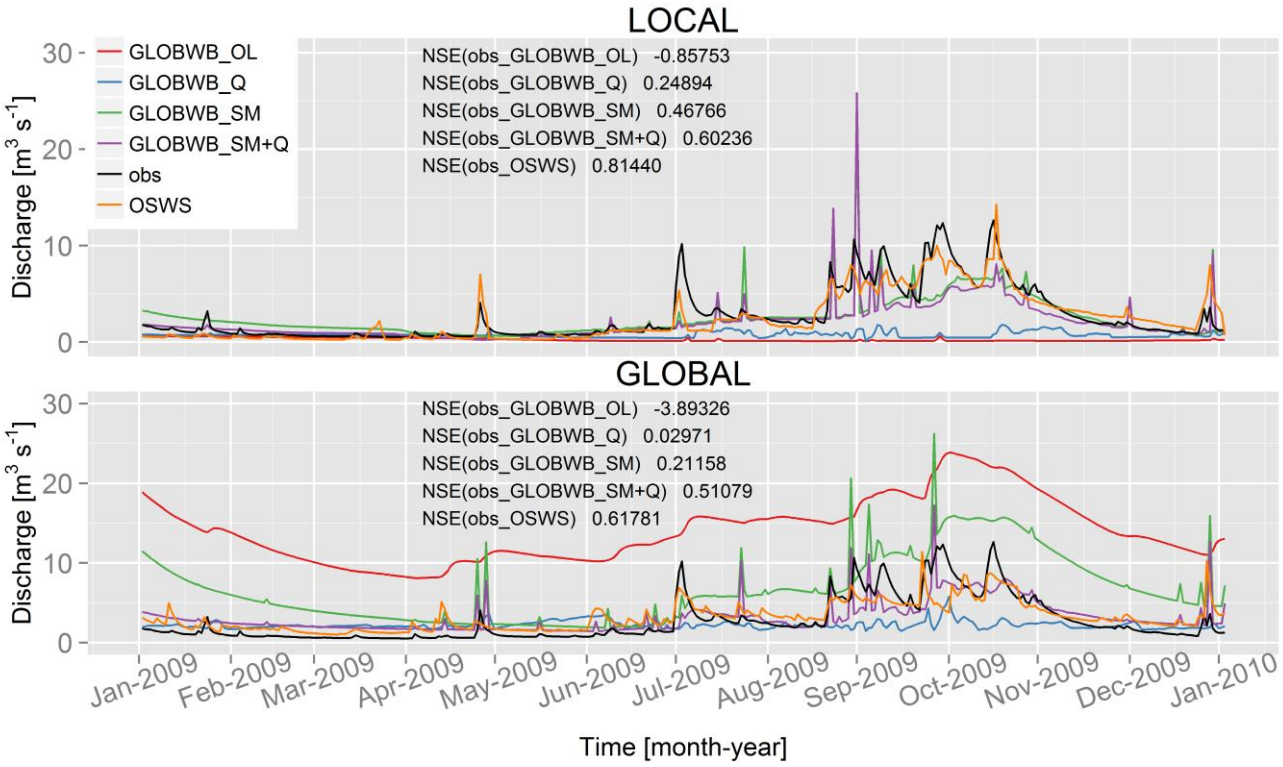


Figure 5.8 Simulated and observed streamflow estimates at 410088 gauging station in a tributary of the Murrumbidgee River for the time period January 2009–January 2010. The upper panel shows streamflow when local data are used as model forcing. Streamflow obtained with the global forced models is shown in the lower panel. Each panel contains results for each data assimilation scenario and the observed streamflow estimates plotted with different coloured lines (OSWS–orange, GLOBWB_OL–red, GLOBWB_Q–blue, GLOBWB_SM–green, GLOBWB_SM+Q–purple and obs–black). The ensemble mean is given for each data assimilation scenario.

To further analyze and quantify the influence of each data assimilation scenario on streamflow estimates, the evaluation metrics ($RMSE, MAE, r$ and NSE) were calculated and included for multiple discharge locations in *Figure 5.9*.

The highest r and NSE and the lowest $RMSE$ and MAE are obtained when models are forced with local meteorological data. The use of global forcing data leads to a reduction in performance, which is more significant for the large-scale than for the local-scale model. Without assimilation, forcing PCR-GLOBWB with local data (LOCAL GLOBWB_OL) instead of global data (GLOBAL GLOBWB_OL) results in an increase of 80% in r and a decrease of 70% in $RMSE$ and 72% in MAE on average. OSWS also improves its streamflow estimates but to a lesser degree, with increases of 7% in r and decreases of 28% in $RMSE$ and 43% in MAE (LOCAL OSWS and GLOBAL OSWS).

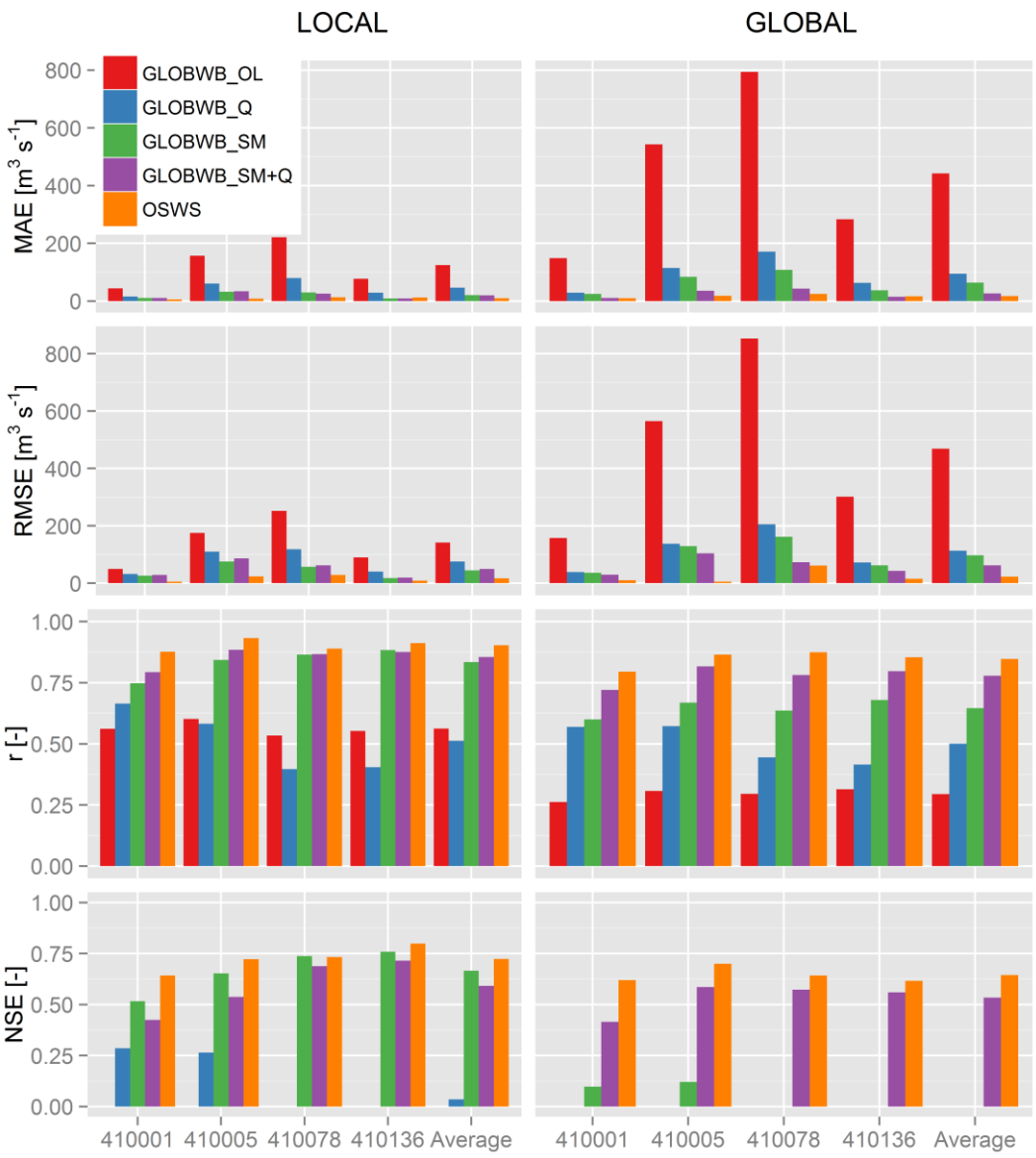


Figure 5.9 Evaluation results for streamflow estimates at 410001, 410005, 410078 and 410136 gauging stations in the Murrumbidgee River. Average values calculated across those locations are shown in the rightmost bar of each histogram. In the rows, four different evaluation metrics are shown, from top to bottom these are: $MAE, RMSE, r$ and NSE. Columns show various forcing data: local and global.

Both observations assimilation, discharge and soil moisture separately (Q and SM), improve streamflow models estimates. The highest improvement is achieved when both observations are assimilated into the model (GLOBWB SM+Q), as expected. The improvement is higher when soil moisture observations (SM) are assimilated than the case of discharge assimilation alone (Q). For example, in terms of r GLOBAL GLOBWB_SM scenario results in an increase of 20% and GLOBAL GLOBWB_Q scenario of 5% relative to GLOBAL GLOBWB_OL scenario at 410107 gauging station. Some possible explanations could be the finer resolution of AMSR-E soil moisture observations and/or the basin hydrological features, which characterize it as a catchment mainly driven by direct runoff,

where the highest contribution to the total runoff comes from the upper soil layer and not from the groundwater zone.

The largest improvements were found at gauging stations in the main channel of the Murrumbidgee River, such as 410001 station where assimilating soil moisture and discharge observations increases r from 0.56 to 0.79 and decreases $RMSE$ and MAE from 49.54 to 28.77 $m^3\ s^{-1}$ and from 43.85 to 10.76 $m^3\ s^{-1}$, i.e. comparing the LOCAL GLOBWB_OL and LOCAL GLOBWB_SM+Q scenarios. This improvement is more significant when the model is forced with global data than with local data. At 410001 station, $RMSE$ varies from 157.62 to 29.01 $m^3\ s^{-1}$, MAE from 148.29 to 10.88 $m^3\ s^{-1}$ and r from 0.26 to 0.72 when GLOBAL GLOBWB_OL and GLOBAL GLOBWB_SM+Q scenarios are compared.

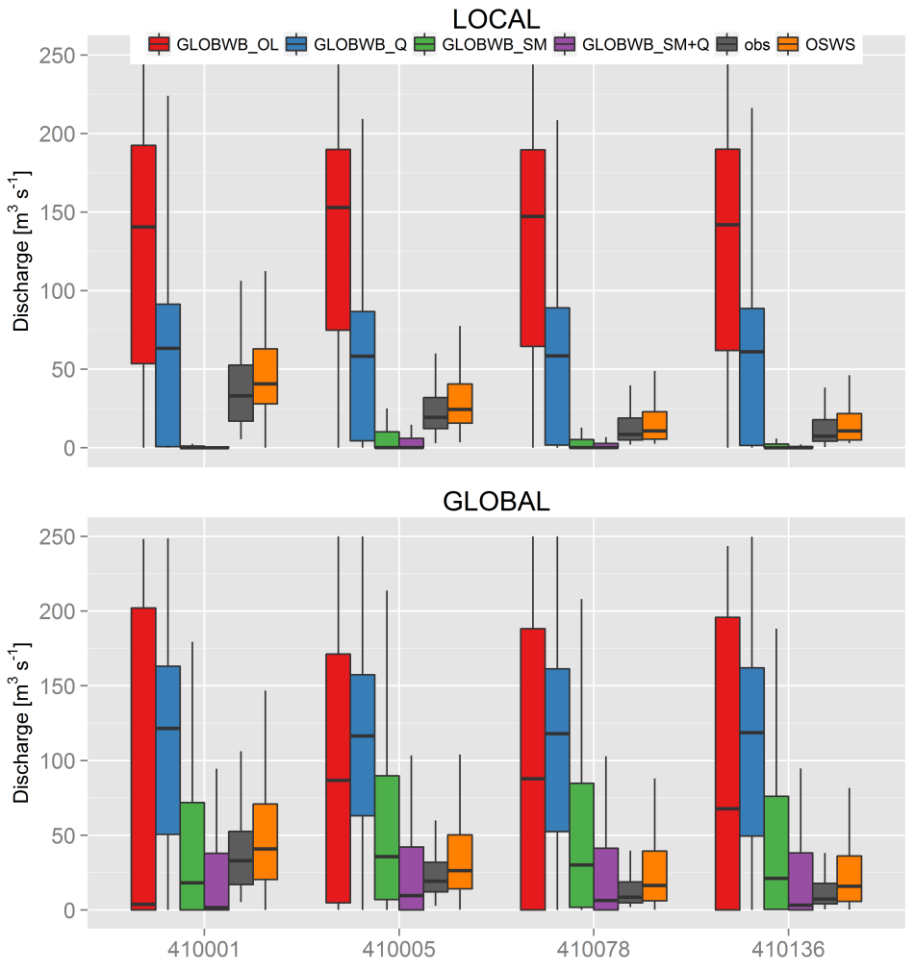


Figure 5.10 Boxplots of streamflow estimates at 410001, 410005, 410078 and 410136 gauging stations in the Murrumbidgee River. The upper panel shows streamflow when local data are used as model forcing. Streamflow obtained with the global forced models is shown in the lower panel. Boxplots of each panel illustrate the first and third quantile ranges (box), the median (dark line) and the maximum–minimum range (whiskers) of streamflow estimates.

Boxplots of streamflow estimates are included in *Figure 5.10*. The results clearly show that, compared to the observed streamflow, the median values of PCR-GLOBWB streamflow without data assimilation are very biased. The greatest amount of spread is observed between OL scenarios and observations, across all stations. Every data assimilation scenario (Q, SM and SM+Q) shows improvement in the statistics of streamflow, correcting not only its median value, but also the overestimation.

In contrast with the soil moisture evaluation results, *Figure 5.8*, *Figure 5.9* and *Figure 5.10* indicate that for streamflow the use of local forcing provides a larger improvement than assimilating soil moisture and discharge observations. Using a global model with local forcing and assimilating satellite soil moisture data yields streamflow predictions comparable to a local model with local forcing along the main river of this catchment. Moreover, also on the main channel, a global model with global forcing may still yield reasonable results as long as both discharge data and soil moisture data are assimilated.

5.7 Discussion

PCR-GLOBWB poorly estimates streamflow and soil moisture, when forced either with high or coarse spatial resolution forcing data, without data assimilation. The derivation of the hydrological model parameters from hydro-geological information at a global scale could be a possible explanation. From the initial scenarios without assimilation, it can be inferred that there is significant space for improvements when discharge and soil moisture observations are assimilated. An alternative route to data assimilation to improve model estimates would be to locally calibrate PCR-GLOBWB using discharge observations from in situ gauging stations. The improvement achieved through model parameters calibration would be possibly even higher than when soil moisture and streamflow observations are assimilated into the model (Wanders et al., 2014b). However, the present study means to be an attempt of providing hydrological estimations with a global model that could be also used in ungauged river basins where scarce in situ data are available.

The joint assimilation of discharge and downscaled satellite soil moisture observations produces the largest improvement on PCR-GLOBWB streamflow estimates (20% reduction in $RMSE$). These results agree with the findings made by Wanders et al. (2014a) in the Upper Danube and it is also in line with the expectations, where more observations lead to a better constrained model simulation.

A major finding of this study is the positive impact of assimilating soil moisture observations on the streamflow estimates, compared to the independent discharge assimilation. The estimations of a large-scale hydrological model driven with coarse resolution forcing data are improved using globally available remotely sensed soil moisture observations. This creates an opportunity for improved global model-based streamflow estimations as a result of the assimilation of remotely sensed soil moisture observations into large-scale hydrological models. Additionally, the adopted methodology has the advantage that it is applicable in river basins all over the world independently of the availability of in situ hydro-meteorological information.

The variable effectiveness of soil moisture assimilation has been previously reported in literature. Whereas some studies found improvements (Draper et al., 2011; Wanders et al., 2014b), others obtained mixed or unsatisfactory results (Crow et al., 2005). The scale of soil moisture observations, the dominant runoff processes in the study basin and the model structure and parameters uncertainties may partly explain this variability. In this particular study, the novel use of a finer spatial resolution satellite soil moisture product together with the climate and hydro-geological characteristics of the catchment could be a possible explanation of the positive impact of soil moisture assimilation. Renzullo et al. (2014) assimilated satellite soil moisture observations from multiple sensors (ASCAT and AMSR-E) obtained at scales coarser than the model (AWRA-L) resolution. In the present study, the scale of soil moisture observations coincides with the model scale. A specific analysis of the impact on streamflow and soil moisture estimates of assimilating non-dowsncaled AMSR-E soil moisture assimilation could be a possible route to further investigate the effect of different spatial resolution soil moisture products. To this end, AMSR-E soil moisture observations at the original spatial resolution were assimilated into PCR-GLOBWB and results analysis presented in this manuscript were reproduced. Results showed that the assimilation of non-downscaled soil moisture observations has a positive impact on soil moisture and discharge estimates, but this improvement is smaller compared to the assimilation of the downscaled soil moisture estimates (*Figure C. 7 and Figure C. 8*). Moreover, runoff in the Murrumbidgee River basin is mainly dominated by direct runoff processes, with reduced contribution from the groundwater zone (Green et al., 2011). These catchment conditions, together with their representation in the model structure are most likely responsible for the added value of assimilating soil moisture. There may be merit in analysing these scenarios in future research studies.

For the assimilation of AMSR-E soil moisture observations, spatial information on the measurement error covariance R (*Equation 5.5*) was based on results from previous studies over Spain (Wanders et al., 2012). Our philosophy was to set the AMSR-E errors to realistic values determined and validated in previous studies, so that all the required information was already available. The determination of AMSR-E observations uncertainty specifically over the Murrumbidgee River basin with physically-based modelling or in situ soil moisture measurements (Su et al., 2013) could be further investigated. In addition, a linear rescaling method was used to match AMSR-E soil moisture observations to the statistics of model states related with soil water. Different matching strategies could be applied in future studies. To account for model uncertainty, stochastic noise in precipitation data was introduced. A sensitivity analysis on model parameters could be another possible approach.

Meteorological data play a key role in soil moisture and discharge model estimates. The various model and data assimilation options were evaluated under both high and low resolution meteorological forcing. In general, the higher spatial resolution and the higher quality of the local forcing data results in better model predictions (PCR-GLOBWB and OSWS) of both soil moisture and streamflow. The coarse resolution of the global forcing could lead to failures to detect extreme rainfall events or differences at specific regions, such as mountainous areas. Recent studies have developed several downscaling procedures, e.g.

geostatistical methods of blending satellite and gauge data (Chappell et al., 2013), that could improve model predictions when global forcing is used.

To improve the representation of the global water cycle using global hydrological models, one could follow multiple strategies. Improve the quality and quantity of ground observation, increase the spatial resolution of the global models or obtain more detailed information on the catchment properties (e.g. soil data). Another way forward is the assimilation of observations and the use of high spatial resolution meteorological data to bridge the gap between the different spatial scales for which large-scale hydrological models are designed and the river basin scale. The advantage of this approach is that it provides a global improvement of the hydrological simulation and the satellite data often have a global coverage. In this study we show the potential gain in model accuracy of using remotely sensed observations. Previous studies on the potential gain of satellite observations for global and continental hydrological models agree with the obtained results in the present manuscript (Andreadis and Lettenmaier, 2006; Lievens et al., 2015).

5.8 Conclusions

The study investigates the influence of discharge and soil moisture assimilation on the accuracy of large-scale model (PCR-GLOBWB) predictions, when driven by meteorological forcing datasets of high and coarse resolution, compared with local-scale model (OSWS) estimations.

Results show poor PCR-GLOBWB streamflow and soil moisture estimates when no observations are assimilated. The assimilation of soil moisture observations results in the largest improvement of the model estimates of streamflow. The joint assimilation of both streamflow and downscaled soil moisture observations leads to further improvement in streamflow simulations (20% reduction in $RMSE$).

In general, the higher spatial resolution of the local forcing data results in better models predictions of both soil moisture and streamflow. The added value of using higher spatial resolution forcing data is more significant for PCR-GLOBWB than for OSWS. When the impact on model accuracy of assimilating observations and forcing the models with higher spatial resolution data are compared, the latter leads to a more substantial improvement of streamflow predictions.

Furthermore, results show that the added contribution of data assimilation, for both soil moisture and streamflow, is more pronounced when the global meteorological data are used to force the models. This is caused by the higher uncertainty and coarser resolution of the global forcing.

The greatest benefit is obtained when local coarse resolution forcing data are used in combination with streamflow and soil moisture observations assimilation into the large-scale hydrological model, PCR-GLOBWB.

In conclusion, the present research study shows that data assimilation of high resolution soil moisture succeeds in resolving short-comings that exist nowadays in global hydrological models and can partly overcome the difference in model performance between a large-scale hydrological model driven by coarse resolution forcing data and a local-scale model forced with higher resolution meteorological data. Moreover, it demonstrates that further investments and improvements in remotely sensed observations, especially in soil moisture products can benefit large-scale hydrological model predictions and bring these closer to those obtained from local-scale hydrological modelling.

6 Assimilating in situ gauged and satellite-based discharge observations for hydrological modelling.

This chapter is based on:

López López P., Weerts H. A., Schellekens J., Van Dijk A. I. J. M. and Rodríguez Sandoval E. A. (in review). Assimilating in situ gauged and satellite-based discharge observations for hydrological modelling. Water Resources Research.

Abstract

In this study, in situ gauged and satellite-based discharge observations were assimilated using an ensemble Kalman filter into a grid-distributed version of the HBV-96 rainfall-runoff model in the Magdalena-Cauca basin from 2006 to 2012. Precipitation based on interpolation of in situ observations and estimates from the global MSWEP dataset were used and compared. Satellite-based discharge were obtained from Satellite Gauging Reaches (SGR) based on water extent estimates from the passive microwave Global Flood Detection System (SGR-GFDS) and from the optical MODIS instrument (SGR-MODIS). The latter data sources are relevant where in situ discharge data are available historically but not in near-real time. Moreover, the impact of model and observation uncertainty on discharge estimates within the data assimilation framework was analysed.

Results showed that timing and flow variability were fairly well captured by unconstrained runs when the model was forced with MSWEP precipitation. Higher agreement between simulated and observed discharge was found when in situ precipitation was used. The largest improvement in discharge estimates was obtained when assimilating in situ gauged discharge observations at locations in the Magdalena and Cauca Rivers. Assimilating satellite-based discharge observations still led to a marked improvement. The output assimilation of SGR-MODIS discharge data increased model discharge performance more than when SGR-GFDS observations were assimilated at most locations. These findings demonstrate the benefit of assimilating satellite-based discharge observations into a hydrological model especially when the model or forcing data are of limited accuracy, which may have a great added value for ungauged or poorly gauged river basins.

6.1 Introduction

Data assimilation techniques optimally combine information from observational data and model simulations to improve model precision and accuracy (Reichle, 2008). Model states are updated using independent observations to obtain the best possible estimate of initial conditions. Data assimilation has been broadly applied in hydrological modelling. Most of previous studies in this field have investigated the impact of assimilating in situ discharge data on model estimates (Pauwels and de Lannoy, 2006; Clark et al., 2008; Seo et al., 2009; Rakovec et al., 2012, 2015; Abaza et al., 2014), since they are often available at high temporal scale and they integrate information of all hydrological processes. In addition to discharge, in situ soil moisture, snowpack and water level data have also been assimilated for improving model simulations (Weerts et al., 2010; Lee et al., 2011).

However, in situ data are not available in many river basins around the world, or even if data are available, they can be sparse, inaccurate or collected at different temporal scales (Sivapalan et al., 2003; Loukas and Vasiliades, 2014). Earth observations provide a promising alternative to in situ data (Van Dijk and Renzullo, 2011). Recent studies have proven the potential benefit for hydrological model predictions of assimilating Earth observations, such as soil moisture (Draper et al., 2011; Wanders et al., 2014b; Lievens et al., 2015; López López et al., 2016), snow (Thirel et al., 2013; Franz et al., 2014) and total water storage (Su et al., 2010; Forman et al., 2012; Van Dijk et al., 2014; Tangdamrongsub et al., 2015; Khaki et al., 2017). However, fewer studies evaluated the impact of assimilating remotely-sensed hydraulic information, such as the extent of surface water bodies or river bathymetry (see Revilla-Romero et al., 2016 for a review). To the authors' knowledge, only two studies (Zhang et al., 2013; Revilla-Romero et al., 2016) have assimilated passive microwave surface water extent derived from the Global Flood Detection System (GFDS) into a hydrological model.

The present study aimed to investigate the impact of assimilating in situ gauged and satellite-based discharge observations on hydrological model simulations in the Magdalena-Cauca basin in Colombia. To this end, eight output assimilation scenarios were analysed, in which discharge data derived from in situ observations and from GFDS water extent as well as from optical MODIS water extent data (Van Dijk et al., 2016) were used. For each scenario, the hydrological model was driven by both local and global precipitation estimates.

In contrast to previous studies that limited their analysis to specific flood events or short time periods up to two years, a seven year period was considered for data assimilation in this study. Moreover, previous studies were focused on the solely output assimilation of satellite-based discharge, whereas in the present study, both in situ and satellite-based were independently assimilated to evaluate their relative potential benefit to improve model predictions.

In the output assimilation process, uncertainties associated with model states, parameters and meteorological forcing as well as observation uncertainties can be taken into account

for a better prediction (Evensen, 2003). As discussed by Weerts and El Serafy (2006) and Seo et al. (2009), all sources of uncertainty should be investigated through sensitivity analysis to evaluate the impact of these assumptions into model simulations. Only few hydrological studies have included this analysis, evaluating the effect of the updating frequency and the observation network density (Lee et al., 2011; Rakovec et al., 2012), the noise specification (Noh et al., 2014), the ensemble size and the localization problem (Rasmussen et al., 2015) and the length of the verification period (Randrianasolo et al., 2014). In the present study, the impact of model (ensemble size, forcing and updating frequency) and observations uncertainty was addressed.

Understanding the potential benefit of assimilating satellite-based discharge observations into a hydrological model driven with global precipitation data is a relevant research opportunity and may have a great added value for ungauged or poorly gauged river basins.

6.2 Study area

The study area is the Magdalena-Cauca basin (*Figure 6.1*), which drains about 25% of the total territory of Colombia (approx. 257,000 km^2), being the primary fluvial system in the country where almost 80% of Colombian population lives. The Magdalena River originates at the Magdalena Lagoon in the Colombian Andes at an elevation of approx. 3,700 m and it flows for about 1,600 km through the western part of Colombia until reaching the Atlantic Ocean (Restrepo et al., 2016). The main tributary of the Magdalena River is the Cauca River. Average annual discharge at the outlet of the basin is approx. 7,200 $m^3\ s^{-1}$, varying from 4,050 $m^3\ s^{-1}$ in March to 10,200 $m^3\ s^{-1}$ in November (Camacho et al., 2008).

Meteorological phenomena in the Magdalena-Cauca basin are profoundly influenced by the effect of El Niño-Southern Oscillation (ENSO; Hoyos et al., 2013), leading to a large inter-annual hydroclimatic variability. In the upper and mid regions of the basin, the climate is characterised by two wet periods (April-May and October-November) and two dry periods (December-March and June-September), whereas in the lower Magdalena-Cauca basin a single wet period occurs (May-November). Average annual precipitation and evapotranspiration are ca. 2,150 mm and 1,630 mm, respectively, with higher values towards the western region of the basin (Poveda et al., 2006).

The Magdalena-Cauca basin is the main water source of the country for human consumption, agricultural and industrial activities and ecosystems support. The basin is extensively regulated, through a dense network of dams. During the last decade, the basin has experienced numerous extreme flood events, such as the 2010-2011 flood which affected four million people and caused economic losses of about US 7.8 billion (Ricaurte et al., 2017) and severe droughts, such as the 2015-2016 drought which caused water shortages in more than 25% of the towns and a substantial amount of fires (Hoyos et al., 2017). All these hydro-meteorological characteristics make the hydrological modelling of the Magdalena-Cauca basin extremely challenging.

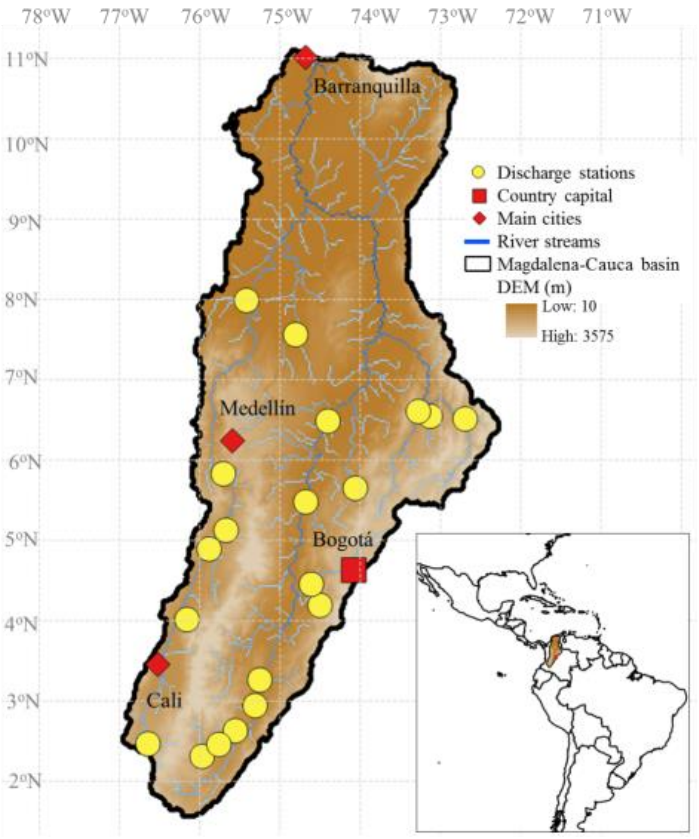


Figure 6.1 Magdalena-Cauca basin in Colombia. Selected discharge stations are shown with yellow points.

6.3 Hydrological model

In this study, the OpenStreams wflow-hbv model (Schellekens et al., 2016) was used. It is a distributed hydrological model programmed in the PCRaster-Python environment (Karssenberger et al., 2010) and based on the lumped HBV-96 model (Lindström et al., 1997).

