Evaluating the Contribution of Remote Sensing Data Products for Regional Simulations of Hydrological Processes in West Africa using a Multi-Model Ensemble

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Thomas Poméon
aus
Mönchengladbach, Deutschland

Bonn, Dezember 2018

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Bernd Diekkrüger
2. Gutachter: Prof. Dr. Karl Schneider
Tag der mündlichen Prüfung: 06.03.2019
Erscheinungsjahr: 2019

To my amazing wife Amélie, for her loving support and never-ending patience. Without you, this would not have been possible.

Acknowledgments

Words can hardly express my gratitude to all those who, through their contributions, large and small, have supported this work.

First and foremost, I would like to deeply thank my supervisor, Prof. Bernd Diekkrüger. He is a person of outstanding integrity whom I greatly admire. Always humble, he is consistently available to discuss scientific ideas or challenges, and easily motivated me to perform to the last. His guidance and advice were invaluable to me as well as this work and I owe him my utmost appreciation. I also thank Prof. Karl Schneider from the University of Cologne as well as Prof. Mariele Evers, Prof. Lothar Schrott and Prof. Jürgen Kusche from the University of Bonn, who have so kindly accepted to be part of the doctoral committee.

Furthermore, I would like to especially thank our project partners from the Institute of Geodesy of the University Bonn, Prof. Jürgen Kusche, Anne Springer and, from the HafenCity University Hamburg, Prof. Annette Eicker. It has been a pleasure working with you on this project for these last few years and I highly appreciate our fruitful project meetings and exchange of ideas and results. Additionally, I would like to acknowledge the great support of both Dr. Rohini Kumar from the Helmholtz Centre for Environmental Research in Leipzig and Dominik Jackisch from the Nanyang Technological University in Singapore, who have helped me immensely in preparing some of the manuscripts.

I owe my sincere thanks to my current and former colleagues at the Hydrology Research Group for the great times we shared and the inspiring exchanges we had. You have made this experience great. A special thanks goes out to Dr. Aymar Bossa, Dr. Yacouba Yira, Dr. Alexandre Danvi and Gero Steup, from whom I have learned so much during these last years. I also dearly thank Inken Rabbel, Dr. Felix op de Hipt, Mouhamed Idrissu, Kristian Näschen, Claudia Schepp, Geofrey Gabiri and Emmanuel Nkundimana for our fruitful discussions, but mostly for the extended coffee breaks, office-football sessions and generally for keeping me sane during stressful times.

During my time at the Hydrology Research Group, we had the pleasure of hosting a number of guest researchers, with whom I had very interesting discussions and pleasant moments in Bonn and West Africa. I would like to especially thank Dr. Djigbo Badou, Dr. Jean Hounkpè, Dr. Charlène Gaba, Dr. Djiby Sambou and Dr. Adama Touré.

Lastly, a big thank you goes out to the staff at the WASCAL Competence Center in Ouagadougou, Burkina Faso, who have done their utmost to facilitate the COAST project workshop in February 2018, especially Dr. Jérôme Tondoh, Dr. Aymar Bossa, Boukare Ouedraogo and Kamil Sanoussi.

This thesis has been funded by the German Research Foundation (Deutsche Forschungsgemeinschaft - DFG) under Grant No. DI443/6-1 through the COAST project. I am very grateful for their support.

Abstract

Water is a crucial resource for human health, agricultural production and economic development. This holds especially true in West Africa, where large parts of the population work as self-sustaining farmers. Accurate knowledge of available water resources is therefore essential to properly manage this valuable commodity. Hydrologic modeling is seen as a key aspect in generating predictions of available resources. However, the overall availability of *in situ* data for model parametrization in West Africa has been steadily declining since the 1990s. When observations are available, they often contain errors and gaps. This lack of data severely hinders the application of hydrologic models in the region. Nowadays, many global and regional remote sensing and reanalysis data products exist which may be used to overcome these problems. A thorough analysis of the contribution of these products to regional simulations of hydrologic processes in West Africa has so far not been conducted. The purpose of this study is to close this gap. The study area spans from 3 to 24° latitude and -18 to 16° longitude and encompasses, among others, the Niger, Volta, and Senegal river basins. This study focuses on three key aspects, namely how the performance of remotely sensed and reanalyzed products can be validated without the availability of *in situ* data for the region; to what extent semi-distributed hydrologic models of the region can be parameterized and validated using these data; and how a fully distributed, grid-based model can be set up, calibrated and validated for sparsely-gauged river basins using multivariate data inputs.

Comparisons of remote sensing and reanalysis precipitation products for the region show strong variability. A hydrologic evaluation was conducted, during which the skill of each precipitation dataset to accurately reproduce observed streamflow in HBV-light simulations was tested. Best results are achieved by products which include satellite infrared and microwave measurements as well as bias-correction based on *in situ* observations. Averaged Nash-Sutcliffe Efficiencies (NSE) of 0.66 were reached during the calibration of the CMORPH CRT and PERSIANN CDR products over six subbasins.

In a next step, three SWAT models were set up for the region using multiple remote sensing and reanalysis data products and then calibrated and validated against observed river discharge with global and local approaches. While streamflow results differ within models and model regions, they are mostly satisfactory with coefficient of determination (R^2) values of 0.52 and 0.51 for calibrations and 0.63 and 0.61 for validations. In a multivariate validation framework, the skill of the model in simulating variables not included in the calibration is further evaluated against remote sensing observations of actual evapotranspiration, soil moisture dynamics, and total water storage anomaly. Here, it has been shown that the models perform robustly and reach a good agreement in relation to observations.

Furthermore, the grid-based mHM model was applied to several river basins in the south of the study area. After the quality of precipitation and evapotranspiration inputs was tested, a multivariate calibration was conducted. Models were calibrated using discharge observations (Q) and, to further constrain model boundary conditions, discharge in combination with remote sensing actual evapotranspiration observations (Q/ET). Finally, the quality of the simulations was tested against streamflow data as well as against remote sensing actual evapotranspiration, soil moisture, and total water storage anomaly data. Streamflow simulations performed well with averaged Kling-Gupta Efficiencies (KGE) of 0.53 for the first (Q) and 0.49 for the second (Q/ET) calibration. Further variables tested during the multiobjective validation were within good predictive ranges, especially during the Q/ET calibration. When SWAT and mHM model results are compared against each other and against external data products, results show that while both models perform robustly, mHM predictions outperform SWAT results.

This study furthers the understanding of the contribution of remote sensing, reanalysis and global data products in regional simulations of hydrologic processes in West Africa. Specific modeling strategies and routines were developed to further increase predictive capabilities of hydrologic models of the region using these freely-available datasets.

Zusammenfassung

Wasser ist ein fundamentaler Rohstoff, dessen Verfügbarkeit die menschliche Gesundheit, agronomische Produktivität and ökonomische Entwicklung beeinflusst. Dies trifft besonders in Westafrika zu, wo große Teile der Bevölkerung von Subsistenzlandwirtschaft leben. Genaue Kenntnisse über verfügbaren Wasserressourcen sind daher unverzichtbar, um eine nachhaltige Bewirtschaftung sicherzustellen. Hydrologische Modellierung wird allgemeinhin als ein wichtiger Faktor betrachtet, um verfügbare Ressourcen zu prognostizieren. Allerdings nimmt die Dichte der *in situ* gemessenen Daten zur Modellparametrisierung in West Afrika seit den 1990ern kontinuierlich ab. Zusätzlich weisen verfügbare Daten teils große Unsicherheiten und Lücken auf. Diese mangelhafte Datenverfügbarkeit erschwert die Anwendung von hydrologischen Modellen in der Region stark. Heutzutage existieren zahlreiche Fernerkundungs- und Reanalysedaten, die genutzt werden können, um das Datenverfügbarkeitsproblem zu umgehen. Allerdings fehlt eine tiefgreifende Analyse, welchen Beitrag diese Datensätze in regionalen Simulationen der hydrologischen Prozesse in Westafrika leisten können. Ziel der vorliegenden Studie ist es, diese Lücke zu schließen. Das Untersuchungsgebiet erstreckt sich von 3 bis 24° geographischer Breite und -18 bis 16° geographischer Länge und umfasst, unter anderem, die Flusseinzugsgebiete von Niger, Volta und Senegal. Drei Kernaspekte liegen im Fokus dieser Studie: Zunächst wird untersucht, wie die Leistung von Fernerkundungs- und Reanalysedatensätzen ohne Zuhilfenahme von *in situ* Daten validiert werden kann. Weiterhin wird ein semidistributives hydrologisches Modell mit diesen Datensätzen parametrisiert und validiert. Darauffolgend wird ein Rastermodell unter Zuhilfenahme von Fernerkundungsdaten aufgesetzt, kalibriert und validiert.

Der Vergleich von Fernerkundungs- und Reanalyseniederschlagsprodukten zeigt eine große Variabilität für die Region auf. Eine hydrologische Evaluierung wurde durchgeführt, wobei die Stärke der Produkte, gemessenen Abfluss in HBV-light Simulationen zu reproduzieren, begutachtet wurde. Es zeigte sich, dass Satellitenprodukte, welche Infrarot- und Mikrowellendaten mit Bodenobservationen vereinen, am besten abschneiden. CMORPH CRT und PERSIANN CDR Produkte erreichen eine über die Teileinzugsgebiete gemittelte Nash-Sutcliffe Effizienz (NSE) von 0.66.

In einem nächsten Schritt wurden drei SWAT Modelle für die Region unter Verwendung verschiedener Datenprodukte aufgesetzt und gegen beobachteten Abfluss in zwei verschiedenen Ansätzen (global und lokal) kalibriert und validiert. Wärend die Abflusssimulationen in der Qualität zwischen Modellen und Regionen variieren, sind die Ergebnisse größtenteils akzeptabel mit einem Bestimmtheitsmaß (R^2) von 0.52 und 0.51 für die Kalibrierung und 0.63 und 0.61 für die Validierung. In einer multivariaten Validierung wurden Modellergebnisse der Variablen tatsächliche Evapotranspiration, Bodenfeuchtedynamik und absolute Wasserspeicheränderung gegen Fernerkundungsprodukte verglichen. Hier wurde gezeigt, dass das Modell diese Variablen robust simuliert und gute Ergebnisse erzielt.

Weiterhin wurde das gegitterte mHM Modell auf Flusseinzugsgebiete im südlichen Untersuchungsgebiet angewendet. Zunächst wurde die Qualität von Niederschlags- und Evapotranspirationseingangsdaten bewertet. Darauffolgend wurde ein multivariater Kalibrieransatz angewandt, bei dem zunächst nur nach Abfluss (Q) und nachfolgend nach Abfluss und tatsächliche Evaporation aus Satellitendaten (Q/ET) kalibriert wurde. Die Qualität der Simulationen wurde gegen Abfluss, tatsächliche Evapotranspiration, Bodenfeuchteanomalie and totale Wasserspeicheranomalie evaluiert. Die Abflussimulationen weisen eine hohe Übereinstimmung mit beobachteten Daten auf. Während der Kalibrierung wurden gemittelte Kling-Gupta Effizienzen (KGE) von 0.53 (Q) und 0.49 (Q/ET) erreicht. Die erweiterte Modellvalidierung während der multivariaten Analyse hat gezeigt, dass Variablen, welche nicht in der Kalibrierung inkludiert waren, ebenfalls eine hohe Übereinstimmung mit Fernerkundungsdaten aufweisen. Dies ist besonders unter der

Q/ET Kalibrierung ersichtlich. Im direkten Vergleich von SWAT und mHM Modellergebnissen und externen Datensätzen zeigt sich, dass während beide Modelle gute Übereinstimmungen gegenüber externen Daten aufweisen, mHM Vorhersagen generell eine höhere Übereinstimmung als SWAT Ergebnisse aufweisen.

Diese Studie trägt zum tieferen Verständnis davon bei, inwiefern Fernerkundungs-, Reanalyse- und globale Datensätze hydrologische Vorhersagen in Einzugsgebieten mit spärlicher Datenverfügbarkeit verbessern können. Im Zuge dessen wurden spezifische Modellierungsstrategien entwickelt, um die Vorhersagekraft bestehender Modelle durch die Nutzung frei verfügbarer Datensätze weiter zu erhöhen.

Contents

List of Figures

List of Tables

List of Abbreviations

95PPU	95 Percent Prediction Uncertainty
AMMA	African Monsoon Multidisciplinary Analysis
AQUASTAT	FAO's Global Water Information System
ASP	Aspect
CATCH	Couplage de l'Atmosphère Tropicale et du Cycle Hydrologique
CCI	Climate Change Initiative
CDC	Climate Data Center
CFSR	Climate Forecast System Reanalysis
CHIRPS	Climate Hazards Group Infrared Precipitation with Station Data
CILSS	Comité permanent Inter-Etats de Lutte contre la Sécheresse dans le Sahel
CLM	Community Land Model
CMORPH	Climate Prediction Center Morphing Technique
CMORPH CRT	CMORPH (Satellite with bias correction)
CMORPH RAW	CMORPH (Satellite only)
COAST	Studying changes of sea level and water storage for coastal regions in West Africa using satellite and terrestrial data sets
CPC	Climate Prediction Center
DDS	Dynamically Dimensioned Search
DEM	Digital Elevation Model
DFG	German Research Foundation
DWD	German Meteorological Agency
EAWAG	Swiss Federal Institute of Aquatic Science and Technology
ECOWAS	Economic Community of West African States
EPSAT-SG	Estimation of Precipitation by SATellite Second Generation
ESA	European Space Agency
ETA	Actual Evapotranspiration
ETP	Potential Evapotranspiration
FAO	Food and Agriculture Organization of the United Nations
GAP	Genetic Algorithm and Powell
GDP	Gross Domestic Product
GIS	Geographic Information System
GLDAS	Global Land Data Assimilation System
GLEAM	Global Land Evaporation Amsterdam Model
GPCC	Global Precipitation Climatology Center
GPCP	Global Precipitation Climatology Project
GR4J	Modèle du Génie Rural à 4 Paramètres Journalier
GRACE	Gravity Recovery and Climate Experiment
GranD	Global Reservoir and Dam database
GRDC	Global Runoff Data Center
GSMAP	Global Satellite Mapping of Precipitation
GTS	Global Telecommunication System
GW	Groundwater

GWSP	Global Water Systems Project
HAR	Hargreaves
HBV	Hydrologiska Byråns Vattenavdelning
HRU	Hydrologic Response Unit
HWSD	Harmonized World Soil Database
HydroSHEDS	Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales
IPCC	Intergovernmental Panel on Climate Change
ISI-MIP	Inter-Sectoral Impact Model Intercomparison Project
ITCZ	Intertropical Convergence Zone
IUSS	International Union of Soil Sciences
IWRM	Integrated Water Resource Management
JRC	European Commission's Joint Research Centre
KGE	Kling-Gupta Efficiency
LAI	Leaf Area Index
LSM	Land Surface Model
LULC	Land Use and Land Cover
MCMC	Monte Carlo Markov Chain
MERRA	Modern-Era Retrospective analysis for Research and Applications
mHM	mesoscale Hydrologic Model
MODIS	MODerate-resolution Imaging Spectroradiometer
MOSAIC	Mosaic Land Surface Model
MPR	Multiscale Parameter Regionalization
MSWEP	Multi-Source Weighted-Ensemble Precipitation
NASA	National Aeronautics and Space Agency
NCEP	National Centers for Environmental Prediction
netCDF	Network Common Data Format
NOAA	National Oceanic and Atmospheric Administration
NSE	Nash-Sutcliffe Efficiency
PBIAS	Percent Model Bias
PERSIANN CDR	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record
Q	Discharge
Q/ET	Discharge and Actual Evapotranspiration
r	Pearson's Correlation Coefficient
R^2	Coefficient of Determination
RA	Reanalysis
RFE	African Rainfall Estimation
RG	Rain Gauge
RMSE	Root Mean Squared Error
SA	Simulated Annealing
SCE	Shuffled Complex Evolution
SCS CN	Soil Conservation Service Curve Number
SM	Soil Moisture
SRFE	Satellite-Based Rainfall Estimate
SRTM	Shuttle Radar Topographic Mission
SUFI	Sequential Uncertainty Fitting
SWAT	Soil and Water Assessment Tool
SWAT CUP	SWAT Calibration and Uncertainty Procedures
SWAT GRID	Grid-Based SWAT Application

TAMSAT	Tropical Applications of Meteorology using Satellite data and ground-based observations
TMPA RT	TRMM-TMPA Real Time (Satellite only)
TRMM-TMPA	Tropical Rainfall Measuring Mission - Multi-Satellite Precipitation Analysis
TWS	Total Water Storage
UNESCO	United Nations Educational Scientific and Cultural Organization
UNICEF	United Nations International Children's Emergency Fund
USDA	United States Department of Agriculture
USGS	United States Geological Survey
VIC	Variable Infiltration Capacity Macroscale Hydrological Model
WGHM	Water GAP Global Hydrology Model
WMO	World Meteorological Organization
WRB	World Reference Base for Soil Ressources
WWF	World Wildlife Fund

1 General Introduction

In the following chapter, a general introduction of the study will be presented. First, the underlying problem statement will be discussed and a brief overview of key studies motivating the research will be given. Following this, the arising research questions will be detailed, as well as the objectives and structure of the study. More comprehensive discussions of the research topics and state of the art can be found in the introduction sections to each of the main study chapters 4, 5, 6 and 7.

This study is part of the COAST project (*Studying changes of sea level and water storage for coastal regions in West Africa using satellite and terrestrial data sets*) of the University of Bonn, supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) under Grant No. DI443/6-1.

1.1 Problem Statement

The climate of West Africa is variable and expected to be highly influenced by climate change. The Intergovernmental Panel on Climate Change (IPCC) predicts a temperature increase of between one and two degrees for West Africa until 2050, depending on the representative concentration pathway (CO_2-equivalent greenhouse gas concentrations) scenario (Oldenborgh et al. 2013). It is expected that this will have a strong impact on the hydrologic cycle and freshwater availability (Bormann and Diekkrüger 2004). Many West African countries already frequently face drought and water stress (Mishra and Singh 2010; Shanahan et al. 2009) and studies have shown droughts to increase in both frequency and severity (Kasei et al. 2010; Masih et al. 2014).

In West Africa, the majority of the labor force works as self-sustaining farmers and generates income by selling surpluses (African Development Bank 2018; Hollinger and Staatz 2015; Jalloh et al. 2013). Since the irrigation infrastructure is poorly developed (FAO 2018), these populations are highly at risk, with limited water availability threatening both food and economic security (Schuol et al. 2008a). Integrated water resource management (IWRM) is seen as a method to ensure sustainable development and to alleviate water crises by managing this valuable resource. One key principle of IWRM as recommended by the Dublin Conference in 1992 is a participatory management approach involving users, planners and policymakers (Rahaman and Varis 2005). However, the quantification of available water resources required to accurately plan water management strategies remains problematic due to limitations of measurement techniques and data availability. Hydrologic models are important tools in this regard and can be used to generate predictions of water resources (Beven 2012). However, many of the West

African river basins are poorly-gauged, with the availability of *in situ* data further declining in recent years (Adjei et al. 2012; Hughes 2006). Even when observations are available, they may include erroneous data or large gaps (Behrangi et al. 2011; Bitew and Gebremichael 2011; Koutsouris et al. 2016). At a small-scale, appropriately funded research projects can establish their own measurement networks to suit modeling needs (e.g. for the modeling of inland valleys in central Benin (Danvi et al. 2017), climate change impact modeling in western Burkina Faso (Yira et al. 2017), or the modeling of hydrologic processes in a headwater catchment in central Benin (Giertz et al. 2006)). At the meso- or large scale, this is not possible and hydrologic modeling exercises are severely hindered by the limited data availability. This lack of observations has even led to the use of statistical weather generators in the region (Schuol and Abbaspour 2006b).

Validating Remotely-Sensed Precipitation Data in Sparsely-Gauged Catchments

Remote sensing is seen as a way to overcome some of the aforementioned problems. While several studies use remote sensing observations of parameters such as elevation or leaf area index (Faramarzi et al. 2010; Schuol et al. 2008a) to parameterize their hydrologic models, the use of remotely-sensed climatological observations is also becoming more pronounced, with many products being available free of charge. However, due to the sometimes large variability between different remote sensing products and limited possibility of validating estimates against ground-based measurements, results may be heavily biased (Thiemig et al. 2013; Tramblay et al. 2016). It is therefore necessary to assess the predictive capabilities of different products and to choose the most appropriate dataset for the study area. Many authors have evaluated the performance of remote sensing precipitation datasets for regions all over the world (Fujihara et al. 2014; Koutsouris et al. 2016; Cohen Liechti et al. 2012; Tobin and Bennett 2014), with some recent articles focusing on the West African domain (Awange et al. 2015; Gosset et al. 2013; Nicholson et al. 2003; Pfeifroth et al. 2016; Thiemig et al. 2012). However, the question arises how performances can be assessed without the availability of *in situ* data. Considering the data scarcity in large parts of the world, it has been suggested that precipitation products may be evaluated in a hydrologic modeling framework, where the performance of the product is assessed based on its ability to accurately simulate streamflow. For this method, streamflow observations must be available. While multiple studies focus on this topic (Behrangi et al. 2011; Bitew and Gebremichael 2011; Bodian et al. 2016; Cohen Liechti et al. 2012; Dile and Srinivasan 2014; Fujihara et al. 2014; Li et al. 2015), only Thiemig et al. (2013) have attempted a hydrologic evaluation for the West African region. In their study, the authors calibrate a spatially-distributed, physically-based hydrologic model using four satellite precipitation- and one reanalysis product. They conclude that remote sensing and reanalysis data can be very useful for hydrologic modeling purposes, especially for water budget calculations. While degraded performance was observed for low flow conditions, this might be attributed to errors in the model structure. The authors stress the fact that precipitation products need to be evaluated for the region (either classically against gauge data or hydrologically using a model) before being used for further modeling purposes. Since new remote sensing data products are continually emerging and older products are improved, an overview of the performance of the most popular precipitation datasets for the

region is currently lacking.