The grid-based hydrological model OpenStreams wflow-hbv determines the different water balance components using three routines: precipitation-snow, soil moisture and runoff response. For each grid cell, the model considers four model states: snow (SN), soil moisture (SM), upper zone storage (UZ) and lower zone storage (LZ). The dynamics of the model states are governed by the following fluxes: precipitation, snowfall, snowmelt, actual evaporation, seepage, capillary rise, percolation, direct runoff (Q_{DR}), interflow from the UZ (Q_{IF}) and baseflow from the LZ (Q_{BF}). The latter three fluxes force the kinematic wave model. This routing scheme calculates the total flow (Q or discharge) using one additional model state, the water level over the drainage network (H). Reservoirs are not included in the model. The structure of the hydrological model is shown in *Figure 6.2*.

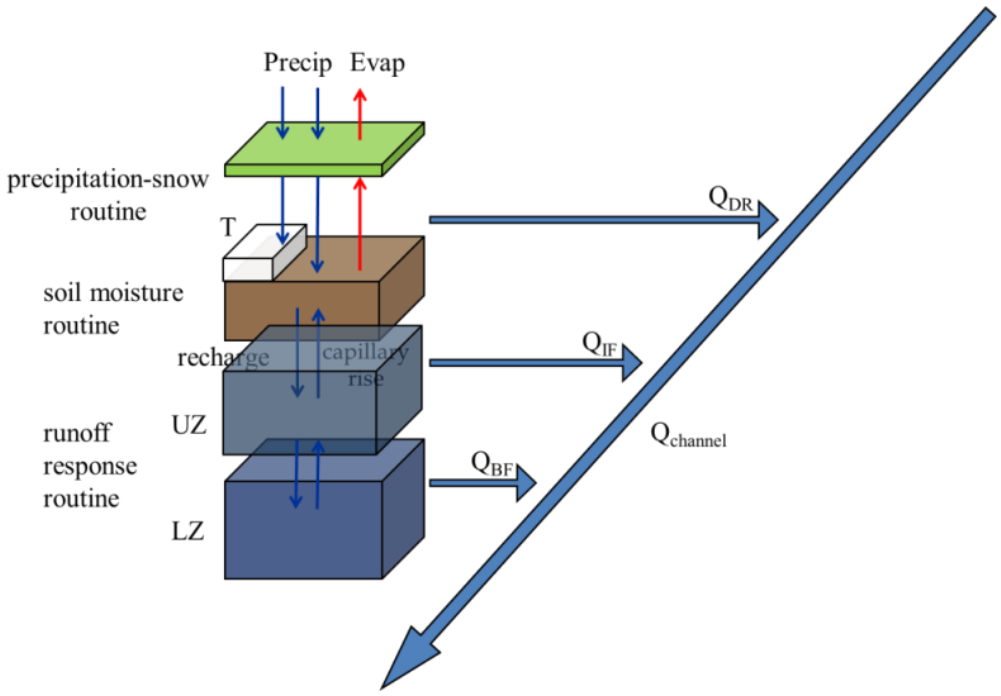


Figure 6.2 OpenStreams wflow-hbv model structure.

OpenStreams wflow-hbv was applied at 2 km spatial resolution and at daily temporal scale for the time period from January 2000 to December 2012. The model was calibrated and validated using interpolated in situ precipitation through the Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, 1992) methodology. Model parameterization of a previously made model for the basin at 1 km grid (López López et al., 2018) was adapted to the new spatial resolution. To spin up the model, the 1 year period 2000 was used. The time period 2001-2005 was used for calibration, and the time period 2006-2012 was used for validation. Traditionally, Nash Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970) is used for model calibration. However, Kling-Gupta efficiency (KGE; Gupta et al., 2009) was used here to avoid underestimating the variability of values.

6.4 Data

6.4.1 Meteorological data

Daily precipitation, air temperature and reference potential evapotranspiration data for 2000-2012 period are needed to force OpenStreams wflow-hbv. Evapotranspiration data were obtained using the FAO Penman-Monteith equation from the WATCH Forcing Data methodology applied to ERA-Interim reanalysis data (WFDEI) at approx. 50 km grid (Weedon et al., 2014). Evapotranspiration was downscaled to a 2 km grid using the e2o-downscaling-tools (Weiland et al., 2015; Schellekens and Weiland, 2017). Temperature data were obtained from 616 weather stations located in the basin. Temperature values at weather stations were interpolated using the inverse distance weighting algorithm to create spatial maps at 2 km spatial resolution. Two different precipitation products were used to analyze the impact of precipitation on discharge estimates with and without data

assimilation: (i) the Multi-Source Weighted-Ensemble Precipitation data (MSWEP; Beck et al., 2017a) and (ii) in situ interpolated precipitation. MSWEP precipitation was downscaled from approx. 25 km to 2 km grid using the e2o-downscaling-tools. In situ precipitation maps at 2 km grid were generated interpolating ground precipitation values at 616 weather stations using the inverse distance weighting algorithm. For a detailed evaluation of precipitation datasets in the Magdalena-Cauca basin, the reader is referred to López López et al. (2018). Temperature and precipitation values of weather stations were provided by the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) of Colombia.

6.4.2 Discharge data

Discharge time series for 20 locations along the Magdalena River and its tributaries were obtained from three different sources: (i) in situ gauged discharge, (ii) discharge derived from passive microwave sensors and (iii) discharge derived from optical MODIS data. The reason behind the selection of the 20 locations is to have a varied representation of sub-basins (topography, area, climatic variability, etc.) and the better performance of the hydrological model and the satellite data, due to their generation process and spatial resolution. *Figure 6.1* shows the locations where discharge data were assimilated and *Table 6.1* includes the location, river and sub-basin area for each discharge site.

Daily in situ discharge data for 20 selected gauging stations were provided by the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) of Colombia (from hereafter referred as IN SITU discharge).

Satellite-based discharge observations were obtained using the Satellite Gauging Reaches (SGRs) method proposed by Van Dijk et al. (2016). In the SGR method, firstly, the differences in the radiation intensities between land and surface water are used to calculate a ratio of the proportion of water and land per grid cell, providing a measurement of water extent. Secondly, a correlation analysis between water extent and discharge is used to select those grid cells with the highest coefficient of correlation. Thirdly, the SGR-based discharge estimates are derived using the cumulative distribution function matching method.

Two data sources were used to derive water extent measurements per grid cell. Firstly, satellite-based water extent data were derived from passive microwave sensors provided by the Global Flood Detection System (GFDS) implemented at the European Commission Joint Research Centre (De Groeve et al., 2015). Data from the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) and the Advanced Microwave Scanning Radiometer for Earth Observation System (AMSR-E) were used to calculate water extent at a spatial resolution of approx. 9 km (Kugler and De Groeve, 2007; De Groeve, 2010). Four days running mean discharge data were used (from hereafter referred as SGR-GFDS discharge) to avoid temporal gaps in the daily time series.

Secondly, satellite-based water extent data were derived from optical MODIS data using MODIS 8 day composites of imagery from NASA provided from the AQUA and TERRA MODIS instruments to calculate water extent at approx. 5 km spatial resolution. Eight days

running mean discharge data were used (from hereafter referred as SGR-MODIS discharge).

SGR-GFDS and SGR-MODIS discharge data were extracted for 2000-2012 period at the locations of the 20 gauging stations. For further details about the discharge time series derived from satellite data, the reader is referred to Van Dijk et al. (2016).

Table 6.1 Discharge stations used for output assimilation.

Station name	Station ID IDEAM	Latitude	Longitude	River	Area [km²]
Mateguadua	26107130	4.02	-76.16	Tuluá	664
El Retiro-Chinchiná	26157020	5.13	-75.67	Chinchiná	986
Villa Losada	21057080	2.31	-75.97	La Plata	1080
Jardín El Hda.	21087070	2.63	-75.56	Yaguará and Iquira	1103
La Playa	21197030	4.20	-74.50	Sumapaz	1259
Borbur	23127010	5.65	-74.06	Carare	1607
San Gil	24027010	6.55	-73.13	Fonce	1849
El Retiro	27037020	7.56	-74.81	Nechí	2551
Paicol	21057060	2.46	-75.76	Paez	4078
Montelíbano-Automat.	25017010	7.99	-75.42	San Jorge	4463
Puente Portillo	21207960	4.46	-74.61	Bogotá	5544
Capitanejo	24037390	6.51	-72.69	Chicamocha	6592
Remolino	24047020	6.61	-73.28	Suárez	9312
Julumito	26017020	2.47	-76.64	Cauca	724
La Virginia	26177030	4.89	-75.88	Cauca	22814
Puente Iglesias	26207030	5.83	-75.71	Cauca	29022
Puente Carretera	21137030	3.26	-75.25	Magdalena	746
Puente Santander-Automat.	21097070	2.94	-75.31	Magdalena	15705
Puerto Salgar-Automat.	23037010	5.48	-74.68	Magdalena	56905
Puerto Berrío-Automat.	23097030	6.49	-74.40	Magdalena	74410

6.5 Methodology

6.5.1 Assimilating in situ gauged and satellite-based discharge observations

6.5.1.1 Ensemble Kalman Filter

For assimilating discharge observations into the model, the ensemble Kalman filter (EnKF) was used (Evensen, 1994, 2003). This data assimilation procedure is based on the Kalman filter (Kalman, 1960) and it uses a Monte Carlo approach that takes into account model and observations errors in a statistically optimal way. Several previous studies on data assimilation into hydrological models have applied the EnKF (Wanders et al., 2014b;

Tangdamrongsub et al., 2015; Revilla-Romero, 2016; López López et al., 2016). In the present study, the EnKF was applied using a generic sequential data assimilation module (OpenDA, 2017) embedded within the Delft-FEWS framework (Werner et al., 2013).

The equation to propagate the model states is defined as

$$x_{t+1} = f(x_t, u_{t+1}, \theta) + \omega_{t+1} \qquad \omega_t \sim N(0, S) \tag{6.1}$$

where x_{t+1} is the model state vector at time $t+1$, f is the hydrological model (i.e. set of equations that represent the hydrological processes in the system which exhibits a non-linear behaviour), u_{t+1} is the model forcing at time $t+1$ (i.e. precipitation, temperature and evapotranspiration), θ are the time-invariant model parameters and ω_{t+1} is the model error which is assumed to be normally distributed with zero mean and covariance S (additive Gaussian white noise). This additive system noise incorporates the overall uncertainties in forcing, model structure and parameters. Model errors can show some spatial patterns to be found in the covariance matrix S. However, as pointed out by Rakovec et al. (2012), quantification of S, for highly nonlinear hydrological systems, is a complicated task and, therefore, it was kept time-invariant.

The discharge observations y (IN SITU, SGR-GFDS and SGR-MODIS) are related to model states by the following equation

$$y_t = h(x_t) + \varepsilon_t \qquad \varepsilon_t \sim N(0, R_t) \tag{6.2}$$

where h is a non-linear function connecting model states and observations (the kinematic wave routing model) and ε_t is the observations error which is normally distributed with zero mean and covariance R_t (additive Gaussian white noise). For spatially independent measurement errors between the observations in vector y_t, R_t is diagonal. The main objective is to find the optimal model state, x_t, based on the model and the observations, y_t.

In each time step of the output assimilation procedure, two different model states vectors can be differentiated: the model states vector before an observation becomes available, which is defined as the forecasted (a priori) model states vector, x_t^-, and the model states vector that takes into account the observed data, which is defined as the updated (a posteriori) model states vector, x_t^+. To obtain the optimal x_t^-, a Monte Carlo approach is used, generating N ensemble members. This implies that the system is represented by N model states vectors instead of only one, which are combined into a matrix

$$X_t^- = (x_t^{-,1}, x_t^{-,2}, x_t^{-,3}, \dots, x_t^{-,N}) \tag{6.3}$$

where

$$x_t^{-,i} = (SN_{1:m}^{-,i}, SM_{1:m}^{-,i}, UZ_{1:m}^{-,i}, LZ_{1:m}^{-,i}, H_{1:m}^{-,i})_t^T, \tag{6.4}$$

$SN^i, SM^i, UZ^i, LZ^i and\ H^i$ are the model states of the ith ensemble member (*6.3 Hydrological model),* m gives the number of grid cells and T is the transpose operator.

Once the model states vector at time t, x_t, has been updated, equation 1 is used to forecast model states at time $t+1$ and hence X_{t+1}^- is obtained as follows

$$X_{t+1}^- = (x_{t+1}^{-,1}, x_{t+1}^{-,2}, x_{t+1}^{-,3}, \dots, x_{t+1}^{-,N}) \tag{6.5}$$

The ensemble mean

$$\overline{x_{t+1}^-} = \frac{1}{N}\sum_{i=1}^{N} x_{t+1}^{-,i} \tag{6.6}$$

is used to calculate the model error for each ensemble member as follows

$$E_{t+1}^- = (x_{t+1}^{-,1} - \overline{x_{t+1}^-}, x_{t+1}^{-,2} - \overline{x_{t+1}^-}, x_{t+1}^{-,3} - \overline{x_{t+1}^-}, \dots, x_{t+1}^{-,N} - \overline{x_{t+1}^-}) \tag{6.7}$$

The ensemble model covariance matrix is calculated from the model error for each ensemble member using the following equation

$$P_{t+1}^- = \frac{1}{N-1} E_{t+1}^- {E_{t+1}^-}^T \tag{6.8}$$

The Kalman gain matrix, which is a weighting factor of model and observations errors, is then obtained:

$$K_{t+1} = P_{t+1}^- H_{t+1}^T (H_{t+1} P_{t+1}^- H_{t+1}^T + R_{t+1})^{-1} \tag{6.9}$$

where $P_{t+1}^- H_{t+1}^T$ is approximated by the forecasted covariance between the model states and the forecasted discharge and $H_{t+1} P_{t+1}^- H_{t+1}^T$ is approximated by the covariance of forecasted discharge (Rakovec et al., 2012):

$$P_{t+1}^- H_{t+1}^T = \frac{1}{N-1}\sum_{i=1}^{N} \left(x_{t+1}^{-,i} - \overline{x_{t+1}^-}\right)\left(h(x_{t+1}^{-,i}) - \overline{h(x_{t+1}^-)}\right)^T \tag{6.10}$$

$$H_{t+1} P_{t+1}^- H_{t+1}^T = \frac{1}{N-1}\sum_{i=1}^{N} \left(h(x_{t+1}^{-,i}) - \overline{h(x_{t+1}^-)}\right)\left(h(x_{t+1}^{-,i}) - \overline{h(x_{t+1}^-)}\right)^T \tag{6.11}$$

where

$$\overline{h(x_{t+1}^-)} = \frac{1}{N}\sum_{i=1}^{N} h(x_{t+1}^{-,i}) \tag{6.12}$$

When observations become available, the model states of the ith ensemble are updated as follows

$$x_{t+1}^{+,i} = x_{t+1}^{-,i} + K_{t+1}(y_{t+1} - h(x_{t+1}^{-,i}) + \varepsilon_t) \quad (6.13)$$

As mentioned in Rakovec et al. (2012), in some previous studies on data assimilation a time delay was noted due to the use of the unit-hydrograph (Weerts and El Serafy, 2006; Pauwels and De Lannoy, 2009). In this study a kinematic wave model for the routing scheme in the distributed hydrological model is used, and hence, the time delay and attenuation are modelled more realistically and the discharge and states x_{t+1} are assumed to depend only on the states x_t (Markov property). Thus, the time delay is explicitly taken into account in the model.

6.5.1.2 Experimental setup

As a first step, satellite-based discharge observations (SGR-GFDS and SGR-MODIS) are evaluated in comparison to IN SITU discharge data. Pearson's correlation coefficient (r) between daily in situ gauged and satellite-based discharge data are calculated at each of the 20 stations to quantify the agreement between both observations. To ensure an independent evaluation, a leave-one-year-out cross-validation analysis was used for 2000-2012 period.

In situ gauged and satellite-based discharge assimilation is performed with the EnKF method updating all model states based on the differences between observed and simulated discharge values. Model parameters are kept constant to avoid too many degrees of freedom, and no uncertainty is considered in the model structure. Instead, ensembles are created by perturbing the precipitation fields using additive correlated Gaussian white noise (i.e. independent of time) with decorrelation length of 30 km (similar to Rakovec et al., 2015). Due to satellite-based discharge data availability (SGR-GFDS: 4-days running mean and SGR-MODIS: 8-days running mean), various running mean periods were tested, and a 14-days running mean was for data assimilation.

Four different scenarios were analysed: (i) Open-loop simulation without data assimilation (OL, reference scenario), (ii) assimilation of in situ gauged discharge (IN SITU), (ii) assimilation of discharge derived from passive microwave sensors (SGR-GFDS) and (iv) assimilation of discharge derived from optical MODIS data (SGR-MODIS). Every scenario is carried out with both in situ interpolated precipitation and MSWEP precipitation; resulting in a total of eight assimilation scenarios.

Initially, for computational reasons, a six month period from June 2008 to December 2008, including a dry and a wet period, was selected to perform various preliminary assimilation experiments related with the ensemble size, the input uncertainty, the discharge observations uncertainty, the updating frequency and the ability of the EnKF to update model states. With these experiments, a better understanding of the effect of model and observations uncertainty on discharge estimates inside the data assimilation framework is being pursued.

Firstly, the impact of assimilation on model states updating was investigated. Secondly, the ensemble size (16, 32 and 64) was analysed to determine the number of ensembles from

which stable model results were obtained with sufficient ensemble spread. Input uncertainty analysis was limited to precipitation uncertainty only. Precipitations (in situ and MSWEP) were perturbed with additive Gaussian noise with standard deviations based on 5%, 10% and 15% of the nominal value added to precipitation. A decorrelation length of 30 km was assumed for precipitation errors following Rakovec et al. (2015). Discharge observation errors were introduced differently to IN SITU, SGR-GFDS and SGR-MODIS discharge data. For IN SITU discharge observations the error covariance was set to zero and standard errors of 5%, 10%, 20% and 30% of the actual discharge were investigated. For SGR-GFDS and SGR-MODIS discharge observations, errors based on the rating of the satellite-based data against station data are calculated and applied at each location. Moreover, the impact of the updating frequency, i.e. how often discharge observations are assimilated into the model, was analysed for updating frequencies of 1, 4, 14 and 30 days. The performance of the data assimilation procedure for each experiment was evaluated calculating KGE for each discharge station.

Finally, once previous aspects of the data assimilation procedure were identified, in situ gauged and satellite-based discharge observations are assimilated for January 2006 to December 2012 time period. The effect of assimilating discharge observations on model performance is evaluated by computing several indicators between observed and simulated discharge, for every scenario and location, including KGE, r and Root Mean Square Error ($RMSE$).

6.6 Results

6.6.1 Evaluating satellite-based discharge observations

To evaluate satellite-based discharge observations, *Figure 6.3* shows r values between daily in situ gauged and satellite-based discharge data for 2000-2012 period. The most successful satellite-based discharge observations ($r > 0.80$, green) are found in the main channel of the Magdalena River. Good agreement between in situ gauged and satellite-based discharge ($0.40 < r < 0.80$) are found at locations in the Cauca River and some smaller tributaries, such as Tuluá and Chicamocha Rivers. The lowest r values ($r < 0.40$) are obtained at locations in sub-basins with smaller area, such as Jardín El Hda. and Julumito.

In the comparison between SGR-GFDS and SGR-MODIS, r values lower than 0.40 are found at 12 locations when evaluating SGR-GFDS observations, whereas at only 3 locations when evaluating SGR-MODIS observations. r values obtained with SGR-MODIS are higher than those obtained with SGR-GFDS at 14 locations. At 3 locations SGR-GFDS and SGR-MODIS observations perform similarly, and at the remaining ones, SGR-GFDS outperforms SGR-MODIS.

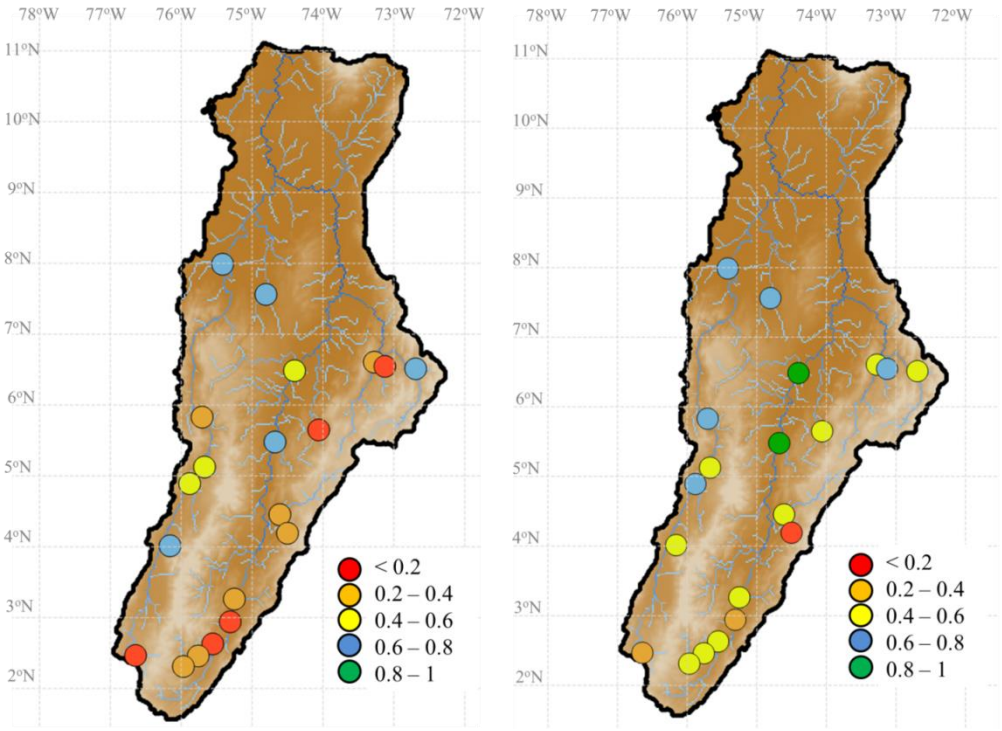


Figure 6.3 r values between daily in situ gauged and satellite-based (SGR-GFDS: left and SGR-MODIS: right) discharge at all locations for 2000–2012 period.

6.6.2 Assimilating in situ gauged and satellite-based discharge observations

6.6.2.1 Impact of assimilation on model states estimates

To better understand the impact of assimilating discharge observations into a spatially distributed hydrological model on model states updating, *Figure 6.4* shows the mean difference between the forecasted and updated model states on 1 December 2008 when IN SITU discharge data were assimilated. This particular day was selected as an example because the model overestimated discharge in most of the stations located in the Magdalena River, whereas discharge simulations were similar to observations in the Cauca River. Because the overestimation of the discharge in the Magdalena River, negative differences are found between the forecasted and updated model states in this part of the basin. The data assimilation framework removes water from that area during the update. The differences are not only observed in terms of discharge, but also in the updated soil moisture and upper zone states, showing a spatial differentiation between the area of the Magdalena River, where the amount of water is decreasing and the area of the Cauca River, where there is hardly any data assimilation increment either positive or negative. The lower zone state is full of water, so no differences were found when comparing the state before and after the updating (and therefore not shown). The discharge overestimation in the Magdalena River is higher when the model is forced with MSWEP precipitation (*Figure 6.4*, second row).

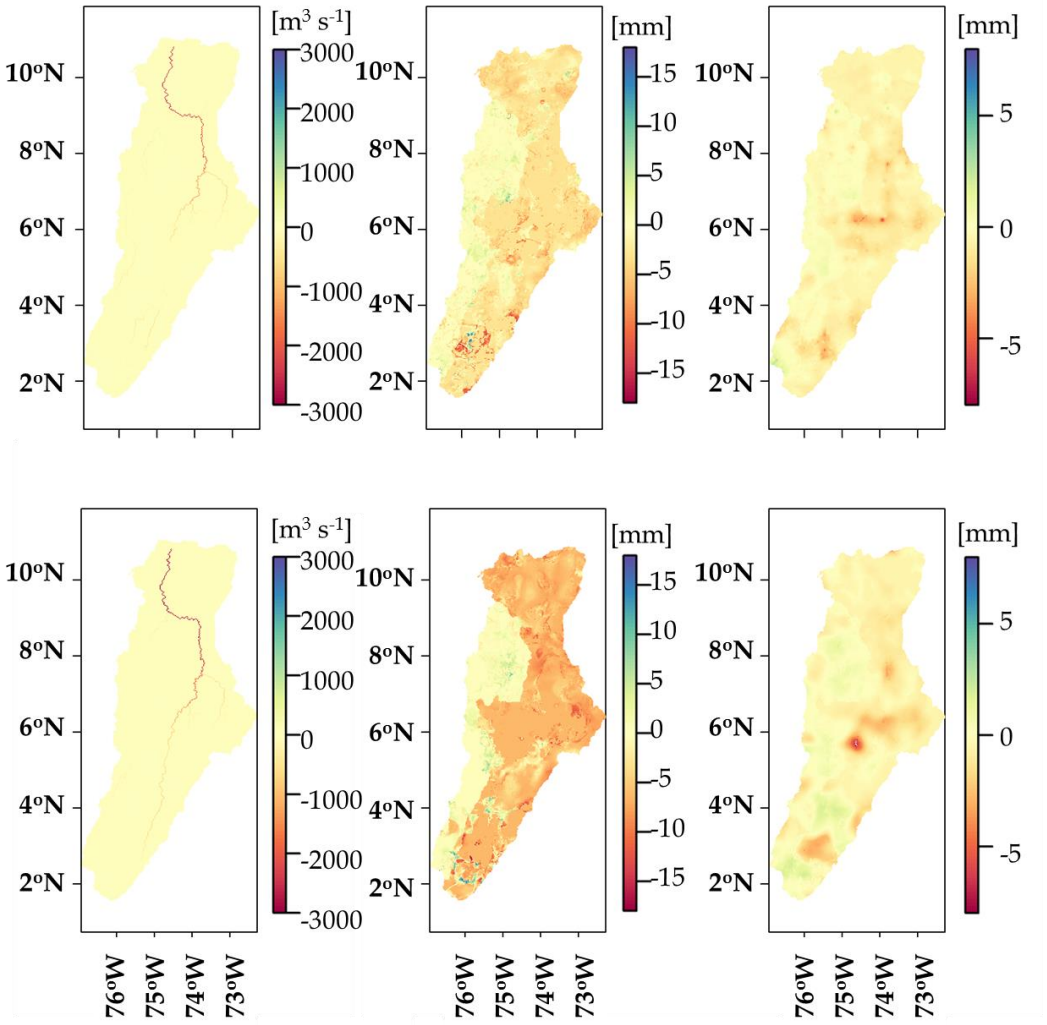


Figure 6.4 Mean difference between updated (X_{t+1}^{+}) and forecasted (X_{t+1}^{-}) mode states on 1 December 2008 when assimilating IN SITU discharge observations into the model forced with in situ (first row) and MSWEP (second row) precipitation products. Three model states are shown: discharge (left column), soil moisture (middle column) and upper zone (right column).

6.6.2.2 Impact of model and observations uncertainty on discharge estimates

Several experiments were performed to assess the impact of methodological choices in the data assimilation framework on discharge estimates, involving assimilation of IN SITU, SGR-GFDS and SGR-MODIS discharge data for June 2008 to December 2008. Results of these experiments are shown in *Figure 6.5*, *Figure 6.6*, *Figure 6.7* and *Figure 6.8*. KGE values between observed and simulated discharge were represented as a function of ensemble size (*Figure 6.5*), precipitation error (*Figure 6.6*), updating frequency (*Figure 6.7*) and in situ gauged discharge error (*Figure 6.8*).

Higher KGE values are obtained when the model is forced with in situ precipitation compared to when MSWEP precipitation is used, with average differences in the order of 0.15. *Figure 6.5* shows that increasing the ensemble size from 16 to 32 leads to an improvement in KGE, whereas when increasing further to 64 ensemble members, almost no improvement is found. These results suggest that 32 ensemble members are sufficient to obtain stable results.

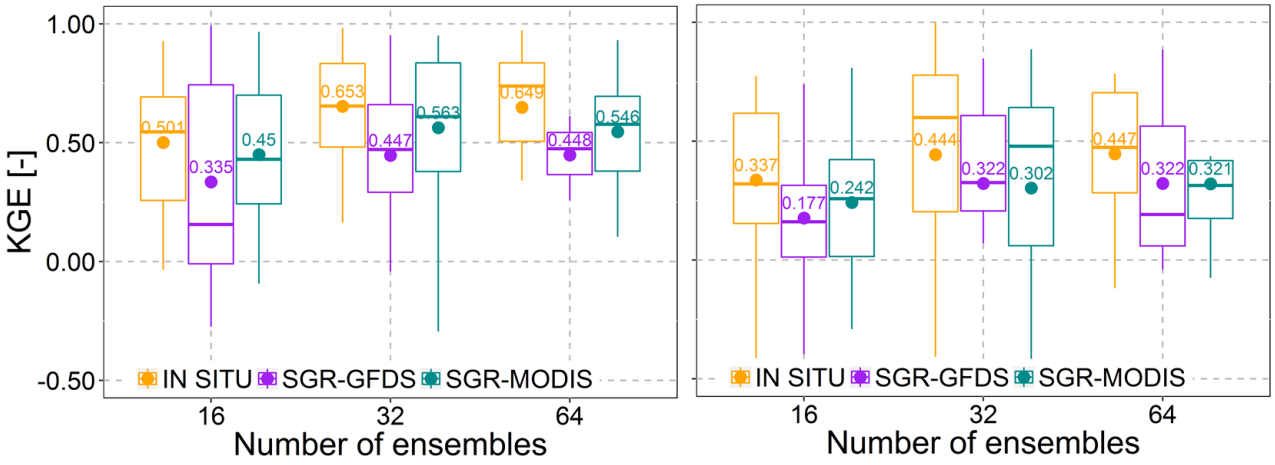


Figure 6.5 KGE values (y-axis) between daily observed and simulated discharge at all stations as function of ensemble size (x-axis) when the model was forced with in situ (left column) and MSWEP (right column) precipitation products for June 2008 to December 2008 period. Different colours indicate different data assimilation scenarios: IN SITU (orange), SGR-GFDS (purple) and SGR-MODIS (blue).

Figure 6.6 shows that the highest KGE values are obtained when precipitation is perturbed with additive Gaussian noise with standard deviations based on 10% of the nominal value. Standard deviations based on 5% and 15% of the nominal value of precipitation show lower discharge performances, although differences are small.

Figure 6.7 shows KGE values obtained when discharge data are assimilated into the model at updating frequencies of 1, 4, 14 and 30 days. The highest KGE values are obtained with updating frequencies of 1 and 4 days, with marginal differences ($KGE < 0.02$ on average). When further increasing the updating frequency, a gradual decrease in KGE for all the cases is found. For updating frequencies of 14 and 30 days, KGE values are reduced an average of 0.05 and 0.15, respectively.

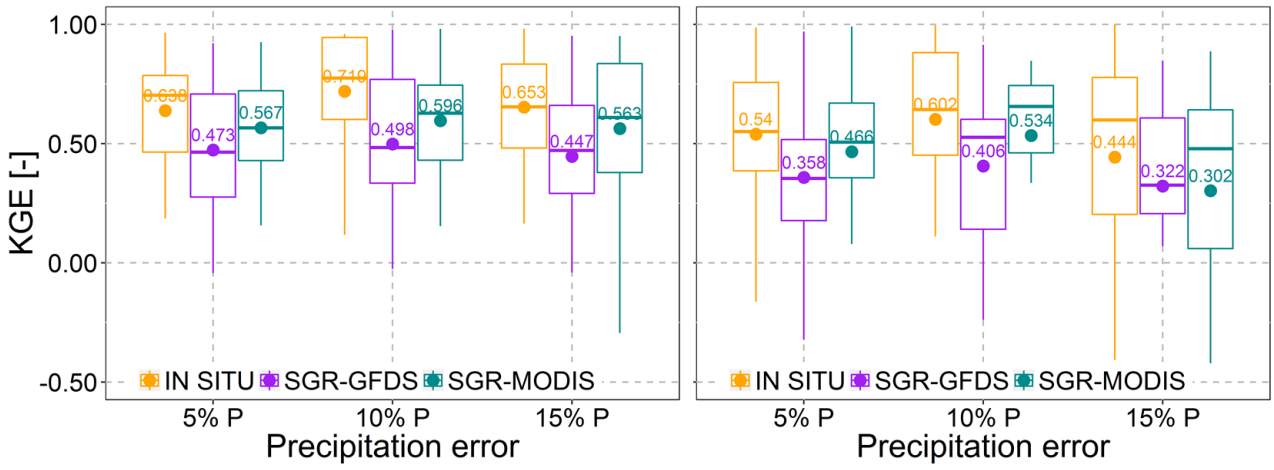


Figure 6.6 KGE values (y-axis) between daily observed and simulated discharge at all stations as function of precipitation error (x-axis) when the model was forced with in situ (left column) and MSWEP (right column) precipitation products for June 2008 to December 2008 period. Different colours indicate different data assimilation scenarios: IN SITU (orange), SGR-GFDS (purple) and SGR-MODIS (blue).

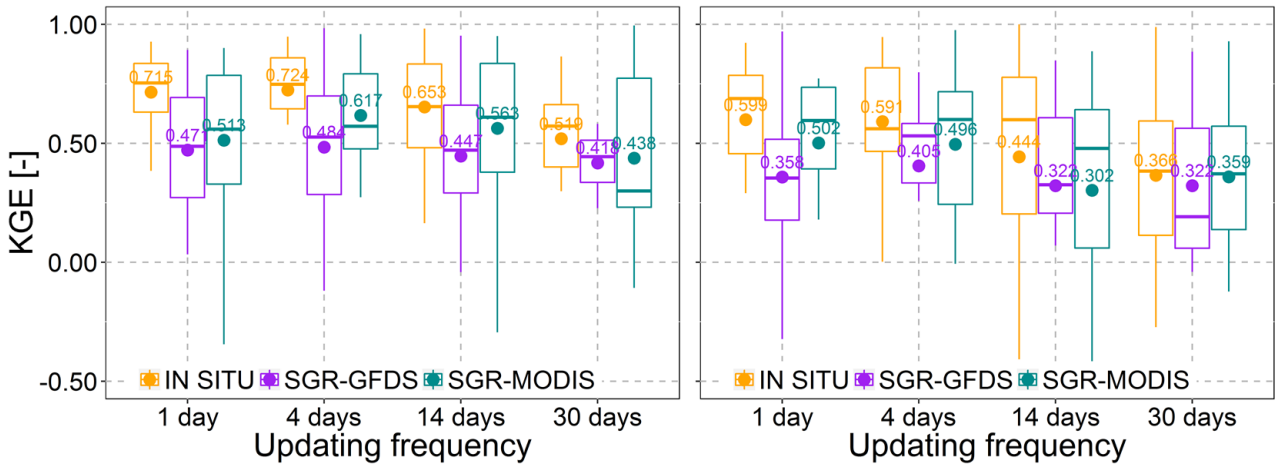


Figure 6.7 KGE values (y-axis) between daily observed and simulated discharge at all stations as function of updating frequency (x-axis) when the model was forced with in situ (left column) and MSWEP (right column) precipitation products for June 2008 to December 2008 period. Different colours indicate different data assimilation scenarios: IN SITU (orange), SGR-GFDS (purple) and SGR-MODIS (blue).

Discharge estimates obtained when assimilating IN SITU observations with a standard error of 10% show the highest performance, as shown in *Figure 6.8*. *KGE* differences obtained with standard errors of 5%, 10%, 20% and 30% vary from 0.01 to 0.09 when in situ precipitation is used and from 0.06 to 0.21 when the model is forced with MSWEP precipitation.

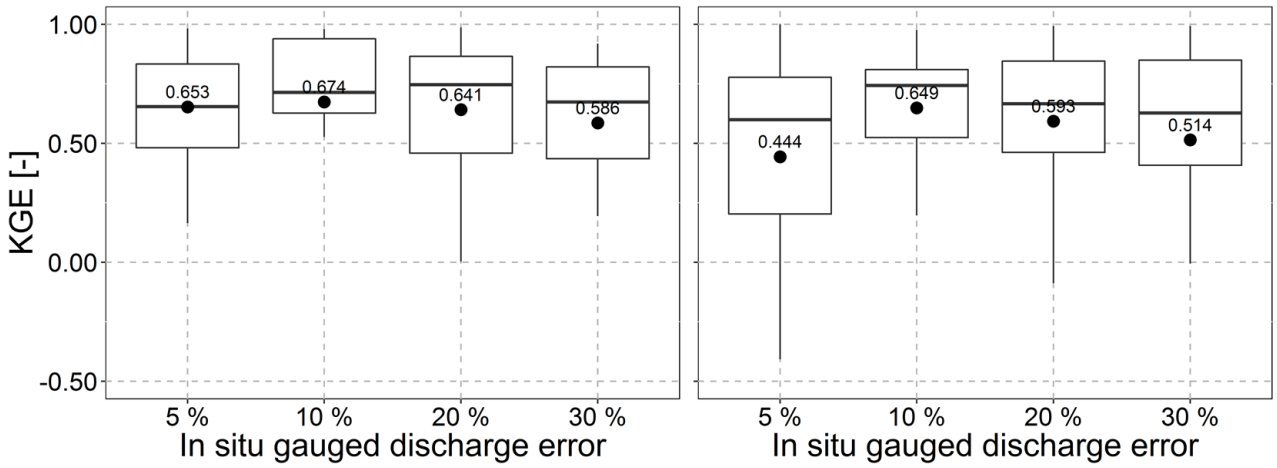


Figure 6.8 KGE values (y-axis) between daily observed and simulated discharge at all stations as function of in situ discharge error (x-axis) when the model was forced with in situ (left column) and MSWEP (right column) precipitation products for June 2008 to December 2008 period.

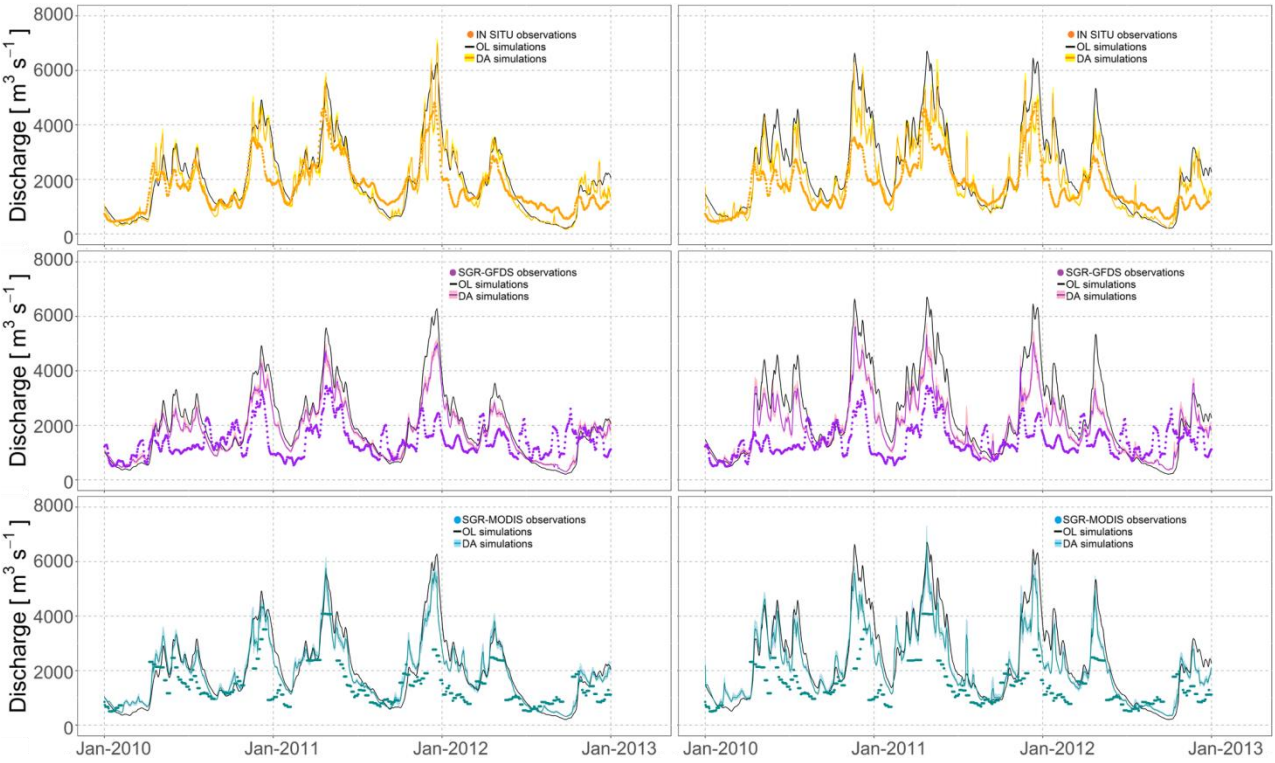


Figure 6.9 Simulated and observed discharge estimates at Puerto Salgar-Automat. in the Magdalena River for January 2010 to January 2013 period when the model was forced with in situ (left column) and MSWEP (right column) precipitation products. Rows show simulated discharge when IN SITU (first row), SGR-GFDS (second row) and SGR-MODIS (third row) discharge data were assimilated.

6.6.2.3 Impact of assimilation on discharge estimates

Once all sources of uncertainty within the data assimilation framework had been analyzed, in situ gauged and satellite-based discharge observations were assimilated into the model for the seven year period from 2006 to 2012. A comparison between observed and simulated discharge without and with IN SITU, SGR-GFDS and SGR-MODIS data assimilation was carried out. *Figure 6.9* shows the time series of discharge observations and simulations from January 2010 to January 2013 at Puerto Salgar-Automat. Simulated discharge without data assimilation (OL, red) shows a good agreement with in situ observations (black) over the entire period. This agreement is higher when forcing the model with in situ precipitation compared to using MSWEP precipitation product. Satellite-based discharge observations (green) capture the temporal variability of flow well, with SGR-MODIS showing better results than SGR-GFDS. The assimilation of SGR-MODIS reduces the differences between observed and simulated discharge time series to a greater extent than when SGR-GFDS data are assimilated (45% versus 15% increase in KGE values).

Boxplots of KGE, r and $RMSE$ values between daily observed and simulated discharge at all stations for the four assimilation scenarios are shown in *Figure 6.10*. Simulated discharge without data assimilation (OL, black), compared to discharge observations, shows reasonable average KGE, r and $RMSE$ values when the model is forced with either in situ ($KGE = 0.58, r = 0.73$ and $RMSE = 79.42\, m^3 s^{-1}$) or MSWEP ($KGE = 0.40$, $r = 0.73$ and $RMSE = 97.03\, m^3 s^{-1}$) precipitation dataset. The largest improvement derived from the discharge data assimilation is achieved when the model is forced with MSWEP precipitation, where the OL simulations are poorer. For example, KGE increases from 0.58 to 0.72 using in situ precipitation, whereas when forcing the model with MSWEP precipitation, KGE increases from 0.40 to 0.61. The highest improvement in KGE and $RMSE$ are found when assimilating IN SITU discharge data. The assimilation of SGR-MODIS discharge improves discharge simulations, increasing KGE values over 0.60. Assimilating SGR-GFDS discharge data also improves discharge estimates, although to a lesser extent. In terms of r, the assimilation of SGR-GFDS and SGR-MODIS discharge data does not show a big potential, leading to similar or even lower r values after the assimilation.

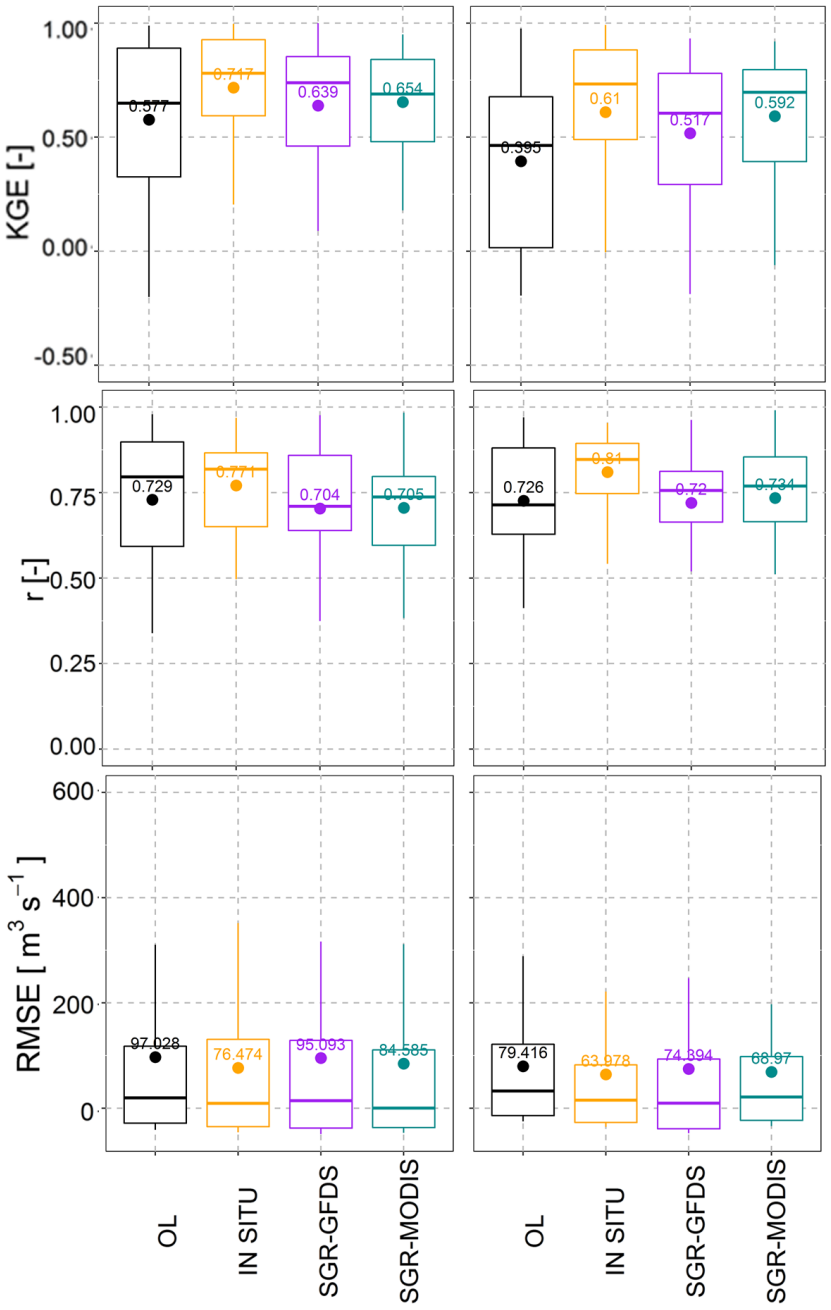


Figure 6.10 KGE, r and $RMSE$ values between daily observed and simulated discharge at all stations when the model was forced with in situ (left column) and MSWEP (right column) precipitation products for 2006 – 2012 period. Different colours indicate the four data assimilation scenarios: OL (black), IN SITU (orange), SGR-GFDS (purple) and SGR-MODIS (blue).

A different spatial pattern is found for the improvements on discharge estimates when assimilating satellite-based discharge observations. *Figure 6.11* shows KGE, r and $RMSE$ variations calculated between each assimilation scenario (IN SITU, SGR-GFDS and SGR-MODIS) and the reference assimilation scenario (OL) at four locations in the Magdalena and Cauca Rivers and their tributaries. Average values calculated across all the

discharge locations are also included in *Figure 6.11*. At locations situated in the main rivers of Cauca and Magdalena, such as La Virginia and Puerto Berrío-Automat., the improvement on simulated discharge after assimilation is significant in bias and flow variability representation, with average *KGE* increases of 0.12 and 0.31, when SGR-MODIS data are assimilated into the model forced with in situ and MSWEP precipitation products, respectively. At locations situated in sub-basins with smaller areas, such as Mateguadua and La Playa, the assimilation of satellite observations improves model discharge estimates to a lesser extent. This pattern may be caused by the lower performance of SGR-GFDS and SGR-MODIS discharge data in rivers of small volume.

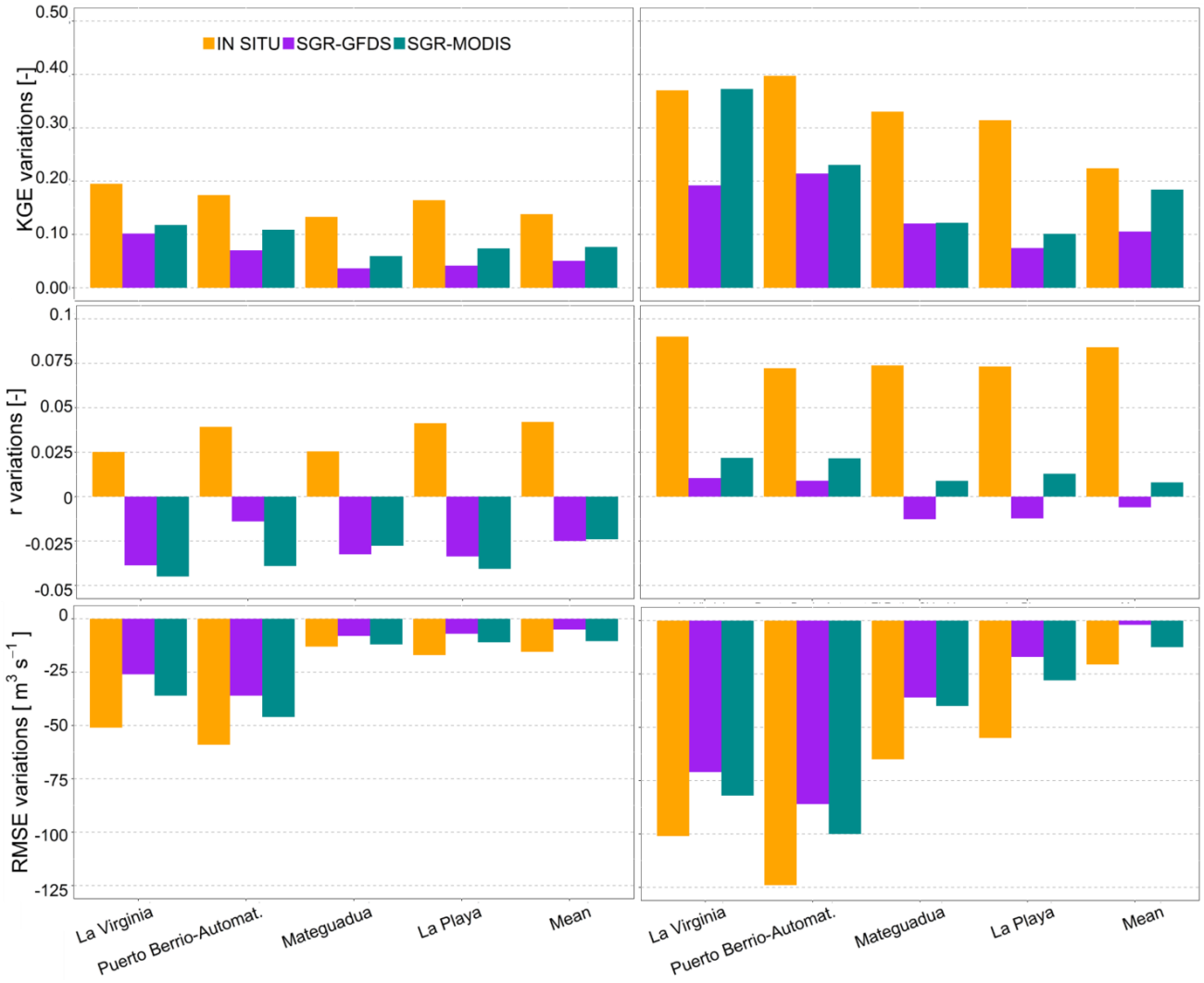


Figure 6.11 KGE (first row), *r* (second row) and *RMSE* (third row) variations comparing daily discharge estimates of assimilation scenarios IN SITU (orange), SGR-GFDS (purple) and SGR-MODIS (blue) with the reference assimilation scenario OL for 2006 – 2012 period. Left and right columns show the variations obtained when the model was forced with in situ and MSWEP precipitation products, respectively.

6.7 Discussion

Initially, the ability to measure discharge using satellite observations in the Magdalena-Cauca basin was evaluated. Topography, river width, floodplain area, land cover types, climatic areas and hydraulic structures, among other variables, impact the relationship between discharge and inundated area and therefore, the performance of the satellite-based discharge observations. In line with previous studies (Revilla-Romero et al., 2014; Van Dijk et al., 2016), the best satellite-based discharge estimates were obtained at locations in large and wide rivers with highly variable discharges, such as the Cauca and Magdalena Rivers (e.g. Puerto Berrío-Automat. and La Virginia). Good agreement between IN SITU, SGR-GFDS and SGR-MODIS data was also found in smaller rivers with lower discharges, but with a proportionally large floodplain (e.g. Tuluá River at Mateguadua). At some locations satellite-based discharge performance was poorer, which could be due to the coarse spatial resolution of the satellite signal (e.g. Julumito) or their situation downstream of a dam (e. g. Puente Portillo) or within highly vegetated areas. To further investigate the main factors that influence satellite-based discharge performance, a more detailed evaluation could be added, for example, categorizing the 20 selected locations as having natural or artificial flow regimes and comparing flow duration curves.

Overall, SGR-MODIS data appeared more suitable to provide discharge observations than SGR-GFDS data. As explained by Van Dijk et al. (2016), this may be due to the higher spatial resolution (approx. 5 km vs. 9 km). Satellite-based discharge data, both SGR-MODIS and SGR-GFDS, were obtained using a similar approach in which IN SITU discharge observations were required. Recently developed satellite products at finer spatial resolutions (e.g. from Sentinel-1 mission) and their combination with hydrological models could be used to overcome this limitation. Satellite-based discharge observations generated in that way would be extremely valuable in poorly gauged or ungauged river basins, where in situ observations are not available, to improve water resources management.

IN SITU, SGR-GFDS and SGR-MODIS discharge observations were assimilated into OpenStreams wflow-hbv model using an EnKF. Alternative data assimilation procedures where more observations can be used with hardly any extra additional computational time, such as the Asynchronous EnKF (Rakovec et al., 2015), may be investigated too. Moreover, the hydrological model may be calibrated using satellite-based discharge, instead of using IN SITU discharge observations (Pedinotti et al., 2014). It was calibrated when forced with in situ precipitation, but further calibration could be done when forced with MSWEP precipitation. In addition, the impact of hydraulic structures on the assimilation procedure may be addressed by including reservoirs in the hydrological model.

When assimilating data into a hydrological model, it would be desirable to objectively constrain all sources of uncertainty (Noh et al., 2014). However, this cannot always be achieved. In the present study, a preliminary analysis aiming to investigate the impact of model and observations uncertainty on discharge estimates was carried out. The effect of the ensemble size was tested, and results showed that increasing the number of ensembles over 32 did not further improve discharge simulations. However, a wider variety of

ensemble sizes should be tested (e.g. 25, 40, 50, 100 and 150) to extract more conclusive remarks. Increasing the number of ensembles over 100 or 200 may lead to a low improvement in model performance, not beneficial enough considering the necessary increase in computational time. Noh et al. (2014) proved that applying localization methods (Fertig et al., 2007; Sakov and Bertino, 2011) can allow considerably reduction of the ensemble size. Moreover, Rasmussen et al. (2015) demonstrated that less ensemble members are required as spatial coverage of discharge observations increases.

Input uncertainty analysis was focused on precipitation errors with results in line with previous assimilation studies (López López et al., 2016). Different noise specification methods to have into account input uncertainty could be tested, such as the rainfall ensemble generator developed by Rakovec et al. (2012).

The impact of the updating frequency at which discharge observations were assimilated was investigated to obtain the most accurate simulations. Results showed that lowering the updating frequency from 1 to 4 days does not decrease the KGE performance, further lowering the frequency to 14 and 30 days lowers the KGE performance.

Previous studies (Di Baldasarre and Montanari, 2009; Revilla-Romero et al., 2016; Wanders et al., 2014b) assumed a standard error for IN SITU discharge observations to be 30% of the actual discharge. In the present study, 5, 10, 20, and 30% standard errors were used. Although slightly higher KGE values were obtained with noise 10%, KGE differences can be considered small, and complementary analysis may be needed.

In situ gauged and satellite-based discharge observations were assimilated into a model forced with in situ and MSWEP precipitation datasets. Assimilating discharge observations improved discharge simulations, but accurate precipitation data are crucial for reproducing observed discharge well in terms of timing and flow variability. The importance of the quality of precipitation data in comparison to data assimilation was also pointed out by Tangdamrongsub et al. (2015).

The data assimilation approach applied here could be tested at other river basins with different hydro-meteorological characteristics or even globally, using large-scale hydrological models, such as PCR-GLOBWB (Sutanudjaja et al., 2018), WR3A (Van Dijk et al., 2014) and LISFLOOD (van der Knijff et al., 2010). Furthermore, satellite-based discharge observations derived from 2013 onwards could be assimilated to compare the impact on near-real-time model discharge estimates. Additional research combining assimilation of satellite-based discharge and soil moisture data could lead to more consistently improved discharge simulations (Wanders et al., 2014b; López López et al., 2016).

6.8 Conclusions

In situ gauged and satellite-based discharge observations were assimilated into a distributed hydrological model, OpenStreams wflow-hbv, forced with in situ and MSWEP precipitation datasets at the Magdalena-Cauca basin in Colombia. Timing and flow variability are fairly well captured by simulated discharge when the model is forced with

MSWEP precipitation (average $KGE = 0.40$). Higher agreement between simulated and observed discharge is found when in situ precipitation is used (average $KGE = 0.58$).

Results show the largest improvement in discharge estimates when assimilating IN SITU observations at locations in the main rivers (average $KGE = 0.72$). Assimilating satellite-based discharge observations still leads to a marked improvement. The assimilation of SGR-MODIS discharge data increases model discharge performance more than when SGR-GFDS observations are assimilated at most locations (average KGE values of 0.60 versus 0.52 after assimilation using MSWEP precipitation).

The added value of assimilation is more pronounced when MSWEP precipitation is used. When comparing the impact of data assimilation and forcing the model with higher quality precipitation data, the latter leads to a more significant improvement of discharge simulations.

In conclusion, this study shows the potential of assimilating satellite-based discharge observations for improving discharge model estimations, when driven by either in situ or global precipitation datasets and possibly model discharge forecasts (i.e. driven by numerical weather prediction). Moreover, it demonstrates that assimilation frameworks like the one here described can be applied in poorly gauged or ungauged river basins to increase model performance.

7 Synthesis

7.1 Introduction

Recent advances in remote sensing technologies have made available a large suite of observations of hydro-meteorological variables worldwide. These satellite-based observations can be integrated into hydrological models to improve water resources management not only at a global scale, but also at a regional or river basin scale. To this end, various approaches can be followed depending mainly on basin characteristics and in situ data availability, which can be classified in three groups: (i) meteorological datasets, (ii) model parameters estimation and data assimilation and (iii) model evaluation. Previous scientific efforts have been carried out following these approaches, however, further research is needed to reach the full potential of satellite-derived observations for increasing the accuracy of large- and local-scale hydrological models.

In this current status, this PhD research aims *to evaluate the applicability of global water resources datasets (including satellite-derived observations, in situ data and models) for hydrological modelling at the river basin scale.* In view of the complexity of this goal, this research focuses on applying and testing techniques that can potentially be implemented in any river basin around the world, depending on its hydro-meteorological characteristics and in situ data richness. Four river basins were selected and different approaches were applied to optimally integrate global datasets into large- and local-scale hydrological models, including the Brahmaputra basin in Bangladesh, the Magdalena-Cauca basin in Colombia, the Oum Er Rbia basin in Morocco and the Murrumbidgee basin in Australia.

In exploring the use of global datasets for hydrological modelling at the river basin scale, the following research questions were defined in this research, and are briefly discussed in the following sections:

1. What are the possibilities and limitations of large-scale hydrological models for water resources and flood estimation?

2. Does spatial precipitation downscaling improve the accuracy of satellite-derived precipitation datasets?

3. What is the impact of precipitation spatial resolution on streamflow model estimations?

4. Does calibration solely based on satellite-derived observations, such as soil moisture and evapotranspiration, improve large-scale model simulations? And how does this relate to calibration based on in situ discharge?