Improving Hydrologic Model Parametrization, Calibration and Validation using Remote Sensing Data

In recent years, there has been a trend towards the use of hydrologic models of ever-increasing complexity (Samaniego et al. 2010). However, particularly in data-scarce catchments, increases in model complexity do not necessarily lead to better simulation results. Multiple very comprehensive studies have been undertaken by Schuol et al. in setting up the physically-based semi-distributed Soil and Water Assessment Tool (SWAT) (Arnold et al. 2012b; Srinivasan et al. 1998) in the West African study area (Schuol and Abbaspour 2006a, 2006b; Schuol et al. 2008a). In their work, the authors furthermore explore the utilization of a weather generator to counteract the lack of *in situ* observations of climatological parameters, as due to the historical period modeled in their study, no remote sensing data was available. After the model was set up, water availability was estimated for the entire West African subcontinent (Schuol et al. 2008a) and afterwards for the whole of Africa (Schuol et al. 2008b). The authors use a wide variety of global and also remote sensing data to parameterize their model, such as a digital elevation model, or a global soil information dataset. Especially spatially distributed models can be over-parameterized, leading to parameter equifinality, where an unlimited number of parameter combinations may lead to identical results (Orth et al. 2015; Samaniego et al. 2013; Samaniego et al. 2010). Hydrologic models are typically calibrated and validated using only discharge observations. While this allows for a confident prediction of runoff, the modeler cannot be certain as to the quality of the estimation of further hydrologic parameters like actual evapotranspiration or total water storage (Rakovec et al. 2016b; Rakovec et al. 2016a; Zink et al. 2018). It therefore becomes necessary to separately validate all variables of interest, which can be achieved by using readily-available remote sensing observations. This has been attempted, in parts, by Xie et al. (2012), who validated a SWAT model of sub-Saharan Africa using satellite total water storage observations. Due to the poor availability of streamflow data for the region, the authors used multi-year average monthly river discharge for the years 1900 to 2001 and a calibration period from 2002-2009. They acknowledge that due to the disagreement in periods of observational data, "[...] *it is difficult to evaluate the model's adequacy in simulating the surface water system*" (Xie et al. 2012). While Xie et al. compartmentalize total water storage deviations into storages such as soil moisture or groundwater storage, no further model results are validated against remote sensing data. The cited studies offer interesting possibilities, especially in light of using remote sensing and global data products to improve model parametrization. However, remote sensing observations of, e.g., actual evapotranspiration, soil moisture and total water storage, offer the potential to further validate the model performance of variables not included in the calibration.

Parameterizing, Calibrating and Validating a Grid-Based Hydrologic Model using Remote Sensing Inputs

As has been stated in the previous section, the increasing complexity of hydrologic models leads to increased equifinality while also becoming more computationally demanding. Especially fully-

and semi-distributed models suffer from this, as during calibration, parameters must be optimized for each modeling unit (i.e. model polygon or grid cell) (Pokhrel et al. 2008; Beven 1993). While a gridded version of the SWAT model exists (SWAT-GRID, Rathjens and Oppelt (2012)), it is unsuited for simulations of mesoscale basins due to its structure and computational constrains (Pignotti et al. 2017). A relatively new grid-based hydrologic model which has been tailored to use remote sensing and global input data is the mesoscale Hydrologic Model (mHM) (Samaniego et al. 2010; Kumar et al. 2013b). The model can be calibrated at a computationally efficient coarse resolution, after which it is scaled to produce outputs at a finer resolution using Multiscale Parameter Regionalization (MPR). mHM has been extensively applied in Europe and the USA (Kumar et al. 2010; Kumar et al. 2013a; Rakovec et al. 2016b; Zink et al. 2017; Samaniego et al. 2013; Thober et al. 2015; Hattermann et al. 2017). Due to the model structure, multivariate calibration and validation combining discharge observations with remote sensing data can be performed. The grid-based nature allows for model outputs to be easily compared against remotely-sensed datasets. mHM has further been applied for hydrologic predictions in a data-scarce basin in India (Samaniego et al. 2011), as well as for soil moisture quantifications (Samaniego et al. 2013), and has also been calibrated using satellite-observed total water storage data to improve partitioning of rainfall into runoff components (Rakovec et al. 2016a). Even though mHM has been applied for climate change impact assessment in the domain of the Upper Niger river (Hattermann et al. 2017), its performance in the West African domain remains largely unknown.

This study contributes to ongoing research efforts to better quantify available water resources of sparsely-gauged basins. Specifically, novel methodological approaches will be developed to validate remote sensing and global precipitation products without the need for *in situ* data, further confirm the performance of hydrologic models using remote sensing data, and set up a grid-based model for parts of the region using remote sensing data inputs and multivariate calibration and validation techniques.

1.2 Research Questions

Considering the problem statement and state of the art, several research questions arise to which this study aims to find answers. Principal and underlying research questions will be described below.

1. **How can the performance of remotely-sensed and reanalyzed precipitation data be assessed for the study area?**

Due to a lack of *in situ* measurements throughout the study area, validating the performance of remotely-sensed and reanalyzed precipitation datasets proves problematic. In this study, a hydrologic evaluation is proposed, where the performance of input data is assessed using a hydrologic modeling framework. This task can be further differentiated into two underlying questions, which will be discussed in chapter 4:

   a) How can differences between ten precipitation datasets be quantified for West African river basins of differing locations and sizes?

b) How robust are these products in accurately simulating streamflow during a hydrologic evaluation?

2. **How can a SWAT model of the study area be set up and multi-objectively validated using remote sensing observations?**

Several studies have explored the possibilities of setting up hydrologic models for parts of- or even the entire study area. The lack of measured data complicates this, with some studies relying on the use of statistical weather generators or an upscaling of locally observed data, which may introduce additional bias. Here, we first establish a routine of setting up the semi-distributed SWAT model using remote sensing and reanalysis inputs. In a second step, a multi-objective validation is conducted, where the model performance is evaluated against streamflow, as well as remotely-sensed actual evapotranspiration, soil moisture and total water storage data. The following underlying questions are addressed in chapter 5:

a) How can a hydrologic modeling framework for West Africa be set up using only freely available data?

b) Which simulation quality can be obtained using these datasets?

c) How can a multi-objective validation be performed?

d) How can the potential and limitations of this approach for assessing the water availability at the regional scale be evaluated?

3. **How can the grid-based mHM model be calibrated and validated in a multivariate framework for the region?**

SWAT is a semi-distributed model relying on the hydrologic response unit approach, during which geolocational information is lost. It was therefore decided to apply a fully-distributed, grid-based model in the region. Here, the mHM model is applied in the study area. The use of mHM is advantageous, as it has been especially developed for predictions at the mesoscale and is designed to work with common remote sensing input formats. mHM further allows for the transferability of parameters between different spatial scales, supporting calibration at a computationally efficient resolution while subsequent runs may be performed at finer resolutions. The model was set up using remotely-sensed and reanalysis data and again validated against streamflow and remote sensing actual evapotranspiration, soil moisture and total water storage estimates. The following underlying questions will be discussed in chapter 6:

a) How can the performance of the mHM model using a multiscale parameter regionalization approach be assessed?

b) How does the model perform under multivariate calibration inputs?

c) How can model parameters be transferred between different spatial scales using multiscale parameter regionalization?

4. **How well do SWAT and mHM simulations of key hydrologic variables compare against each other and against remote sensing, as well as global model results?**

For this last research question, model outputs of both SWAT and mHM will be extracted for an overlapping area and compared against each other and against external data. Selected variables include actual evapotranspiration, soil moisture anomaly, total water storage anomaly, groundwater recharge and water yield. This research question is discussed in chapter 7.

1.3 Objectives

Following the problem statement and research questions, the main objectives of this study will be outlined briefly. The principle objective is to increase the performance and predictive capabilities of select hydrologic models in the West African domain by implementing remote sensing observations for more rigorous model calibration and/or validation. Modeling frameworks developed in this study have been communicated to the scientific public through the publication of research articles and disseminated to West African partner organizations, as has already been done during a project workshop in Ouagadougou, Burkina Faso, in February 2018. In order to entice a maximum of researchers to reproduce our modeling approaches for their studies, all software and data used in this study are easily accessible and free of charge.

Specifically, this study aims to:

1. **Analyze and quantify uncertainties of remote sensing precipitation datasets over West Africa using a hydrologic evaluation framework.**

2. **Create a framework to set up and validate the semi-distributed hydrologic SWAT model for the study area using remotely-sensed and reanalyzed data.**

3. **Create a robust, computationally efficient, multivariate calibration and validation framework for the application of the grid-based mHM hydrologic model in the study area.**

4. **Assess the comparability of select SWAT and mHM outputs against each other and against external data.**

1.4 Structure of this Study

This study is structured into eight chapters, starting with the general introduction as chapter 1. In chapter 2, the study area will be described in detail, focusing on physical (topography, climate, geology, soils, hydrology and land use), as well as economic (agriculture, population and economy) aspects. An overview of the three hydrologic models used in this thesis (HBV-light, SWAT and mHM) and their routines is given in chapter 3. In chapter 4, the possibility of evaluating the performance of ten precipitation datasets using the HBV-light model is explored. Chapter 5 focuses on the multi-objective validation of multiple SWAT models set up for the study area using remote sensing datasets. Special attention is paid to the evaluation of discharge, actual evapotranspiration, soil moisture and total water storage performance. Chapter

6 presents a modeling framework for the computationally efficient multivariate calibration and validation of the grid-based mHM model. Results of the simulations using SWAT and mHM are compared in detail in chapter 7 and a general conclusion is given in chapter 8. Chapters 4, 5, and 6 have been previously published. The articles have been formatted to fit the style of this study. Corresponding citations shall be briefly presented:

Chapter 4: Poméon, T., Jackisch, D., Diekkrüger, B., 2017. Evaluating the performance of remotely sensed and reanalysed precipitation data over West Africa using HBV light. Journal of Hydrology. 547, 222–235. doi:10.1016/j.jhydrol.2017.01.055.
https://www.sciencedirect.com/science/article/pii/S0022169417300653

Chapter 5: Poméon, T., Diekkrüger, B., Springer, A., Kusche, J., Eicker, A., 2018. Multi-Objective Validation of SWAT for Sparsely-Gauged West African River Basins-A Remote Sensing Approach. Water 10, 22. doi:10.3390/w10040451.
http://www.mdpi.com/2073-4441/10/4/451

Chapter 6: Poméon, T., Diekkrüger, B., Kumar, R. 2018. Computationally Efficient Multi-Parameter Calibration and Validation of a Grid-Based Hydrological Model in Sparsely-Gauged West African River Basins. Water. 10, 26. doi:10.3390/w10101418.
https://www.mdpi.com/2073-4441/10/10/1418

2 Study Area

2.1 Location and Topography

The study area is located in West Africa, spanning from 3 to 24° latitude and -18 to 16° longitude. The distinction of West Africa from other regions is not sharp (Iloeje 1986; Morgan and Pugh 1969), with the term "West Africa" itself having been coined by Europeans during the colonial era (Church 1957; Morgan and Pugh 1969). It has been described as the "[...] *area lying west of the plateaux and peaks of the Cameroons and Adamawa, and south of the Sahara* [...]" and as "[...] *the combined area of the former British and French West Africa together with Portuguese Guinea, Liberia, the Cape Verde Islands, and the islands of Fernando Po, São Tomé and Principe and Annobon.*" (Morgan and Pugh 1969). Nowadays, West Africa is considered to consist of the 15 member states of the Economic Community of West African States (ECOWAS): Benin, Burkina Faso, Cape Verde, Ivory Coast, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Niger, Nigeria, Senegal, Sierra Leone, and Togo; as well as Mauritania. The region covers an area of five million km^2 (Jalloh et al. 2013). An overview is presented in Figure 2.1.

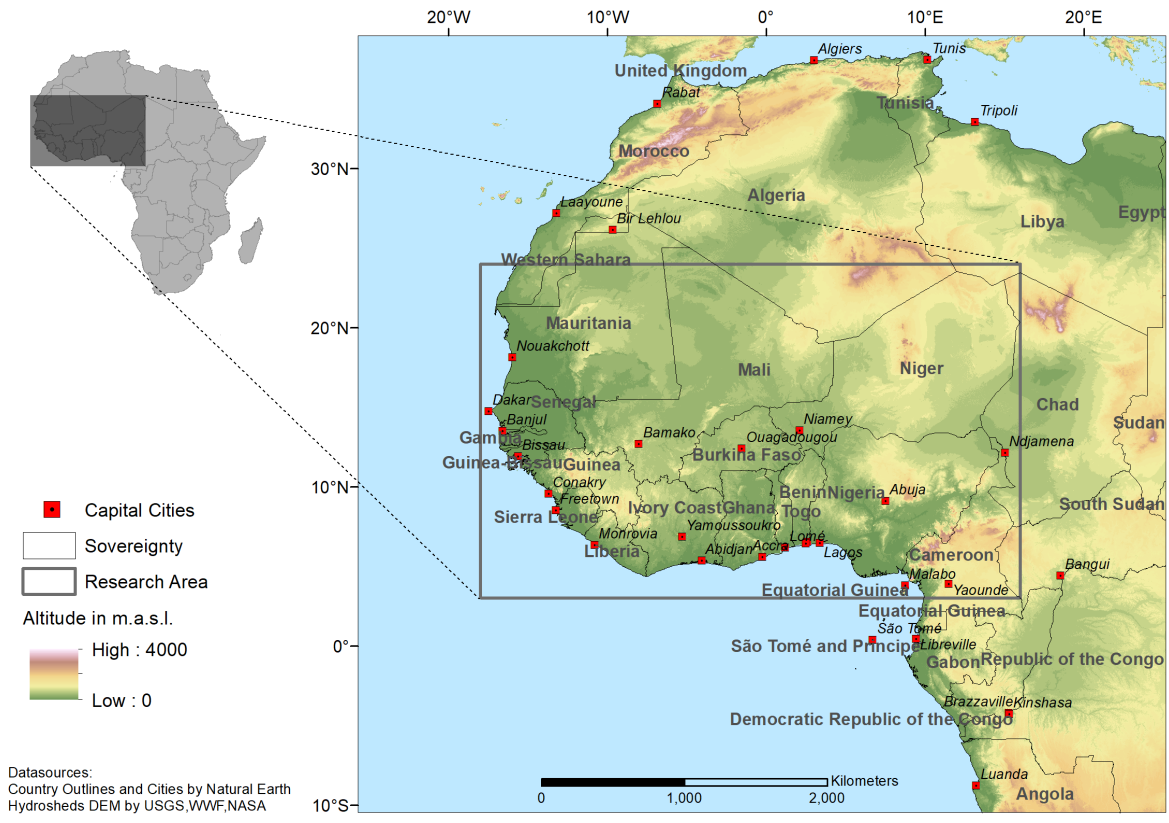

Figure 2.1: Study area.

This study focuses exclusively on continental West Africa without considering islands. Unlike other major regions of Africa, the relief is mainly flat and low, and does not interrupt zonal climate and vegetation patterns (CILSS 2016). Some mountainous regions do, however, exist and shall be briefly mentioned. The Fouta Djallon mountains in Guinea rise 1000 to 1500 m and are located in the path of southwesterly winds. Consequentially, precipitation is high and thus the mountains form an important source for several streams and rivers (Church 1957; Iloeje 1986). The Akwapim-Atakora range spans from southern Ghana (Akwapim Hills) through Togo (Togo Mountains) and into northern Benin (Atakora Mountains), and peaks can reach heights of almost 1000 m according to digital elevation model data (CILSS 2016; Lehner et al. 2008; Lehner et al. 2013). The Jos Plateau of central Nigeria is 1200 to 1500 m high, separated from the Hausaland High Plain by a 600 m scarp. Some Niger tributaries originate here (Iloeje 1986). In western Cameroon, the 4096 m tall active Mount Cameroon volcano is the highest peak in western Africa (Encyclopaedia Britannica 2018b). Towards northern Niger, the Aïr Mountains rise between 1000 and 1500 m over a length of 400 km. While rainfalls are higher than in the surrounding area, they are erratic, mostly falling in extreme events causing severe erosion. The drainage direction is westwards (Church 1957).

2.2 Climate

The climate of West Africa is strongly influenced by the prevailing northeast and southwest airstreams. While the northeasterlies are of dry continental origin, the southwesterlies originate in the Atlantic and are warm and humid. Seasonal changes between these airstreams produce a limited monsoon, leading to contrasting dry and wet seasons over the course of the year. In West Africa, the airstreams are convergent in the Intertropical Convergence Zone (ITCZ) located slightly north of the equator but moving south and north during the year. Because of their vapor pressure differential, the ITCZ forms a boundary between the humid and dry airmasses, with the dry, warmer air from the Sahara climbing over the denser, moister air originating in the Atlantic. This phenomenon is known as the Monsoon Trough (Buckle 1996). Areas north of 8° latitude experience a single annual rainfall maximum between July and August, while along the coast between Liberia and the Niger delta, an anomalous, short dry season prevails during this time. Further east towards the central Nigerian coast, this dry season is not experienced (Buckle 1996; Ojo 1977).

The climate in the region is diverse, spanning from equatorial monsoonal in the south to arid desert in the north. Figure 2.2 shows an updated Köppen and Geiger climate classification map for the study area based on the work of Kottek et al. (2006) and Rubel et al. (2017). Exemplary climate graphs are based on station data from the German Meteorological Office – Climate Data Center (DWD 2018) and the World Meteorological Organization (WMO 2018).

Equatorial monsoonal climate prevails along the southwest and southeast West African coast. Here, the climate is humid for most parts of the year. In Port-Harcourt, Nigeria, annual rainfalls can reach over 2300 mm and the temperature is very stable with a mean of 26.7°C. The climate in the south and up to central Burkina Faso is mostly equatorial winter dry. Examples here are Abidjan and Cotonou. Clearly visible is the previously described second dry period in August. Maximum rainfalls are reached in June during the primary, and October/November

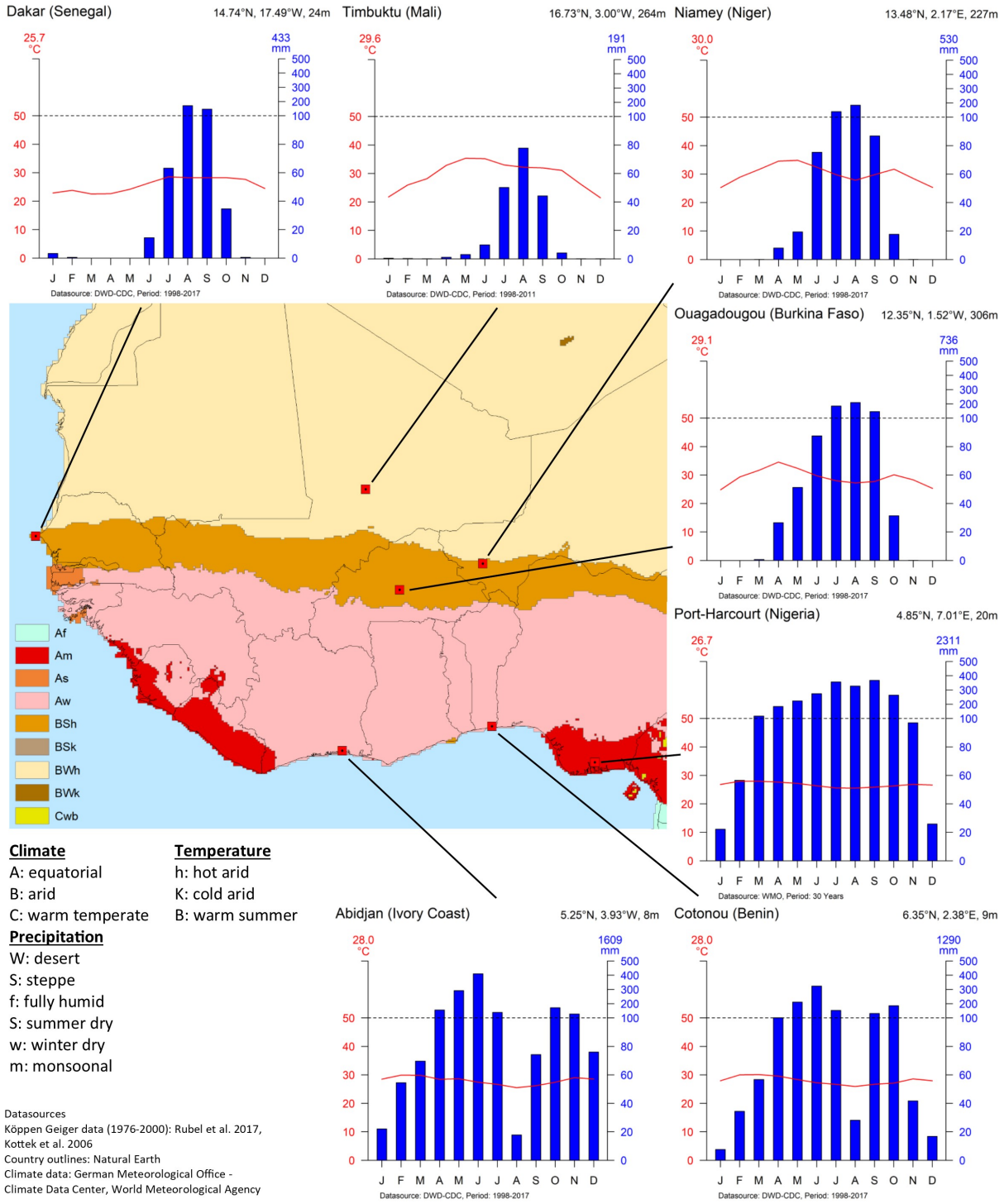

Figure 2.2: Köppen and Geiger climate classification of the study area with examples.

during the secondary rainy season. Rainfall sums vary by 300 mm between the two cities, with Abidjan reaching 1600 mm and Cotonou 1290 mm, while mean temperatures are equal at 28°C. Towards the north, precipitation decreases. The hot arid steppe zone covers central Senegal, southern Mali and northern Burkina Faso and Nigeria. Examples here are Dakar, Niamey and Ouagadougou. During the single rainy season, maximum precipitation falls in August in all cities. Out of the three, Ouagadougou has the highest annual precipitation with 736 mm, as opposed to 530 mm in Niamey and 433 mm in Dakar. Temperatures are similar for Ouagadougou

and Niamey with 29 and 30°C on average. Peak temperatures are reached in August, with lower temperatures following during the rainy season. A second peak is reached after the end of the rainy season in October. Temperatures in oceanic Dakar are lower with an average of 25.7°C, but rise during the second half of the year. To the north, the climate changes to hot arid desert. In the example of Timbuktu, Mali, annual rainfalls only reach 191 mm with a mean temperature of 29.6°C. The rainy season is short with maximum precipitation reached in August, and temperatures are high from April to October.