5. Is it possible to bring closer large-scale model simulations to those obtained with locally calibrated models with the assimilation of satellite-derived soil moisture observations?

6. Is it possible to improve large- and local-scale model simulations assimilating satellite-derived observations, such as soil moisture and discharge? And how does this relate to assimilating in situ discharge?

7. What is the improvement on model estimations that can be achieved assimilating satellite-derived observations in comparison to increasing the accuracy of meteorological forcing data (precipitation)?

This final chapter summarizes the main findings regarding these research questions, discusses their implications for the application of global water resources datasets for improving hydrological model simulations at river basin scale and provides a number of recommendations for future research.

7.2 Main results

In Chapter 2 several large-scale hydrological models and global datasets were evaluated for water resources and flood estimation in the Brahmaputra River basin. Five large-scale hydrological models (HTESSEL, LISFLOOD, PCR-GLOBWB, SURFEX-TRIP and WATERGAP3) were forced with global meteorological datasets (based on satellite-derived observations, in situ data and models) and their discharge model simulations and those derived from two multi-model ensemble combinations of them (MEAN and BS-MEAN) were compared with those obtained with a locally calibrated hydrological model (NAM+MIKE BASIN). WFDEI and MSWEP precipitation products were evaluated with in situ data and used to drive all hydrological models. This experimental set up aims to address research question 1.

In terms of meteorological datasets, results showed that both global precipitation datasets underestimated rainfall in the pre-monsoon, monsoon and post-monsoon periods (April to October), with lower underestimation with the MSWEP precipitation product. When integrating global meteorological datasets into large-scale and local-scale hydrological models, simulated and observed discharge showed fairly reasonable agreement. Similar or even better discharge performance was found for some large-scale models and the multi-model simulations when compared with local-scale model estimates. Flood evaluation results revealed that, in spite of magnitude differences, all hydrological models could give acceptable estimates of peak flows. These results are promising for large-scale and global meteorological datasets applications to support water resources and flood management, which is particularly relevant for poorly gauged or ungauged river basins.

In Chapter 3 the impact of the spatial resolution of four globally available precipitation products (MSWEP, TRMM, CMORPH and PERSIANN; based on satellite-derived observations, in situ data and models) on discharge estimates (with OpenStreams wflow-hbv) was evaluated in the Magdalena-Cauca River basin. A downscaling methodology was applied to increase the spatial resolution of satellite-derived precipitation estimates from 25 km to 1 km using a geographically weighted regression algorithm with auxiliary information from vegetation response (EVI), elevation, aspect and slope. In situ

precipitation measurements were also used in combination to global datasets to derive a precipitation product merging the benefits of both datasets. This experimental set up aims to address research questions 2 and 3.

The four global precipitation products performed quite well at their original spatial resolution, with MSWEP and TRMM showing higher correlations and lower biases. The downscaling methodology improved precipitation estimates, which was also reflected in discharge model simulations, increasing efficiencies in the order of 0.10 to 0.50 (KGE). To sum up, the increase of precipitation spatial resolution improved discharge model simulations which emphasizes the significance of having accurate and precise meteorological datasets for hydrological modelling.

In Chapter 4 a large-scale hydrological model (PCR-GLOBWB) and three global precipitation datasets (EI, WFDEI and MSWEP; based on satellite-derived observations, in situ data and models) were evaluated for water resources estimation in the Oum Er Rbia River basin. Moreover, the model was calibrated using satellite-derived soil moisture and evapotranspiration observations and five calibration scenarios were investigated: (i) reference scenario using the standard parameterization, (ii) calibration using in situ discharge, (iii) calibration using GLEAM actual evapotranspiration, (iv) calibration using ESA CCI surface soil moisture and (v) step-wise calibration using GLEAM actual evapotranspiration and ESA CCI surface soil moisture. This experimental set up aims to address research question 4.

In terms of meteorological datasets, similar performances were found between the three global precipitation products, with WFDEI showing the lowest performance. When calibrating the model to satellite-derived soil moisture and evapotranspiration observations, results showed reasonable discharge estimates with efficiencies from 0.50 to 0.75 (NSE). The combined use of both satellite-derived observations allows calibrating parameters that represent different hydrological processes in the model, leading to considerable improvements in model simulations. These findings open up the possibility to use large-scale hydrological models driven with global meteorological datasets and calibrated with satellite-derived soil moisture and evapotranspiration observations for water resources estimation for ungauged river basins.

In Chapters 5 and 6, the impact of assimilating satellite-derived soil moisture and discharge observations on model simulations was investigated. In Chapter 5 an Ensemble Kalman filter was successfully applied to assimilate in situ discharge and satellite-derived soil moisture observations (AMSR-E) into a large-scale hydrological model (PCR-GLOBWB) in the Murrumbidgee River basin. Moreover, a global meteorological dataset (WFDEI; based on satellite-derived observations, in situ data and models) was evaluated to in situ data and was used as model forcing. Four different assimilations scenarios were investigated: (i) open-loop simulation without data assimilation (reference scenario), (ii) assimilation of in situ discharge observations, (iii) assimilation of satellite-derived soil moisture observations and (iv) assimilation of in situ discharge and satellite-derived soil moisture observations. Large-scale model simulations of every assimilation scenario were compared to a locally

calibrated hydrological model (OpenStreams wflow-sbm) driven with in situ and global meteorological datasets. This experimental set up aims to address research questions 5, 6 and 7.

In terms of meteorological datasets, WFDEI showed good agreement with in situ data, with larger differences (magnitude and spatial variation) with local observations in mountainous regions. Moreover, it was shown that the assimilation of satellite-derived soil moisture observations resulted in the largest improvement of discharge model estimates. The combined assimilation of in situ discharge and satellite-derived soil moisture observations improved model simulations even further (20% reduction in $RMSE$), coming closer to estimates of the locally calibrated model. These results showed that integrating global meteorological datasets, large-scale hydrological models and satellite-derived observations can lead to similar performances at the river basin scale as derived from a local-scale model driven with in situ data.

In Chapter 6 a similar approach to Chapter 5 was followed, but in this experiment in situ gauged and satellite-derived discharge observations were assimilated into a local-scale hydrological model (OpenStreams wflow-hbv) of the Magdalena-Cauca River basin using an Ensemble Kalman filter. Based on the results obtained in Chapter 3, the MSWEP global precipitation product was selected to force the hydrological model to compare the impact of precipitation to data assimilation on model simulations. Four assimilation scenarios were analysed: (i) open-loop simulation without data assimilation (reference scenario), (ii) assimilation of in situ gauged discharge (IN SITU), (iii) assimilation of discharge derived from passive microwave sensors (SGR-GFDS) and (iv) assimilation of discharge derived from optical MODIS data (SGR-MODIS). This experimental set up aims to address research questions 6 and 7.

Results showed that assimilating satellite-based discharge observations improved discharge model simulations. The assimilation of SGR-MODIS discharge data increased model discharge performance more than when SGR-GFDS observations were assimilated at most locations (32% versus 21% increase in average KGE values). These findings proved the potential of assimilating satellite-derived discharge observations for improving estimations of hydrological models driven with global precipitation datasets for ungauged river basins.

7.3 Key messages

When integrating global datasets into hydrological models, the theoretical background is present, however, researchers are often hampered by a lack of knowledge about their practical implementation at the river basin scale. This PhD thesis aimed to contribute to this research gap.

In all the chapters of this thesis, it was shown that good quality meteorological datasets, particularly precipitation, are crucial for accurate model simulations. Meteorological input affects model simulations more than calibrating model parameters or assimilating additional observations. Several global meteorological datasets were inter-compared and evaluated against in situ measurements and results were promising for their future

application in hydrological modelling on other river basins around the world. Moreover, in spite of biases detected in some areas and periods, global meteorological datasets were useful to estimate water resources at the river basin scale, using either large- or local-scale hydrological models. Large-scale hydrological models were used in Chapters 2, 4 and 5. Local models were used in Chapters 3, 5 and 6. Local and global meteorological datasets were used in all Chapters.

Several techniques, including downscaling of global meteorological datasets, model parameters estimation and data assimilation using satellite-derived observations, were investigated in combination with large- and local-scale hydrological models to evaluate the applicability of global water resources datasets (based on satellite-derived observations, in situ data and models) for hydrological modelling at the river basin scale. The combined use of satellite-derived observations of various hydrological variables, such as soil moisture and evapotranspiration for model parameters calibration in Chapter 4, may lead to similar model performances as when using in situ datasets. Moreover, when combining observations of different variables, all possible sources of information should be used, including satellite, in situ and models datasets, such as was done in Chapter 5 with the assimilation of satellite-derived soil moisture and in situ discharge observations.

This PhD thesis has shown the great potential of using globally available datasets and models for hydrological modelling in ungauged river basins. It evaluated a suite of approaches that could be applied in other river basins worldwide, especially in those with limited in situ data available. These results could bring major advances in the field of hydrological modelling and water resources management in general.

7.4 Future perspectives

7.4.1 Additional satellite-derived observations

Scientific efforts are made to improve existent satellite-derived observations and to release new products. Future research studies may explore the use of other readily-available satellite-derived observations in hydrological models by implementing some of the methodologies described in this thesis. For example, the possibility to use other satellite-derived soil moisture observations, such as the Soil Moisture Ocean Salinity (SMOS; Kerr et al., 2001), evapotranspiration observations, such as the latest version of the GLEAM product (GLEAM_v3.1; Martens et al., 2017), or precipitation observations, such as the Global Precipitation Measurement mission product (GPM), should be explored. The integration into large- and local-scale hydrological models of additional combinations of satellite-derived observations of soil moisture, evapotranspiration and discharge (and other hydro-meteorological variables) could be also further analyzed. Alternative satellite-derived observations, such as snow water equivalent (Global Snow Monitoring for Climate Research, GlobSnow v2.0; Takala et al., 2011) and flooding extent (SENTINEL-1 Synthetic Aperture Radar, SAR), and their practical implementation to improve the accuracy of hydrological models at the river basin scale should be also investigated. Other possibility would be to explore the use of CubeSat for ultrahigh-resolution applications, such as the detection of vegetation health and condition, the crop water use, the delineation of surface

water changes and changes in lake and river extent, etc. (McCabe et al., 2017). In addition, most of recently developed global water resources datasets result from the combination of not only satellite observations, but also in situ data and models. Further investigation on disentangling the impact of satellite observations solely may have broader implications for their application in hydrological modelling.

7.4.2 Additional large- and local-scale hydrological models

Chapter 2 showed that large-scale hydrological models, such as WATERGAP3 and PCR-GLOBWB, are suitable for water resources estimations and flood detection. Moreover, Chapters 4 and 5 showed reasonable performance of PCR-GLOBWB when calibrating model parameters and assimilating satellite-derived observations. However, large-scale hydrological models are under a continuous process of improvement and several advances have been done in the last years. For example, a new version of PCR-GLOBWB (PCR-GLOBWB 2; Sutanudjaja et al., 2018) at 5 arc min spatial resolution was developed. Future studies should evaluate the impact on the new model simulations, at finer spatial scale and thus with an improved representation of the hydrological processes, when implementing some of the techniques described in this thesis.

Moreover, a large variety of local-scale hydrological models exist, with different model concept and structures. In Chapter 2, the NAM+MIKE BASIN model was used, whereas in Chapters 3, 5 and 6, the OpenStreams wflow model was applied. Similar methodologies to those presented in this research could be alternatively applied into other hydrological models to quantify and compare the impact of integrating satellite-derived observations on model performance.

7.4.3 Additional river basins

A selection of four river basins was made to test the applicability of global water resources datasets for hydrological modelling at the river basin scale, including the Brahmaputra basin in Bangladesh, the Magdalena-Cauca basin in Colombia, the Oum Er Rbia basin in Morocco and the Murrumbidgee basin in Australia. They show a diversity of hydro-meteorological characteristics, in situ data availability and water management conditions. For example, climate varies from tropical in the Magdalena-Cauca (Chapters 3 and 6) basin to semiarid in the Oum Er Rbia basin (Chapter 4). In situ data availability also differs between the basins, with a dense network of stations in the Murrumbidgee basin (Chapter 5) and a limited number of stations in the Brahmaputra basin (Chapter 2). However, this selection does not cover all different types of basins existent in the world and thus, this PhD research should be extended to other river basins, for example to basins located in temperate and polar climatic regions and/or with different levels of in situ data availability.

The extension of this PhD research to additional river basins, large- and local-scale hydrological models and satellite-derived observations would help to give a global structured assessment of the impact of each of these individual components on hydrological estimations which is currently lacking and would facilitate the results transferability.

7.4.4 In situ observations

Although satellite-derived observations have shown a high potential for hydrological modelling, in situ measurements are of great importance for hydrological studies. In all the chapters of this thesis a consistent comparison of the impact of in situ gauged and satellite-derived meteorological datasets was carried out, showing better model estimates when in situ information was used. Moreover, Chapter 4 showed that in situ discharge data were needed to calibrate model parameters and Chapter 5 and 6 showed that the assimilation of in situ discharge data resulted in the highest improvement in model estimates. Hence, it is still important to encourage initiatives, such as those done by the Global Runoff Data Centre (GRDC) to make available a global archive of in situ discharge observations. Moreover, in situ data of other variables different to discharge, such as soil moisture evapotranspiration and groundwater levels, are also needed for the validation of satellite-derived observations and field campaigns should be organised to gather as much local information as possible. Furthermore, promising alternatives to gather in situ data could be the crowd sourcing techniques (Lowry and Fienen, 2013) encouraging citizens to collect hydrometric measurements and make them accessible with mobile phones.

7.4.5 Incorporation of human influence

Human influence, e.g. reservoirs, diversions, irrigation, surface and groundwater extractions, etc. was not specifically considered for any of the river basins. In some of the river basins analyzed in this research, the water cycle is heavily influenced by human activities (e.g. the Magdalena-Cauca River basin in Chapters 3 and 6) which should not be ignored in their hydrological modelling. Various studies have analyzed the human influences on the hydrological cycle comparing several large-scale hydrological models and have emphasized the importance of taking into account impacts of human interventions for hydrological modelling (Haddeland et al., 2013; Wada et al., 2014). Therefore, hydrological models used in this research should be improved by incorporating human influence.

7.4.6 Opportunities for improving decision making and prediction

Despite the large benefits that could be derived by the integration of satellite-derived observations into large- and local-scale hydrological models at the river basin scale, these global datasets are underutilized by local end users and decision makers in water resources. Some possible barriers for the optimal implementation of global datasets for practical applications are their coarse spatial resolution, the lack of local end users' knowledge (handling data formats, using information on observational error, etc..), unawareness of the new technical possibilities, high computational requirements and financial limitations. An example of the current scientific initiatives in this respect is the REC project, "Crop irrigation management by multi-sensor remote sensing approach", funded by the European Commission Horizon 2020 Programme for Research and Innovation (H2020). This project aims to optimize on-farm irrigation management by adjusting irrigation to crop water requirements along the growing season developing a tool at the field scale which combines land surface models and satellite-derived soil moisture observations and which could be applied in real agricultural practices by decision makers and end users.

Moreover, a more user-oriented approach to assess how satellite-derived observations can contribute to more efficient and reliable decision making is needed (Linés et al., 2017; Kaune et al., 2018). For example, Kaune et al. (2017) propose a tool to evaluate the degree of availability of hydrological information, including in situ and satellite-derived datasets, in a basin to optimize water allocation decisions in irrigation districts. To this end, identifying relevant policies and operational processes and developing user-oriented metrics to test the socio-economic value of global water resources estimates are also needed.

In addition, global water resources datasets could be integrated into hydrological models aiming to improve predictions for hydrological extremes, such as floods and droughts. That integration would allow increasing our understanding of the relationships between climate variability and the occurrence (intensity and frequency) of extreme events, and would help quantifying the impact of changing precipitation, evapotranspiration and other atmospheric variables on the water dynamics. For example, data assimilation could be used to include satellite-derived observations, such as MODIS or GFDS discharge data used in Chapter 6, into flood forecasting systems to improve the forecasts by reducing errors in initial conditions and/or parameters (Moradkhani et al., 2012; Rakovec et al., 2012; Vrugt et al., 2013; Li et al., 2015).

Appendices

Appendix A

Table A. 1 ***r*** values between basin average MSWEP precipitation and EVI considering lag times between 1 week and 3 months at 25 km spatial resolution.

Lag time [weeks]	1	2	3	4	5	6	7	8	9	10	11	12
***r* [-]**	-0.006	0.176	0.253	0.580	0.534	0.423	0.362	0.101	0.098	0.088	0.171	0.322

Table A. 2 Model performance for calibration and validation periods (2001–2004 and 2005–2012) at the 22 discharge stations.

Station name	***KGE* [-]**		***RMSE* [m³ s⁻¹]**		***r* [-]**	
	Cal	**Val**	**Cal**	**Val**	**Cal**	**Val**
La Nueva	0.653	0.660	5.260	10.441	0.759	0.729
Mateguadua	0.484	0.478	15.885	10.529	0.736	0.706
Puente Negro	0.583	0.599	25.458	19.239	0.750	0.759
El Cóndor	0.463	0.471	52.978	39.569	0.643	0.650
Puente Anori	0.653	0.642	32.873	41.716	0.684	0.659
San Gil	0.475	0.476	38.649	48.160	0.734	0.741
Cartago	0.593	0.584	50.519	68.595	0.794	0.771
Puerto Araujo-Automat.	0.434	0.430	166.823	156.057	0.698	0.695
La Ceiba	0.516	0.506	83.332	133.874	0.774	0.731
Piedras Cobre-Automat.	0.489	0.479	200.607	205.741	0.594	0.595
Remolino	0.536	0.536	150.782	153.458	0.551	0.549
El Jordán	0.691	0.690	50.339	70.445	0.722	0.723
La Esperanza	0.762	0.751	250.100	238.811	0.921	0.902
La Virginia	0.483	0.487	337.873	307.204	0.859	0.854
Puente Iglesias	0.594	0.583	420.165	406.124	0.853	0.820
Apavi	0.519	0.515	610.843	588.930	0.887	0.884
Barbosa	0.635	0.620	1273.132	1352.012	0.855	0.845
Puente Carretera	0.435	0.431	19.677	18.681	0.633	0.631
Puente Santander-Automat.	0.469	0.478	280.334	275.236	0.555	0.565
Puerto Salgar-Automat.	0.765	0.755	500.432	552.372	0.800	0.811
Puerto Berrío-Automat.	0.855	0.868	520.109	585.250	0.896	0.881
Regidor	0.481	0.480	1280.011	1329.066	0.858	0.856

Appendix B

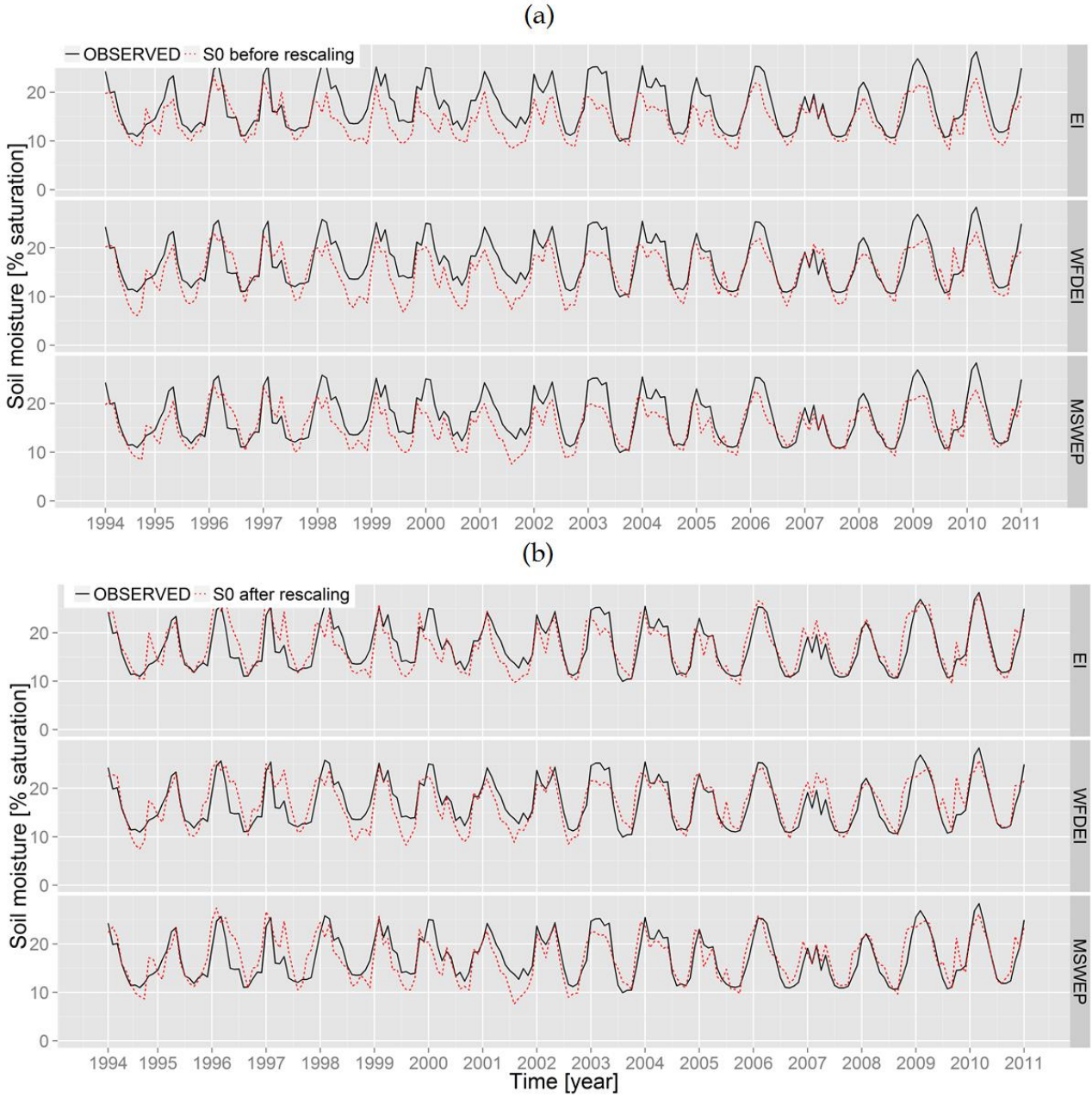


Figure B. 1 Comparisons between monthly ESA CCI soil moisture (black) and estimated soil moisture (red) time series over the Oum Er Rbia basin for the validation time period. Rows show the three global precipitation products. The red dashed lines represent soil moisture estimates from calibration scenario S0 (reference scenario) a) before the mean-standard deviation matching is applied and b) after the mean-standard deviation matching is applied.

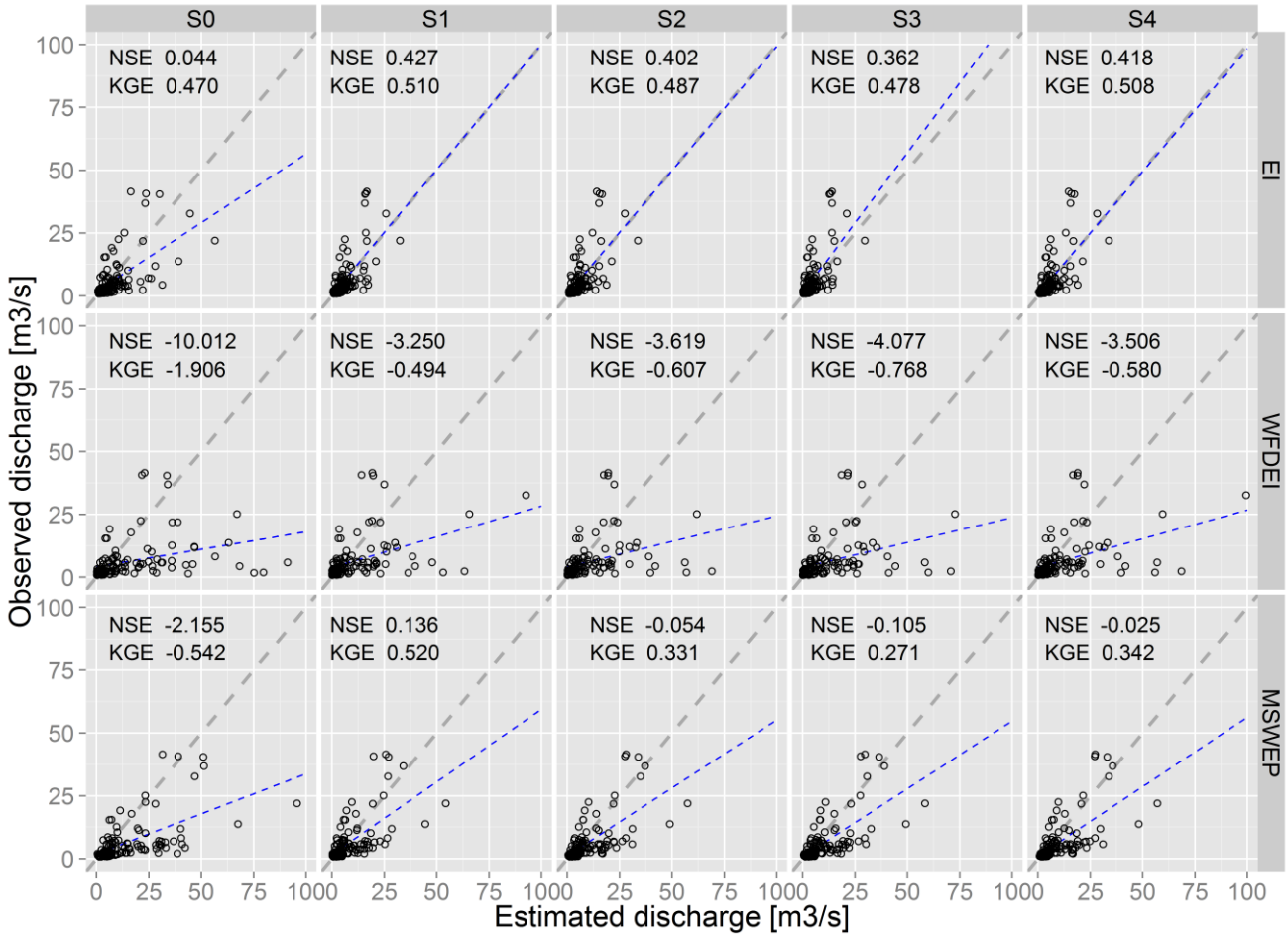


Figure B. 2 Scatterplots of monthly estimated discharge (x-axis) and observed discharge (y-axis) at Ait Ouchene. Rows show the three global precipitation products and columns show the five calibration scenarios.

Appendix C

Table C. 1 Evaluation results of the catchment daily means of soil moisture estimates (the exact values of *Figure 5.6* are included in this table for clarification).

	Local			**Global**		
	RMSE **[$m^3\ m^{-3}$]**	***MAE*** **[$m^3\ m^{-3}$]**	***r*** **[-]**	***RMSE*** **[$m^3\ m^{-3}$]**	***MAE*** **[$m^3\ m^{-3}$]**	***r*** **[-]**
OSWS	0.07571	0.05716	0.79865	0.07854	0.06082	0.77603
GLOBWB_OL	0.09735	0.07520	0.43071	0.10094	0.07662	0.40715
GLOBWB_Q	0.09738	0.07523	0.43041	0.10095	0.07664	0.40652
GLOBWB_SM	0.09085	0.06699	0.53452	0.09372	0.06811	0.50230
GLOBWB_SM+Q	0.09032	0.06609	0.52085	0.09302	0.06845	0.49623

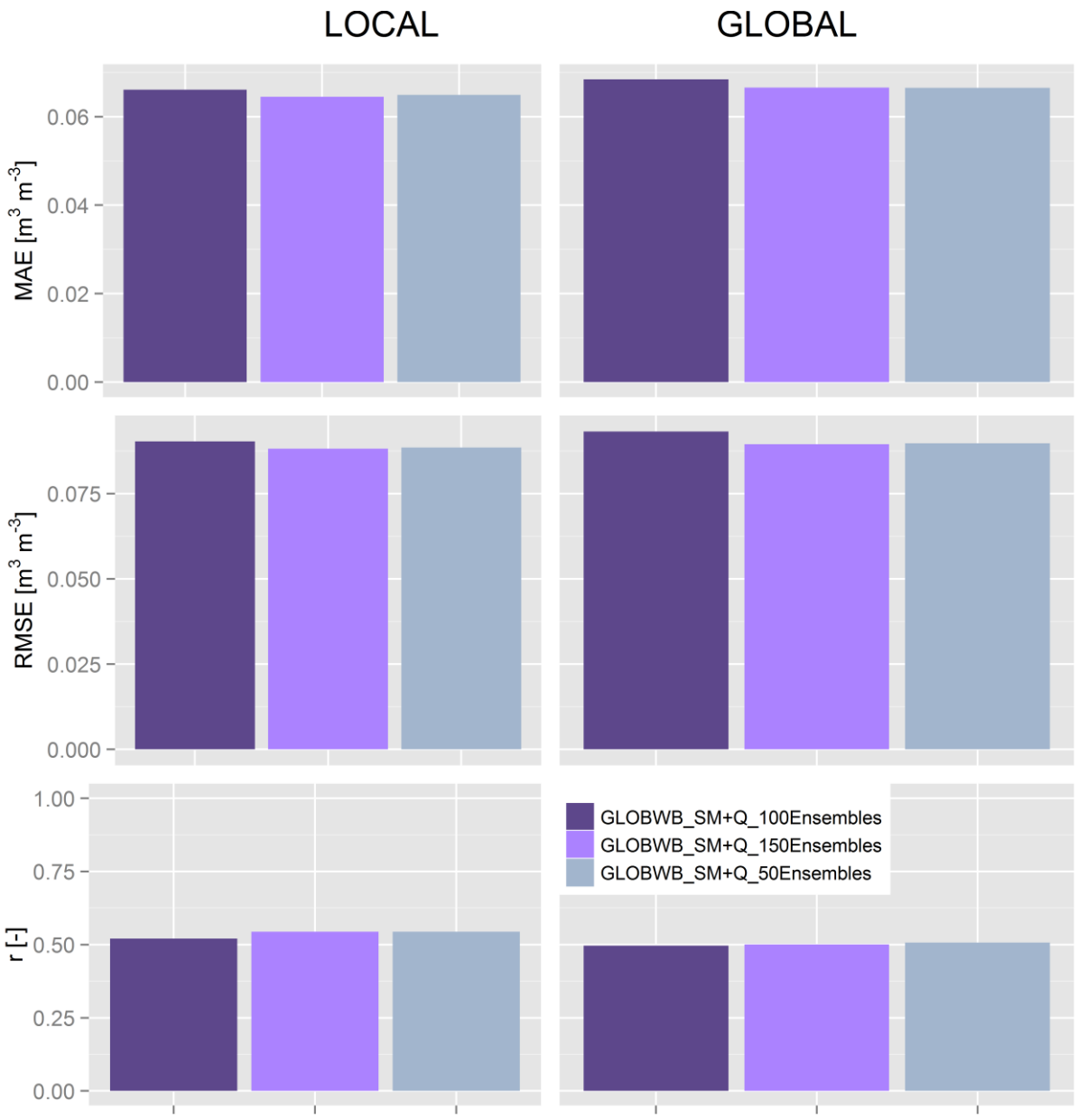


Figure C. 1 Comparison of evaluation results of the catchment daily means of soil moisture in the Murrumbidgee River basin. Assimilation scenario GLOBWB_SM_Q has been reproduced using 50, 100 and 150 ensemble members.