West Africa is prone to drought, with a period of severe and long-lasting drought beginning in the 1960s and reaching into the mid-1970s (Mishra and Singh 2010; Shanahan et al. 2009). From the 1980s onwards, droughts again occurred frequently (Mishra and Singh 2010). Changing sea surface temperatures have been identified as the main cause of the variability of the West African monsoon leading to lower rainfalls (Shanahan et al. 2009). Especially in a region where subsistence farmers rely on regular rainfalls for crop production, droughts can, and have in the past, affect the lives of millions of people (Fafchamps et al. 1998; Shanahan et al. 2009). Studies have shown West African droughts to increase in frequency, severity and area affected (Kasei et al. 2010; Masih et al. 2014). Besides impacting water availability and food production, droughts can also trigger further economic, societal, and environmental change (Gautier et al. 2016). Based on the analysis of 3000 years of sedimentation in Lake Bosumtwi, Ghana, it has been suggested that the multidecadal droughts faced since the 1960s are in fact not anomalous, indicating the risk for even longer and severe occurrences (Shanahan et al. 2009).

2.3 Geology

The West African basement complex consists mainly of Pre-Cambrian rocks, like the rest of the continent. Formations are folded and aligned in a north-south direction. Approximately one third of these rocks, which can be up to 3000 million years old, is exposed at the surface. In the other areas, as shown in Figure 2.3, the Pre-Cambrian basement complex is covered by later rock formations and sedimentary basins which formed during times when large parts of West Africa were covered by lakes and shallow seas (Ahn 1970; Church 1957; Schlüter 2006). The Pre-Cambrian basement complex is characterized by metamorphic rocks from earlier formations with later volcanic intrusions and consists mainly of schists, phyllites, quartzites, granites and gneisses. The origin of most minerals found in the region, such as diamonds, iron, chromium, manganese and gold, is associated with Pre-Cambrian materials (Ahn 1970). Evenly eroded Pre-Cambrian formations provided a level floor for the advance and retreat of shallow seas, which deposited the materials now covering the basement complex. Large areas, including much of the Sahara, were covered by sea during the Ordovician and later Silurian periods. During the Devonian, marine and continental phases alternated and in the Lower Carboniferous, the area was again covered by sea. Since the Eocene or Oligocene, continental conditions prevail over West Africa (Church 1957; Schlüter 2006). During the Secondary and Tertiary, intrusions and extrusions of volcanic outpourings, associated with tectonic uplift processes, further influenced the geology (Ahn 1970; Church 1957). Volcanic activities continued into the Quaternary, forming among others the still active Mount Cameroon (Church 1957). Erosion of sandstone is a key process during the Quaternary, producing vast quantities of sand, which was distributed by the

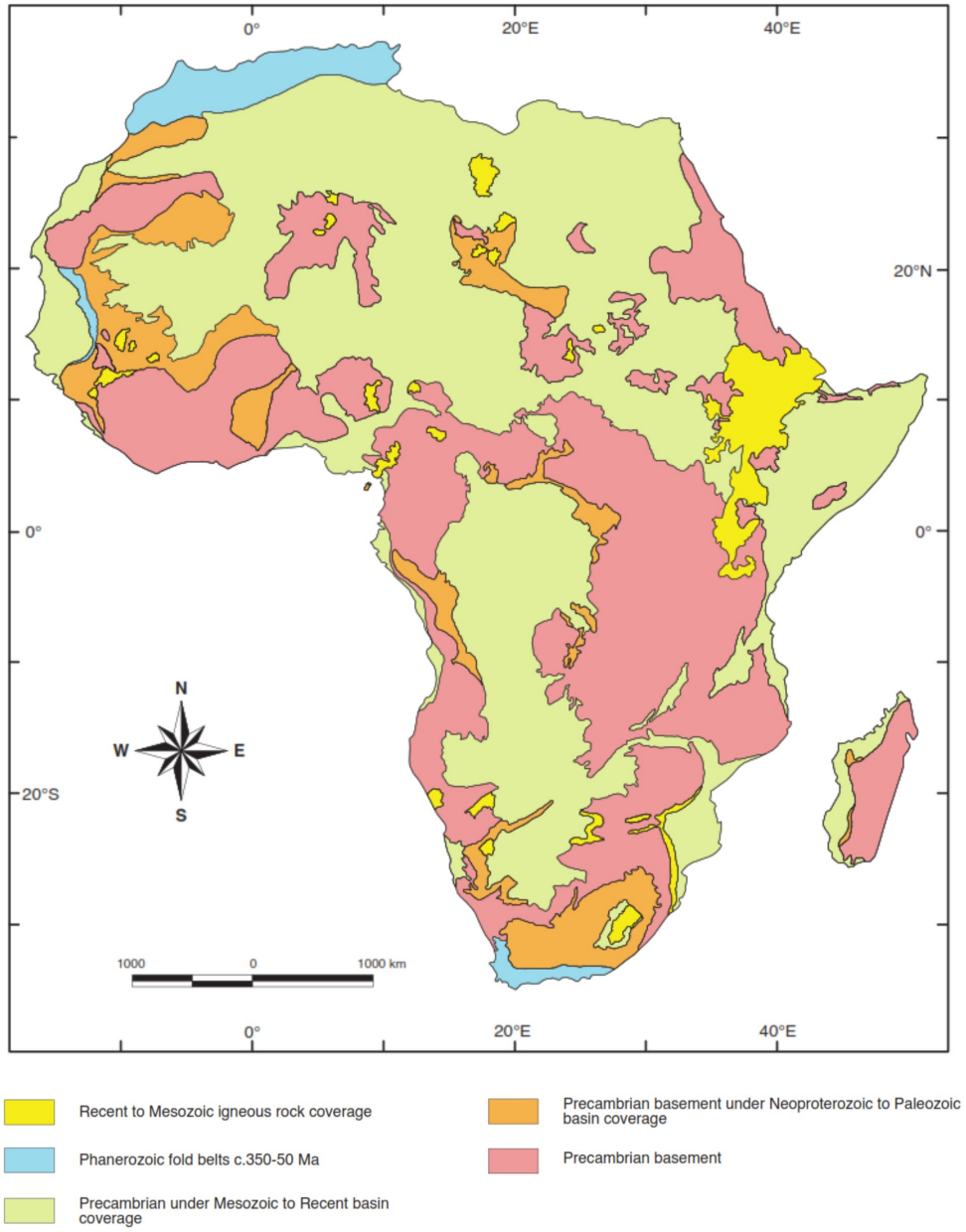

Figure 2.3: Exposure of pre-cambrian rocks (up to 3000 mio. years old), partly under younger coverage. Schlüter (2006), after Key (1992).

wind and, in conjunction with a change in the global climate, facilitated the establishment of the Sahara Desert (Church 1957; Schlüter 2006). Lastly, the period is defined by the expansion and contraction of vegetation belts in result to climate changes (Schlüter 2006).

2.4 Soils

Soils in West Africa have developed in subtropical and tropical conditions under high temperatures and variable humidity (Gaiser et al. 2010). While soil parent materials are defined by the underlying geological structures, soil development is also influenced by the climate and vegetation belts present in the region (Ahn 1970). Desert and semi-desert soils are generally much

younger than their counterparts in the humid tropics. Soils in the semi-arid zones have been developed since the last glacial between 20,000 and 30,000 years ago, while tropical soils have been developed in the last five million years. Climatic changes over the last two million years have led to increases and decreases in precipitation. This variability has facilitated the translocation of soil material, removing either parts of, or the entire soil and exposing parent material, thus restarting the soil development anew. Contrarily, in humid regions, the land surface remained stable, which led to the development of older soils (Gaiser et al. 2010). An overview of the distribution of dominant soil groups is given in Figure 2.4.

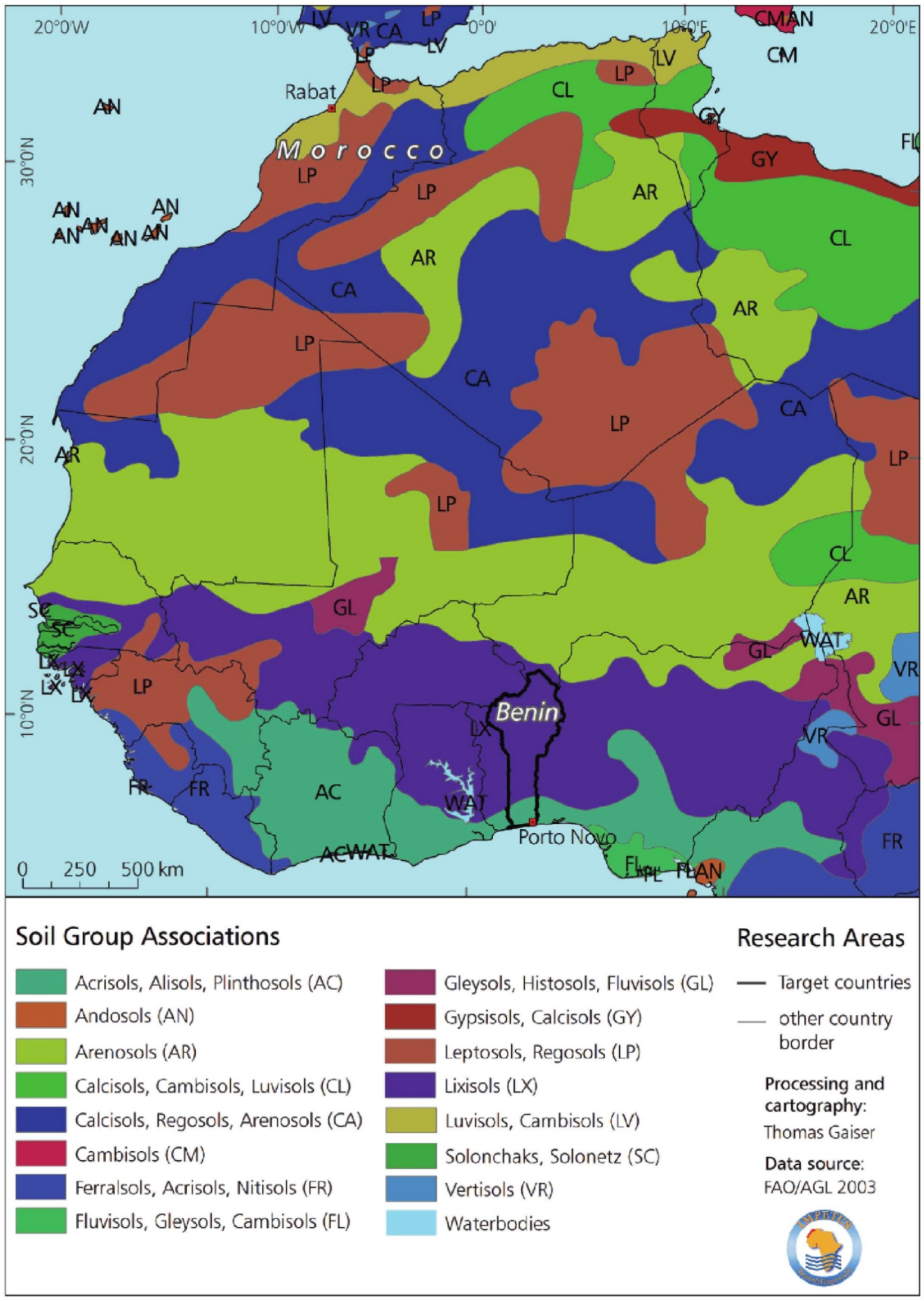

Figure 2.4: Soil associations with dominant soil groups across northwestern and western Africa, Gaiser et al. (2010), based on FAO/AGL (2003).

Calcisols, Regosols, Arenosols and Leptosols constitute the arid and hyper-arid parts of the Sahara Desert (Gaiser et al. 2010). Calcisols are generally defined by the accumulation of calcic materials and may host sparse vegetation, such as shrubs, grasses and herbs, and are used for grazing. Arenosols offer a low nutrient storage capacity and the sparse vegetation can be used for extensive nomadic grazing. While agricultural use is possible on these soils, improper management may cause the soil to become unstable and to revert back to shifting sands. Regosols, on the other hand, are minimally developed and have only marginal agricultural significance; while Leptosols are very shallow and extremely stony, allowing for wet season grazing with a high risk of erosion (IUSS Working Group WRB 2006). On the semi-arid southern fringe, acidic Arenosols developed from carbonate-free aeolian sand deposits. While these sandy soils may be deeply rooted, their water retention capacity is low. Towards the sub-humid savannah in the south, slightly acidic and increasingly loamy Lixisols and Luvisols prevail (Gaiser et al. 2010). Lixisols host savannah and open woodland vegetation, which is mostly used for grazing. The soils are prone to erosion as degraded soils have a low aggregate stability. Tillage of wet Lixisols or the use of heavy machinery compacts the soil and leads to structural deterioration. The cultivation of tubers and groundnuts further increases the erosion risk. Deteriorated Lixisols regenerate only slowly. Luvisols are characterized by clay migration into the subsoil. The soils are fertile and suitable for varying agricultural uses. If a high silt content is present, Luvisols are also prone to erosion (IUSS Working Group WRB 2006). In the west, soils of Gambia and the Casamance region of Senegal consist mostly of salt-accumulating Solonchaks, for which cultivation methods need to be adapted (Gaiser et al. 2010; IUSS Working Group WRB 2006). The natural vegetation of this estuary zone is composed of mudflats and mangrove swamps (CILSS 2016).

Approaching the humid areas towards the southern coast, weathered Acrisols and Ferralsols with low base saturation are predominant (Gaiser et al. 2010). Acrisols need to be carefully managed to retain their agricultural value. The shifting cultivation (slash and burn) performed for centuries in the region takes this into account by cultivating an area for one or a few years with the subsequent regeneration period spanning several decades. This way, limited resources are preserved (IUSS Working Group WRB 2006; Church 1957). However, this traditional system has been mostly replaced by crop fallow rotations (Janssen et al. 2010). Ferralsols are deeply weathered red and yellow soils of the tropics. They have good physical properties and erosion resistance. However, their chemical fertility is low and needs to be overcome with fertilizer application for agricultural production (IUSS Working Group WRB 2006).

2.5 Hydrology

Hydrologic conditions strongly vary in the region, with northern countries like Mali and Niger facing considerably dryer conditions than their southern counterparts, due to the rainfall distribution in the Monsoon Trough (Buckle 1996). An overview of key hydrologic aspects for the West African countries is given in Table 2.1. Annual precipitation may be as low as 92 (Mauritania) or 151 mm (Mali) and can reach up to 2500 and 2400 mm in Sierra Leone and Liberia. Highest surface- and groundwater generation is observed for Guinea, with the Niger, Senegal and Gambia rivers rising in the Fouta Djallon highlands (Balek 1977; CILSS 2016; FAO-AQUASTAT

2018). Many downstream countries are highly dependent on external renewable water resources, such as Mali (50%), Benin (61%), Gambia (62.5%), Niger (89.7%) and especially Mauritania (96.5%). Only Guinea and Sierra Leone do not receive any external water (FAO-AQUASTAT 2018).

Table 2.1: Key water resource availability indicators. Based on FAO-AQUASTAT (2018).

Country	Precipitation in mm/year	Surface Water in 10^9 m^3/year	Groundwater in 10^9 m^3/year	External Water in 10^9 m^3/year	Dependency Ratio in %
Benin	1039	10	1.8	16.1	61
Burkina Faso	748	8	9.5	1	7.4
Ivory Coast	1348	74	37.8	7.3	8.7
Gambia	836	3	0.5	5	62.5
Ghana	1187	29	26.3	25.9	46.1
Guinea	1651	226	38	0	0
Guinea-Bissau	1577	12	14	15.4	49
Liberia	2391	200	45	32	13.8
Mali	282	50	20	60	50
Mauritania	92	0.1	0.3	11	96.5
Niger	151	1	2.5	30.6	89.7
Nigeria	1150	214	87	65.2	22.8
Senegal	686	23.8	3.5	13.2	33.8
Sierra Leone	2526	150	25	0	0
Togo	1168	10.8	5.7	3.2	21.8

Based on the availability of *in situ* data, eleven river basins were chosen for analysis in this study, namely the Niger, Senegal, Volta, Comoé, Gambia, Ouémé, Mono, Pra, Ankobra, Couffo and Ayensu basins. The location of the basins is shown in Figure 2.5. Further details on the basin areas are given in Table 2.2. The larger and (to this study) most relevant rivers Niger, Senegal, Volta, Comoé, Gambia and Oueme will be described in more detail below.

Niger The Niger river basin is the largest in West Africa, covering an area of approximately 2.3 Mio. km^3 (Lehner et al. 2013; Lehner et al. 2008). The source of the river is located in the Fouta Djallon highlands in Guinea, from where it flows northwards and seasonally floods the flat alluvial plain of the Inner Niger Delta in Mali (Balek 1977; CILSS 2016). Flooding occurs along a length of 380 km and can cover up to 40,000 km^2, making the Inner Niger Delta the largest floodplain in West Africa. The seasonal flooding also plays an important role for the pasture and rice farming. An estimated one million people depend on the delta's resources (CILSS 2016). Further towards the south, in Nigeria, two reservoirs formed by the Kainji and Jebra dams have been constructed for hydropower generation (Lehner et al. 2011). Downstream, the Niger meets the Benue, its largest tributary with a length of 1400 km. 130 km before reaching the Atlantic, the Niger flows into a delta region consisting of lakes, wetlands and channels in the south of Nigeria (Balek 1977; CILSS 2016). In total, the river is 4100 km long (Balek 1977).

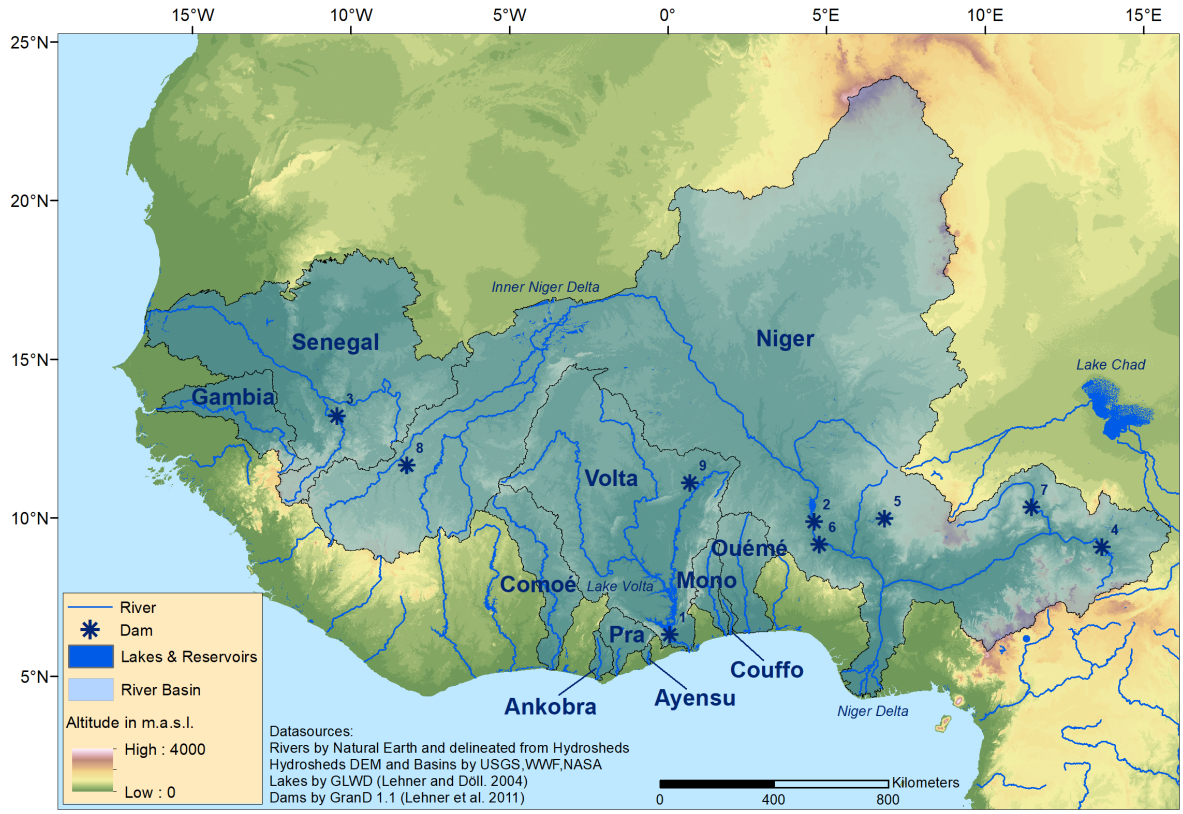

Figure 2.5: River basins chosen for analysis. See also Table 2.2. For details on reservoirs exceeding 2 km^3 storage capacity, see Table 2.3.

Table 2.2: Selected river basins. Drainage areas calculated based on Hydrosheds basin shapefiles (Lehner et al. 2008; Lehner et al. 2013).

River Basin	Area in km^2	Countries in Watershed
Niger	2,270,667	Algeria, Benin, Burkina Faso, Cameroon, Chad, Guinea, Ivory Coast, Mali, Mauritania, Niger, Nigeria
Senegal	480,809	Guinea, Mali, Mauritania, Senegal
Volta	425,425	Benin, Burkina Faso, Ghana, Ivory Coast, Mali, Togo
Comoé	84,675	Burkina Faso, Ghana, Ivory Coast, Mali
Gambia	76,376	Gambia, Guinea, Senegal
Ouémé	60,965	Benin, Nigeria, Togo
Mono	24,451	Benin, Togo
Pra	23,479	Ghana
Ankobra	8523	Ghana
Couffo	3703	Benin, Togo
Ayensu	1725	Ghana

Senegal The Senegal river flows from the western side of the Fouta Djallon highlands towards the northwest (Balek 1977). At their confluence, the main tributaries Bafing (south) and Bakoye (north) contribute 50% and 20% of the total flow. Overall, the Senegal is 1800 km long and is dammed shortly before reaching the sea at St. Louis by the Diama dam (OMVS et al. 2012).

Volta The Volta river has three main tributaries, which are (from west to east) the Black Volta (Mouhoun), the White Volta (Nakambe, with the Red Volta (Nazinon) as its main tributary), and the Oti (Pendjari). The Black Volta originates in the Kong mountains of southwestern Burkina Faso. It is estimated to be 1363 km long. The White Volta flows from the north of Burkina Faso 1136 km southwards before it joins the Black Volta in northern Ghana. The Red Volta, which originates near Ouagadougou in Burkina Faso, is the White Volta's main tributary with a length of 393 km. The Oti river flows from the Atakora Mountains in northern Benin, where it is known as the Pendjari river. It continues southwards through Togo and Ghana, where it is named Oti. It flows for 936 km, before joining the Volta at Kete Krachi (Mul et al. 2015). 80 km from its mouth, the Volta is dammed by the Akosombo dam, creating Lake Volta, the world's largest artificial lake by area with a surface area of close to 8500 km^2 (Balek 1977; CILSS 2016). The reservoir has a storage capacity of 148 km^3 and is mainly used for hydropower generation (Lehner et al. 2011).

Comoé The Comoé, which originates in Mali and flows southwards through Burkina Faso and Ivory Coast, is the longest river of Ivory Coast, measuring 1160 km in length (Yéo et al. 2016). While deforestation and an increase in agricultural activities pose a major threat to the ecosystem along the river, 11,500 km^2 of area in northern Ivory Coast have been declared a national park in 1968 and were designated a UNESCO World Heritage Site in 1983 (CILSS 2016).

Gambia The river Gambia is 1120 km long and also rises in the Fouta Djallon highlands of Guinea, from where it winds westwards towards the coast at Banjul. It is the only river in West Africa which is accessible by oceangoing ships. Swampy areas along the banks and regular saltwater inundation up to 200 km upstream make the western reaches a poor area for agricultural production. Agriculture is mostly conducted on flats along the mid- and upstream portions of the river (CILSS 2016; Encyclopaedia Britannica 2018a).

Ouémé The Ouémé river, which originates in the Atacora Mountains, is approximately 510 km long. Its main tributaries are the Zou and Okpara rivers, which are 150 and 200 km long. Shortly before it would reach the Atlantic in southern Benin, it drains into Lake Nokoué (150 km^2 area). The lake is connected to a coastal lagoon system which drains into the sea. The annual discharge closely follows rainfall trends and is highly variable (Diekkrüger et al. 2010).

Several reservoirs are located in the selected basins. An overview of the largest ones is given in Table 2.3. The larger reservoirs are mainly used for hydropower generation, but also as a source for irrigation water, while smaller reservoirs are almost exclusively used for irrigation (Lehner et al. 2011). The Akosombo dam in Ghana was built in the 1960s and Lake Volta formed from 1962 to 1966. The main purpose of the dam is the generation of electricity for the nearby aluminum smelting industry, and the reservoir covers an area of 8500 km^2, which annually fluctuates due to rising and falling lake levels by 1000 km^2 (Balek 1977; CILSS 2016; Gyau-Boakye 2001). Its immense area and storage capacity have significantly impacted the region and influence the seismicity, sediment loads, morphology, microclimate and prevalence of

vector-borne diseases like Schistosomiasis and Malaria, in addition to social impacts caused by resettlements (Gyau-Boakye 2001). Other large reservoirs in the area are significantly smaller, with the largest, Kainji (Nigeria) and Manantali (Mali) having storage volumes of 15 and 11.3 km^3 respectively. Further large dams do not exceed a volume of 8 km^3. Most large dams in West Africa are located in the Niger river basin both on the Niger itself and on its tributaries, the exception being Akosombo and Kompienga (Volta river basin) and Manantali (Senegal river basin) (Lehner et al. 2011).

Table 2.3: Reservoirs in the study area with storage capacities exceeding 2 km^3. Based on Lehner et al. (2011).

No.	Name	Country	River Basin	River	Capacity in km^3	Main Use
1	Akosombo	Ghana	Volta	Volta	148	Hydropower
2	Kainji	Nigeria	Niger	Niger	15	Hydropower
3	Manantali	Mali	Senegal	Bafing	11.3	Irrigation
4	Lagdo	Cameroon	Niger	Benue	7.8	Irrigation
5	Shiroro	Nigeria	Niger	Kaduna	7	Hydropower
6	Jebra	Nigeria	Niger	Niger	3.6	Hydropower
7	Dadin Kowa	Nigeria	Niger	Gongola	2.9	Irrigation
8	Selingue	Mali	Niger	Sankarani	2.2	Irrigation
9	Kompienga	Burkina Faso	Volta	Ouale	2	Hydropower

2.6 Agriculture, Land Use and Land Cover

Agriculture plays an important role in the livelihoods of the West African population. While 60% of the labor force are employed in the agricultural sector, it contributes only 35% to the Gross Domestic Product (GDP). Consequentially, it can be assumed that farmers are poor and produce mostly at subsistence levels (African Development Bank 2018; Jalloh et al. 2013). On the other hand, West Africa is emerging as a growing market due to population and income growth, urbanization, dietary diversification and increasing prices for agricultural products. At the same time, food demand has shifted from bulk commodities towards more differentiated and refined products, offering producers the possibility of value addition (Hollinger and Staatz 2015; Jalloh et al. 2013). However, rapid growth may also lead to problems due to limitations of natural resources such as agricultural land and water (Hollinger and Staatz 2015; Janssen et al. 2010). An overview of the increase in agricultural areas from 1961 to 2015 for the main countries within the study region is given in Figure 2.6.

Percentages of agricultural area vary strongly by country, with the wetter southern countries (e.g. Ghana, Guinea, Ivory Coast, Nigeria, Togo) reaching higher levels than the dryer northern ones (e.g. Mali, Niger). While 39% of the respective countries' area were used for agriculture in 1961, this increased to 50% in 2015. All countries (except for Guinea) have seen an increase in agricultural area. Strongest increases can be seen in Benin (20%), Nigeria (19%), Sierra Leone (18%) and Ghana (17%). In contrast, worldwide agricultural areas increased from 33 to 36% of the total area (FAO 2018).

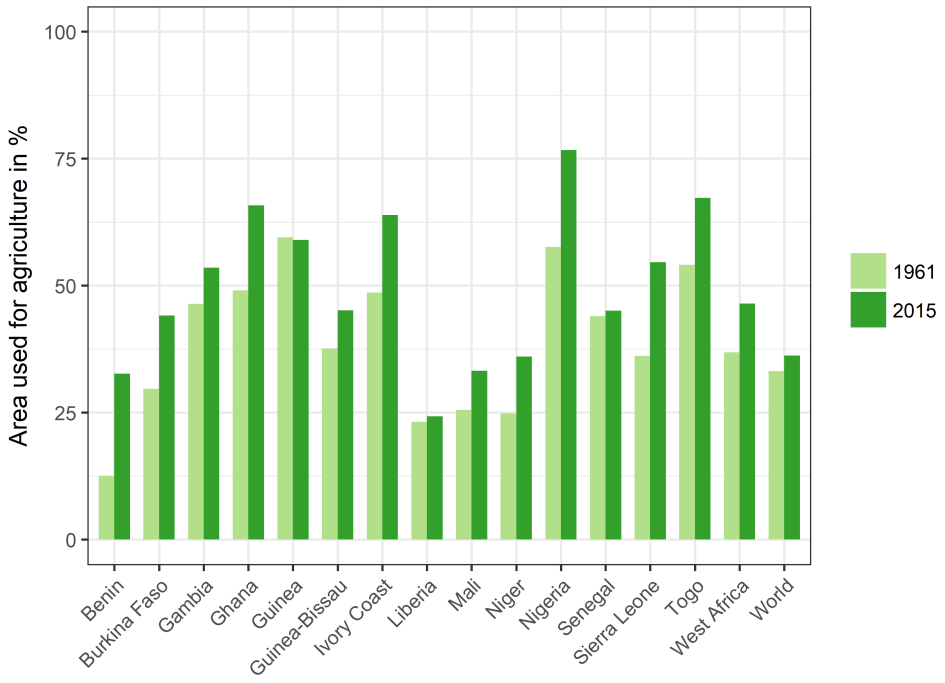

Figure 2.6: Area used for agriculture in %. Based on FAO (2018).

In rural areas, smallholder agriculture is traditionally performed at the family level, with husband, wife (or wives) and children working together. For more labor-intensive tasks, e.g. tillage, additional help may be sought, as observed by the author in central Benin. Families commonly cultivate main crops on land which belongs to the community, with cultivation rights granted to individual farmers. Cropped areas may be as small as 0.2 ha and rarely exceed 8 ha. Traditional shifting agriculture has nowadays been replaced by crop fallow rotation (Janssen et al. 2010; Morgan and Pugh 1969). A separate "garden" for vegetable cultivation is often located on communal land (Morgan and Pugh 1969). In Benin and Burkina Faso, the author has observed these gardens to be mostly located close to both family accommodations and water sources (improved and unimproved) for irrigation purposes, and protected from livestock and wild animals by wooden fences, see also Figure 2.7. Interactions between semi-nomadic animal husbandry and settled agriculture are poor in the region (Janssen et al. 2010). The author witnessed several conflicts in central Benin in 2013, where crops were destroyed by livestock due to a lack of coordination of harvest and animal grazing.

The major food crops in West Africa can be differentiated into cereals (sorghum, millet, maize, and rice), tubers (cassava, sweet potatoes, yams) and legumes (cowpeas and groundnuts). Cocoa, coffee and cotton constitute the main cash crops (Jalloh et al. 2013). In some countries, such as Benin, rice is also considered a cash crop due to its comparatively high value (Igué 2000).

Figure 2.7: Traditional agricultural practices in West Africa. (a): tillage with ox and plough, (b): hoes used for digging, (c): rice harvest using a sickle, (d): fenced household vegetable "garden". Pictures (a), (b) and (c): Benin, 2013; picture (d): Burkina Faso, 2015. All photos by the author.

The distribution of the main crops closely follows precipitation patterns, with tubers cultivated mainly in the south, rice in the southwest and cereals in the north. In central West Africa, both tubers and cereals prevail (Janssen et al. 2010; Morgan and Pugh 1969). In some central areas, previously unused, seasonally waterlogged and highly suitable inland valleys are increasingly used for rice production (Windmeijer and Andriesse 1993; Danvi et al. 2016). The evolution of areas cultivated with the main crops in West Africa is shown in Figure 2.8. An increase can be observed for all crops. Mainly produced are cereals such as millet and sorghum, which show a strong increase from 1982 to 1994 and maize, which was cultivated less in the later part of the 1990s before again increasing from the 2000s onwards. A sudden decrease in the years 2008 and 2009 can be attributed to drought (Kasei et al. 2010). Tubers (cassava, yams), legumes (groundnuts) and rice are cultivated in roughly equal parts. Cultivated areas sharply increase for all crops from 2010 onwards, except for millet and groundnuts. Strongest increases in cultivated area can be observed for yams (8-fold increase), cassava (6.6-fold), rice (5.8-fold) and maize (4.5-fold). The area cultivated with sorghum almost doubles (1.9-fold increase), groundnuts and millet areas are increased 1.7- and 1.6-fold, respectively (FAO 2018).

Agriculture in West Africa is almost exclusively rainfed (Janssen et al. 2010). While the area equipped for irrigation in West Africa has increased threefold (420,000 ha to 1.93 mio. ha) from 1961 to 2015, irrigated land remains marginal in comparison to total agricultural area, with only Guinea-Bissau and Senegal reaching more than one percent, in comparison to the world average of seven percent, as is depicted in Figure 2.9. Consequentially, farmers remain dependent on rainfalls and are vulnerable to water stress (Jalloh et al. 2013).

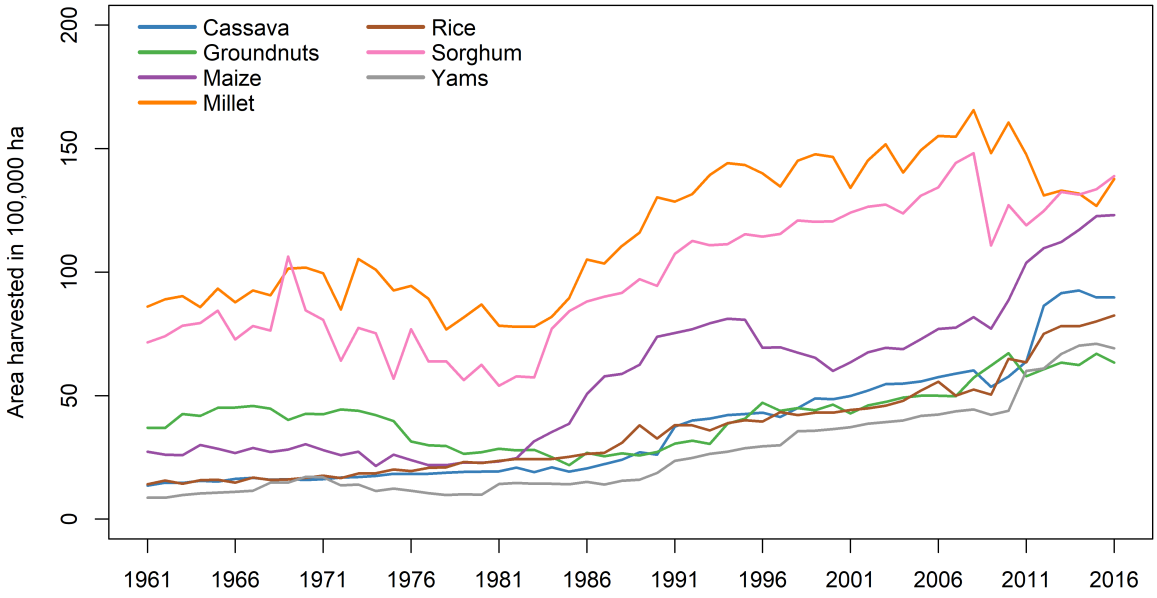

Figure 2.8: Area harvested for main crops in West Africa. Based on FAO (2018).

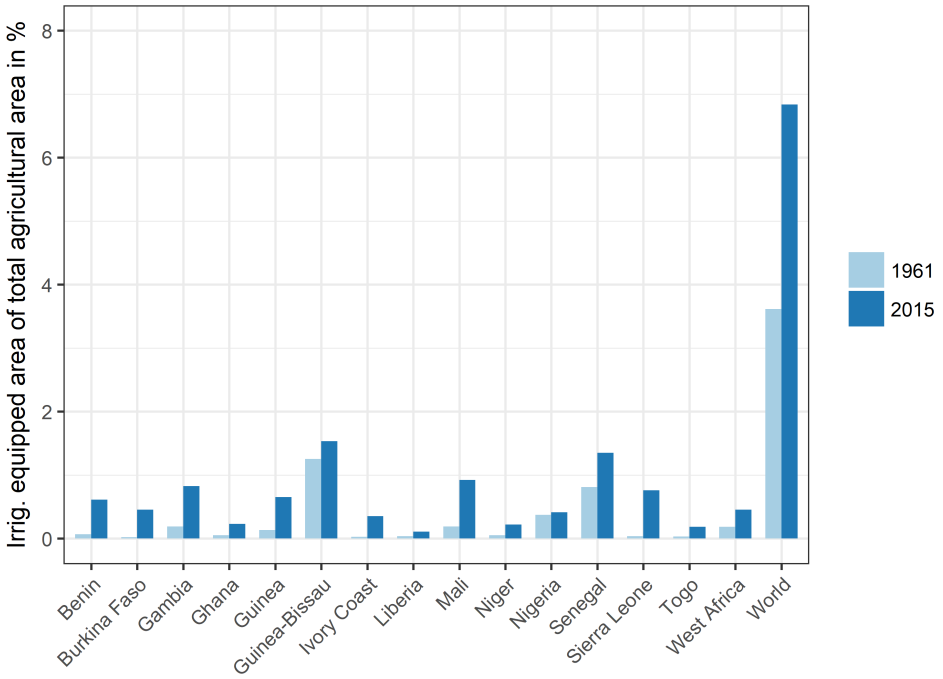

Figure 2.9: Agricultural area equipped for irrigation. Based on FAO (2018).

Two land use and land cover (LULC) maps were utilized in this study: Globcover 2.3 (Figure 2.10) and Landscapes of West Africa (Figure 2.11). The Globcover 2.3 raster map depicts the land use of the year 2009 in a 300 m resolution and is produced by the European Space Agency (ESA) (Bontemps et al. 2011). The Landscapes of West Africa LULC raster dataset is produced by the *Comité permanent Inter-Etats de Lutte contre la Sécheresse dans le Sahel*

(CILSS) for the years 1975, 2000 and 2013 and is available in a 2 km resolution (CILSS 2016). In this study, the 2013 map was used.

Vegetation patterns are similar in both maps. While Globcover is a global product, the CILSS map is only available for CILSS member states and includes neither Cameroon, nor the largely barren areas towards the Sahara Desert in the north. Both maps, however, show clear patterns of land cover change from the humid south to the arid north. In the south, the land is dominated by agriculture, forest and shrubland/savannah. In the central parts, shrubland/savannah and agriculture are predominant. Towards the north, they give way to agriculture and grassland, which border on the sparsely vegetated and sandy areas of the Sahara Desert.

Figure 2.12 shows the change in land use from 1975 to 2000 and 2013 in the research basins based on the CILSS data. This was achieved by first reclassifying the 27 CILSS classes into seven new classes: agriculture, bare area, forest, grassland, savannah/shrubland, settlements and wetlands. For a complete overview, Globcover data was included in this analysis, where the 22 Globcover classes were similarly reclassed. The conversion of savannah/shrubland into agricultural area is immediately obvious. From 1975 to 2013, savannah decreased by 13.6%, forest by 2.2% and grassland by 2.66%. On the other hand, agricultural area increased by 15.8%. Bare areas also increased by 1.1%. Further increases can be seen for settlements (0.4%) and wetlands (0.2%). Globcover predicts a far higher percentage of bare area than CILSS. This is because the product covers the whole domain, including desert areas not covered by CILSS. While estimates on agriculture, grassland and savannah/shrubland areas are similar to CILSS results, they are slightly underpredicted. This may be due to the high percentage of bare area skewing the results. Forest area is predicted to be higher, which may be due to forests not being identically classified in Globcover and CILSS maps. Globcover settlement and wetland classes are also estimated to occur less than CILSS results. In addition to the previous interpretation approaches, this may also be influenced by pixel sizes, since Globcover offers a finer resolution of 300 m compared to the 2 km of CILSS, allowing for more accurate representation.