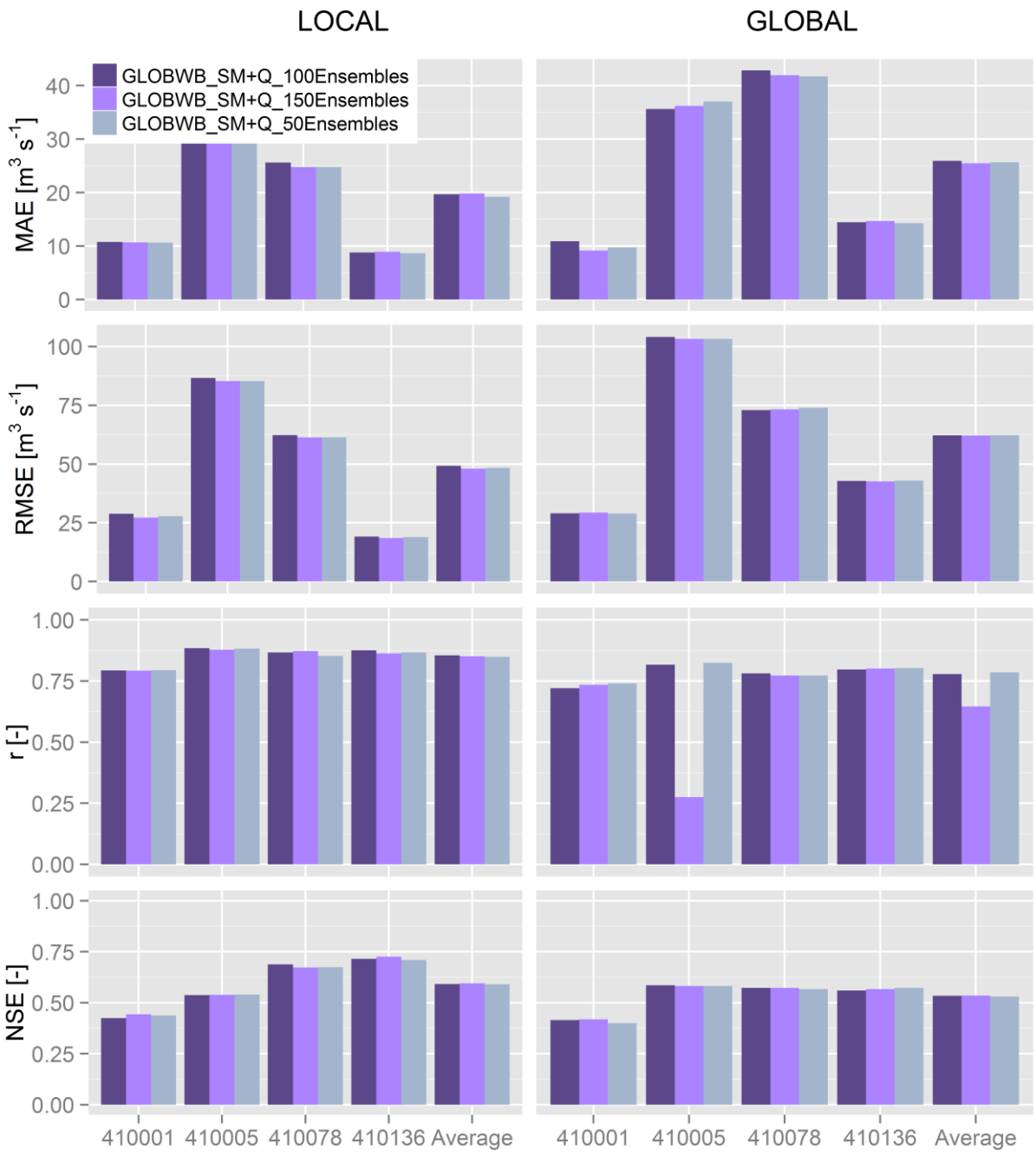


Figure C. 2 Comparison of evaluation results for streamflow estimates at 410001, 410005, 410078 and 410136 locations in the Murrumbidgee River. Assimilation scenario GLOBWB_SM_Q has been reproduced using 50, 100 and 150 ensemble members.

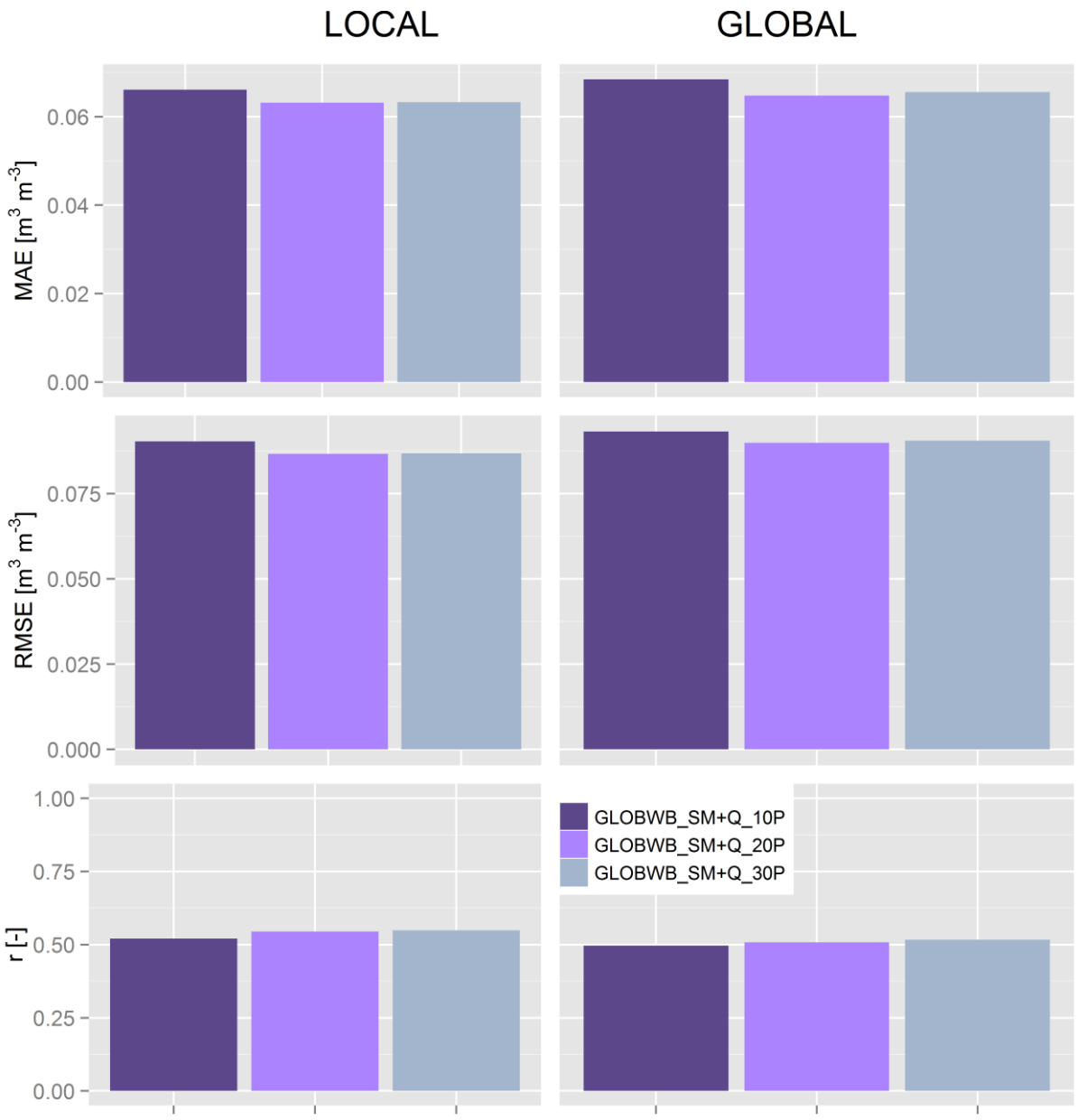


Figure C. 3 Comparison of evaluation results of the catchment daily means of soil moisture in the Murrumbidgee River basin. Assimilation scenario GLOBWB_SM_Q has been reproduced perturbing the precipitation with additive Gaussian white noise with standard deviation of 10%, 20% and 30% of the nominal value.

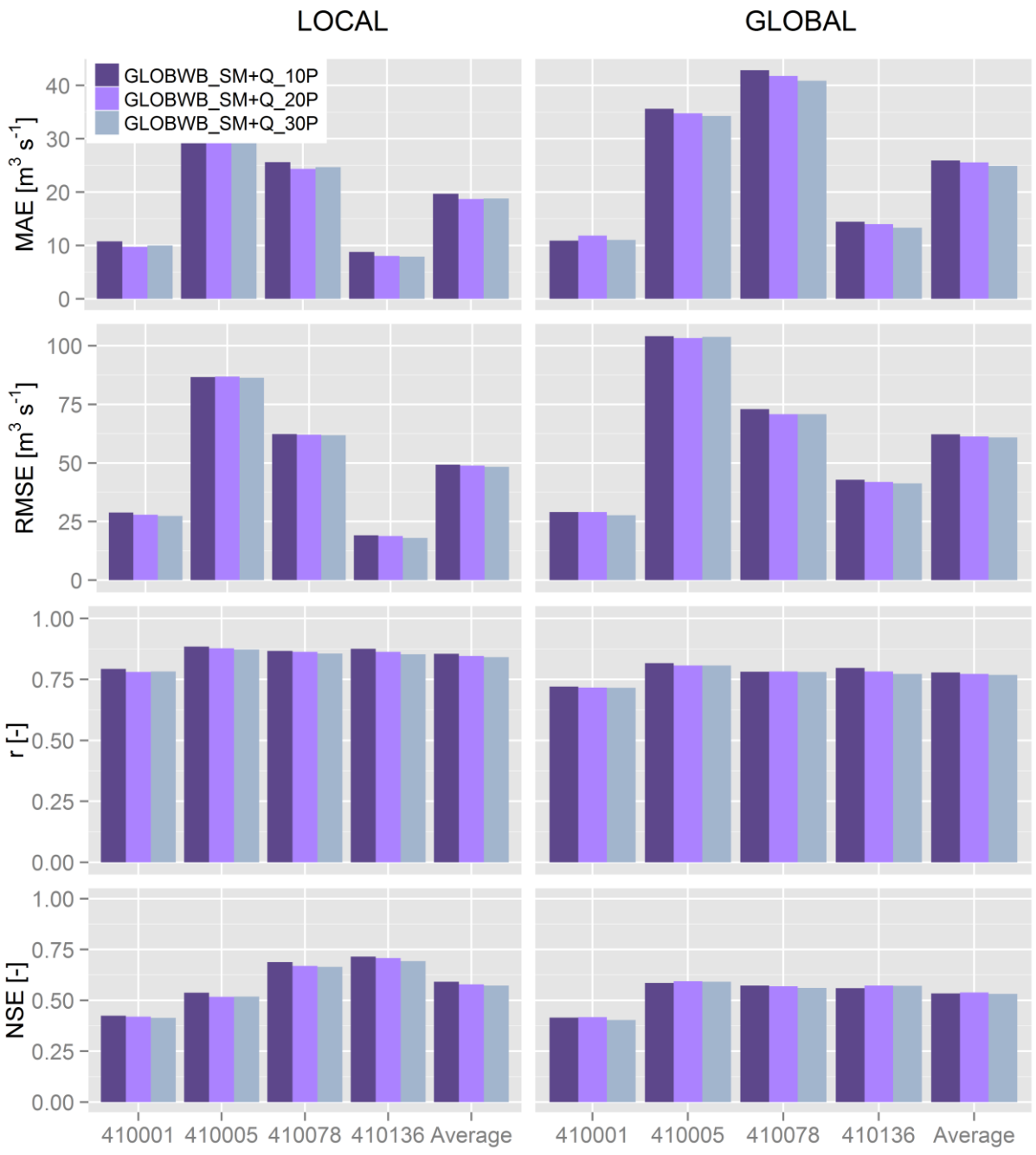


Figure C. 4 Comparison of evaluation results for streamflow estimates at 410001, 410005, 410078 and 410136 locations in the Murrumbidgee River. Assimilation scenario GLOBWB_SM_Q has been reproduced perturbing the precipitation with additive Gaussian white noise with standard deviation of 10%, 20% and 30% of the nominal value.

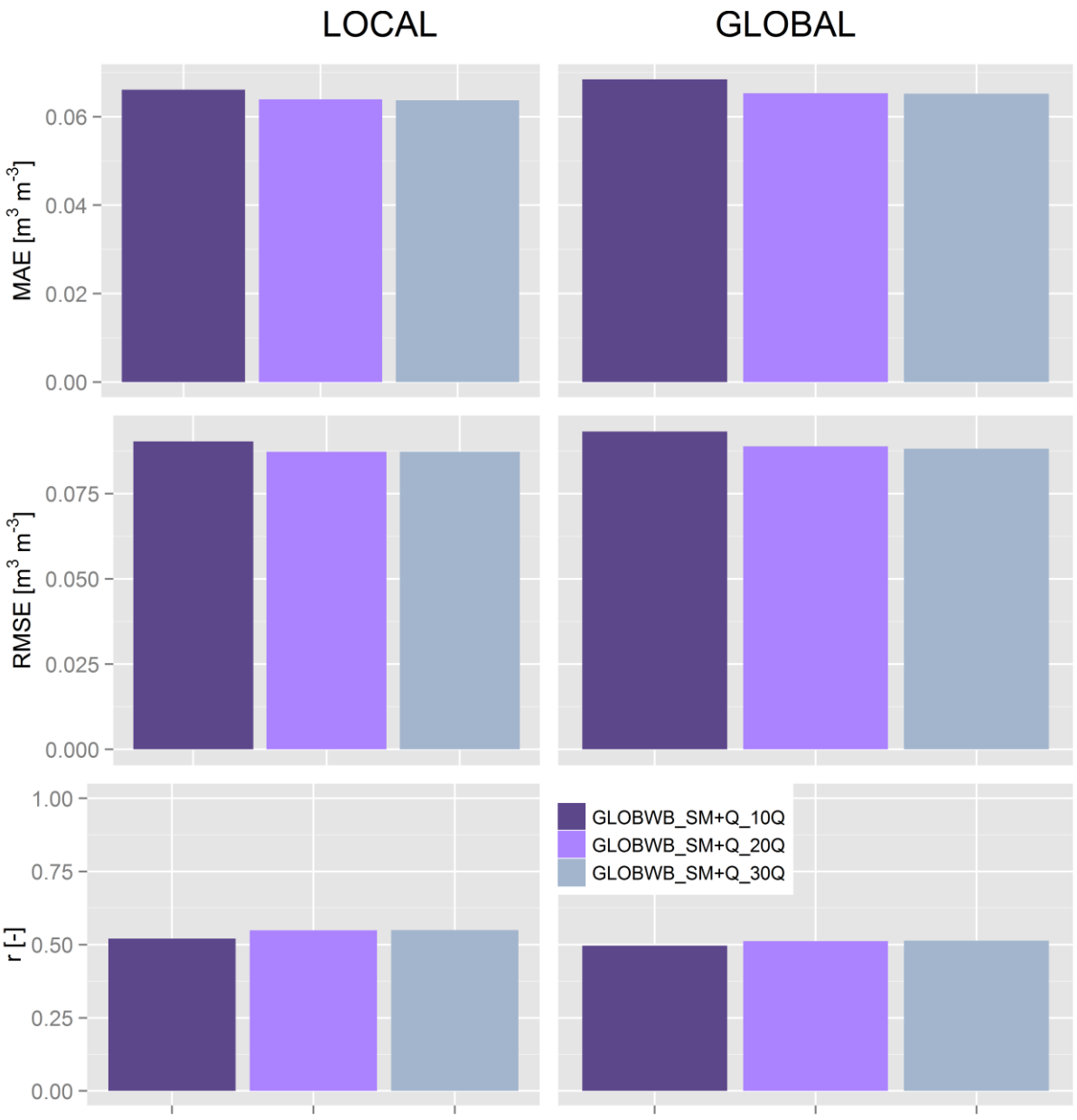


Figure C. 5 Comparison of evaluation results of the catchment daily means of soil moisture in the Murrumbidgee River basin. Assimilation scenario GLOBWB_SM_Q has been reproduced setting a standard error for the discharge observations of 10%, 20% and 30% of the discharge.

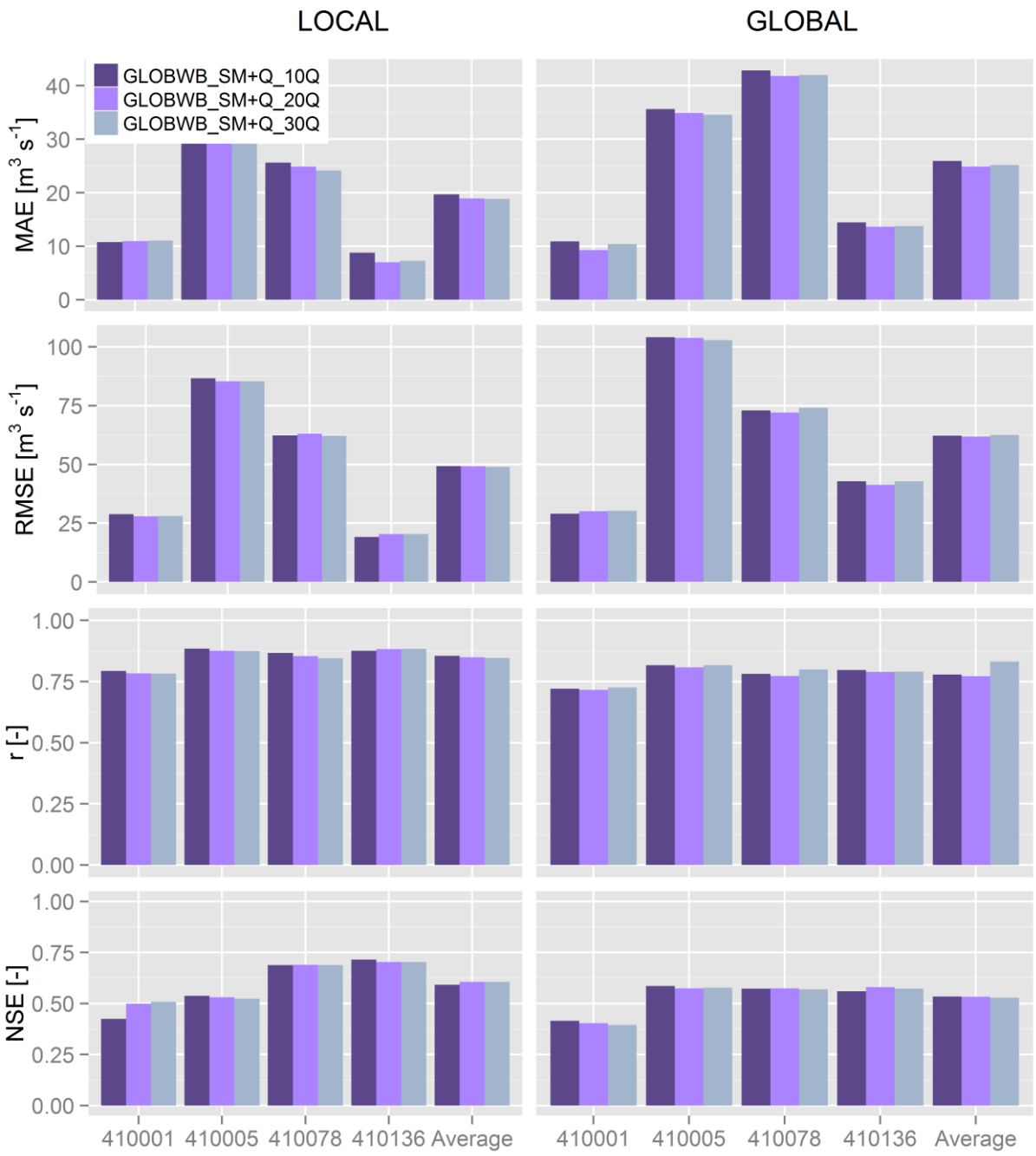


Figure C. 6 Comparison of evaluation results for streamflow estimates at 410001, 410005, 410078 and 410136 locations in the Murrumbidgee River. Assimilation scenario GLOBWB_SM_Q has been reproduced setting a standard error for the discharge observations of 10%, 20% and 30% of the discharge.

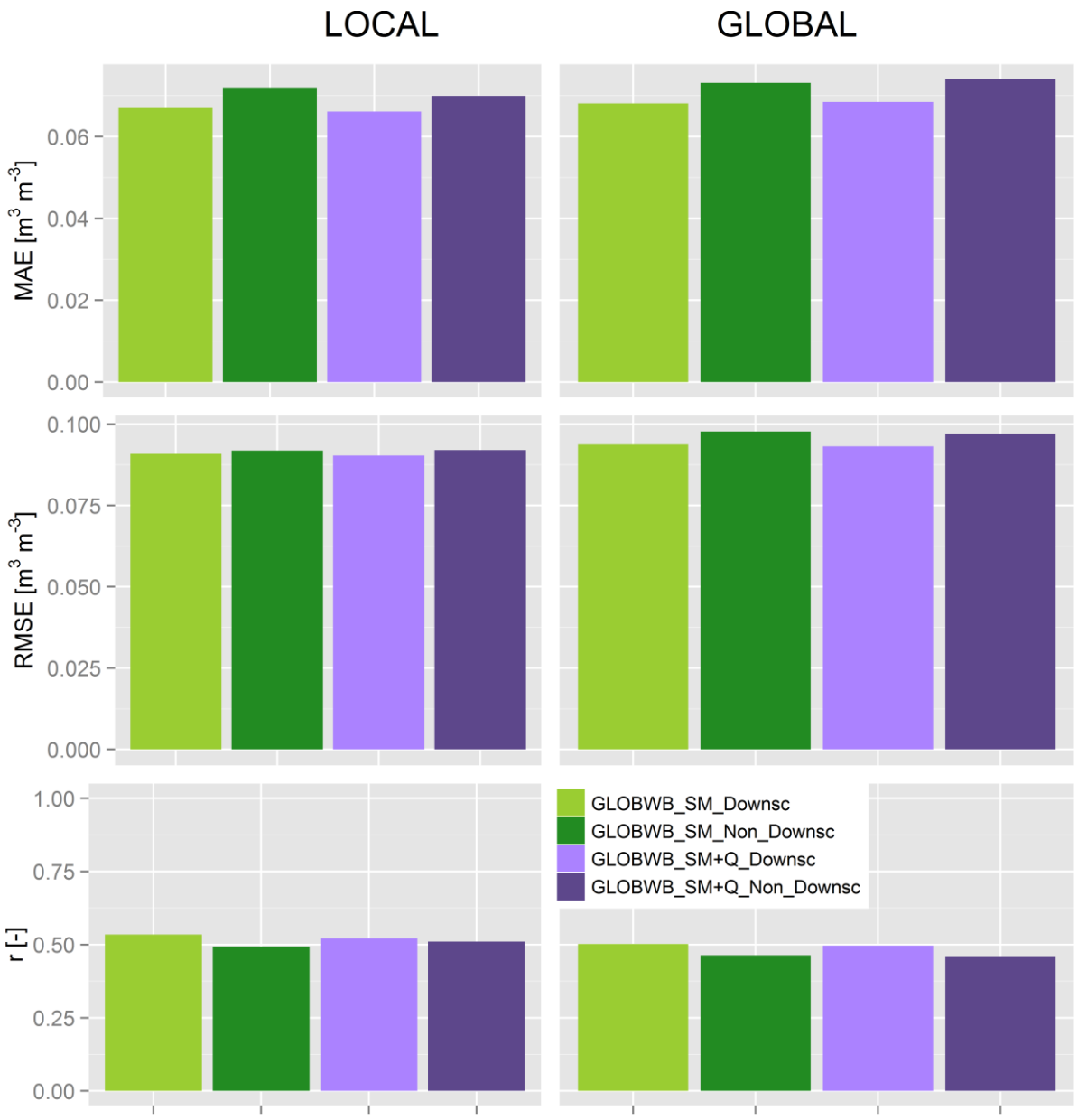


Figure C. 7 Comparison of evaluation results of the catchment daily means of soil moisture in the Murrumbidgee River basin. Assimilation scenarios GLOBWB_SM (green) and GLOBWB_SM+Q (purple) are shown when downscaled and non-downscaled soil moisture observations are assimilated.

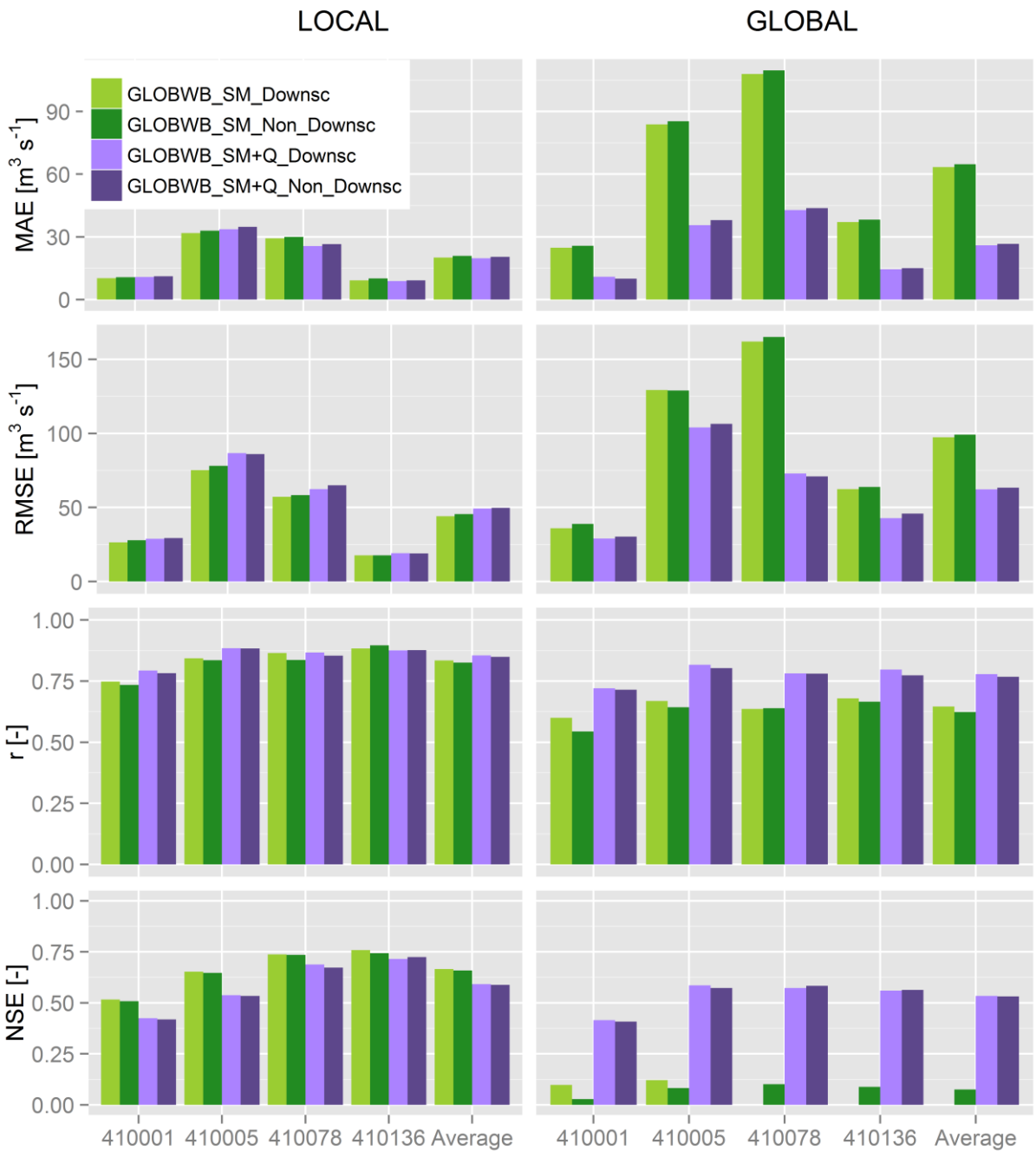


Figure C. 8 Comparison of evaluation results for streamflow estimates at 410001, 410005, 410078 and 410136 locations in the Murrumbidgee River. Assimilation scenarios GLOBWB_SM (green) and GLOBWB_SM+Q (purple) are shown when downscaled and non-downscaled soil moisture observations are assimilated.

References

Abaza, M., Anctil, F., Fortin, V., and Turcotte, R. (2014). Sequential streamflow assimilation for short-term hydrological ensemble forecasting. *Journal of Hydrology*, 519, 2692-2706. https://doi.org/10.1016/j.jhydrol.2014.08.038.

Adam, J. C. and Lettenmaier, D. P. (2003). Adjustment of global gridded precipitation for systematic bias. *Journal of Geophysical Research-Atmospheres*, 108, 4257. http://dx.doi.org/10.1029/2002JD002499.