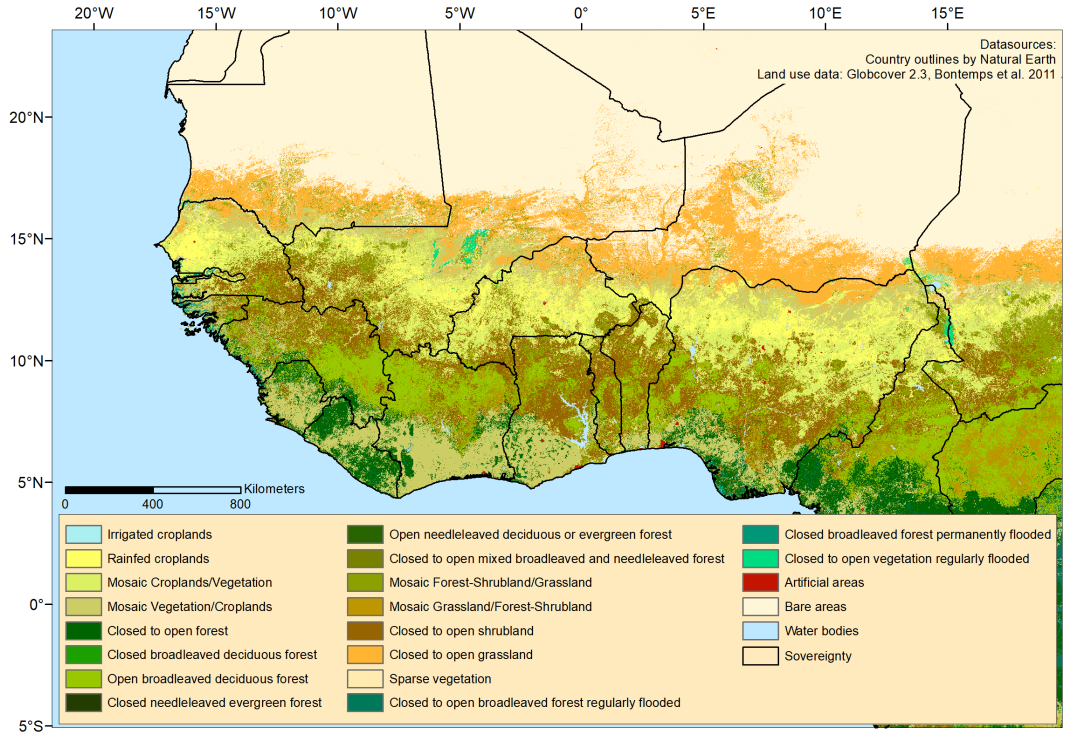

Figure 2.10: Globcover 2.3 land use and land cover of the year 2009 (Bontemps et al. 2011).

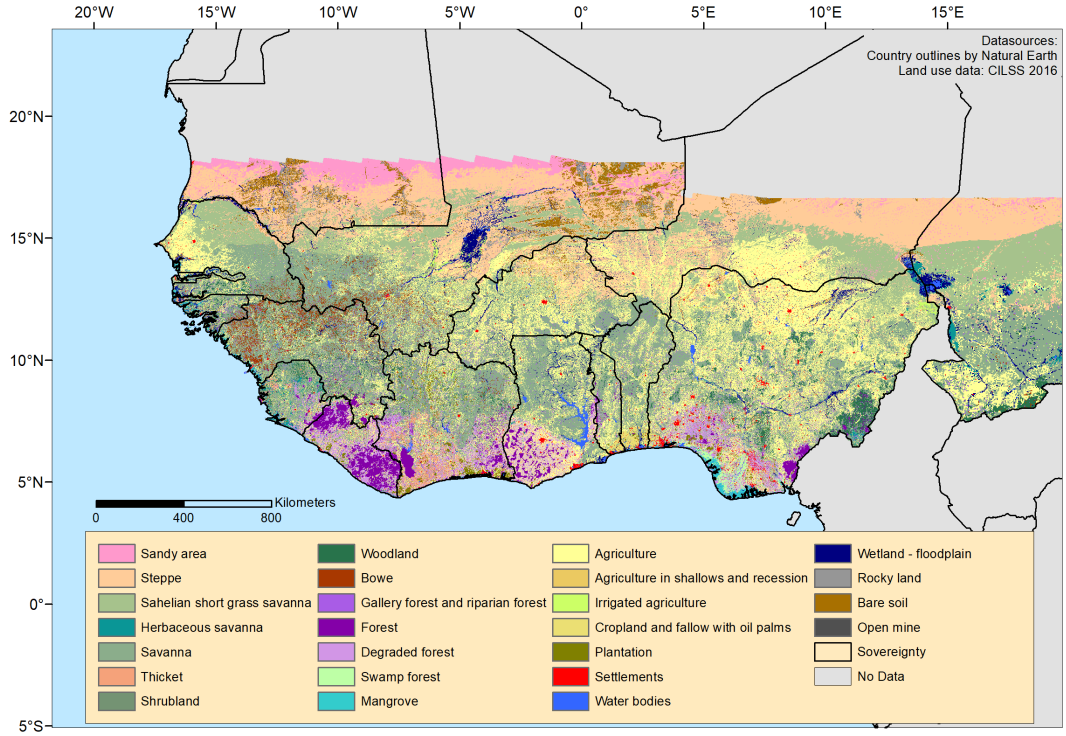

Figure 2.11: CILSS land use and land cover of the year 2013 (CILSS 2016).

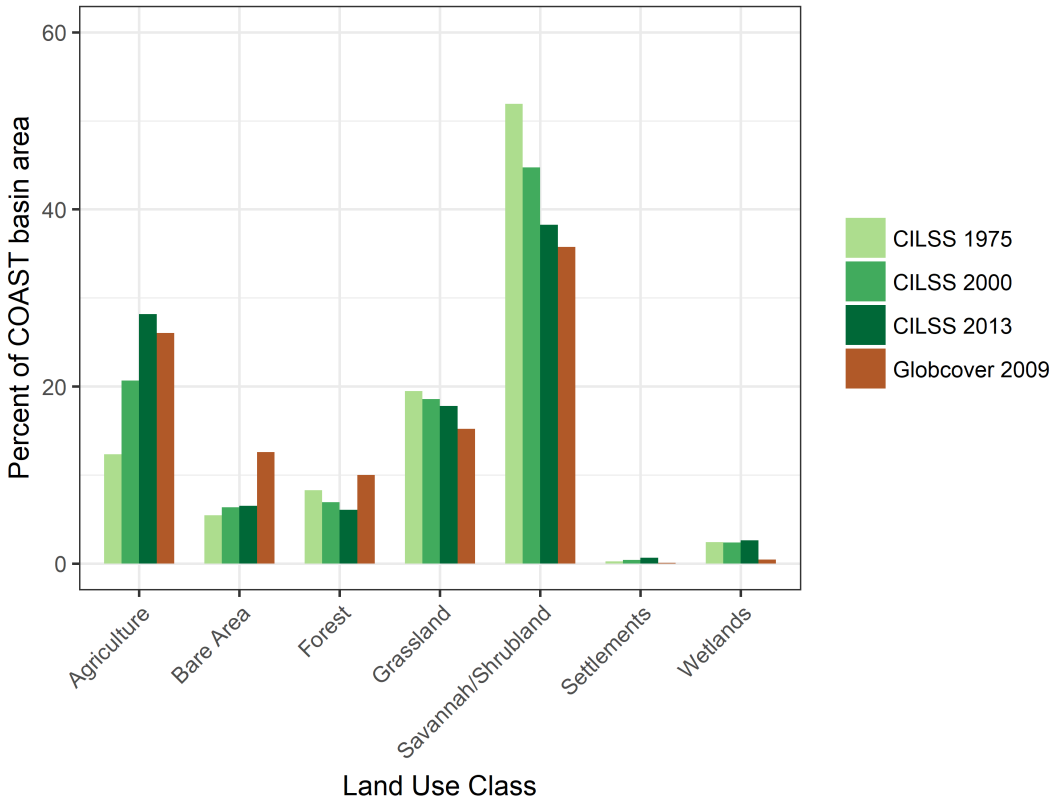

Figure 2.12: Land use and land cover change according to the CILSS maps of 1975, 2000 and 2013 (CILSS 2016). Globcover product (2009) (Bontemps et al. 2011) added for further reference.

2.7 Population and Economy

Over the last 30 years, the population of West Africa has doubled, growing approximately by 2.7% annually. In poorer countries, growth is higher, and birth rates have lowered only in very few states. The population is young, with 44% below the age of 15 (Hollinger and Staatz 2015). The estimated population growth for both rural and urban populations is shown in Figure 2.13. While West Africa hosted only 2.8% of the world population in the 1950s, with no significant urbanization, this has dramatically changed. It is estimated that West Africa currently hosts 4.9% of the world population, with rural and urban populations having equalized. Projections show a constant increase in growth of the urban population, while the growth of the rural population remains stable. Towards 2100, an estimated 14.1% of the world population could live in West Africa (FAO 2018). Strong migratory movements exist from rural to urban areas and, on a larger scale, from landlocked Sahelian countries to wealthier southern states with coastal access (Hollinger and Staatz 2015).

Except for Mauritania, all West African countries are member states of the Economic Community of West African States (ECOWAS) (Jalloh et al. 2013). Most of West Africa experiences economic growth, which was especially strong between 2012 and 2015, then slowed in 2016, and has recently regained momentum, with a 2.5% increase in the Gross Domestic Product (GPD) in 2017, which is predicted to rise to 3.6% in 2018. Main drivers are the recovery of the oil price

and increased production in Nigeria and Ghana, as well as good agricultural performance in the region. In most countries, services constitute the dominant sector, except for Sierra Leone and Liberia, where agriculture makes up most of the GDP. Manufacturing remains the lowest contributing sector and is limited to light industry and production of consumer goods (African Development Bank 2018).

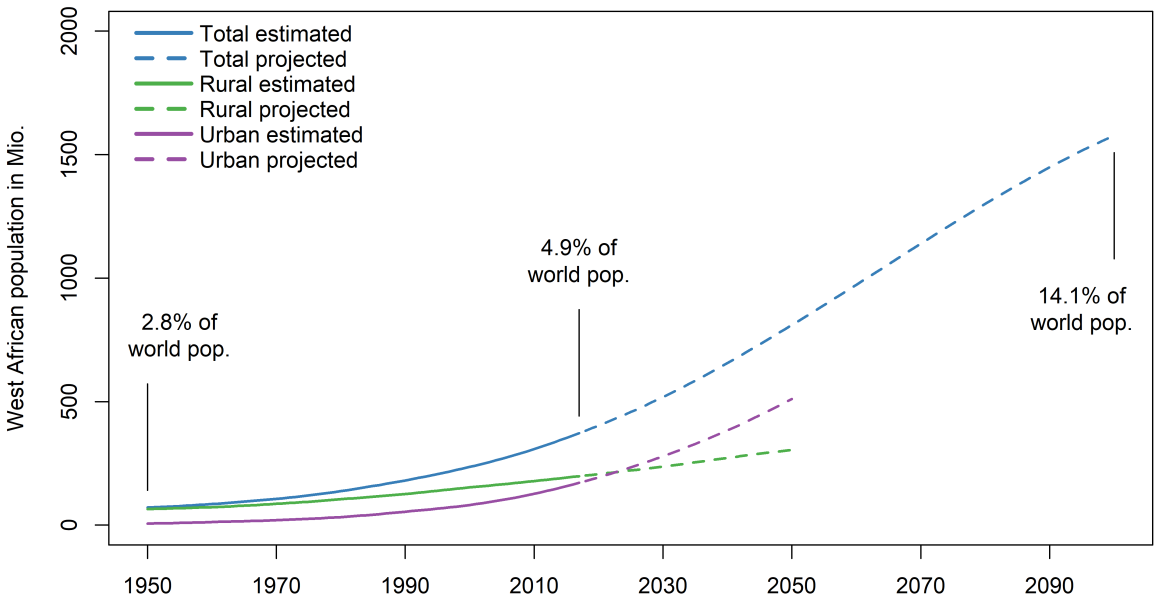

Figure 2.13: Estimated and projected population growth in West Africa. Based on FAO (2018).

An overview of the GDP increase from 2000 to 2016 is given in Figure 2.14. While low in a worldwide comparison, the GPD has increased strongly in West Africa, except for Gambia and Liberia. However, apart from Ghana, Ivory Coast and Nigeria, West African countries did not exceed 2500 US $ per capita in 2016. Strongest increases from 2000 to 2016 have occurred in Ghana (from 2260 to 3980 US $ per capita) and Nigeria (from 2850 to 5440 US $ per capita) (FAO, 2018). On average, the per capita GDP in West Africa has increased from 2310 to 3900 US $ (FAO 2018).

Access of the population to improved water sources has also rapidly progressed (Figure 2.15), from 57% in 2000 to 73% in 2015 (FAO 2018). Unimproved water sources include open wells and streams, while improved sources include potentially safe water sources, such as piped water, protected wells and springs, and rainwater harvesting (WHO and UNICEF 2017). For an example of improved and unimproved water sources, see Figure 2.16. Accessibility to improved water sources in Niger, Sierra Leone, Togo and Nigeria is limited to less than 70% of the population, while in the other countries, accessibility is above 75% and in Ghana and Gambia, it reaches more than 88%, which is on par with the world average of 91% (FAO 2018).

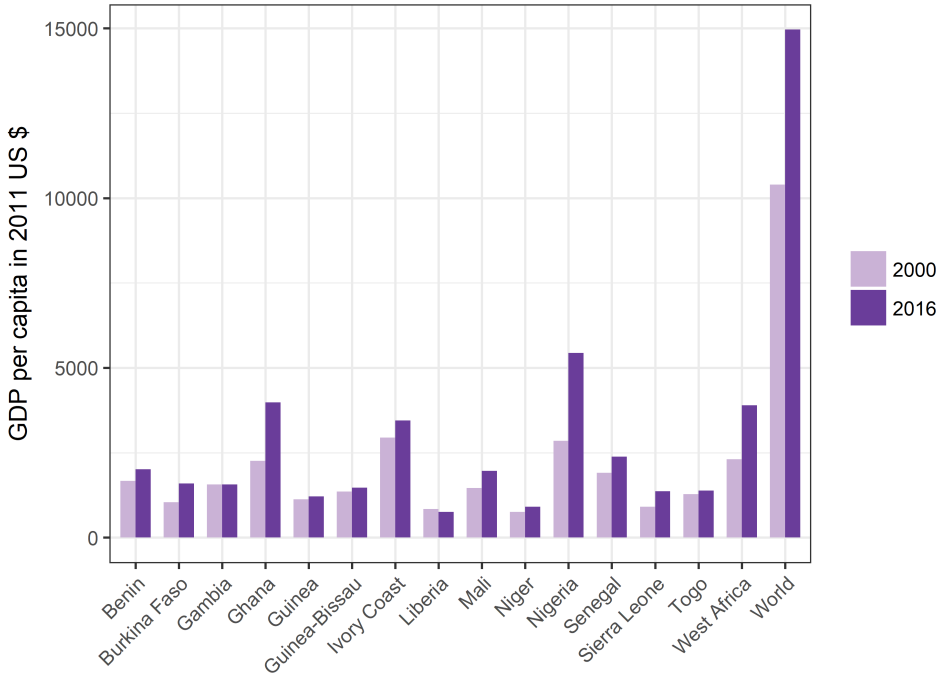

Figure 2.14: Gross domestic product per capita in 2011 US $. Based on FAO (2018).

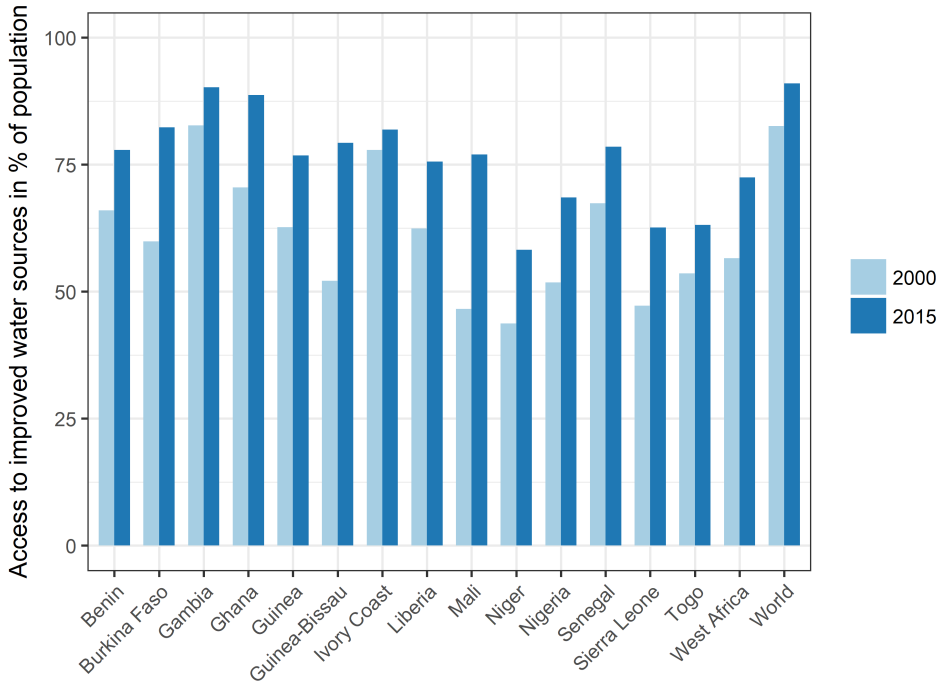

Figure 2.15: Access to improved water sources. Based on FAO (2018).

Figure 2.16: Examples of unimproved and improved water sources. (a): Open well in Burkina Faso, 2015; (b): Covered borehole with a foot-operated pump in Benin, 2013. All photos by the author.

3 Model Choice and Description

In accordance with the problem statement and accompanying research questions, three models were identified to fit the scope of this study. In the following chapter, the identification of suitable models as well as their structure and underlying routines will be described in detail. Further information on the selected models can be found in chapters 4, 5, and 6.

3.1 HBV-light

To perform a hydrologic evaluation, as proposed under research question one, an appropriate model first needs to be chosen. Several key requirements were defined. First of all, the model had to be of a conceptual nature without the need for large amounts of input data. Secondly, model setup and calibration should be easily and quickly performed, as a total of 60 models needed to be calibrated and validated. It would furthermore be advantageous to use a model that has already been applied in data-scarce regions with good results. In order to increase reproducibility, the model should also be widely accessible free of charge. Two models that fit all necessary criteria were considered: the French Modèle du Génie Rural à 4 Paramètres Journalier (GR4J, Perrin et al. (2003)) and the Swedish Hydrologiska Byråns Vattenavdelning-light model (HBV-light, Bergström (1992) and Seibert and Vis (2012)). HBV-light was chosen due to its higher number of parameters (10 as opposed to 4) and additional calibration options. The model has also previously been applied in data-scarce basins (Hattermann et al. 2017; Bárdossy 2007; Kebede et al. 2014; Zelelew and Alfredsen 2013; Bitew and Gebremichael 2011; Jackisch et al. 2014).

The conceptual model was developed in the 1970s by Sven Bergström from the Swedish Meteorological and Hydrological Institute (Bergström 1992; Seibert and Vis 2012). The light version of HBV (Seibert 2005; Seibert and Vis 2012) is based on the 1992 HBV-6 formulation (Bergström 1992). Changes include the addition of a warm-up period, as well as the option to include groundwater level observations and a choice of response routines with a delay parameter (Seibert 2005). Due to the its ease of application, HBV-light is often used to introduce students to hydrologic modeling (Seibert and Vis 2012). Its strengths are the low demand for input data and the limited number of parameters (Rientjes et al. 2013; Rusli et al. 2015). An overview of the model structure is given in Figure 3.1. It includes four main routines: snow, soil, routing, and response; and simulates daily discharge based on precipitation, temperature and potential evapotranspiration timeseries (Seibert and Vis 2012). The model routines will be described in more detail below, based on Seibert and Vis (2012) and Seibert (2005). Due to the lack of snowfall in the study area, the snow routine will not be covered in this description.

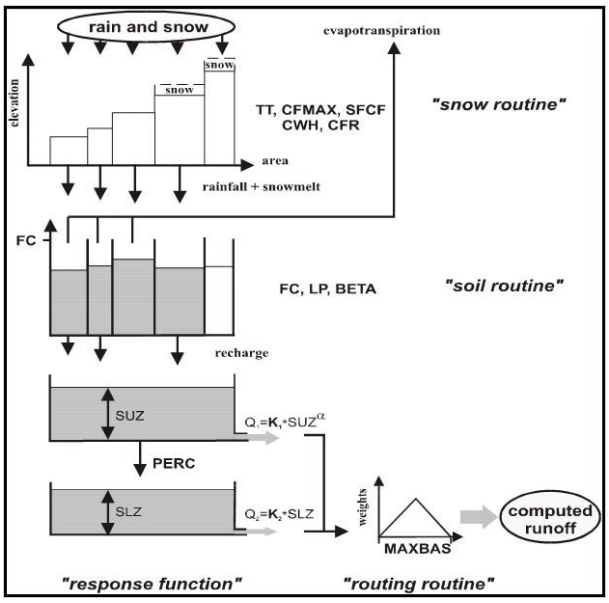

Figure 3.1: HBV model structure (Seibert 2000). See also Figure 4.2.

Within the soil routine, the groundwater recharge is computed based on the amount of liquid precipitation entering the soil (Equation 3.1). Water is partitioned based on the ratio between the current and maximum water content of the soil box. Concerning evapotranspiration (Equation 3.2), as long as S_{SOIL}/P_{FC} is above $P_{FC} * P_{LP}$, actual evapotranspiration equals potential evapotranspiration. If S_{SOIL}/P_{FC} is below this threshold, a linear reduction is used. Recharge is first added to the upper groundwater box and a percolation parameter defines the maximum percolation between the upper and the lower box. Runoff is then computed based on either two or three linear outflow equations (P_{K0}, P_{K1} and P_{K2} in d^{-1}). Two outflow equations are used if the recharge added to the groundwater box S_{LZ} is above a threshold value P_{UZL} and three if it is below (Equation 3.3). Runoff from the groundwater boxes is then transformed into streamflow based on a triangular weighting function governed by the P_{MAXBAS} parameter (Equation 3.4).

$$\frac{F(t)}{I(t)} = \left(\frac{S_{SOIL}(t)}{P_{FC}}\right)^{P_{BETA}} \tag{3.1}$$

where $I(t)$ is the amount of liquid water entering the soil profile in mm d^{-1}, $F(t)$ is the flux to the groundwater in mm d^{-1}, $S_{SOIL}(t)$ is the current amount of water in the soil box in mm, P_{FC} is the maximum value of water in the soil box in mm and P_{BETA} is a dimensionless parameter determining the relative contribution of liquid precipitation to runoff.

$$E_{act} = E_{pot} * min\left(\frac{S_{SOIL}(t)}{P_{FC} * P_{LP}}, 1\right) \tag{3.2}$$

where E_{act} is the actual evapotranspiration in mm d^{-1}, E_{pot} is the potential evapotranspiration in mm d^{-1}, $S_{SOIL}(t)$ is the current amount of water in the soil box in mm, P_{FC} is the maximum value of water in the soil box in mm and $P_{FC} * P_{LP}$ is the soil moisture value above which E_{act} reaches E_{pot} in mm.

$$Q_{GW}(t) = P_{K2} * S_{LZ} + P_{K1} * S_{UZ} + P_{K0} * max(S_{UZ} - P_{UZL}, 0) \tag{3.3}$$

where $Q_{GW}(t)$ is the flow from the groundwater boxes in mm d^{-1}, P_{K0}, P_{K1} and P_{K2} are recession coefficients per d^{-1}, S_{UZ} is the recharge added to the upper groundwater box, S_{LZ} is the recharge added to the lower groundwater box and P_{UZL} is a threshold parameter in mm.

$$Q_{sim}(t) = \sum_{i=1}^{P_{MAXBAS}} c(i) * Q_{GW}(t-i+1)$$
$$\text{where } c(i) = \int_{i-1}^{i} \frac{2}{P_{MAXBAS}} - \left| u - \frac{P_{MAXBAS}}{2} \right| * \frac{4}{P_{MAXBAS}^2} du \tag{3.4}$$

where $Q_{sim}(t)$ is the simulated streamflow in mm d^{-1} and P_{MAXBAS} is a triangular weighting function parameter governing transformation of flow from the groundwater boxes to streamflow.

3.2 SWAT

In this study, it was further evaluated how well a physically-based semi-distributed hydrologic model can be parameterized and validated using remote sensing observations. Additional requirements were that the model had already been applied in the area and that it was widely used and free of charge. The selected Soil and Water Assessment Tool (SWAT) (Arnold et al. 1998; Arnold et al. 2012b) fits all these criteria, and has been intensively used for the region in multiple studies (Schuol and Abbaspour 2006a, 2006b; Schuol et al. 2008b; Schuol et al. 2008a; Xie et al. 2012).

SWAT includes hydrology, weather, sedimentation, soil temperature and properties, crop growth, nutrients, pesticides and agricultural management routines (Arnold et al. 1998; Arnold et al. 2012b). The first SWAT version was developed by the United States Department of Agriculture (USDA) and released in the 1990s as SWAT 94.2 (Arnold et al. 2012b; Gassman et al. 2007), with a GIS (Geographic Information System) interface first introduced in 1998 (Arnold et al. 1998). In this study, SWAT 2012 was used. The model divides a watershed into subbasins and the subbasins further into Hydrologic Response Units (HRUs): areas with identical land use, management, topographical- and soil characteristics. While subbasins are spatially located within the watershed, HRUs represent only fractions of the watersheds and are no longer spatially identified, thus rendering SWAT a semi-distributed model. However, as done in this study, watersheds can be set up to contain only the dominant HRU, thereby preserving spatial distribution (Arnold et al. 2012b). SWAT has the advantage of being usable in GIS environments (ArcGIS and Q-GIS) and thus having a graphical user interface for data integration and model manipulation. Also, a dedicated calibration and uncertainty estimation program named SWAT-CUP (Abbaspour et al. 2007) is freely available. Existing drawbacks are mainly due to the model structure and physically-based nature of the model. For instance, model in- and outputs are internally handled *via* text files, leading to a large number of very small files in simulations with many subbasins. This approach can be computationally inefficient, as the computer system will be limited by read and write speeds of the hard drive, and not by processing power. Due to its physically-based nature, SWAT requires a multitude of input

parameters which may not be readily available in data-scarce basins. By estimating parameters where no observations are present, additional bias is introduced.

SWAT simulates the hydrologic cycle based on the water balance equation (Equation 3.5, Neitsch et al. (2011)). An overview of simulated processes of the land phase is given in Figure 3.2.

$$SW_t = SW_0 + \sum_{i=1}^{t}(R_{day} - Q_{surf} - E_a - w_{seep} - Q_{GW}) \tag{3.5}$$

where SW_t is the final soil water content in mm on day i, SW_0 is the initial soil water content on day i, t is the time in days, R_{day} is the precipitation on day i in mm, Q_{surf} is the surface runoff on day i in mm, E_a is the actual evapotranspiration on day i in mm, w_{seep} is the water entering the vadose zone from the soil profile on day i in mm and Q_{GW} is the return flow on day i in mm.

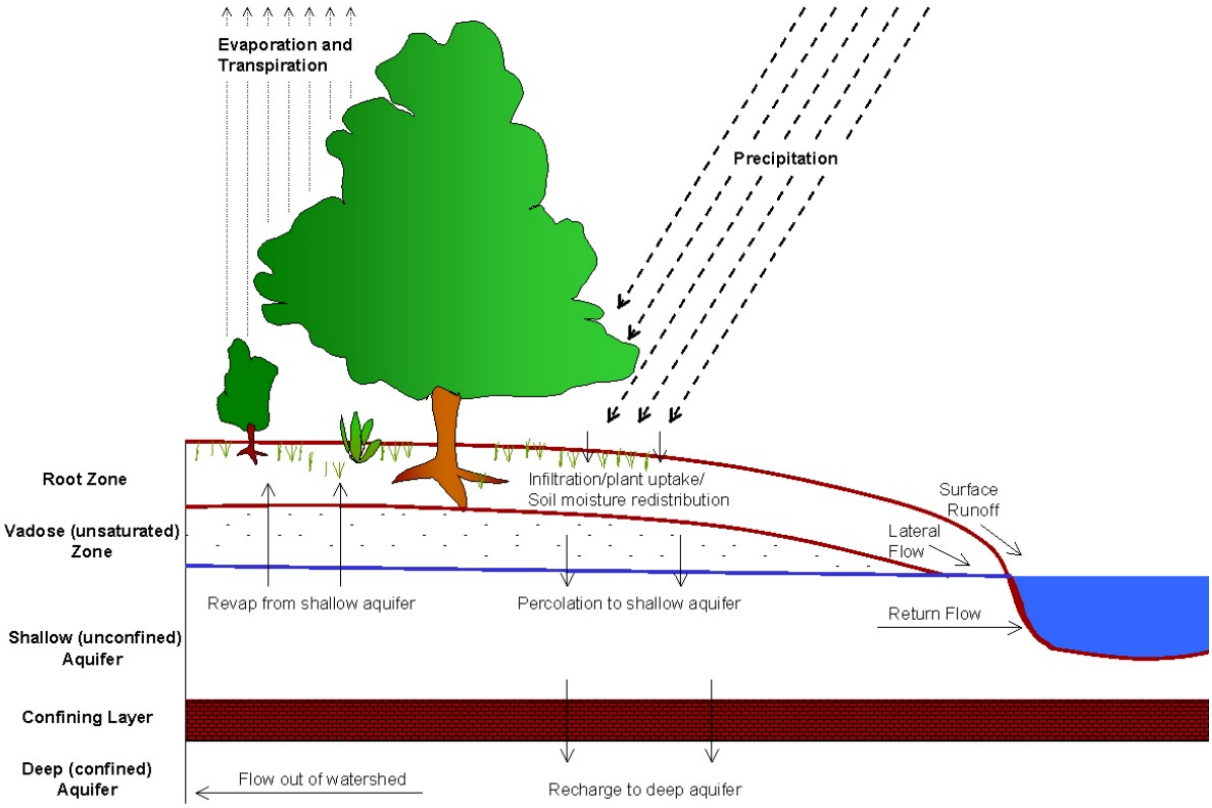

Figure 3.2: Schematic representation of the hydrologic cycle as simulated by SWAT (Neitsch et al. 2011).

SWAT calculates surface runoff volume based on the USDA Soil Conservation Service Curve Number (SCS CN) (USDA SCS 1972) approach. The curve number rates hydrologic performance based on soil type, land use, management, slope and soil water content (DeLiberty and Legates 2003), shown in Equation 3.6. The variable retention parameter s, given in the equation with a constant 254 mm, may change due to fluctuations in soil water content (Equation 3.7) (Arnold et al. 1998).

$$
\begin{aligned}
Q = \frac{(R - 0.2s)^2}{R + 0.8s} \quad &\text{if} \quad R > 0.2s \\
\text{and} \quad Q = 0.0 \quad &\text{if} \quad R \leq 0.2s \\
\text{where} \quad s = 254 \left(\frac{100}{CN} - 1 \right)&
\end{aligned}
\tag{3.6}
$$

where Q is the surface runoff in mm d^{-1}, R is the precipitation in mm d^{-1}, and s is a variable retention parameter related to the curve number CN.

$$
\begin{aligned}
s = s_1 \left(1 - \frac{FFC}{FFC + exp[w_1 - w_2(FFC)]} \right) \\
\text{where} \quad FFC = \frac{SW - WP}{FC - WP}
\end{aligned}
\tag{3.7}
$$

where s is a variable retention parameter related to the curve number, s_1 is the value of s associated with CN_1, FFC is the fraction of field capacity, w_1 and w_2 are shape parameters, SW is the soil water content of the root zone in mm, WP is the wilting point water content in mm and FC is the water content at field capacity in mm.

Percolation is calculated for each soil layer using a storage routing technique coupled to a crack-flow model (Arnold et al. 1998). If the water content exceeds field capacity and the layer below is not saturated, water can percolate. The amount of water percolating is calculated according to Equation 3.8. Water percolating past the root zone enters the unsaturated vadose zone, from where it either forms groundwater recharge or becomes return flow in a downstream subbasin (Neitsch et al. 2011).

$$
\begin{aligned}
SW_{ly,excess} = SW_{ly} - FC_{ly} \quad &\text{if} \quad SW_{ly} > FC_{ly} \\
\text{and} \quad SW_{ly,excess} = 0 \quad &\text{if} \quad SW_{ly} \leq FC_{ly}
\end{aligned}
\tag{3.8}
$$

where $SW_{ly,excess}$ is the drainable amount of water in the soil layer on a given day in mm, SW_{ly} is the water content of the soil layer on a given day in mm and FC_{ly} is the water content of the soil layer at field capacity in mm.

Lateral flow occurs when water cannot percolate past an impenetrable layer in the soil. This depends on the amount of water in the soil, saturated hydraulic conductivity, slope, porosity and length of the hillslope (Equation 3.9, Neitsch et al. (2011)).

$$
\begin{aligned}
Q_{lat} = 0.024 * \left(\frac{2 * SW_{ly,excess} * K_{sat} * slp}{\phi_d * L_{hill}} \right) \\
\text{where} \quad \phi_d = \phi_{soil} - \phi_{fc}
\end{aligned}
\tag{3.9}
$$

where Q_{lat} is the water discharged from the hillslope outlet in mm d^{-1}, $SW_{ly,excess}$ is the drainable amount of water in the soil layer on a given day in mm, K_{sat} is the saturated hydraulic conductivity in mm h^{-1}, slp is the increase in elevation per unit distance, ϕ_d is the drainable

porosity of the soil layer in mm/mm, ϕ_{soil} is the total porosity of the soil layer in mm/mm, ϕ_{fc} is the porosity of the soil layer at field capacity in mm/mm and L_{hill} is the length of the hillslope in m.

SWAT separates the groundwater zone into a shallow and a deep aquifer. Baseflow from the shallow aquifer is calculated using a linear storage approach. While water from the shallow aquifer may also revap into the unsaturated zone by evaporative demand or be removed by deep-rooted plants. Water entering the deep aquifer does not contribute to flow within the subbasin. The water balance for the shallow aquifer is given in Equation 3.10 (Neitsch et al. 2011).

$$aq_{sh,i} = aq_{sh,i-1} + w_{rchrg,sh} - Q_{gw} - w_{revap} - w_{pump,sh} \tag{3.10}$$

where $aq_{sh,i}$ is the water stored in the shallow aquifer on day i in mm, $aq_{sh,i-1}$ is the water stored in the shallow aquifer on day $i-1$ in mm, $w_{rchrg,sh}$ is the recharge entering the shallow aquifer on day i in mm, Q_{gw} is the baseflow into the main channel on day i, w_{revap} is the water moving into the overlying soil zone due to water deficiencies on day i in mm and $w_{pump,sh}$ is the water removed by abstraction (e.g. pumping) on day i in mm.

3.3 mHM

It was furthermore decided to study the effects of multivariate calibration and validation of a grid-based model using remote sensing data for the modeling of hydrologic processes in West Africa. Two main requirements were that the model had to be grid-based and capable of multivariate calibration. Furthermore, as all software employed in this thesis, it needed to be free of charge. The choice was made to apply the relatively new grid-based mesoscale Hydrologic Model (mHM) (Samaniego et al. 2010; Kumar et al. 2013b) in this case. Further considerations for the use of mHM were its native capability to handle netCDF multidimensional gridded in- and outputs, easing the data preparation and result evaluation process. A major drawback is that data preparation can be difficult to accomplish, as input data requirements are very strict in relation to netCDF cell size and projection formatting. Here, mHM diverges from commonly accepted conventions. Some programming experience in preparing spatial data using e.g. R or Python is therefore necessary.

mHM is a spatially explicit hydrologic model and was developed at the Helmholtz Centre for Environmental Research in Leipzig, Germany. mHM simulates daily and hourly distributed predictions of hydrologic variables. Numerical approximations of dominant hydrologic processes are based on the HBV (Bergström 1976, 1992) and VIC (Liang et al. 1994) model formulations. The model accounts for canopy interception, snow accumulation and melting, soil moisture dynamics, infiltration, surface runoff, discharge generation, evapotranspiration, subsurface storage, deep percolation, baseflow, and flood routing.

At the mesoscale, processes of the hydrologic cycle span different spatial regimes, with input data separated by several orders of magnitude. To better represent this spatial variability, mHM is differentiated into three levels. The first level, L0, is the most detailed level and describes the small-scale properties of the basin (e.g. elevation, slope, soil characteristics and land cover).

At the second level, L1, dominant mesoscale hydrologic processes are described. L1 represents the resolution of the hydrologic model routines and output and only requires the regionalized fields of model parameters as input. The final level, L2, spatially discretizes the meteorological forcings (e.g. precipitation and temperature) (Samaniego et al. 2010; Samaniego et al. 2017b). A schematic representation of the model for a single grid cell is given in Figure 3.3. Precipitation P falls in the grid cell. Actual evapotranspiration from the canopy E_1, water bodies E_2 and surface E_3 is then calculated based on the potential evapotranspiration (ETP). ETP can be derived either from input data or by calculation (e.g. Hargreaves-Samani (Hargreaves and Samani 1985) and Penman-Monteith (Allen et al. 1998) equations), using further meteorological parameters like temperature or radiation R_s. Excess water is then partitioned into vertical and lateral flow. Vertically, water percolates into the soil, drains below the root zone into the unsaturated zone and further recharges groundwater in the saturated zone I, C, K. Laterally, runoff q_1, fast interflow q_2, slow interflow q_3 and baseflow q_4 are differentiated (Samaniego et al. 2010).

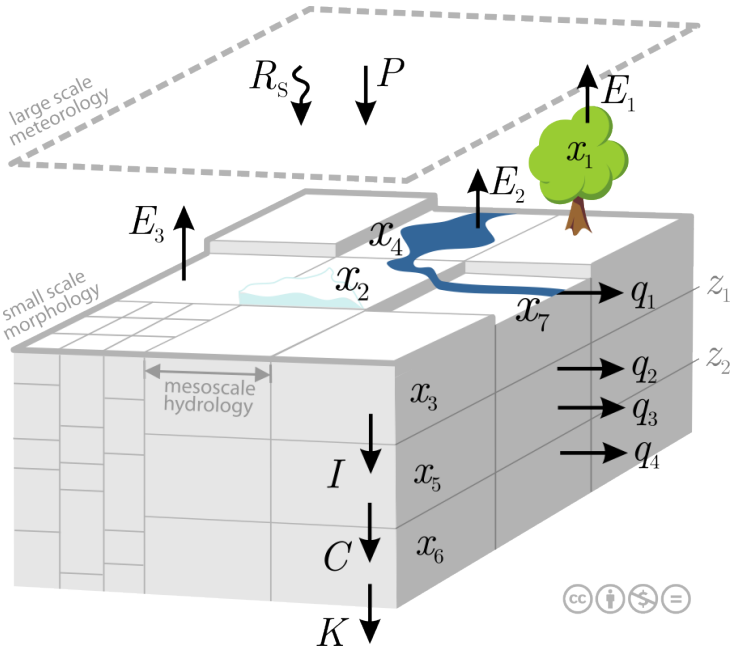

Figure 3.3: Schematic representation of the mHM model (Samaniego et al. 2017b). See also Figure 6.2.

River basins are open, natural systems and are, especially at the mesoscale, very heterogeneous. They have no clear, but rather fuzzy boundary conditions. Since the spatial heterogeneity is mainly described through discrete variables (land cover, soil texture, geology, etc.), it is problematic to assume continuity of the input variables. Therefore, the evolution of the state variables given at a specific location (pixel) i within the domain Ω is described using a system of ordinary differential equations (Samaniego et al. 2017b). Seven states are calculated by mHM: canopy storage, snowpack, soil moisture content in the root zone, impounded water, storage in a subsurface reservoir, storage in a groundwater reservoir and storage in the channel (Table 3.1). The ordinary differential equations are given in Equations 3.11 - 3.17. The respective model inputs, fluxes, outputs and indices are listed in Table 3.2.

Table 3.1: mHM states. After Samaniego et al. (2017b).

State	Description (all water depths in mm)	Equation
x_1	Depth of the canopy storage	3.11
x_2	Depth of the snowpack	3.12
x_3	Depth of the soil moisture content in the root zone	3.13
x_4	Depth of water impounded in reservoirs, water bodies, or sealed areas	3.14
x_5	Depth of the water storage in the subsurface reservoir	3.15
x_6	Depth of the water storage in the groundwater reservoir	3.16
x_7	Depth of the water storage in the channel reservoir	3.17

$$\dot{x}_{1i} = P_i(t) - F_i(t) - E_{1i}(t) \tag{3.11}$$

$$\dot{x}_{2i} = S_i(t) - M_i(t) \tag{3.12}$$

$$\dot{x}^l_{3i} = (1 - p^l) I^{l-1}_i(t) - E^l_{3i}(t) - I^l_i(t) \tag{3.13}$$

$$\dot{x}_{4i} = p^1(R_i(t) + M_i(t)) - E_{2i(t)} - q_{1i}(t) \tag{3.14}$$

$$\dot{x}_{5i} = I^L_i(t) - q_{2i}(t) - q_{3i}(t) - C_i(t) \tag{3.15}$$

$$\dot{x}_{6i} = C_i(t) - q_{4i}(t) \tag{3.16}$$

$$\dot{x}_{7i} = \hat{Q}^0_i(t) - \hat{Q}^1_i(t) \tag{3.17}$$

For all equations: $\forall i \in \Omega$

Table 3.2: mHM inputs, fluxes, outputs and indices (mHM state equations 3.11-3.17). After Samaniego et al. (2017b).