Adler, R. F., Kidd, C., Petty, G., Morissey, M., and Goodman, H. M. (2001). Intercomparison of global precipitation products: The third precipitation intercomparison project (PIP–3). *Bulletin of the American Meteorological Society*, 82(7), 1377-1396. https://doi.org/10.1175/1520-0477(2001)082<1377:IOGPPT>2.3.CO;2

Ahmed, A. U., and Mirza, M. M. Q. (1998). Review of causes and dimensions of floods with particular reference to flood'98: national perspectives. *Perspectives on flood*, 67-84.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S. (2003). Development and testing of the WaterGAP 2 global model of water use and availability. *Hydrological Sciences Journal*, 48, 317–337. https://doi.org/10.1623/hysj.48.3.317.45290

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). Crop evapotranspiration – Guidelines for computing crop water requirements, FAO Irrigation and Drainage Paper 56, *FAO, Rome*, 300 pp.

Alvarez-Garreton, C., Ryu, D., Western, A. W., Su, C.-H., Crow, W. T., Robertson, D. E., and Leahy, C. (2015). Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: comparison between lumped and semi-distributed schemes. *Hydrology and Earth System Sciences*, 19, 1659–1676, doi:10.5194/hess-19-1659- 2015.

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., and Lavabre, J. Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. (2001). *Journal of Hydrology*, 250(1), 206-223. https://doi.org/10.1016/S0022-1694(01)00437-1.

Andréassian, V., Perrin, C., and Michel, C. (2004). Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models. *Journal of Hydrology*, 286, 19–35. https://doi.org/10.1016/j.jhydrol.2003.09.030.

Andreadis, K. M. and Lettenmaier, D. P. (2006). Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, 29, 872–886. https://doi.org/10.1016/j.advwatres.2005.08.004.

Arino, O., Ramos, J., Kalogirou, V., Defourny, P. and Achard, F. (2010). GlobCover 2009. *Proceedings of the living planet Symposium*, SP-686.

Arnaud, P., Bouvier, C., Cisner, L., Dominguez, R. (2002). Influence of rainfall spatial variability on flood prediction. *Journal of Hydrology*, 260, 216–230, 2002. 2176, 2177. https://doi.org/10.1016/S0022-1694(01)00611-4.

Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., and Javelle, P. Sensitivity of hydrological models to uncertainty in rainfall input. (2011). *Hydrological Sciences Journal*, 56(3), 397-410. https://doi.org/10.1080/02626667.2011.563742

Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P. (2015). PERSIANN-CDR. Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96(1), 69-83. https://doi.org/10.1175/BAMS-D-13-00068.1.

Atkinson, P. M. (2013). Downscaling in remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 22, 106–114. https://doi.org/10.1016/j.jag.2012.04.012.

Aubert, D., Loumagne, C., and Oudin, L. (2003). Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall runoff model, *Journal of Hydrology*, 280, 145–161. https://doi.org/10.1016/S0022-1694(03)00229-4.

Bajracharya, S. R., Palash, W., Shrestha, M. S., Khadgi, V. R., Duo, C., Das, P. J., and Dorji, C. (2015). Systematic evaluation of satellite-based rainfall products over the Brahmaputra basin for hydrological applications. *Advances in Meteorology*, http://dx.doi.org/10.1155/2015/398687.

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K. (2009). A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the integrated forecast system. *Journal of Hydrometeorology*, 10, 623–643. https://doi.org/10.1175/2008JHM1068.1.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F. (2015). ERAInterim/Land: a global land surface reanalysis data set. *Hydrology and Earth System Sciences*, 19, 389–407. https://doi.org/10.5194/hess-19-389-2015.

Bardossy, A. and Das, T. (2008). Influence of rainfall observation network on model calibration and application. *Hydrology and Earth System Sciences*, 12, 77–89, doi:10.5194/hess-12-77-2008.

Barrett, D. J. and Renzullo, L. J. (2009). On the efficacy of combining thermal and microwave satellite data as observational constraints for root-zone soil moisture estimation. *Journal of Hydrometeorology*, 10, 1109–1127. https://doi.org/10.1175/2009JHM1043.1.

Bastiaanssen, W., Menenti, M., Feddes, R., and Holtslag, A. (1998). A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. *Journal of Hydrology*, 212, 198–212. https://doi.org/10.1016/S0022-1694(98)00253-4.

Beck, H. E., de Jeu, R. A., Schellekens, J., Van Dijk, A. I., and Bruijnzeel, L. A. (2009). Improving curve number based storm runoff estimates using soil moisture proxies. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2, 250–259. doi: 10.1109/JSTARS.2009.2031227.

Beck, H. E., Van Dijk, A. I., de Roo, A., Miralles, D. G., McVicar, T. R., Schellekens, J., and Bruijnzeel, L. A. (2016). Global-scale regionalization of hydrologic model parameters. *Water Resources Research*, 52, 3599–3622. https://doi.org/10.1002/2015WR018247.

Beck, H. E., Van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A. (2017a). MSWEP: 3-hourly 0.25◦ global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21, 589-615, https://doi.org/10.5194/hess-21-589-2017.

Beck, H. E., Van Dijk, A. I., de Roo, A., Dutra, E., Fink, G., Orth, R., and Schellekens, J. (2017b). Global evaluation of runoff from 10 state-of-the-art hydrological models. *Hydrology and Earth System Sciences*, 21(6), 2881. https://doi.org/10.5194/hess-21-2881-2017.

Bell, V. A., and Moore, R. J. (2000). The sensitivity of catchment runoff models to rainfall data at different spatial scales. *Hydrology and Earth System Sciences Discussions*, European Geosciences Union, 2000, 4 (4), pp.653-667.

Berne, A., Delrieu, G., Creutin, J. D., and Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, 299(3), 166-179. https://doi.org/10.1016/j.jhydrol.2004.08.002.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. *Geoscientific Model Development*, 4, 677–699, doi:10.5194/gmd-4-677-2011.

Beven, K., and Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, 6(3), 279-298. http://dx.doi.org/10.1002/hyp.3360060305.

Bezak, N., Brilly, M., and Šraj, M. (2014). Comparison between the peaks-over-threshold method and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, 59(5), 959-977. https://doi.org/10.1080/02626667.2013.831174.

Bierkens, M. F. P. (2015). Global hydrology 2015: state, trends, and directions. *Water Resources Research*, 51, 4923–4947. https://doi.org/10.1002/2015WR017173.

Bohn, T. J., Sonessa, M. Y., and Lettenmaier, D. P. (2010). Seasonal hydrologic forecasting: do multimodel ensemble averages always yield improvements in forecast skill? *Journal of Hydrometeorology*, 11, 1358– 1372. https://doi.org/10.1175/2010JHM1267.1.

Bohnenstengel, S. I., Schlünzen, K. H., and Beyrich, F. (2011). Representativity of in situ precipitation measurements–A case study for the LITFASS area in North-Eastern Germany. *Journal of Hydrology*, 400(3), 387-395. https://doi.org/10.1016/j.jhydrol.2011.01.052.

BoM – Bureau of Meteorology: Australian Government, Water Data Online, available at: http://www.bom.gov.au/waterdata/, last access: 25 September 2015.

Bouchaou, L., Michelot, J., Qurtobi, M., Zine, N., Gaye, C., Aggarwal, P., Marah, H., Zerouali, A., Taleb, H., and Vengosh, A. (2009). Origin and residence time of groundwater in the Tadla basin (Morocco) using multiple isotopic and geochemical tools. *Journal of Hydrology*, 379, 323–338. https://doi.org/10.1016/j.jhydrol.2009.10.019.

Bradley, A. V., Gerard, F. F., Barbier, N., Weedon, G. P., Anderson, L. O., Huntingford, C., Aragao, L. E. O. C., Zelazowski, P., and Arai, E. (2011). Relationships between phenology, radiation and precipitation in the Amazon region. *Global Change Biology*, 17(6), 2245-2260. doi: 10.1111/j.1365-2486.2011.02405.x.

Breusch, T. S., and Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, 1287-1294. doi: 10.2307/1911963.

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., and Bittelli, M. (2011). Soil moisture estimation through ASCAT and AMSR-E sensors: an intercomparison and validation study across Europe. *Remote Sensing of Environment*, 115, 3390– 3408. https://doi.org/10.1016/j.rse.2011.08.003.

Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., and Hahn, S. (2012). Assimilation of surface-and root-zone ASCAT soil moisture products into rainfall–runoff modelling. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 2542–2555. doi: 10.1109/TGRS.2011.2177468.

Burgers, G., van Leeuwen, P. J., and Evensen, G. (1998). Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126, 1719– 1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Camacho, L. A., Rodríguez, E., and Pinilla, G. (2008). Modelación dinámica integrada de cantidad y calidad del agua del Canal del Dique y su sistema lagunar, Colombia. *In XXIII Latinamerican Congress on Hydraulic (IARH).*

Campo, L., Caparrini, F., and Castelli, F. (2006). Use of multi-platform, multi-temporal remote-sensing data for calibration of a distributed hydrological model: an application in the Arno basin, Italy. *Hydrological Processes*, 20, 2693–2712.

Cannon, A. J. (2011). Quantile regression neural networks: implementation in R and application to precipitation downscaling. *Computers and Geosciences*, 37, 1277–1284. https://doi.org/10.1016/j.cageo.2010.07.005.

Ceccherini, G., Ameztoy, I., Hernández, C. P. R., and Moreno, C. C. (2015). High-resolution precipitation datasets in South America and West Africa based on satellite-derived rainfall, enhanced vegetation index and digital elevation model. *Remote Sensing*, 7(5), 6454-6488. doi:10.3390/rs70506454.

Chappell, A., Renzullo, L. J., Raupach, T. H., and Haylock, M. (2013). Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia. *Journal of Hydrology*, 493, 105–114. https://doi.org/10.1016/j.jhydrol.2013.04.024

Cheema, M. J. M., and Bastiaanssen, W. G. (2012). Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin. *International Journal of Remote Sensing*, 33(8), 2603-2627. https://doi.org/10.1080/01431161.2011.617397.

Chen, F., Crow, W. T., Starks, P. J., and Moriasi, D. N. (2011). Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. *Advances in Water Resources*, 34, 526–536. https://doi.org/10.1016/j.advwatres.2011.01.011.

Chen, F., Liu, Y., Liu, Q., and Li, X. (2014). Spatial downscaling of TRMM 3B43 precipitation considering spatial heterogeneity. *International Journal of Remote Sensing*, 35(9), 3074-3093. https://doi.org/10.1080/01431161.2014.902550.

Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Jochen, S. M., and Uddstrom, M. J. (2008). Hydrological data assimilation with the ensemble Kalman filter: use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources*, 31, 1309–1324. https://doi.org/10.1016/j.advwatres.2008.06.005.

Collischonn, B., Collischonn, W., and Tucci, C. E. M. (2008). Daily hydrological modelling in the Amazon basin using TRMM rainfall estimates. *Journal of Hydrology*, 360(1), 207-216. https://doi.org/10.1016/j.jhydrol.2008.07.032.

Crow, W. T., Bindlish, R., and Jackson, T. J. (2005). The added value of spaceborne passive microwave soil moisture retrievals for forecasting rainfall–runoff partitioning. *Geophysical Research Letters*, 32, L18401, doi:10.1029/2005GL023543.

De Jeu, R. A. M., Wagner, W., Holmes, T. R. H., Dolman, A. J., Van De Giesen, N. C., and Friesen, J. (2008). Global soil moisture patterns observed by space borne microwave radiometers and scatterometer. *Surveys in Geophysics*, 29, 399–420. https://doi.org/10.1007/s10712-008-9044-0.

De Goncalves, L., Shuttleworth, W. J., Nijssen, B., Burke, E. J., Marengo, J. A., Chou, S. C., ... and Toll, D. L. (2006). Evaluation of model-derived and remotely sensed precipitation products for continental South America. *Journal of Geophysical Research-Atmospheres*, 111(D16). doi: 10.1029/2005JD006276.

De Groeve, T. (2010). Flood monitoring and mapping using passive microwave remote sensing in Namibia. *Geomatics, Natural Hazards and Risk,* 1(1), 19–35. https://doi.org/10.1080/19475701003648085.

De Groeve, T., G. R. Brakenridge, and S. Paris. (2015). Global flood detection system: Data product specifications (version 2015), Tech. Rep. JRC97421, 16 pp., Eur. Comm. Joint Res. Cent., Ispra, Italy, 2015.

Decharme, B., Alkama, R., Douville, H., Becker, M., and Cazenave, A. (2010). Global evaluation of the ISBA-TRIP continental hydrological system. Part II: Uncertainties in river routing simulation related to flow velocity and groundwater storage. *Journal of Hydrometeorology*, 11, 601–617, https://doi.org/10.1175/2010JHM1212.1.

Decharme, B., Martin, E., and Faroux, S. (2013). Reconciling soil thermal and hydrological lower boundary conditions in land surface models. *Journal of Geophysical Research-Atmospheres*, 118, 7819–7834. doi: 10.1002/jgrd.50631.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. doi: 10.1002/qj.828.

Devia, G. K., Ganasri, B. P., and Dwarakish, G. S. (2015). A review on hydrological models. *Aquatic Procedia*, 4, 1001-1007. https://doi.org/10.1016/j.aqpro.2015.02.126.

Dewan, T. H. (2015). Societal impacts and vulnerability to floods in Bangladesh and Nepal. *Weather and Climate Extremes*, 7, 36-42. https://doi.org/10.1016/j.wace.2014.11.001.

DHI Water and Environment, (2015). MIKE BASIN – Short Introduction and Tutorial. Available at: https://www.tu-braunschweig.de/Medien-DB/geooekologie/mike-11-short-introduction-tutorial.pdf.

Di Baldassarre, G., and Montanari, A. (2009). Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences*, 13(6), 913. www.hydrol-earth-syst-sci.net/13/913/2009/.

Digital Soil Map of the World. (2007). Food and Agriculture Organization. Available at: http://www.fao.org/geonetwork/srv/en/metadata.show?id=14116.

Dinku, T., Connor, S. J., and Ceccato, P. (2010). Comparison of CMORPH and TRMM-3B42 over mountainous regions of Africa and South America. *Satellite Rainfall Applications for Surface Hydrology* (pp. 193-204). Springer Netherlands. https://doi.org/10.1007/978-90-481-2915-7_11.

Döll, P., Fiedler, K., and Zhang, J. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrology and Earth System Sciences*, 13, 2413–2432. www.hydrol-earth-syst-sci.net/13/2413/2009/.

d'Orgeval, T., Polcher, J., and de Rosnay, P. (2008). Sensitivity of the West African hydrological cycle in ORCHIDEE to in- filtration processes. *Hydrology and Earth System Sciences*, 12, 1387–1401. doi:10.5194/hess-12-1387-2008.

Dorigo, W., Gruber, A., De Jeu, R., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R., and Kidd R. (2015). Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment*, 162, 380–395. https://doi.org/10.1016/j.rse.2014.07.023.

Draper, C. S., Walker, J. P., Steinle, P. J., De Jeu, R. A. M., and Holmes, T. R. H. (2009). An evaluation of AMSR–E derived soil moisture over Australia. *Remote Sensing of Environment,* 113, 703–710, https://doi.org/10.1016/j.rse.2008.11.011.

Draper, C., Mahfouf, J. F., Calvet, J. C., Martin, E., and Wagner, W. (2011). Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France. *Hydrology and Earth System Sciences*, 15(12), 3829-3841. https://doi.org/10.5194/hess-15-3829-2011.

Duan, Z., and Bastiaanssen, W. G. M. (2013). First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure. *Remote Sensing of Environment,* 131, 1-13. https://doi.org/10.1016/j.rse.2012.12.002Get rights and content.

Durbin, J., and Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, 38(1/2), 159-177. doi: 10.2307/2332325.

Ebtehaj, A. M. and Foufoula-Georgiou, E. (2013). On variational downscaling, fusion, and assimilation of hydrometeorological states: A unified framework via regularization, *Water Resources Research*, 49, 5944–5963, doi:10.1002/wrcr.20424.

Emmanuel, I., Andrieu, H., Leblois, E., and Flahaut, B. (2012). Temporal and spatial variability of rainfall at the urban hydrological scale. *Journal of Hydrology*, 430, 162-172. https://doi.org/10.1016/j.jhydrol.2012.02.013.

Escobar C. R., Restrepo J.D., Brakenridge G.R., Kettner A.J. (2017). Satellite-Based Estimation of Water Discharge and Runoff in the Magdalena River, Northern Andes of Colombia. *Remote Sensing of Hydrological Extremes*. *Springer Remote*

Sensing/Photogrammetry. Springer, Cham. https://doi.org/10.1007/978-3-319-43744-6_1.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research-Oceans*, 99(C5), 10143-10162.

Evensen, G. The ensemble Kalman filter: Theoretical formulation and practical implementation. (2003). *Ocean dynamics*, 53(4), 343-367. https://doi.org/10.1007/s10236-003-0036-9.

Ezzine, H., Bouziane, A., Ouazar, D., and Driss, M. (2017). Downscaling of open coarse precipitation data through spatial and statistical analysis, integrating NDVI, NDWI, elevation and distance from Sea. *Advances in Meteorology*. https://doi.org/10.1155/2017/8124962.

Fang, J., Du, J., Xu, W., Shi, P., Li, M., and Ming, X. (2013). Spatial downscaling of TRMM precipitation data based on the orographical effect and meteorological conditions in a mountainous area. *Advances in Water Resources*, 61, 42-50. https://doi.org/10.1016/j.advwatres.2013.08.011.

Fang, Z., Bogena, H., Kollet, S., Koch, J., and Vereecken, H. (2015). Spatio-temporal validation of long-term 3D hydrological simulations of a forested catchment using empirical orthogonal functions and wavelet coherence analysis. *Journal of Hydrology*, 529, 1754–1767. https://doi.org/10.1016/j.jhydrol.2015.08.011.

Fekete, B. M., Vörösmarty, C. J., Roads, J. O., and Willmott, C. J. (2004). Uncertainties in precipitation and their impacts on runoff estimates. *Journal of Climate*, 17(2), 294-304. https://doi.org/10.1175/1520-0442(2004)017<0294:UIPATI>2.0.CO;2.

Fenicia, F., Savenije, H. H., Matgen, P., and Pfister, L. (2007). A comparison of alternative multiobjective calibration strategies for hydrological modelling, *Water Resources Research*, 43, W03434, https://doi.org/10.1029/2006WR005098.

Fernandez Prieto, D., Berry, P. A. M., Benveniste, J., Courault, D., Cretaux, J. F., Van Dijk, A. I. J. M., Doell, P., Gambacorta, A., De Jeu, R. A. M., Kerr, Y., Kidd, K., Koike, T., Kustas, W., Levizzani, V., McCabe, M. F., Montanari, A., Menenti, M., Njoku, E. G., van Oevelen, P., Prigent, C., Roebeling, R., Schulz, J., Su, Z. B., Wagner, W., and Wood, E. (2009). Earth Observation and Water Cycle Science: towards a multi-mission observation strategy, Earth Observation and Water Cycle Science, Frascati, Italy, 18–20 November, 2009.

Fertig, E. J., Hunt, B. R., Ott, E., and Szunyogh, I. (2007). Assimilating non-local observations with a local ensemble Kalman filter. *Tellus A*, 59(5), 719-730. doi: 10.1111/j.1600-0870.2007.00260.x

Forman, B. A., Reichle, R. H., and Rodell, M. (2012). Assimilation of terrestrial water storage from GRACE in a snow-dominated basin. *Water Resources Research*, 48(1). doi: 10.1029/2011WR011239.

Franz, K. J., Hogue, T. S., Barik, M., and He, M. (2014). Assessment of SWE data assimilation for ensemble streamflow predictions. *Journal of Hydrology*, 519, 2737-2746. https://doi.org/10.1016/j.jhydrol.2014.07.008.

Fu, S., Sonnenborg, T. O., Jensen, K. H., and He, X. (2011). Impact of precipitation spatial resolution on the hydrological response of an integrated distributed water resources model. *Vadose Zone Journal*, 10(1), 25-36. doi:10.2136/vzj2009.0186.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Hoell, A., and Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. doi:10.1038/sdata.2015.66.

Gascon, T., Vischel, T., Lebel, T., Quantin, G., Pellarin, T., Quatela, V., Leroux, D., and Galle, S. (2015). Influence of rainfall space-time variability over the Ouémé basin in Benin. *Proceedings of the International Association of Hydrological Sciences*, 368, 102-107. https://doi.org/10.5194/piahs-368-102-2015.

Gash, J. H. C. (1979). An analytical model of rainfall interception by forests. *Quarterly Journal of the Royal Meteorological Society*, 105, 43–55. doi: 10.1002/qj.49710544304.

Gessner, U., Naeimi, V., Klein, I., Kuenzer, C., Klein, D., and Dech, S. (2013). The relationship between precipitation anomalies and satellite-derived vegetation activity in Central Asia. *Global and Planetary Change*, 110, 74-87. https://doi.org/10.1016/j.gloplacha.2012.09.007.

Gevaert, A. I., Parinussa, R. M., Renzullo, L. J., Van Dijk, A. I. J. M., and de Jeu, R. A. M. (2015). Spatio-temporal evaluation of resolution enhancement for passive microwave soil moisture and vegetation optical depth. *International Journal of Applied Earth Observation and Geoinformation*, 45, 235–244. https://doi.org/10.1016/j.jag.2015.08.006.

Ghatak, M., Kamal, A., and Mishra, O. P. (2012). Background paper flood risk management in South Asia. *Proceedings of the SAARC workshop on flood risk management in South Asia.*

Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228(1), 113-129. https://doi.org/10.1016/S0022-1694(00)00144-X.

Goswami, D. C. (2008). Managing the wealth and woes of the river Brahmaputra. Ishani, 2(4).

Green, D., Petrovic, J., Moss, P., and Burrell, M. (2011). Water resources and management overview: Murrumbidgee catchment, NSW Office of Water, Sydney, 2011.

Gupta, H. V., Sorooshian, S., and Yapo, P. O. (1998). Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resources Research*, 34, 751–763. doi: 10.1029/97WR03495.

Gupta, H. V., Wagener, T., and Liu, Y. (2008). Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, 22, 3802–3813. doi: 10.1002/hyp.6989.

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

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., Stacke, T., Tessler, Z. D., Wada, Y. and Wisser, D. (2014). Global water resources affected by human interventions and climate change. *Proceedings of the National Academy of Sciences of the United States of America*, 111(9), 3251-3256. https://doi.org/10.1073/pnas.1222475110.

Hafeez, M., van de Giesen, N., Bardsley, E., Seyler, F., Pail, R., and Taniguchi, M. (2011). GRACE, remote sensing and ground-based methods in multi-scale hydrology: proceedings of symposium JHO1 held during IUGG2011*, IAHS Publications*. ISBN 978-1-907161-18-6.

Hagemann, S. and Gates, L. D. (2003). Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations. *Climate Dynamics*, 21, 349–359. https://doi.org/10.1007/s00382-003-0349-x.

Halbert, K., Nguyen, C. C., Payrastre, O., and Gaume, E. (2016). Reducing uncertainty in flood frequency analyzes: A comparison of local and regional approaches involving information on extreme historical floods. *Journal of Hydrology*, 541, 90-98. https://doi.org/10.1016/j.jhydrol.2016.01.017.

Hall, D. K., Salomonson, V. V., and Riggs, G. A. (2006). MODIS/Terra snow cover daily L3 global 500m grid. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media. (Version 5. [Tile h09v04]).

Haque, A., and Jahan, S. (2015). Impact of flood disasters in Bangladesh: A multi-sector regional analysis. *International Journal of Disaster Risk Reduction*, 13, 266-275. https://doi.org/10.1016/j.ijdrr.2015.07.001.

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

Havnø, K., Madsen, M.N., Dørge. (1995). J. MIKE 11 — a generalized river modelling package. In: Singh, V.P. (Ed.). Computer Models of Watershed Hydrology, *Water Resources Publications,* Colorado, pp. 733–782.

Hijmans, R. J., Came.ron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965-1978. doi: 10.1002/joc.1276.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki R., Nakamura, K., and Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, 95(5), 701-722. https://doi.org/10.1175/BAMS-D-13-00164.1.

Houdret, A. (2008). Les conflits autour de l'eau au Maroc: origines sociopolitiques et écologiques et perspectives pour transformation des conflits, Thèse de doctorat: Sciences politiques, Universität Duisburg-Essen, 72 pp.

Hoyos, N., Escobar, J., and Restrepo, J.C. (2013). Impact of the 2010-2011 La Niña phenomenon in Colombia, South America: The human toll of an extreme weather event. *Applied Geography*, 39:16–25. doi: 10.1016/j.apgeog.2012.11.018.

Hoyos, N., Correa-Metrio, A., and Sisa, A. (2017). The environmental envelope of fires in the Colombian Caribbean. *Applied Geography*, 84:42–54. doi: 10.1016/j.apgeog.2017.05.001.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C. (2013). A decade of Predictions in Ungauged Basins (PUB) – a review. *Hydrological Sciences Journal*, 58, 1198–1255. https://doi.org/10.1080/02626667.2013.803183.

Hsu, K. L., Gao, X., Sorooshian, S., and Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, 36(9), 1176-1190. https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1), 195-213. https://doi.org/10.1016/S0034-4257(02)00096-2.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, 8(1), 38-55. https://doi.org/10.1175/JHM560.1.

Hunink, J. E., Immerzeel, W. W., and Droogers, P. (2014). A High-resolution Precipitation 2-step mapping Procedure (HiP2P): Development and application to a tropical mountainous area. *Remote Sensing of Environment*, 140, 179-188. https://doi.org/10.1016/j.rse.2013.08.036.

IDEAM: Atlas climatológico de Colombia (1981-2010). Available at: http://atlas.ideam.gov.co/visorAtlasClimatologico.html. Last access: 23-August-2017.

Immerzeel, W. W., Quiroz, R. A., and De Jong, S. M. (2005). Understanding precipitation patterns and land use interaction in Tibet using harmonic analysis of SPOT VGT-S10 NDVI time series. *International Journal of Remote Sensing*, 26(11), 2281-2296. https://doi.org/10.1080/01431160512331326611.

Immerzeel, W. (2008). Historical trends and future predictions of climate variability in the Brahmaputra basin. *International Journal of Climatology*, 28(2), 243-254. doi: 10.1002/joc.1528.

Immerzeel, W. and Droogers, P. (2008). Calibration of a distributed hydrological model based on satellite evapotranspiration. *Journal of Hydrology*, 349, 411–424. https://doi.org/10.1016/j.jhydrol.2007.11.017.

Immerzeel, W. W., Rutten, M. M., and Droogers, P. (2009). Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sensing of Environment*, 113(2), 362-370. https://doi.org/10.1016/j.rse.2008.10.004.

Isenstein, E. M., Wi, S., Yang, Y. E., and Brown, C. (2015). Calibration of a distributed hydrologic model using streamflow and remote sensing snow data. *World Environmental and Water Resources Congress 2015*, ASCE, 973–982.

Islam, M. M., and Sado, K. (2000). Development of flood hazard maps of Bangladesh using NOAA-AVHRR images with GIS. *Hydrological Sciences Journal*, 45(3), 337-355. https://doi.org/10.1080/02626660009492334.

Jacobs, J. M., Myers, D. A., and Whitfield, B. M. (2003). Improved rainfall/runoff estimates using remotely sensed soil moisture. *Journal of the American Water Resources Association*, 39, 313–324. doi: 10.1111/j.1752-1688.2003.tb04386.x.

Janis, M. J., Hubbard, K. G., and Redmond, K. T. (2004). Station density strategy for monitoring long-term climatic change in the contiguous United States. *Journal of Climate*, 17(1), 151-162.bardossy. https://doi.org/10.1175/1520-0442(2004)017<0151:SDSFML>2.0.CO;2.

Javanmard, S., Yatagai, A., Nodzu, M. I., BodaghJamali, J., and Kawamoto, H. (2010). Comparing high-resolution gridded precipitation data with satellite rainfall estimates of TRMM_3B42 over Iran. *Advances in Geosciences*, 25, 119-125. https://doi.org/10.5194/adgeo-25-119-2010.

Jia, S., Zhu, W., Lű, A., and Yan, T. (2011). A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. *Remote Sensing of Environment*, 115(12), 3069-3079. https://doi.org/10.1016/j.rse.2011.06.009.

Jones, A., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Dewitte, O., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Le Roux, P., Micheli, E., Montanarella, L., Spaargaren, O., Thiombiano, L., Van Ranst, E., Yemefack, M., and Zougmoré, R. (2013). Soil Atlas of Africa, European Commission, Luxembourg, https://doi.org/10.2788/52319.

Jones, D. A., Wang, W., and Fawcett, R. (2009). High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal*, 58.4, 233–248.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5(3), 487-503. https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35-45. doi:10.1115/1.3662552.

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

Kaune, A., Werner, M., Rodríguez, E., Karimi, P., and De Fraiture, C. (2017). A novel tool to assess available hydrological information and the occurrence of sub-optimal water allocation decisions in large irrigation districts. *Agricultural Water Management*, 191, 229-238. https://doi.org/10.1016/j.agwat.2017.06.013.

Kaune, A., López López, P., Gevaert, A., Veldkamp, T., Werner, M. and de Fraiture, C. The benefit of using an ensemble of global hydrological models in surface water availability for irrigation area planning. *Water Resources Management* (in review).

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J., and Berger, M. (2001). Soil moisture retrieval from space: The Soil Moisture and Ocean Salinity (SMOS) mission. *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1729–1735. doi: 10.1109/36.942551.

Kerr, Y. H., Waldteufel, P., Richaume, P., Wigneron, J. P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Jugles, S. E., Leroux, D., Mialon, A., and Delwart, S. (2012). The SMOS soil moisture retrieval algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 50(5), 1384-1403. doi: 10.1109/TGRS.2012.2184548.

Khaki, M.., Hoteit, I., Kuhn, M.., Awange, J., Forootan, E., Van Dijk, A. I. J. M.., Schumacher, M. and Pattiaratchi, C. (2017). Assessing sequential data assimilation techniques for integrating GRACE data into a hydrological model. *Advances in Water Resources*, 107, 301-316. https://doi.org/10.1016/j.advwatres.2017.07.001.