Inputs	Description	Unit
P	Daily precipitation depth	mm d^{-1}
E_p	Daily potential evapotranspiration	mm d^{-1}
T	Daily mean air temperature	°C
Fluxes	**Description**	**Unit**
S	Snow precipitation depth	mm d^{-1}
R	Rain precipitation depth	mm d^{-1}
M	Melting snow depth	mm d^{-1}
E_p	Potential evapotranspiration	mm d^{-1}
F	Throughfall	mm d^{-1}
E_1	Actual evaporation from the canopy	mm d^{-1}
E_2	Actual evapotranspiration	mm d^{-1}
E_3	Actual evaporation from free water bodies	mm d^{-1}
I	Recharge, infiltration, or effective precipitation	mm d^{-1}
C	Percolation	mm d^{-1}
q_1	Surface runoff from impervious areas	mm d^{-1}
q_2	Fast interflow	mm d^{-1}
q_3	Slow interflow	mm d^{-1}
q_4	Baseflow	mm d^{-1}
Outputs	**Description**	**Unit**
$\hat{Q}_i^0$	Simulated discharge entering the river stretch at cell i	m^3 s^{-1}
$\hat{Q}_i^1$	Simulated discharge leaving the river stretch at cell i	m^3 s^{-1}
Indices	**Description**	**Unit**
l	Index denoting a root zone horizon	
t	Time index for each Δt interval	
p^l	Overall influx fraction accounting for the impervious cover within a cell	

Please note that in Figure 3.3, also after Samaniego et al. (2017b), E_2 describes actual evaporation from free water bodies and E_3 actual evapotranspiration.

In mHM, Multiscale Parameter Regionalization (MPR) is applied to estimate the regionalized fields of model parameters at the L0 resolution and upscaling operators are used to generate effective parameters at the L1 resolution. The regionalization is performed by linking model parameters to catchment attributes such as terrain, slope, aspect, soil texture and land cover *via* a set of pedotransfer functions and free calibration parameters (Samaniego et al. 2010; Kumar et al. 2013b). The *a priori* relationships of the transfer functions and upscaling operators are defined through process understanding and empirical evidence (Samaniego et al. 2010; Samaniego et al. 2017b). An example is given in Figure 3.4. Here, u are the basin predictors such as land cover class or flow direction, β are predictors used in the regionalization functions and γ is a vector of transfer parameters. For more information, see Samaniego et al. (2010).

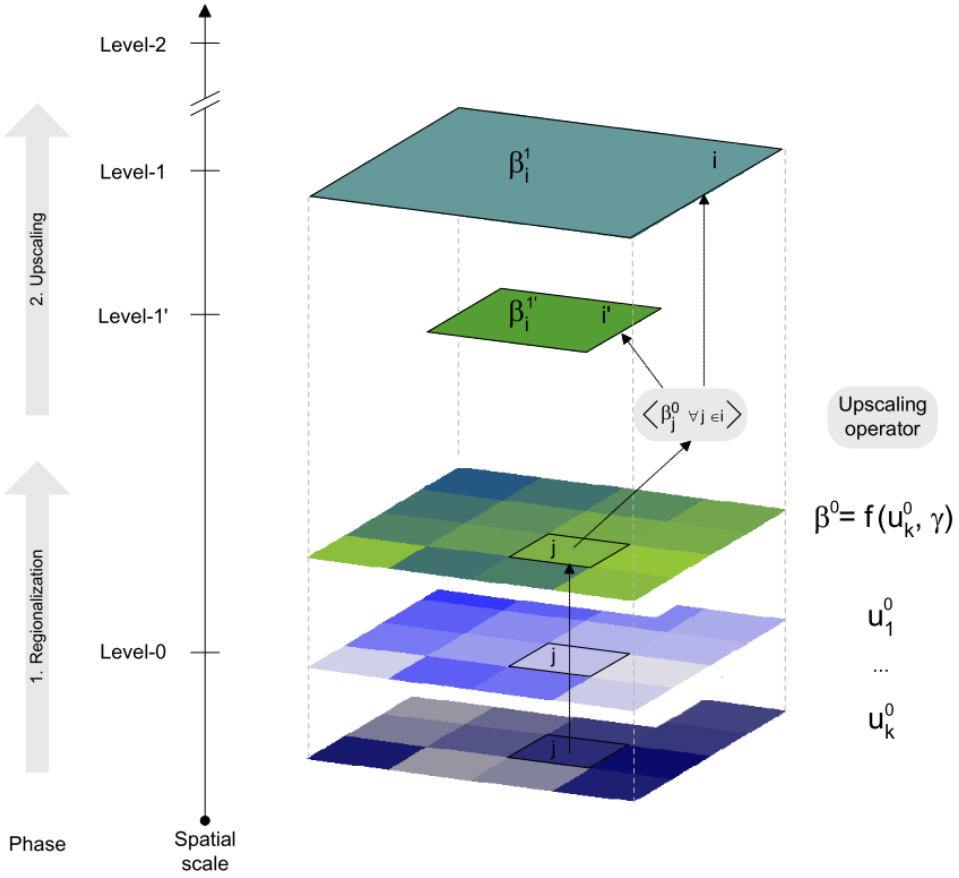

Figure 3.4: Schematic representation of the multiscale parameter regionalization (Samaniego et al. 2017b).

4 Evaluating the Performance of Remotely Sensed and Reanalysed Precipitation Data over West Africa using HBV light

This chapter has been published as: Poméon, T., Jackisch, D., Diekkrüger, B., 2017. Evaluating the performance of remotely sensed and reanalysed precipitation data over West Africa using HBV light. Journal of Hydrology. 547, 222–235. doi:10.1016/j.jhydrol.2017.01.055.
https://www.sciencedirect.com/science/article/pii/S0022169417300653

Abstract: Water is a crucial resource in West Africa, where large parts of the population rely on rainfed agriculture. Therefore, accurate knowledge of the water resources is of the utmost importance. Due to the declining number of rain gauging stations, the use of satellite and reanalysis precipitation datasets in hydrological modeling is steadily rising. However, accurate information on the benefits and deficits of these datasets is often lacking, especially in the West African subcontinent. For validation purposes, these products are commonly compared to freely available rain gauge data, which has in some cases already been used to bias correct the products in the first place. We therefore explored the possibility of a hydrological evaluation, where a model is calibrated for each dataset using streamflow as the observed variable. In this study, ten freely available satellite and reanalysis datasets (CFSR, CHIRPS, CMORPHv1.0 CRT, CMORPHv1.0 RAW, PERSIANN CDR, RFE 2.0, TAMSAT, TMPA 3B42v7, TMPA 3B42 RTv7 and GPCC FDDv1) were thus evaluated for six differently sized and located basins in West Africa. Results show that while performances differ, most datasets manage to somewhat accurately predict the observed streamflow in a given basin. Best results were achieved by datasets which use a multitude of input data, namely infrared and microwave satellite data, as well as observations from rain gauges (usually GPCC) for bias correction. If considering only the Nash Sutcliffe Efficiency averaged for all six basins during the calibration phase, best results were achieved by CMORPH CRT and PERSIANN CDR (both 0.66), followed by TAMSAT, CHIRPS and TMPA 3B42 (all three 0.64). Average results were achieved by RFE 2.0 (0.63), GPCC (0.61) and TMPA 3B42 RT (0.54). CMORPH RAW and CFSR performed worst (0.36 and -0.34 on average).

Keywords: Precipitation datasets; Satellite precipitation estimates; Hydrological evaluation; Remote sensing

4.1 Introduction

Precipitation data is one of the most important drivers of hydrological models (Hughes 2006; Thiemig et al. 2013). However, especially in developing countries, rain gauge (RG) networks are sparse or nonexistent and often include erroneous data or large gaps (Behrangi et al. 2011; Bitew and Gebremichael 2011; Koutsouris et al. 2016). This situation is aggravated by a further decline in the number of rainfall stations due to financial or maintenance problems (Adjei et al. 2012; Hughes 2006). A large number of satellite-based rainfall estimate (SRFE) and reanalysis (RA) products with high spatial and temporal resolutions are freely available and have the potential to complete gaps or even replace rain gauge measurements (Fujihara et al. 2014; Koutsouris et al. 2016; Thiemig et al. 2013). The uncertainties of these products over western Africa are largely unknown and only recently scientists have started validating SRFE and RA products for the sub-continent (Awange et al. 2015). These validations are mostly conducted by using statistical measures to compare SRFEs to RG point data.

With regard to the fact that most of the existing scientific papers only include and analyse a few precipitation datasets (Cohen Liechti et al. 2012; Thiemig et al. 2013; Tobin and Bennett 2014), this study further contributes to the ongoing debate on satellite derived precipitation products' performances by focusing on eight different state of the art satellite precipitation datasets, one primarily reanalysis product, and one rain gauge dataset.

SRFE and reanalysis product validation studies over West Africa were carried out, among others, by Awange et al. (2015), Gosset et al. (2013), Nicholson et al. (2003), Pfeifroth et al. (2016) and Thiemig et al. (2012). All of the aforementioned use at least one source of RG data in order to assess the performance of the SRFE or RA product. A major source of RG data is the Global Precipitation Climatology Center (GPCC) product, which has been used to validate precipitation datasets in Africa, e.g. by Adeyewa and Nakamura (2003). However, large data gaps exist over Africa, and the number of gauges sending data varies over time (Awange et al. 2015). Figure 4.1 gives an overview of GPCC pixels including station data for at least 80% of days for the period from 2003 to 2013. Also, gauges connected to the Global Telecommunication System (GTS) are frequently used for bias correction of satellite data, creating merged satellite rain gauge datasets like TMPA 3B42 or PERISANN CDR (Ashouri et al. 2015; Huffman et al. 2007). It remains questionable whether the performance of bias corrected SRFEs should be evaluated against data from the same rain gauges used for the bias correction since the datasets are not independent (Awange et al. 2015; Maidment et al. 2013; Novella and Thiaw 2013; Thiemig et al. 2012).

A new approach which has recently gained popularity is the so called hydrological evaluation of SRFE and RA products. In this approach, the performance of the datasets is evaluated by calibrating a hydrological model for different precipitation products and evaluating the ability of each product to reproduce the observed streamflow. This method on the one hand circumvents the need for reliable RG information, but on the other hand introduces the need for streamflow information as a variable for model calibration (Behrangi et al. 2011; Thiemig et al. 2013). Since the hydrological evaluation takes place at the watershed scale, no problem of scale discrepancy arises as with the validation by using point source RG data. Also, the usefulness of the SRFEs and RA products to generate streamflow is directly evaluated within the hydrological model of

choice (Bitew and Gebremichael 2011; Thiemig et al. 2013). It should however also be considered that the model concept and process representation introduces uncertainties into the analysis. So can the results of excess or below average precipitation for example be dampened by the parameters governing infiltration and evaporation (Seibert 1997). Nevertheless, when using one single model concept which has been shown to be applicable to the climatic conditions, the results may be biased but this bias would be consistent over all precipitation products and therefore does not affect the interpretation and evaluation.

Studies using this method in recent years are plentiful, e.g. Behrangi et al. (2011), Bitew and Gebremichael (2011), Bodian et al. (2016), Cohen Liechti et al. (2012), Dile and Srinivasan (2014), Fujihara et al. (2014) and Li et al. (2015). However, most of these studies focus on areas outside of West Africa, where only few hydrological evaluations were conducted so far, covering a limited number of different products and climatological regions. A need for more studies focusing on West Africa and featuring more products is apparent.

Some of the first authors to validate SRFE and RA products in West Africa were Thiemig et al. (2013), who evaluated CMORPH (Climate Prediction Center Morphing Technique), RFE 2.0 (African Rainfall Estimation), TMPA 3B42 (Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis), PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks) and ERA-Interim over the Volta basin using the LISFLOOD model. They concluded that the hydrological evaluation using this model is a well-suited approach to validate precipitation datasets over sparsely or ungauged catchments, with their model producing better results in the high flow than during the low flow periods. During this study, some products were bias corrected for a second simulation (e.g. PERSIANN and CMORPH, which exhibited biases over lowland areas) using gauge measurements, and generally produced better results afterwards. They concluded that the best dataset is the one with the best intrinsic data quality needing the least amount of preprocessing and bias correction. For the Volta basin, RFE 2.0 and TMPA 3B42 were asserted as best performing. Also in 2013, Gosset et al. validated PERSIANN, CMORPH, TMPA 3B42 real time, GSMaP (Global Satellite Mapping of Precipitation) MVK and real time, GPCP (Global Precipitation Climatology Project), TMPA 3B42, RFE 2.0 and EPSAT-SG (Estimation of Precipitation by SATellite Second Generation) using high resolution gauge data from two AMMA-CATCH sites in Niger and Benin as well as two hydrological models. The Soil Conservation Service method (SCS) was used to generate runoff for the Niamey site and the lumped conceptual Modèle du Génie Rural à 4 paramètres Journalier (GR4J) was used to generate streamflow for the Benin site. For the statistical evaluation, the authors concluded that TMPA 3B42 performed best on both sites with very low biases. RFE 2.0 exhibited low biases but also low correlations to gauge data with an overestimation of low rain rates. Concerning the hydrological evaluation in the Niamey area, the two products best able to generate streamflow are TMPA 3B42 (slight overestimation) and RFE 2.0 (low bias but underestimated streamflow). The evaluation conducted for the Ouémé area revealed RFE 2.0 and EPSAT to perform best with a slight underestimation of the discharge. TMPA 3B42 was found to have a high inter-annual variability of the biases and did not provide adequate results.

In 2014, Fujihara et al., using a conceptual hydrological model, assessed that a bias-corrected

SRFE dataset outperforms RA datasets over the White Volta basin. Results showed that while the reanalysis products exhibited acceptable seasonal accuracy, annual variations were not reproduced. The only merged SRFE/RG product used (GPCP) delivered good results and proved to be superior. Very recently, Bodian et al. (2016) conducted a hydrological evaluation of TMPA 3B42 precipitation estimates over the Upper Sénégal river basin, also using the GR4J model. Results showed good agreement with measured discharge.

As stated before, the current study is a contribution to the ongoing debate on satellite derived and reanalysis precipitation products' performances. Specifically, we aim at exploring the differences between the considered datasets for basins of different locations and sizes and evaluating the robustness of the products in accurately simulating streamflow during a hydrological evaluation. Ten state of the art datasets which were generated according to different methods were chosen for the evaluation. For the first time, the conceptual HBV model is used in a hydrological evaluation, generating knowledge in its applicability in exploring the different capabilities of precipitation datasets to generate streamflow. HBV was chosen due to it being user-friendly, easily accessible and free of charge while having been proven to deliver robust results (Seibert and Vis 2012), in order to entice a maximum of researchers to adopt this validation approach. Conceptual models like HBV light were already successfully applied in simulating discharge in West Africa, e.g. by Cornelissen et al. (2013).

4.2 Materials and Methods

4.2.1 Study Area

The basins used to evaluate the performance of the selected precipitation estimates in generating daily streamflow were chosen within the COAST (Studying changes of sea level and water storage for coastal regions in West-Africa, using satellite and terrestrial data sets) project study area in West Africa. Large (about 100,000 km^2) as well as small ($\sim$2500 km^2) sized basins were identified both in the south, which is characterized by a bimodal rainy season, and the north, which is subject to a unimodal rainy season, in order to evaluate differences arising from the diverse rainfall distributions as well as the increase in aridity from south to north Fink et al. (2010). The research area (5°N-15°N and 6°W-4°E) is depicted in Figure 4.1. Specific basins were chosen based on the availability and reliability of discharge data, an overview of the chosen basins is given in Table 4.1. Subwatersheds within the selected watershed were delineated using the ArcSWAT model with a threshold area of at least 2500 km^2. This seems a reasonable subbasin size if one wants to model the whole of the subcontinent, as is planned for the future. Schuol et al. applied the SWAT model to estimate the freshwater availability in West Africa in 2008 and suggested a minimum drainage area for the subbasins of 10,000 km^2.

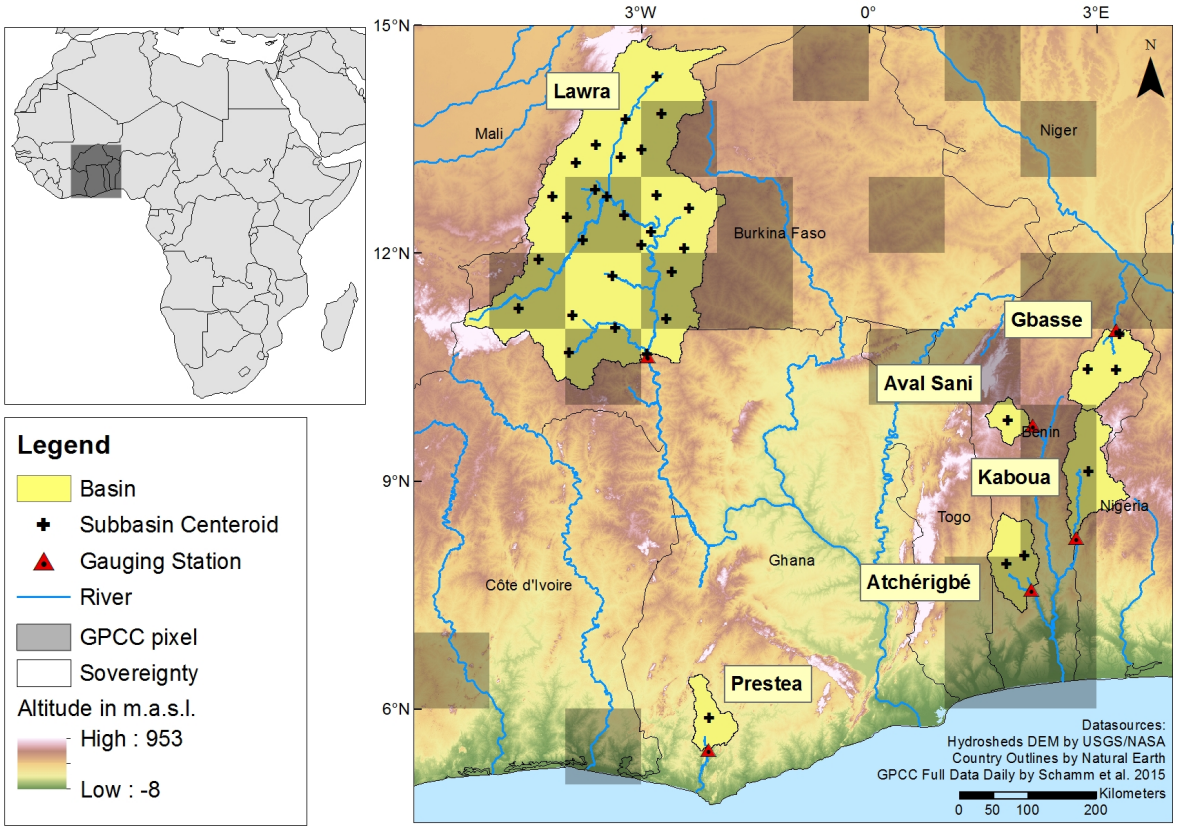

Figure 4.1: Research area. Pixels of precipitation products including the subbasin centroids were used for data extraction. Highlighted GPCC pixels have station data for at least 80% of days of the period 2003–2013.

Table 4.1: Selected basins.

Subbasin	River	Country	Area	River Basin	Discharge Observations
Prestea (PR)	Ankobra	Ghana	4305 km^2	Ankobra	1998–2006
Lawra (LA)	Black Volta	Burkina Faso, Ghana, Mali	106,074 km^2	Volta	1998–2006
Aval Sani (AS)	Ouémé	Benin	2594 km^2	Ouémé	1999–2011
Kaboua (KA)	Okpara	Benin, Nigeria	9714 km^2	Ouémé	1998–2010
Atchérigbé (AT)	Zou	Benin, Togo	7115 km^2	Ouémé	2001–2010
Gbasse (GB)	Sota	Benin	8504 km^2	Niger	2003–2006

4.2.2 Products

Ten precipitation products were identified for evaluation during this study. While this study focuses on satellite precipitation datasets, the GPCC product, which relies on station data, was included since it is often used to compare remotely sensed data to observations. We furthermore included the CFSR reanalysis product because it is often used for SWAT simulations as data can be downloaded in SWAT format. Although SWAT-ready CFSR is widely used by modelers in

Africa in our experience, we remain convinced that an evaluation of the data quality is necessary before use. Each product will be briefly described in this section with an overview being given in Table 4.2.

Table 4.2: Selected precipitation products. RA: reanalysis, MW: microwave imager, IR: infrared, RG: rain gauges (used for bias correction), pres: present, FDDv1: Full Data Daily Version 1, cov: coverage, res: resolution.

Product	Instrument	Spatial cov.	Spatial res.	Temp. cov.	Provider	Reference
CFSR	RA, RG	90°N-90°S	0.3125°	1979–2014	NOAA-NCEP	Saha et al. (2010)
CHIRPS	MW,IR,RG	50°N-50°S	0.05°	1981-pres.	NOAA-CPC	Funk et al. (2015)
CMORPH v1.0 RAW	MW,IR	60°N-60°S	0.25°	1998-pres.	NOAA-CPC	Joyce et al. (2004)
CMORPH v1.0 CRT	MW,IR,RG	60°N-60°S	0.25°	1998-pres.	NOAA-CPC	Joyce et al. (2004), Xie et al. (2011)
PERSIANN CDR	MW,IR,RG	60°N-60°S	0.25°	1983-pres.	NOAA-NCDC	Ashouri et al. (2015)
RFE 2.0	MW,IR,RG	AFRICA	0.1°	2000-pres.	NOAA-CPC	The NOAA Climate Prediction Center (2002)
TAMSAT	IR,RG	AFRICA	0.0375°	1983-pres.	University of Reading	Maidment et al. (2014), Tarnavsky et al. (2014)
TMPA 3B42 v7	MW,IR,RG	50°N-50°S	0.25°	1998-pres.	NASA/JAXA	Huffman et al. (2007)
TMPA 3B42RT v7	MW,IR	60°N-60°S	0.25°	2000-pres.	NASA/JAXA	Huffman et al. (2007)
GPCC FDDv1	RG	90°N-90°S	1°	1988–2013	DWD	Schamm et al. (2015)

CFSR: The CFSR (Climate Forecast System Reanalysis) coupled atmosphere-ocean-land system is developed at NOAA-NCEP and provides a variety of atmospheric parameters including precipitation, temperature, wind speed, relative humidity and radiation, available from 1979 to present (Blacutt et al. 2015; Dile and Srinivasan 2014; Saha et al. 2010). CFSR is based on historical and operational archives and incorporates various data sources such as radiosondes, surface observations including rain gauge information or data from satellite instruments (Saha et al. 2010). Furthermore, precipitation estimates are updated every 6 h in near real time (Fuka et al. 2013).