Khu, S. T. and Madsen, H. (2005). Multiobjective calibration with Pareto preference ordering: An application to rainfallrunoff model calibration. *Water Resources Research*, 41, W03004, https://doi.org/10.1029/2004WR003041.

Kite, G. and Droogers, P. (2000). Comparing evapotranspiration estimates from satellites, hydrological models and field data. *Journal of Hydrology*, 229, 3–18. https://doi.org/10.1016/S0022-1694(99)00195-X.

Koch, J., Jensen, K. H., and Stisen, S. (2015). Toward a true spatial model evaluation in distributed hydrological modelling: Kappa statistics, Fuzzy theory, and EOF-analysis benchmarked by the human perception and evaluated against a modelling case study. *Water Resources Research*, 51, 1225–1246. doi: 10.1002/2014WR016607.

Koch, J., Mendiguren, G., Mariethoz, G., and Stisen, S. (2017). Spatial sensitivity analysis of simulated land-surface patterns in a catchment model using a set of innovative spatial performance metrics. *Journal of Hydrometeorology*, 18, 1121–1142. https://doi.org/10.1175/JHM-D-16-0148.1

Koren, V. I., Finnerty, B. D., Schaake, J. C., Smith, M. B., Seo, D. J., and Duan, Q. Y. (1999). Scale dependencies of hydrologic models to spatial variability of precipitation. *Journal of Hydrology*, 217(3), 285-302. https://doi.org/10.1016/S0022-1694(98)00231-5.

Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Takayabu, Y. N., Ushio, T., Nakagawa, K., Iwanami, K., Kachi, M., and Okamoto, K. (2007). Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Transactions on Geoscience and Remote Sensing*, 45(7), 2259-2275. doi: 10.1109/TGRS.2007.895337.

Kugler, Z., and T. De Groeve. (2007). The global flood detection system, Off. For Off. Publ. of the Eur. Commun., Luxembourg, Ispra, Italy, 2007.

Kumar, R., Samaniego, L., and Attinger, S. (2013). Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations. *Water Resources Research*, 49, 360–379. doi: 10.1029/2012WR012195.

Lang, M., Ouarda, T. B. M. J., and Bobée, B. (1999). Towards operational guidelines for over-threshold modelling. *Journal of Hydrology*, 225(3), 103-117. https://doi.org/10.1016/S0022-1694(99)00167-5.

Langella, G., Basile, A., Bonfante, A., and Terribile, F. (2010). High-resolution space–time rainfall analysis using integrated ANN inference systems. *Journal of Hydrology*, 387(3), 328-342. https://doi.org/10.1016/j.jhydrol.2010.04.027.

Lanza, L. G., Schultz, G. A., and Barrett, E. C. (1997). Remote sensing in hydrology: some downscaling and uncertainty issues. *Physics and Chemistry of the Earth*, 22, 215–219. https://doi.org/10.1016/S0079-1946(97)00135-3.

Lee, H., Seo, D. J., and Koren, V. (2011). Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the

data and initial model soil moisture states. *Advances in Water Resources*, 34(12), 1597-1615. https://doi.org/10.1016/j.advwatres.2011.08.012.

Lee, H., Seo, D.-J., Liu, Y., Koren, V., McKee, P., and Corby, R. (2012). Variational assimilation of streamflow into operational distributed hydrologic models: effect of spatiotemporal scale of adjustment. *Hydrology and Earth System Sciences*, 16, 2233–2251, doi:10.5194/hess-16-2233-2012.

Lehner, B., Döll, P., Alcamo, J., Henrichs, T., and Kaspar, F. (2006). Estimating the impact of global change on flood and drought risks in Europe: a continental, integrated analysis. *Climatic Change*, 75(3), 273-299. https://doi.org/10.1007/s10584-006-6338-4.

Li, Y., Ryu, D., Western, A. W., and Wang, Q. J. (2013). Assimilation of stream discharge for flood forecasting: The benefits of accounting for routing time lags. *Water Resources Research*, 49(4), 1887-1900. doi: 10.1002/wrcr.20169.

Liang, X., Lettenmaier, D. P., Wood, E. F., and Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research-Atmospheres*, 99, 14415–14415. doi: 10.1029/94JD00483.

Liang, X., Wood, E. F., and Lettenmaier, D. P. (1996). Surface soil moisture parameterization of the VIC-2L model: evaluation and modification. *Global and Planetary Change*, 13, 195–206. https://doi.org/10.1016/0921-8181(95)00046-1.

Lievens, H., Tomer, S. K., Al Bitar, A., De Lannoy, G. J. M., Drusch, M., Dumedah, G., Hendricks Franssen, H.-J., Kerr, Y. H., Martens, B., Pan, M., Roundy, J. K., Vereecken, H., Walker, J. P., Wood, E. F., Verhoest, N. E. C. and Pauwels, V. R. N. (2015). SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. *Remote Sensing of Environment*, 168, 146–162. https://doi.org/10.1016/j.rse.2015.06.025.

Linés, C., Werner, M. and Bastiaanssen, W. (2017). The predictability of reported drought events and impacts in the Ebro Basin using six different remote sensing data sets. *Hydrology and Earth System Sciences*, 21(9), 4747. https://doi.org/10.5194/hess-21-4747-2017.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S. (1997). Development and test of the distributed HBV-96 hydrological model. *Journal of Hydrology*, 201(1-4), 272-288. https://doi.org/10.1016/S0022-1694(97)00041-3.

Linsley Jr, R. K., Kohler, M. A., and Paulhus, J. L. (1975). Hydrology for engineers. Mc Graw Hill, New York.

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., Van Dijk, A. I. J. M., McCabe, M. F., and Evans, J. P. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrology and Earth System Sciences*, 15, 425–436. doi:10.5194/hess-15-425-2011.

Liu, Y., Dorigo, W. A., Parinussa, R., de Jeu, R. A., Wagner, W., McCabe, M. F., Evans, J., and Van Dijk, A. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment*, 123, 280–297. https://doi.org/10.1016/j.rse.2012.03.014.

Liu, J., Duan, Z., Jiang, J., and Zhu, A. (2015). Evaluation of three satellite precipitation products TRMM 3B42, CMORPH, and PERSIANN over a subtropical watershed in China. *Advances in Meteorology*, http://dx.doi.org/10.1155/2015/151239.

Lo, M.-H., Famiglietti, J. S., Yeh, P.-F., and Syed, T. (2010). Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data. *Water Resources Research,* 46, W05517, https://doi.org/10.1029/2009WR007855.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C. (2014). When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, 18(2), 575. doi:10.5194/hess-18-575-2014.

Long, Y., Zhang, Y., and Ma, Q. (2016). A merging framework for rainfall estimation at high spatiotemporal resolution for distributed hydrological modelling in a data-scarce area. *Remote Sensing*, 8(7), 599. doi:10.3390/rs8070599.

Looper, J. P., Vieux, B. E., and Moreno, M. A. (2012). Assessing the impacts of precipitation bias on distributed hydrologic model calibration and prediction accuracy. *Journal of Hydrology*, 418, 110–122. https://doi.org/10.1016/j.jhydrol.2009.09.048.

López López, P., Wanders, N., Schellekens, J., Renzullo, L. J., Sutanudjaja, E. H., and Bierkens, M. F. (2016). Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. *Hydrology and Earth System Sciences*, 20(7), 3059-3076. doi:10.5194/hess-20-3059-2016.

López López, P., Immerzeel W.W., Rodriguez Sandoval E.A., Sterk G., and Schellekens J. (2018). Impact of high spatial resolution precipitation on streamflow simulations. *Frontiers in Earth Science*s, in review.

Loukas, A. and Vasiliades, L. (2014). Streamflow simulation methods for ungauged and poorly gauged watersheds. *Natural Hazards and Earth System Sciences*, 14, 1641–1661, https://doi.org/10.5194/nhess-14-1641-2014.

Lowry, C. S., and Fienen, M. N. (2013). CrowdHydrology: crowdsourcing hydrologic data and engaging citizen scientists. *GroundWater*, 51(1), 151-156. doi: 10.1111/j.1745-6584.2012.00956.x.

Lu, J., Sun, G., McNulty, S. G., and Amatya, D. M. (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. *Journal of the American Water Resources Association*, 41, 621–633. doi: 10.1111/j.1752-1688.2005.tb03759.x

Martens, B., Miralles, D., Lievens, H., Fernández-Prieto, D., and Verhoest, N. (2016). Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture. *International Journal of Applied Earth Observation and Geoinformation*, 48, 146–162.

Martens, B., Miralles, D., Lievens, H., van der Schalie, R., de Jeu, R., Fernández-Prieto, D., and Verhoest, N. (2016). GLEAM v3: updated land evaporation and root-zone soil moisture datasets. *EGU General Assembly Conference Abstracts*, vol. 18, p. 4253.

Martens, B., Gonzalez Miralles, D., Lievens, H., van der Schalie, R., de Jeu, R. A., Fernández-Prieto, D., Beck, H., Dorigo, W. A. and Verhoest, N. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5), 1903-1925. doi:10.5194/gmd-10-1903-2017.

Mascaro, G., Vivoni, E. R., and Méndez-Barroso, L. A. (2015). Hyperresolution hydrologic modelling in a regional watershed and its interpretation using empirical orthogonal functions. *Advances in Water Resources*, 83, 190–206. https://doi.org/10.1016/j.advwatres.2015.05.023.

Massari, C., Brocca, L., Tarpanelli, A., and Moramarco, T. (2015). Data assimilation of satellite soil moisture into rainfall-runoff modelling: a complex recipe? *Remote Sensing*, 7, 11403–11433. doi:10.3390/rs70911403.

McCabe, M. F., Aragon, B., Houborg, R., and Mascaro, J. (2017). CubeSats in hydrology: Ultrahigh-resolution insights into vegetation dynamics and terrestrial evaporation. *Water Resources Research*, 53. https://doi.org/10.1002/ 2017WR022240.

Mediero, L., Santillán, D., Garrote, L., and Granados, A. (2014). Detection and attribution of trends in magnitude, frequency and timing of floods in Spain. *Journal of Hydrology*, 517, 1072-1088. https://doi.org/10.1016/j.jhydrol.2014.06.040

Mediero, L., Kjeldsen, T. R., Macdonald, N., Kohnova, S., Merz, B., Vorogushyn, S., Wilson, D., Alburquerque, T., Bloschl, G., Bogdanowicz, E., Castellarini, A., Hall, J., Kobold, M., Kriauclunie, J., Lang, M., Madsen, H., Onusluel, G., Perdigao, R. A. P., Roald, L. A., Salinas, J. L., Toumazis, A. D., Veijalainen, N., and ÓðinnÞórarinssonr. (2015). Identification of coherent flood regions across Europe by using the longest streamflow records. *Journal of Hydrology*, 528, 341-360. https://doi.org/10.1016/j.jhydrol.2015.06.016.

Meesters, A., De Jeu, R. A. M., and Owe, M. (2005). Analytical derivation of the vegetation optical depth from the microwave polarization difference index. *IEEE Geoscience and Remote Sensing Letters*, 2, 121. doi: 10.1109/LGRS.2005.843983

Mei, Y., Anagnostou, E. N., Nikolopoulos, E. I., and Borga, M. (2014). Error analysis of satellite precipitation products in mountainous basins. *Journal of Hydrometeorology*, 15(5), 1778-1793. https://doi.org/10.1175/JHM-D-13-0194.1.

Merlin, O., Chehbouni, A., Boulet, G., and Kerr, Y. (2006). Assimilation of disaggregated microwave soil moisture into a hydrologic model using coarse-scale meteorological data. *Journal of Hydrometeorology*, 7, 1308–1322. https://doi.org/10.1175/JHM552.1.

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., and Dolman, A. J. (2011). Magnitude and variability of land evaporation and its components at the global scale. *Hydrology and Earth System Sciences*, 15,967–981, https://doi.org/10.5194/hess-15-967-2011.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based

observations. *Hydrology and Earth System Sciences*, 15, 453–469. doi:10.5194/hess-15-453-2011.

Mirza, M. M. Q., Warrick, R. A., Ericksen, N. J., and Kenny, G. J. (2001). Are floods getting worse in the Ganges, Brahmaputra and Meghna basins? *Global Environmental Change Part B: Environmental Hazards,* 3(2), 37-48. https://doi.org/10.1016/S1464-2867(01)00019-5

Mirza, M. M. Q. (2003). Three recent extreme floods in Bangladesh: a hydro-meteorological analysis. *Flood problem and management in South Asia* (pp. 35-64). Springer Netherlands. https://doi.org/10.1007/978-94-017-0137-2_2.

Mirza, M. M. Q. (2011). Climate change, flooding in South Asia and implications. *Regional Environmental Change*, 11(1), 95-107. https://doi.org/10.1007/s10113-010-0184-7.

Mitchell, T. D. and Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, 25, 693–712. doi: 10.1002/joc.1181

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R. (2005). Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, 28, 135–147. https://doi.org/10.1016/j.advwatres.2004.09.002.

Moradkhani, H. (2008). Hydrologic remote sensing and land surface data assimilation. *Sensors*, 8, 2986–3004. doi:10.3390/s8052986.

Moradkhani, H., C. M. DeChant, and S. Sorooshian. (2012). Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. *Water Resources Research*, 48, W02501, doi:10.1029/2012WR012144.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L.,Harmel, R. D., and Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers*, 50, 885–900. doi: 10.13031/2013.23153.

Mulvaney, T.J. (1850). On the use of self-registering rain and flood gauges in making observations on the relations of rainfall and flood discharges in a given catchment. *Transactions of Institution of Civil Engineers Ireland*, 4, 18-31.

Murray, V., and Ebi, K. L. (2012). IPCC special report on managing the risks of extreme events and disasters to advance climate change adaptation (SREX). http://dx.doi.org/10.1136/jech-2012-201045.

Nagy, B. K., Mohssen, M., and Hughey, K. F. D. (2017). Flood frequency analysis for a braided river catchment in New Zealand: Comparing annual maximum and partial duration series with varying record lengths. *Journal of Hydrology*, 547, 365-374. https://doi.org/10.1016/j.jhydrol.2017.02.001.

Nash, J. E., and Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282-290. https://doi.org/10.1016/0022-1694(70)90255-6

Naeimi, V., Scipal, K., Bartalis, Z., Hasenauer, S., and Wagner, W. (2009). An improved soil moisture retrieval algorithm for ERS and METOP scatterometer observations. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 1999–2013. doi: 10.1109/TGRS.2008.2011617.

New, M., Lister, D., Hulme, M., and Makin, I. (2002). A high-resolution data set of surface climate over global land areas. *Climate Research*, 21, 1–25. http://www.jstor.org/stable/24866815

Nielsen, S.A., and Hansen, E. (1973). Numerical simulation of the rainfall-runoff process on a daily basis. Hydrology Research, 4(3), 171-190.

Nishida, K. (2003). Validation of prototype MODIS evapotranspiration (MOD16) in the eastern Asia. *AGU Fall Meeting Abstracts*, http://adsabs.harvard.edu/abs/2003AGUFM.B21E0765N.

Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V. (2003). Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 215–229. doi: 10.1109/TGRS.2002.808243.

Nkiaka, E., Nawaz, N. R., and Lovett, J. C. (2017). Evaluating Global Reanalysis Datasets as Input for Hydrological Modelling in the Sudano-Sahel Region. *Hydrology*, 4(1), 13. doi:10.3390/hydrology4010013.

Noh, S. J., Rakovec, O., Weerts, A. H., and Tachikawa, Y. (2014). On noise specification in data assimilation schemes for improved flood forecasting using distributed hydrological models. *Journal of Hydrology*, 519, 2707-2721. https://doi.org/10.1016/j.jhydrol.2014.07.049.

Ouatiki, H., Boudhar, A., Tramblay, Y., Jarlan, L., Benabdelouhab, T., Hanich, L., El Meslouhi, M. R., and Chehbouni, A. (2017). Evaluation of TRMM 3B42 V7 Rainfall Product over the Oum Er Rbia Watershed in Morocco. *Climate*, 5, 1, https://doi.org/10.3390/cli5010001.

Owe, M., De Jeu, R., and Holmes, T. (2008). Multisensor historical climatology of satellite-derived global land surface moisture. *Journal of Geophysical Research-Earth Surface*, 113, F01002, doi:10.1029/2007JF000769.

Paul, B. K. (1997). Flood research in Bangladesh in retrospect and prospect: a review. *Geoforum*, 28(2), 121-131. https://doi.org/10.1016/S0016-7185(97)00004-3.

Parajka, J., Naeimi, V., Blöschl, G., Wagner, W., Merz, R., and Scipal, K. (2006). Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale. *Hydrology and Earth System Sciences*, 10, 353–368, www.hydrol-earth-syst-sci.net/10/353/2006/.

Park, N. W. (2013). Spatial downscaling of TRMM precipitation using geostatistics and fine scale environmental variables. *Advances in Meteorology*. http://dx.doi.org/10.1155/2013/237126.

Parrish, D. F., and Derber, J. C. (1992). The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8), 1747-1763. https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Pauwels, V. R., and De Lannoy, G. J. (2006). Improvement of modelling soil wetness conditions and turbulent fluxes through the assimilation of observed discharge. *Journal of Hydrometeorology*, 7(3), 458-477. https://doi.org/10.1175/JHM490.1.

Pearson, K. (1896). Mathematical Contributions to the Theory of Evolution – On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs. *Proceedings of the Royal Society in London*, 60, 489–498. doi: 10.1098/rspl.1896.0076.

Peischl, S., Walker, J. P., Rüdiger, C., Ye, N., Kerr, Y. H., Kim, E., Bandara, R., and Allahmoradi, M. (2012). The AACES field experiments: SMOS calibration and validation across the Murrumbidgee River catchment. *Hydrology and Earth System Sciences*, 16, 1697– 1708, doi:10.5194/hess-16-1697-2012.

Pedinotti, V., Boone, A., Ricci, S., Biancamaria, S., and Mognard, N. (2014). Assimilation of satellite data to optimize large-scale hydrological model parameters: a case study for the SWOT mission. *Hydrology and Earth System Sciences*, 18(11), 4485-4507. https://doi.org/10.5194/hess-18-4485-2014

Poveda, G., Waylen, P.R., Pulwarty, R.S. (2006). Annual and inter-annual variability of the present climate in northern South America and southern Mesoamerica. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 234:3–27. doi: 10.1016/j.palaeo.2005.10.031.

Pulliainen, J. (2006). Mapping of Snow Water Equivalent and Snow Depth in Boreal and Sub-Arctic Zones by Assimilating Space-Borne Microwave Radiometer Data and Ground-Based Observations. *Remote Sensing of Environment*, 101(2): 257–269. doi: 10.1016/j.rse.2006.01.002.

QGIS. (2017). A free and open source geographic information system. Available at: http://www.qgis.org/en/site/. Last access: 21-August-2017.

Quiroz, R., Yarlequé, C., Posadas, A., Mares, V., and Immerzeel, W. W. (2011). Improving daily rainfall estimation from NDVI using a wavelet transform. *Environmental Modelling and Software*, 26(2), 201-209. https://doi.org/10.1016/j.envsoft.2010.07.006.

Rakovec, O., Weerts, A. H., Hazenberg, P., Torfs, P. J. J. F., and Uijlenhoet, R. (2012). State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy. *Hydrology and Earth System Sciences*, 16, 3435–3449, doi:10.5194/hess-16- 3435-2012.

Rakovec, O., Weerts, A. H., Sumihar, J., and Uijlenhoet, R. (2015). Operational aspects of asynchronous filtering for flood forecasting. *Hydrology and Earth System Sciences*, 19(6), 2911-2924. http://dx.doi.org/10.5194/hess-19-2911-2015.

Randrianasolo, A., Thirel, G., Ramos, M. H., and Martin, E. (2014). Impact of streamflow data assimilation and length of the verification period on the quality of short-term ensemble hydrologic forecasts. *Journal of Hydrology*, 519, 2676-2691. https://doi.org/10.1016/j.jhydrol.2014.09.032.

Rasmussen, J., Madsen, H., Jensen, K. H., and Refsgaard, J. C. (2015). Data assimilation in integrated hydrological 195odelling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance. *Hydrology and Earth System Sciences*, 19(7), 2999-3013. www.hydrol-earth-syst-sci.net/19/2999/2015/.

Reichle, R. H. and Koster, R. D. (2004). Bias reduction in short records of satellite soil moisture. *Geophysical Research Letters*, 31, L19501, https://doi.org/10.1029/2004GL020938.

Reichle, R. H. (2008). Data assimilation methods in the Earth sciences. *Advances in Water Resources,* 31(11), 1411-1418. https://doi.org/10.1016/j.advwatres.2008.01.001.

Reichle, R. H., De Lannoy, G. J., Forman, B. A., Draper, C. S., and Liu, Q. (2014). Connecting satellite observations with water cycle variables through land data assimilation: examples using the NASA GEOS-5 LDAS. *Surveys in Geophysics*, *The Earth's Hydrological Cycle, Springer Netherlands,* 577–606. https://doi.org/10.1007/s10712-013-9220-8.

Renzullo, L. J., Van Dijk, A. I. J. M., Perraud, J. M., Collins, D., Henderson, B., Jin, H., and McJannet, D. L. (2014). Continental satellite soil moisture data assimilation improves root-zone moisture analysis for water resources assessment. *Journal of Hydrology*, 519, 2747– 2762. https://doi.org/10.1016/j.jhydrol.2014.08.008.

Restrepo, J. D., and Kjerfve, B. (2000). Magdalena River: interannual variability (1975–1995) and revised water discharge and sediment load estimates. *Journal of Hydrology*, 235(1), 137-149. https://doi.org/10.1016/S0022-1694(00)00269-9.

Restrepo, J. D., and Syvitski, J. P. (2006). Assessing the effect of natural controls and land use change on sediment yield in a major Andean river: the Magdalena drainage basin, Colombia. *Ambio: a Journal of the Human Environment*, 35(2), 65-74. https://doi.org/10.1579/0044-7447(2006)35[65:ATEONC]2.0.CO;2.

Restrepo JD, Escobar R, Tosic M. (2016). Fluvial fluxes from the Magdalena River into Cartagena Bay, Caribbean Colombia: Trends, future scenarios, and connections with upstream human impacts. *Geomorphology,* 1951–1952. doi: 10.1016/j.geomorph.2016.11.007.

Revilla-Romero, B., Thielen, J., Salamon, P., De Groeve, T., and Brakenridge, G. R.(2014). Evaluation of the satellite-based Global Flood Detection System for measuring river

discharge: influence of local factors. *Hydrology and Earth System Sciences*, 18(11), 4467-4484. https://doi.org/10.5194/hess-18-4467-2014

Revilla-Romero, B., Wanders, N., Burek, P., Salamon, P., and de Roo, A. (2016). Integrating remotely sensed surface water extent into continental scale hydrology. *Journal of Hydrology*, 543, 659-670. https://doi.org/10.1016/j.jhydrol.2016.10.041.

Ricaurte, L.F., Olaya-Rodríguez, M.H. and Cepeda-Valencia, J. (2017). Future impacts of drivers of change on wetland ecosystem services in Colombia. *Global Environmental Change*, 44:158–169. doi: 10.1016/j.gloenvcha.2017.04.001.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., and Woollen, J. (2011). MERRA: NASA's modern-era retrospective analysis for research and applications. *Journal of Climate*, 24, 3624–3648. https://doi.org/10.1175/JCLI-D-11-00015.1

Rientjes, T., Muthuwatta, L. P., Bos, M., Booij, M., and Bhatti, H. (2013). Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration. *Journal of Hydrology*, 505, 276–290. https://doi.org/10.1016/j.jhydrol.2013.10.006.

Roy, A., Royer, A., and Turcotte, R. (2010). Improvement of springtime streamflow simulations in a boreal environment by incorporating snow-covered area derived from remote sensing data. *Journal of Hydrology*, 390, 35–44. https://doi.org/10.1016/j.jhydrol.2010.06.027.

Sælthun, N. R. (1996). The Nordic HBV model. *Norwegian Water Resources and Energy Administration Publication,* 7, 1-26.

Sahoo, A. K., De Lannoy, G. J., Reichle, R. H., and Houser, P. R. (2013). Assimilation and downscaling of satellite observed soil moisture over the Little River Experimental Watershed in Georgia, USA. *Advances in Water Resources*, 52, 19–33. https://doi.org/10.1016/j.advwatres.2012.08.007.

Sakov, P., and Bertino, L. (2011). Relation between two common localisation methods for the EnKF. *Computational Geosciences*, 15(2), 225-237. https://doi.org/10.1007/s10596-010-9202-6.

Samaniego, L., Kumar, R., and Attinger, S. (2010). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46, W05523, https://doi.org/10.1029/2008WR007327.

Sangati, M., and Borga, M. (2009). Influence of rainfall spatial resolution on flash flood modelling. *Natural Hazards and Earth System Sciences*, 9(2), 575-584. https://doi.org/10.5194/nhess-9-575-2009.

Schellekens, J., Van Verseveld, W., Euser, T., Winsemius, H., Thiange, C., Bouaziz, L., Tollenaar, D. and De Vries, S. (2016). Openstreams/wflow: 2016.04 Test release, http://doi.org/10.5281/zenodo.167057.

Schellekens, J. and Weiland, F. S. (2017). earth2observe/downscaling-tools. Pre-release, https://doi.org/10.5281/zenodo.545779.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., Van Dijk, A., Weiland, F. S., Minvielle, M., Calvet, J. C., Decharme, B., Eisner, S., Fink, G., Dlorke, M., Pessenteiner, S., Van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W. and Weedon, G. P. (2017). A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset. *Earth System Science Data*, 9(2), 389. https://doi.org/10.5194/essd-9-389-2017.

Schellekens, J. (2017). Available at: https://publicwiki.deltares.nl/display/INFORMED/WRR2+design Last access: 25 October 2017.

Schmugge, T. J., Kustas, W. P., Ritchie, J. C., Jackson, T. J., and Rango, A. (2002). Remote sensing in hydrology. *Advances in Water Resources*, 25, 1367–1385. https://doi.org/10.1016/S0309-1708(02)00065-9.

Schultz, G. and Engman, E. (2000). Remote sensing in hydrology and water management, Springer Science & Business Media.

Schuurmans, J. M., and Bierkens, M. F. P. (2007). Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model. *Hydrology and Earth System Sciences*, 3 (4), pp.2175-2208. https://doi.org/10.5194/hess-11-677-2007.

Segond, M. L., Wheater, H. S., and Onof, C. (2007). The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK. *Journal of Hydrology*, 347(1), 116-131. https://doi.org/10.1016/j.jhydrol.2007.09.040.

Seibert, J. and Beven, K. J. (2009). Gauging the ungauged basin: how many discharge measurements are needed? *Hydrology and Earth System Sciences*, 13, 883–892, https://doi.org/10.5194/hess-13-883-2009.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Review*, 99, 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004.

Seo, D. J., Cajina, L., Corby, R., and Howieson, T. (2009). Automatic state updating for operational streamflow forecasting via variational data assimilation. *Journal of Hydrology*, 367(3), 255-275. https://doi.org/10.1016/j.jhydrol.2009.01.019.

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika,* 52(3/4), 591-611. doi: 10.2307/2333709.

Shi, Y., and Song, L. (2015). Spatial downscaling of monthly TRMM precipitation based on EVI and other geospatial variables over the Tibetan Plateau From 2001 to 2012.

Mountain Research and Development, 35(2), 180-194. https://doi.org/10.1659/MRD-JOURNAL-D-14-00119.1.

Simpson, J., Kummerow, C., Tao, W. K., and Adler, R. F. (1996). On the tropical rainfall measuring mission (TRMM). *Meteorology and Atmospheric Physics*, 60(1), 19-36. https://doi.org/10.1007/BF01029783.

Sivapalan, M., Takeuchi, K., Franks, S., Gupta, V. K., Karambiri, H., Lakshimi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E. (2003). IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrological Sciences Journal,* 48, 857–880. https://doi.org/10.1623/hysj.48.6.857.51421.

Sivapalan, M. (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes*, 17(15), 3163-3170. doi: 10.1002/hyp.5155.

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., and Richter, H. (2012). The Murrumbidgee soil moisture monitoring network data set. *Water Resources Reseach*, 48, W07701, doi:10.1029/2012WR011976.

Smith, M. B., Koren, V. I., Zhang, Z., Reed, S. M., Pan, J. J., and Moreda, F. (2004). Runoff response to spatial variability in precipitation: an analysis of observed data. *Journal of Hydrology*, 298(1), 267-286. https://doi.org/10.1016/j.jhydrol.2004.03.039.

Stasinopoulos, D. M., and Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, 23(7), 1-46.

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., and Braithwaite, D. (2000). Evaluation of PERSIANN system satellite–based estimates of tropical rainfall. *Bulletin of the American Meteorological Society*, 81(9), 2035-2046. https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.

Su, C. H., Ryu, D., Young, R. I., Western, A. W., and Wagner, W. (2013). Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia. *Remote Sensing of Environment*, 134, 1–11. https://doi.org/10.1016/j.rse.2013.02.016.

Su, H., Yang, Z. L., Dickinson, R. E., Wilson, C. R., and Niu, G. Y. (2010). Multisensor snow data assimilation at the continental scale: The value of Gravity Recovery and Climate Experiment terrestrial water storage information. *Journal of Geophysical Research-Atmospheres*, 115(D10). doi: 10.1029/2009JD013035.

Sun, C., Walker, J. P., and Houser, P. R. (2004). A methodology for snow data assimilation in a land surface model. *Journal of Geophysical Research-Atmosphere*, 109, D08108, doi:10.1029/2003JD003765.

Sun, G., Ranson, K. J., Kharuk, V. I., and Kovacs, K. (2003). Validation of surface height from shuttle radar topography mission using shuttle laser altimeter. *Remote Sensing of Environment*, 88(4), 401-411. https://doi.org/10.1016/j.rse.2003.09.001.

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P. (2011). Large-scale groundwater modelling using global datasets: a test case for the Rhine-Meuse basin. *Hydrology and Earth System Sciences*, 15, 2913–2935, https://doi.org/10.5194/hess-15-2913-2011.

Sutanudjaja, E., Van Beek, L., De Jong, S., Van Geer, F., and Bierkens, M. (2014). Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data. *Water Resources Research*, 50, 687–705. doi: 10.1002/2013WR013807.

Sutanudjaja, E., van Beek, R., Wada, Y., Bosmans, J., Drost, N., de Graaf, I., de Jong, K., Lopez Lopez, P., Pessenteiner, S., Oliver, S., Straatsma, M., Wanders, N., Wisser, D., and Bierkens, M. (2017). PCR-BLOBWB_model, available at: https://github.com/UU-Hydro/PCR-GLOBWB_model, last access: 17 April 2017.

Sutanudjaja, E.H., van Beek, L.P.H., Bosmans, J.H.C., Drost N., de Graaf, I.E.M., van der Ent, R.J., Hoch, J.M., de Jong, K., Karssenberg, D., Lopez Lopez, P., Pessenteriner, S., Schmitz, O., Straatsma, M.W., Vannametee, E., Wada, Y., Wanders, N., Wisser, D. and Bierkens, M.F.P. (2018). PCR-GLOBWB 2.0: a 5 arc-minute global hydrological and water resources model. *Geoscientific Model Development Discussions*. https://doi.org/10.5194/gmd-2017-288.

SWBD, Shuttle Radar Topography Mission Water Body Data. (2017). Available at: https://dds.cr.usgs.gov/srtm/version2_1/SWBD/. More information can be found in: https://lta.cr.usgs.gov/srtm_water_body_dataset. Last data access: 27 July, 2017.

Takala, M., K. Luojus, J. Pulliainen, C. Derksen, J. Lemmetyinen, J. P. Kärnä, J. Koskinen, B. Bojkov. (2011). Estimating Northern Hemisphere Snow Water Equivalent for Climate Research Through Assimilation of Spaceborne Radiometer Data and Ground-Based Measurements. *Remote Sensing of Environment*, 115(12): 3517–3529. doi: 10.1016/j.rse.2011.08.014.

Tangdamrongsub, N., Steele-Dunne, S. C., Gunter, B. C., Ditmar, P. G., and Weerts, A. H. (2015). Data assimilation of GRACE terrestrial water storage estimates into a regional hydrological model of the Rhine River basin. *Hydrology and Earth System Sciences*, 19(4), 2079-2100. www.hydrol-earth-syst-sci.net/19/2079/2015/.

Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C. (2004). The gravity recovery and climate experiment: Mission overview and early results. *Geophysical Research Letters*, 31, L09607, https://doi.org/10.1029/2004GL019920.

Te Linde, A. H., Aerts, J. C. J. H., Hurkmans, R. T. W. L., and Eberle, M. (2008). Comparing model performance of two rainfall-runoff models in the Rhine basin using different

atmospheric forcing data sets. *Hydrology and Earth System Sciences*. www.hydrol-earth-syst-sci.net/12/943/2008/.

Teng, H., Shi, Z., Ma, Z., and Li, Y. (2014). Estimating spatially downscaled rainfall by regression kriging using TRMM precipitation and elevation in Zhejiang Province, southeast China. *International Journal of Remote Sensing*, 35(22), 7775-7794. https://doi.org/10.1080/01431161.2014.976888.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V., and De Roo, A. (2012). Validation of satellite-based precipitation products over sparsely gauged African river basins. Journal of Hydrometeorology, 13(6), 1760-1783, https://doi.org/10.1175/JHM-D-12-032.1.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., and De Roo, A. (2013). Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. Journal of Hydrology, 499, 324-338, https://doi.org/10.1016/j.jhydrol.2013.07.012.

Thirel, G., Salamon, P., Burek, P., and Kalas, M. (2013). Assimilation of MODIS snow cover area data in a distributed hydrological model using the particle filter. *Remote Sensing*, 5(11), 5825-5850. doi:10.3390/rs5115825.

Todini, E. (2007). Hydrological catchment modelling: past, present and future. *Hydrology and Earth System Sciences*, 11(1), 468-482. https://doi.org/10.5194/hess-11-468-2007.

Tramblay, Y., Bouvier, C., Ayral, P. A., and Marchandise, A. (2011). Impact of rainfall spatial distribution on rainfall-runoff modelling efficiency and initial soil moisture conditions estimation. *Natural Hazards and Earth System Sciences*, 11(1), 157-170. https://doi.org/10.5194/nhess-11-157-2011.

Tramblay, Y., Bouaicha, R., Brocca, L., Dorigo, W., Bouvier, C., Camici, S., and Servat, E. (2012). Estimation of antecedent wetness conditions for flood modelling in northern Morocco. *Hydrology and Earth System Sciences*, 16, 4375–4386, https://doi.org/10.5194/hess-16-4375-2012.

Tramblay, Y., Thiemig, V., Dezetter, A., and Hanich, L. (2016). Evaluation of satellite-based rainfall products for hydrological modelling in Morocco. *Hydrological Sciences Journal*, 61, 2509–2519. https://doi.org/10.1080/02626667.2016.1154149.

Van Beek, L. P. H. and Bierkens, M. F. P. (2009). The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification, Tech. Rep., Department of Physical Geography, Utrecht University, Utrecht, the Netherlands. Available at: http: //vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf. Last access: 28 February 2011.

Van Beek, L., Wada, Y., and Bierkens, M. F. (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources Research*, 47, W07517, https://doi.org/10.1029/2010WR009791.

Van Der Knijff, J. M., Younis, J., and 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 Science*, 24, 189–212, https://doi.org/10.1080/13658810802549154.

Van Dijk, A. I. J. M. and Renzullo, L. J. (2009). The Australian Water Resources Assessment system: blending water cycle observations and models at local and continental scale. GEWEX/iLEAPS 'Water in a Changing Climate'; 24-28 August 2009; Melbourne. 2009.

Van Dijk, A. I. J. M. (2010). The Australian water resources assessment system, Version 0.5. Available at: http: //www.clw.csiro.au/publications/waterforahealthycountry/ 2010/wfhc-awras-evaluation-against-observations.pdf. Last access: 8 December 2015.

Van Dijk, A. I. J. M. and Renzullo, L. J. (2011). Water resource monitoring systems and the role of satellite observations. *Hydrology and Earth System Sciences*, 15, 39–55, doi:10.5194/hess-15-39-2011.

Van Dijk, A. I. J. M., Renzullo, L. J., Wada, Y., and Tregoning, P. (2014). A global water cycle reanalysis (2003–2012) merging satellite gravimetry and altimetry observations with a hydrological multi-model ensemble. *Hydrology and Earth System Sciences*, 18, 2955–2973, doi:10.5194/hess-18-2955-2014.

Van Dijk, A. I., Brakenridge, G. R., Kettner, A. J., Beck, H. E., De Groeve, T., and Schellekens, J. (2016). River gauging at global scale using optical and passive microwave remote sensing. *Water Resources Research*, 52(8), 6404-6418. doi: 10.1002/2015WR018545.

Vereecken, H., Huisman, J., Bogena, H., Vanderborght, J., Vrugt, J., and Hopmans, J. (2008). On the value of soil moisture measurements in vadose zone hydrology: A review. *Water Resources Research*, 44, W00D06, https://doi.org/10.1029/2008WR006829.

Vertessy, R. A. and Elsenbeer, H. (1999). Distributed modelling of storm flow generation in an Amazonian rain forest catchment: effects of model parameterization. *Water Resources Research*, 35, 2173–2187. doi: 10.1029/1999WR900051.

Villarini, G., Smith, J. A., Serinaldi, F., Bales, J., Bates, P. D., and Krajewski, W. F. (2009). Flood frequency analysis for nonstationary annual peak records in an urban drainage basin. *Advances in Water Resources*, 32(8), 1255-1266. doi: 10.1029/1999WR900051.

Viney, N. R., Bormann, H., Breuer, L., Bronstert, A., Croke, B. F. W., Frede, H., Gräffe, T., Hubrechts, L., Jakeman, A. J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindström, G., Seibert, J., Sivapalan, M., and Willems, P. (2009). Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) II: Ensemble combinations and predictions. *Advances in Water Resources*, 32, 147–158. https://doi.org/10.1016/j.advwatres.2008.05.006.

Vischel, T., and Lebel, T. (2007). Assessing the water balance in the Sahel: Impact of small scale rainfall variability on runoff. Part 2: Idealized modelling of runoff sensitivity. *Journal of Hydrology*, 333(2), 340-355. https://doi.org/10.1016/j.jhydrol.2006.05.020.

Vörösmarty, C. J., Fekete, B. M., Meybeck, M., and Lammers, R. (2000). A simulated topological network representing the global system of rivers at 30 min spatial resolution (STN-30). *Global Biogeochemical Cycles*, 14, 599–621.

Vrugt, J. A., Gupta, H. V., Nualláin, B., and Bouten, W. (2006). Real-time data assimilation for operational ensemble streamflow forecasting. *Journal of Hydrometeorology*, 7, 548–565. https://doi.org/10.1175/JHM504.1

Vrugt, J. A., C. J. F. terBraak, C. G. H. Diks, and G. Schoups. (2013). Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. *Advances in Water Resources*, 51, 457–478, doi:10.1016/j.advwatres.2012.04.002.

Wagener, T., Gupta, H., Yatheendradas, S., Goodrich, D., Unkrich, C., and Schaffner, M. (2007). Understanding sources of uncertainty in flash-flood forecasting for semi-arid regions. *IAHS PUBLICATION*, 313, 204.

Wagner, W., Verhoest, N. E. C., Ludwig, R., and Tedesco, M. (2009). Editorial "Remote sensing in hydrological sciences". *Hydrology and Earth System Sciences*, 13, 813–817, doi:10.5194/hess-13-813-2009.

Wada, Y., Wisser, D., and Bierkens, M. F. P. (2014). Global modelling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics*, 5, 15–40, doi:10.5194/esd-5-15-2014.

Wanders, N., Karssenberg, D., Bierkens, M., Parinussa, R., de Jeu, R., van Dam, J., and de Jong, S. (2012). Observation uncertainty of satellite soil moisture products determined with physically-based modelling. *Remote Sensing Environment*, 127, 341–356, doi:10.1016/j.rse.2012.09.004.

Wanders, N., Karssenberg, D., de Roo, A., de Jong, S. M., and Bierkens, M. F. P. (2014a). The suitability of remotely sensed soil moisture for improving operational flood forecasting. *Hydrology and Erath System Sciences*, 18, 2343–2357, doi:10.5194/hess-18-2343- 2014.

Wanders, N., Bierkens, M. F., de Jong, S. M., de Roo, A., and Karssenberg, D. (2014b). The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models. *Water Resources Research*, 50, 6874–6891, doi:10.1002/2013WR014639.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, 50(9), 7505-7514. doi: 10.1002/2014WR015638.

Weerts, A. H., and El Serafy, G. Y. (2006). Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research*, 42(9). doi: 10.1029/2005WR004093.

Weerts, A. H., El Serafy, G. Y., Hummel, S., Dhondia, J., and Gerritsen, H. (2010). Application of generic data assimilation tools (DATools) for flood forecasting purposes. *Computers and Geosciences*, 36(4), 453-463. https://doi.org/10.1016/j.cageo.2009.07.009.

Weiland, F. S., Lopez, P., Van Dijk, A., and Schellekens, J. (2015). Global high-resolution reference potential evaporation. MODSIM 2015, *Conference Proceedings, Broadbeach*, Queensland, Australia.

Werner, M., Schellekens, J., Gijsbers, P., Van Dijk, A., van den Akker, O., and Heynert, K. (2013). The Delft-FEWS flow forecasting system. *Environmental Modelling Software,* 40, 65–77. https://doi.org/10.1016/j.envsoft.2012.07.010.

Wesseling, C. G., Karssenberg, D. J., Burrough, P. A., and Deursen, W. (1996). Integrating dynamic environmental models in GIS: the development of a dynamic modelling language. *Transactions in GIS*, 1, 40–48. doi: 10.1111/j.1467-9671.1996.tb00032.x.

Winsemius, H., Schaefli, B., Montanari, A., and Savenije, H. (2009). On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. *Water Resources Research*, 45, W12422, https://doi.org/10.1029/2009WR007706.

Wu, H. and Li, Z.-L. (2009). Scale issues in remote sensing: a review on analysis, processing and modelling. *Sensors,* 9, 1768–1793. doi:10.3390/s90301768.

Xu, H., Xu, C. Y., Chen, H., Zhang, Z., and Li, L. (2013). Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China. *Journal of Hydrology*, 505, 1-12. https://doi.org/10.1016/j.jhydrol.2013.09.004.

Xu, G., Xu, X., Liu, M., Sun, A. Y., and Wang, K. (2015). Spatial downscaling of TRMM precipitation product using a combined multifractal and regression approach: Demonstration for South China. *Water,* 7(6), 3083-3102. doi:10.3390/w7063083.

Yang, H., Piao, S., Zeng, Z., Ciais, P., Yin, Y., Friedlingstein, P., Sitch, S., Ahlström, A., Guimberteau, M., Huntingford, C., Levis, S., Levy, P. E., Huang, M., Li, Y., Li, X., Lomas, M. R., Peylin, P., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Zhao, F., and Wang, L. (2015). Multicriteria evaluation of discharge simulation in dynamic global vegetation models. *Journal of Geophysical Research-Atmospheres*, 120, 7488–7505. doi: 10.1002/2015JD023129.

Zaitchik, B. F., Rodell, M., and Reichle, R. H. (2008). Assimilation of GRACE terrestrial water storage data into a land surface model: results for the Mississippi River basin. *Journal of Hydrometeorology*, 9, 535–548. https://doi.org/10.1175/2007JHM951.1.

Zeweldi, D. A., and Gebremichael, M. (2009). Evaluation of CMORPH precipitation products at fine space–time scales. *Journal of Hydrometeorology*, 10(1), 300-307. https://doi.org/10.1175/2008JHM1041.1.

Zhang, Y., Hong, Y., Wang, X., Gourley, J. J., Gao, J., Vergara, H. J., and Yong, B. (2013). Assimilation of passive microwave streamflow signals for improving flood forecasting: A first study in Cubango River Basin, Africa. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(6), 2375-2390. DOI: 10.1109/JSTARS.2013.2251321.

Zhao, F., Zhang, L., Chiew, F. H., Vaze, J., and Cheng, L. (2013). The effect of spatial rainfall variability on water balance modelling for south-eastern Australian catchments. *Journal of Hydrology*, 493, 16-29. https://doi.org/10.1016/j.jhydrol.2013.04.028.

Zoccatelli, D., Borga, M., Viglione, A., Chirico, G. B., and Blöschl, G. (2011). Spatial moments of catchment rainfall: rainfall spatial organisation, basin morphology, and flood response. *Hydrology and Earth System Sciences*, 15(12), 3767-3783. https://doi.org/10.5194/hess-15-3767-2011.

Summary

During the last decades, the number and diversity of water related challenges have significantly increased. Moreover, climate change has intensified these challenges limiting water resources available in many river basins around the world. Hydrological models in combination with in situ hydro-meteorological information can be used to overcome these challenges. However, many regions worldwide are poorly gauged or ungauged with limited or non-available in situ data. Recently, several scientific initiatives have emerged aiming to provide global datasets as a promising complement and/or alternative to in situ datasets. These globally available datasets can be integrated into hydrological models to improve water resources management at the river basin scale. To this end, several integration techniques exist which can be classified in three groups: (i) meteorological datasets, (ii) model parameters estimation and data assimilation and (iii) model evaluation. The theoretical background of these integration techniques is present, however, researchers are often hampered by a lack of knowledge about their practical implementation at the river basin scale. In this context, the main objective of this thesis was placed: *to evaluate the applicability of global water resources datasets (including satellite-derived observations, in situ data and models) for hydrological modelling at the river basin scale.*

In view of the complexity of this goal, this research focused on applying and testing techniques that can potentially be implemented in any river basin around the world, depending on its hydro-meteorological characteristics and in situ data richness. Four river basins were selected (the Brahmaputra basin in Bangladesh, the Magdalena-Cauca basin in Colombia, the Oum Er Rbia basin in Morocco and the Murrumbidgee basin in Australia) and different approaches were applied to optimally integrate global datasets into large- and local-scale hydrological models, starting with those on model application (Chapters 2 and 3), moving to those on model parameters estimation and calibration (Chapter 4) and ending with those on data assimilation (Chapters 5 and 6).

In all chapters of this thesis, several global meteorological datasets (based on satellite-derived observations, in situ data and/or models), including EI, WFDEI, MSWEP, TRMM, CMORPH and PERSIANN, were evaluated by comparison to in situ data and used to force large- and local-scale hydrological models at the four selected river basins. Results of this thesis showed that good quality meteorological datasets, particularly precipitation, are crucial for accurate model simulations. Moreover, in spite of biases detected in some areas and periods, global meteorological datasets were useful to estimate water resources at the river basin scale, using either large- or local-scale hydrological models.

Five different large-scale hydrological models (HTESSEL, LISFLOOD, PCR-GLOBWB, SURFEX-TRIP and WATERGAP3) were evaluated for water resources and flood estimation in the Brahmaputra River basin, by comparing their discharge model simulations with those obtained with a locally calibrated hydrological model (Chapter 2). Results of this thesis showed that discharge simulations by the large–scale hydrological models were similar to those provided by the local-scale model and a fairly reasonable agreement was found between modelled and observed discharge values. In addition, in spite of magnitude

differences, all hydrological models could give acceptable estimates of peak flows. These results were used for understanding the potentialitites and limitations of large-scale hydrological models and globally available datasets for watere resources management at the river basin scale.

Various global precipitation products (MSWEP, TRMM, CMORPH and PERSIANN) were downscaled in combination with in situ precipitation measurements using a geographically weigthed regression algorithm with auxiliary information from vegetation response, elevation, aspect and slope. The impact of spatial resolution of precipitation on discharge model simulations was evaluated in the Magdalena-Cauca River basin (Chapter 3). Results of this thesis showed that the increase of precipitation spatial resolution improved discharge model estimates which emphasizes the significance of having accurate and precise meteorological datasets for hydrological modelling.

The large-scale hydrological model PCR-GLOBWB was calibrated using satellite-based soil moisture (ESA CCI) and evapotranspiration (GLEAM) observations at the Oum Er Rbia basin (Chapter 4). Results of this thesis showed that the combined use of both satellite-based observations allows calibrating parameters that represent different hydrological processes in the model, leading to considerable improvements in model simulations. These findings open up the possibility to use large-scale hydrological models driven with global meteorological datasets and calibrated with satellite-derived observations for water resources estimation at the river basin scale.

Satellite-based soil moisture (AMSR-E) and discharge (SGR-GFDS and SGR-MODIS) observations were assimilated into a large- (PCR-GLOBWB) and a local-scale (OpenStreams wflow-hbv) hydrological model using an Ensemble Kalman filter in the Murrumbidgee and Magdalena-Cauca basins and a total of sixteen assimilation scenarios combining in situ and satellite-based observations were investigated (Chapters 5 and 6). Results of this thesis showed that assimilating satellite-based observations improved discharge model simulations. These findings proved the potential of assimilating satellite-based soil moisture and discharge observations for improving estimations of hydrological models driven with global precipitation datasets at the river basin scale.

Although there is room for improvement, the results of this thesis show that globally available datasets can improve large- and local-scale hydrological model estimations at the river basin scale, which is especially relevant in data-poor regions where water management is constrained by a lack of in situ information.

Samenvatting

Gedurende de laatste decennia is het aantal en de diversiteit van water gerelateerde uitdagingen aanzienlijk toegenomen. Bovendien heeft klimaatverandering deze uitdagingen geïntensiveerd waardoor de beschikbare watervoorraden in veel stroomgebieden over de hele wereld worden beperkt. Hydrologische modellen in combinatie met lokaal beschikbare hydrologische en meteorologische informatie kunnen worden gebruikt om deze uitdagingen te analyseren en mogelijke oplossingen te ontwikkelen. Een probleem hierbij is dat in veel stroomgebieden de lokale gegevens ontbreken doordat er weinig tot niet gemeten wordt. Onlangs zijn verschillende wetenschappelijke initiatieven opgestart om wereldwijde datasets te ontwikkelen die gebruikt kunnen worden als lokale gegevens ontbreken. Deze wereldwijd beschikbare datasets kunnen worden geïntegreerd in hydrologische modellen om het waterbeheer op de schaal van het stroomgebied te verbeteren met behulp van een verscheidenheid aan technieken. De theoretische achtergrond van deze integratietechnieken is beschikbaar, maar door een gebrek aan kennis is de praktische toepassing op stroomgebiedsschaal beperkt. Dit proefschrift heeft tot doel het evalueren en testen van wereldwijde hydrologische datasets en integratietechnieken voor hydrologische modellering van stroomgebieden.

Acknowledgements

First and foremost, I would like to deeply thank Jaap Schellekens, Geert Sterk and Marc Bierkens for offering me the opportunity to carry out this PhD research in collaboration with Deltares and Utrecht University. Your guidance and scientific advice was invaluable. Thank you for letting me being part of such an amazing project and community in hydrological modelling and Earth observations. I am really grateful for your constant support, encouragement and guidance, which have made me grow as a researcher.

Special thanks to all members of the evaluation committee. It is a great honour to have all of you assessing this PhD work. Thanks to all the scientific experts who reviewed the publications included in this thesis helping me to improve my research.

I gratefully acknowledge Deltares and Utrecht University for funding this research project, including all the trips, conferences and publications, through the European Union Seventh Framework Programme, eartH2Observe project (Global Earth Observations for integrated water resources assessment). Being part of the eartH2Observe project has given me the opportunity to work with great and amazing experts from all around the world. Thanks to the incredible team of the Water Resources Research Group of National University of Colombia (GIREH) in Bogotá, especially to Erasmo, Inés, Albeiro, David, Nicolás and Pedro Felipe. Special thanks to Stefan, Mira, Mohammed, Vinay and Gianni at the International Centre for Agricultural Research in the Dry Areas (ICARDA) for their great work and for making me feel part of the "Morocco team" from the beginning. Thanks to the people at the Commomwealth Scientific and Industrial Research Organisation (CSIRO) and the Australian National University (ANU) in Canberra, Australia, who hosted me for a couple of months and made me feel welcome during my stay: Luigi, Albert, Juan Pablo and Marta. Thanks to the Insitute of Water Modelling (IWM) in Dhaka, Bangladesh, for their great engagement and collaboration during this work, especially to Liza, Saleh, Sohel Masud and Abdulla Hel Kafi. You took me on a boat trip crossing the Brahamaputra River and that is an experience that I will never forget. I had a very enjoyable time working with you all and there are no words that can express my huge gratitude for letting me learn from you.

The contribution of a number of colleagues of Deltares and Utrecht University was crucial for the completion of this work. Special thanks to Frederiek, Albrecht, Walter, Rens and Micha for your insightful comments and suggestions, invaluable assistance and hepful discussions. I also want to thank Laurene, Edwin, Niko, Jannis, Joyce, Stefanie, Ruud, Marjolein, Yvonne and Anne for being my family at work.

Anouk and Alex, we have been "PhD warriors" during the past four years and that has meant a lot to me. Anouk, we started this PhD trip almost at the same time and since then we have been travelling together sharing all our adventures. Thank you especially for our time together in Australia and for not thinking I was crazy when I jumped to pick my phone from the sea! Thank you also for coming to Deltares, I was very happy working with you and I have learned a lot from you. Alex, your enthusiasm, motivation and ideas gave me energy during the last part of the PhD. I will always remember our talks at the Unesco-

IHE canteen about the “added value of hydrological information”, especially those moments when I could not understand what you were talking about! Thanks Alex.

My time in The Netherlands was greatly enjoyable mostly due to many friends that became into my family abroad. Special thanks to my housemates, Marco, Panos, Andrei, Edwin, Petra, Arun, Fieke, Xrisa and Becca. Living with you during the first years of the PhD and knowing that you would be at home after work meant a lot to me. I will never forget our dinners together. Lydia, Laura and Tara, you have been my big support all these years. Your friendship is priceless and without you I can not imagine my time in The Netherlands. Thanks a lot. Thanks to the Flood Risk Management crew who has been always there for me: Duc, Katya, Shristi, Nilay, Richard, Siobbhan, Kate, Maria, Anthony, Atiqul and Tamim. Thanks to all my friends in The Netherlands, or wherever you are in the world, thank you for all the laughs and good moments: Nikeh, Robert, Tiaravanni, Juan Carlos, Miguel, Patricia, Alifta, Andrea, Leah, Alex Kerensky, Marianne, Joanne, Shabana, Lorinc, Neiler, Eskedar, Saffa, Hadi, Aki, Danish, Aaashish, Audry, Maurizio, Feroz, Zahrah, Vittorio, Milk, Okan, Siyuan and Clara.

To my best Spanish friends, Ana, Noelia and Vicky, thank you for your support and for always remaining by my side and making me feel very near in spite of the distance. During the last years, our lives have followed different paths, but we have always found the way to continue being together, sharing and enjoying our new experiences. That means a lot to me, thank you so much.

Mamá, papá y Fernan (y Rasta y Tuno!), vosotros siempre habéis estado ahí para apoyarme ya fuera por Skype o Whatsapp desde la distancia o viniendo a verme a cada uno de los sitios en los que he vivido en los últimos años. Cada día me habéis dado fuerzas para seguir y para creer en mí. Muchas gracias por todo el apoyo, los consejos y por cuidar siempre de mí, sin vosotros no habría podido hacerlo. Os quiero muchitooooo!!

Thanks to Pepe, my best friend and love. You have been there while I was spending hours talking about data assimilation and hydrological models. Thanks for listening to me and for your encouragement and support during the last years. Even in those moments when I was thinking to give up, you have always managed to cheer me up and to make me smile. Like you always say: no todo es tan malo! Thank you for every moment together. Te quiero millones!

About the author

Patricia López López was born on 2 November 1987 in Granada, Andalucía, Spain. In 2005 she started with her Bachelor of Science (BSc) in Civil Engineering (5-year degree) at Granada University. During the last year of her BSc, she worked on characterizing of biomass ashes and evaluating their application in mortars and self-compacting concrete as part of a National Research project funded by the Spanish Ministry of Science and Innovation.

She continued her studies with a Master of Science (MSc) in Flood Risk Management funded by the European Comission at four different European universities: Technical University of Dresden in Germany, IHE Delft Insitute for Water Education in The Netherlands, Technical University of Catalunya in Spain and University of Ljubljana in Slovenia. For her internship, she worked at Deltares, Delft in The Netherlands where she focused on the estimation, use and verification of predictive hydrological uncertainty in flood forecasting systems. She completed her study in 2013 with the thesis "Improving the estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales)".

In 2014 Patricia started as a PhD researcher at the Department of Physical Geography, Faculty Geosciences at Utrecht University in collaboration with the Catchment Hydrology Department at Deltares. She worked on the application and evaluation of global water resources datasets for hydrological modelling at the river basin scale. Her PhD research was carried out in the framework of the European Union project eartH2Observe (Global Earth Observations for integrated water resources assessment, FP7). During her PhD research, she collaborated with the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Canberra, Australia (with a 3 months internship); the Institute of Water Modelling (IWM) in Dhaka, Bangladesh; the National University of Colombia in Bogota, Colombia and the International Center for Agricultural Research in the Dry Areas (ICARDA) in Rabat, Morocco.

List of publications

Peer-reviewed publications

López López, P., Wanders, N., Schellekens, j., Renzullo, L. J. Sutanudjaja, E. H. and Bierkens, M. F. P. (2016). Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. *Hydrology and Earth System Sciences*, 20, 3059-3076, https://doi.org/10.5194/hess-20-3059-2016.

López López, P., Sutanudjaja, E. H., Schellekens, J., Sterk, G. and Bierkens, M. F. P. (2017). Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products. *Hydrology and Earth System Sciences*, 21, 3125-3144, https://doi.org/10.5194/hess-21-3125-2017.

López López, P., Weerts, A. H., Schellekens, J., Van Dijk, A. I. J. M. and Rodríguez Sandoval, E. A. Assimilating in situ gauged and satellite-based discharge observations for hydrological modelling. *Submitted to Water Resources Research*.

López López, P., Immerzeel, W. W., Rodríguez Sandoval, E. A., Sterk, G. and Schellekens, J. Impact of high spatial resolution precipitation on streamflow simulations. *Submitted to Frontiers in Earth Sciences.*

López López, P., Sultana, T., Abdulla Hel Kafi, Md., Shahadat Hossain, Md., Saeh Khan, A. and Sohel Masud, Md. Evaluation of a global ensemble of hydrological models for water resources and flood estimation in the Brahmaputra River basin. *Submitted to Water Resources Management.*

Conference abstracts and presentations

López López, P., Wanders, N., Sutanudjaja, E., Renzullo, L. J., Sterk, G., Schellekens, J. and Bierkens, M. F. P. Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. Oral presentation at 2015 General Assembly, EGU (European Geosciences Union), Vienna, Austria, 12-17 April.

López López, P., Wanders, N., Sutanudjaja, E. H., Renzullo, L. J., Sterk, G., Schellekens, J. and Bierkens, M. F. P. Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. Oral presentation at 2015 IAHR (International Association for Hydro-Environment Engineering and Research) World Congress, Den Haag, The Netherlands, 28 June-3 July.

López López, P., Weerts, A. H., Schellekens, J., Sterk, G., Kockx, A. C., de Jeu, R. and Van Dijk, A. I. J. M. The impact of assimilation of streamflow and downscaled satellite

soil moisture observations for hydrological forecasting. Oral presentation at 2015 MODSIM (International Congress on Modelling and Simulation), Gold Coast, Queensland, Australia, 29 November–4 December.

López López, P., Wanders, N., Schellekens, J., Renzullo, L. J., Sutanudjaja, E. and Bierkens, M. F. P. Improved large-scale hydrological modelling through the assimilation of streamflow and downscaled satellite soil moisture observations. Oral presentation at 2015 General Assembly, AGU (American Geosciences Union), San Francisco, California, 14-18 December.

López López, P., Strohmeier, S., Haddad M., Sutanudjaja, E., Karrou, M., Sterk, G., Schellekens J. and Bierkens, M. F. P. Application of Earth observation products for hydrological modelling of the Oum Er Rbia River basin. Oral presentation at 2016 General Assembly, EGU (European Geosciences Union), Vienna, Austria, 17-22 April.

López López, P., Strohmeier, S., Sutanudjaja, E., Haddad, M., Karrou, M., Sterk, G., Schellekens, J. and Bierkens, M. F. P. Calibration of hydrological models based on remotely sensed soil moisture and evapotranspiration. Oral presentation at 2016 General Assembly, AGU (American Geosciences Union), San Francisco, California, 12-16 December.