CHIRPS: CHIRPS (Climate Hazards Group Infrared Precipitation with Station data) is a dataset developed especially for drought monitoring purposes. The data is generated from 1981 to present (Funk et al. 2015). Compared to other precipitation products, its main characteristic is its very fine spatial resolution of 0.05° (Katsanos et al. 2016). CHIRPS uses several data sources, such as the monthly precipitation climatology CHPclim, infrared measurements from geostationary satellites, and information from the TMPA 3B42 product. The precipitation estimates are merged with *in situ* gauge data from several archives including GTS to reduce biases (Funk et al. 2015).

CMORPHv1.0 RAW and CRT: CMORPH (Climate Prediction Center Morphing Technique) is developed by NOAA-CPC and is a global precipitation analysis algorithm, with data available from 1998 to present 18 h after real time observation. Originally providing precipitation estimates for the period 2002 onwards, two new versions of CMORPH have been released recently and labeled as CMORPHv1.0 RAW and CRT. The former CMORPH dataset has been renamed to CMORPHv0.x (Koutsouris et al. 2016). Two versions are available: CMORPH RAW derives

half hourly precipitation estimates which are based on satellite microwave (MW) and infrared (IR) information. The precipitation estimates are morphed through a time-weighted linear interpolation (Joyce et al. 2004; Ramarohetra et al. 2013). CMORPH RAW is a satellite-only product, since no RG information is incorporated (Joyce et al. 2004; Thiemig et al. 2013). CMORPH CRT is similar to CMORPH RAW, but has been bias-corrected using historical and real-time gauge data (Xie et al. 2011).

PERSIANN CDR: PERSIANN CDR (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record) is maintained by the University of California together with NOAA and provides data from 1983 to present (Ashouri et al. 2015). Precipitation estimates are calculated from IR and MW satellite data (Ashouri et al. 2015; Jobard et al. 2011). Unlike the PERSIANN product (Hsu et al. 1997), which is available in near real time and solely based on satellite measurements (Awange et al. 2015; Jobard et al. 2011; Ramarohetra et al. 2013), PERSIANN CDR incorporates GPCP gauge data into its precipitation estimates (Ashouri et al. 2015).

RFE 2.0: The RFE (African Rainfall Estimation) 2.0 algorithm was developed by the NOAA CPC. Daily rainfall estimates are available from late 2000 to present (The NOAA Climate Prediction Center 2002). The dataset is based on IR and MW data and incorporates rain gauge information from the GTS network, which is used to correct biases (Cohen Liechti et al. 2012; Thiemig et al. 2012). Up to 1000 GTS stations are merged to derive the final precipitation product (Ramarohetra et al. 2013). The product was originally developed for drought monitoring purposes on the African continent and is available within 24 h after real time (Jobard et al. 2011).

TAMSAT: TAMSAT (Tropical Applications of Meteorology using Satellite data and ground-based observations) was developed at the University of Reading specifically for continental Africa with a spatial resolution of 0.0375°, which makes it the finest resolution among all the datasets considered. Rainfall estimates from TAMSAT are available from 1983 to present. The TAMSAT method uses IR imagery from METEOSAT (Grimes et al. 1999; Maidment et al. 2014; Tarnavsky et al. 2014). Rain rates are then calculated based on a cloud temperature threshold, which varies according to time and location (Jobard et al. 2011). Contrary to other merged products, TAMSAT does not use GTS data but historical data of about 4000 stations from various African agencies mostly acquired since the early 1990s (Maidment et al. 2014).

TRMM TMPA 3B42v7: The TRMM (Tropical Rainfall Measuring Mission) Multi-satellite Precipitation Analysis (TMPA) product is produced at NASA's Goddard Space Flight Center (Huffman et al. 2007), with TMPA 3B42v7 (hereafter named TMPA) being the latest version. Originally developed for rainfall retrievals in the tropics, it has been extended to a quasi-global coverage. Rainfall estimates are available for the period of 1998 to present. The TMPA algorithm integrates rainfall estimates from multiple satellites and a variety of sensors. Passive MW data is derived from several low earth orbiting satellite sensors additionally to satellite IR data (Jobard et al. 2011). Rain gauge observations from the GPCC are used to bias correct the precipitation estimates (Moazami et al. 2013; Worqlul et al. 2014).

TRMM TMPA 3B42v7 RT: Similar to the TMPA 3B42v7 dataset, the TMPA 3B42v7RT (Real-Time, hereafter named TMPA RT) product is produced by NASA. It provides daily precipitation estimates from March 2000 to the present day. The product is based on the same algorithm as TMPA 3B42v7, which combines MW and IR estimates. Unlike 3B42v7, no rain gauge data is used for bias-correction due to the real-time nature of the product. However, a climatological adjustment is performed (Huffman et al. 2007). Its precipitation estimates are available nine hours after overflight (Huffman et al. 2007; Jobard et al. 2011).

GPCC FDDv1: The GPCC FDDv1 product (Global Precipitation Climatology Centre Full Data Daily 1°) developed at the Deutscher Wetterdienst is a global land-surface precipitation dataset that covers the period from 1988 to 2013. It is based on station data from up to 29,000 gauge stations per month, including GTS stations with automated quality control. The GPCC FDDv1 precipitation estimates, hereafter named GPCC, provide a temporal resolution of one day, while the spatial resolution is 1°, making it the coarsest precipitation dataset among the analyzed products. Interpolation through a block Kriging scheme with a global variogram is carried out for each station, producing the daily rainfall totals (Schamm et al. 2015; Schneider et al. 2015).

4.2.3 Extraction of Precipitation Data

The chosen SRFEs and RA rainfall estimates were acquired *via* various online sources, all being freely available to the general public. With the exception of CFSR, all datasets were distributed in the common netCDF data format. In order to extract the precipitation data for the research basins, we followed the approach utilized by the semi-distributed, physically based Soil and Water Assessment Tool (SWAT) hydrological model (Arnold et al. 1998), which uses weather data from the climate station located closest to the center of each subbasin. This was accomplished by deriving the centroid for each subbasin and extracting the data from the precipitation pixel covering this point (Winchell et al. 2010). For larger basins, namely Atchérigbé, Gbasse and Lawra, precipitation data was likewise derived for each subbasin and in a second step area-averaged to generate the data for the whole basin. This way, the performance of the approach for single-cell as well as area-averaged data can be evaluated. When we compared the SWAT centroid method to calculating the area-average of all pixels in the respective basins for the TMPA dataset from 1998 to 2013, the correlation of the daily data for both methods was high, ranging from $r = 0.87$ to $r = 0.99$ with a mean r of 0.91. A comparison of the total rainfall sums generated by both methods revealed an r value of 0.99. The average monthly rainfall sums were also highly correlated with an r of between 0.998 and 0.999 for all six considered basins. However, when we considered the amount of rainy days, the basin-averaged method included between 10.6 and 29.4% (average: 17.8%) more rainy days than the distributed method. Therefore, the distributed approach was chosen in order to preserve a representative account of the number of rainy days while at the same time not increasing or decreasing total rainfall sums. This method was also chosen by Bitew and Gebremichael for basins of 299 and 1656 km^2 size over complex topography in Ethiopia; Dile and Srinivasan for 5–30 km^2 basins, also in Ethiopia and Tobin; and Bennett for basins of between 175 and 200 km^2 in Texas (Bitew et al. 2012;

Bitew and Gebremichael 2011; Dile and Srinivasan 2014; Tobin and Bennett 2009).

4.2.4 Model

The hydrological evaluation was carried out using the conceptual HBV light model (Seibert 2000). HBV light is a development of the HBV model developed in the 1970s by the Swedish Meteorological and Hydrological Institute (SMHI) and its current version was made available in 2010 (Seibert and Vis 2012; Singh 1995). Compared to other hydrological models, the relatively low demand for input data and the limited number of parameters are strengths of the HBV model (Rientjes et al. 2013; Rusli et al. 2015). The structure of HBV consists of four main routines (snow-, soil-, routing- and response routine) and simulations of daily discharge are based on time series of precipitation and temperature, as well as potential evapotranspiration (Seibert and Vis 2012). Since no snowfalls occur in the research area, the snow routine was not considered in this study. As shown in Figure 4.2, liquid precipitation contributes to soil moisture and runoff depending on the actual moisture status. The remaining precipitation contributes to the soil moisture storage, which can still evaporate as long as there is enough soil water available. This means that in the soil routine, groundwater recharge and actual evaporation are simulated as functions of actual water storage (Seibert and Beven 2009). Three components, namely runoff, interflow and baseflow contribute to the discharge at the outlet of the modeled basin (Aghakouchak and Habib 2010). The response routine uses three linear reservoir equations and the final routing routine uses a triangular weighting function to compute discharge (Seibert and Beven 2009).

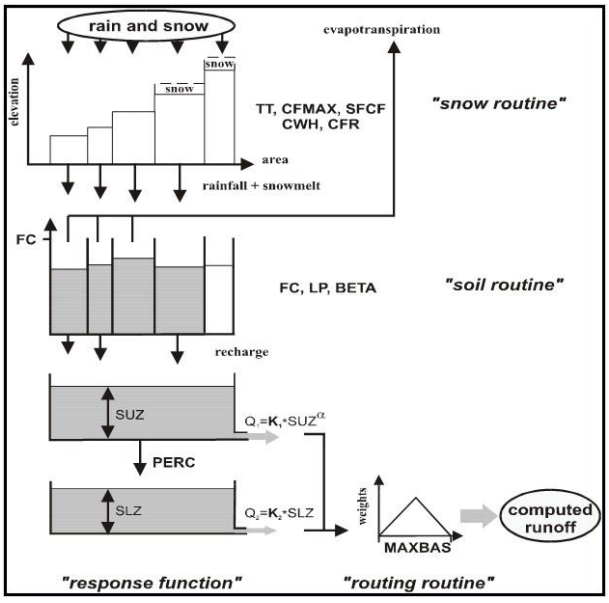

Figure 4.2: HBV model structure (Seibert 2000).

The Genetic Algorithm and Powell (GAP) optimization was used to calibrate the model. The default model parameter ranges were chosen for the calibration, as they are described to be realistic by the developer (Seibert 2000; Uhlenbrook et al. 1999) and similar sets have been used in a multitude of studies (Radchenko et al. 2014; Seibert 1999, 1997). Increasing the parameter

range bears the risk to generate a good simulation by using unrealistic parameter sets. This is especially the case when rainfall is extremely over- or underestimated and a good discharge simulation is obtained at the expenses of a realistic simulation of evapotranspiration. Since the goal was to compare the simulation results of the precipitation products among each other and not to reach the absolute highest model efficiency, these ranges were accepted for every model run. The calibrated parameters and their initial ranges are shown in Table 4.3.

The model was calibrated for each precipitation product and basin separately in 10,000 GAP runs, calibration and validation periods are supplied in Table 4.4. This method was applied to explore the independence of the results from reference data (see also Behrangi et al. (2011), Bodian et al. (2016), and Tobin and Bennett (2009)). However, many studies evaluating the efficiency of SRFE and RA products initially calibrate a model using observed rain gauge data (Bitew et al. 2012; Gourley et al. 2011; Thiemig et al. 2013). Because no rain gauge data was available for this study, the only alternative was to use GPCC data. Since some of the evaluated products were bias-corrected using GPCC data, the question arises whether this method does not discriminate products not including GPCC data in their estimates. The same holds true for a classical comparison between GPCC and GPCC-corrected datasets. To test this hypothesis, the model was also initially calibrated with GPCC data for the Lawra basin and the initial parameters were applied to the other precipitation products.

Table 4.3: Parameters and their initial (default) ranges in the gap optimization runs.

Parameter	Description	Unit	Min.	Max.
Soil moisture routine				
F_c	Maximum soil moisture storage	mm	100	550
L_p	Soil moisture threshold for reduction of evaporation	mm	0.3	1
BETA	Shape coefficient		1	5
Response routine				
PERC	Maximum flow from upper to lower groundwater box	mm d^{-1}	0	4
UZL	Threshold parameter for K_0 outflow		0	70
K_0	Recession coefficient	d^{-1}	0.1	0.5
K_1	Recession coefficient	d^{-1}	0.01	0.2
K_2	Recession coefficient	d^{-1}	0.00005	0.1
Routing routine				
MAXBAS	Routing, length of weighting function	d	1	2.5

Table 4.4: Calibration and validation periods.

	Prestea/Lawra	Aval Sani	Kaboua	Atchérigbé	Gbasse
Warm-up	2001	2001-2002	2001	2001-2002	2003
Calibration	2002-2005	2003-2008	2002-2008	2003-2007	2004-2005
Validation	2006	2009-2011	2009-2010	2008-2010	2006

As the objective functions, the Nash Sutcliffe efficiency (NSE) and Percent BIAS (PBIAS) were chosen. The NSE is calculated as shown in Equation 4.1.

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n} (Y_i^{obs}) - (Y_i^{sim})^2}{\sum_{i=1}^{n} (Y_i^{obs}) - (Y^{mean})^2} \right] \tag{4.1}$$

where Y_i^{obs} is the i^{th} observation of the variable to be evaluated, Y_i^{sim} is the i^{th} simulation of the variable to be evaluated, Y^{mean} is the mean of the observed variables and n is the number of observations. The NSE can take values between -∞ and 1. The performance ratings proposed by Moriasi et al. (2007) were applied in this study. The result of a simulation is unsatisfactory if the NSE is ≤ 0.5, satisfactory if between > 0.5 and ≤ 0.65, good if between > 0.65 and ≤ 0.75 and very good if between > 0.75 and ≤ 1.

PBIAS expresses the deviation of the data to be evaluated in percent and is calculated according to Equation 4.2.

$$PBIAS = \frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim}) * 100}{\sum_{i=1}^{n} (Y_i^{obs})} \tag{4.2}$$

where Y_i^{obs} is the i^{th} observation of the variable to be evaluated and Y_i^{sim} is the i^{th} simulation of the variable to be evaluated. It should be noted that positive values indicate an underestimation and negative values an overestimation (Moriasi et al. 2007).

Since no data on the potential evapotranspiration was available for the chosen basins, ETP was calculated from the CFSR reanalysis data using the Penman-Monteith approach (Allen et al. 1998) shown in Equation 4.3:

$$ETP = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \tag{4.3}$$

The parameters refer to "A hypothetical reference crop with an assumed crop height of 0.12 m, a fixed surface resistance of 70 s m^{-1} and an albedo of 0.23." (Allen et al. 1998). Where ETP is the reference evapotranspiration in mm/day, Δ is the slope vapor pressure curve in kPa per °C, R_n is the net radiation at the crop surface in MJ/m^2/day, G is the soil heat flux density in MJ/m^2/day, γ is the psychrometric constant in kPa/°C, T is the mean daily air temperature at 2 m height in °C, u_2 is the wind speed at 2 m height in m/s and $e_s - e_a$ is the saturation vapor pressure deficit in kPa.

Since ten precipitation products were evaluated for six basins, 60 models were calibrated and a validation run was performed for each product and basin. In the last step of the evaluation, a ranking system was developed in order to gain a one-glance overview of the performance of each product using NSE and PBIAS values. Products received one point for a very good NSE, two points for a good NSE, three points for a satisfactory NSE, and four points for an unsatisfactory NSE. The same was applied for PBIAS with one point being given for a very low PBIAS ($<$ 5%), two points for a low PBIAS (5 $<$ 15%), three points for medium PBIAS (15 $<$ 30%) and four points for a high PBIAS ($>$ 30%). Products were considered as performing very good if the score was between two and three, good if between four and five, satisfactory if between six and seven and unsatisfactory if the score was higher than seven.

4.3 Results and Discussion

First, the GPCC Full Data Daily (FDD) 1° product gauge coverage was evaluated for the research area (5°N-15°N and 6°W-4°E, see Figure 4.1). Since the number of stations reporting for each grid cell are supplied in the GPCC dataset on a daily basis, cells with actual station data for the period of interest can be derived, as previously shown by Nikulin et al. (2012). The percentage of pixels with reporting station per observation days (in percentage) over the period of 2003–2013 is shown in Figure 4.3. While a maximum of 48% of pixels include a station for 5% of observation days, the pixel count quickly reduces if longer timelines are considered. For 50% of the period, only 33% of the pixels have reporting data included and for 80% of days, this reduces to 23%. The results show large areas of irregular gauge measurements and interpolation. Therefore, the GPCC product should by no means be considered as a high resolution gauge dataset in this region.

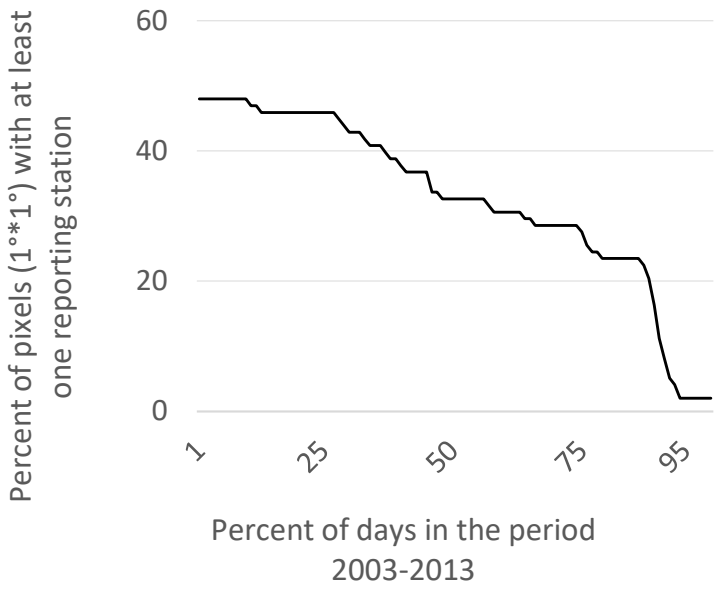

Figure 4.3: Overview of GPCC station data in percent of pixels with reporting station for the research area of 5°N-15°N and 6°W-4°E (see Figure 4.1).

4.3.1 Comparison of Precipitation Estimates

In a preliminary data exploration, the precipitation estimates derived from all ten products for the six basins were compared. This was conducted in two steps. In the first step, the average yearly and monthly sums for the eleven-year period from 2003 to 2013 (the only period where data from all ten precipitation estimates was available) were calculated for each product and basin and compared in Figure 4.4.

Results show that in all basins except for Prestea, the average yearly sums of PERSIANN CDR, TMPA, CHIRPS, CMORPH CRT, RFE 2.0 and GPCC are similar. Average deviations from the mean are 0.1% for PERSIANN CDR, +2.2% for TMPA, -0.8% for CHIRPS, -5.9% for CMORPH CRT, -6.1% for RFE 2.0 and +1.1% for GPCC. In the Prestea Basin, RFE 2.0 and

TAMSAT estimate less precipitation than the average. The CFSR reanalysis product generally estimates low precipitation rates in the central basins of Aval Sani, Kaboua and Gbasse with an average of 25.8% less than the mean. In the northernmost basin of Lawra, the average annual precipitation is decidedly lower with a deviation from the average of -63.9%, the lowest estimation of all products. In the southern basins however, CFSR tends to deliver higher precipitation amounts with values for Atchérigbé being close to the mean (-5.6%). In the southernmost Prestea basin, CFSR seems to overestimate the annual precipitation, deviating from the mean by 31.1%.

For the CMORPH RAW product annual averages seem high, with the highest yearly precipitation of all products being reached in all basins except Prestea, where it comes second after CFSR. Deviations from the mean are between +18.0% in Prestea and +48.3% in Lawra, with an average overestimation of 35.8%. TAMSAT generally shows below average annual precipitation, with values for Aval Sani, Kaboua, Atchérigbé and Gbasse being similar to CFSR values. In the Lawra basin, values are also below the average (-23.2%) but higher than CFSR. In Prestea, TAMSAT estimates a lower than average precipitation of -22.2%, which is similar to RFE 2.0 but -760 mm/y lower than CFSR.

The yearly sums generated by the non-bias corrected, real time TMPA RT product are higher than the mean estimates by an average of 16.1% except for the Kaboua basin, where the sums are similar. The overestimation of CMORPH RAW in this region as well as the similar performance of TMPA and RFE 2.0 have also been observed by Thiemig et al. (2012) for the Volta basin, and Gosset et al. (2013) for Benin and Niger. Pfeifroth et al. (2016) and Thiemig et al. (2012) both observe an overestimation of the precipitation by the PERSIANN product, which is not observed in this study, most probably because the PERSIANN CDR product used is created using different input data as well as bias adjustment.

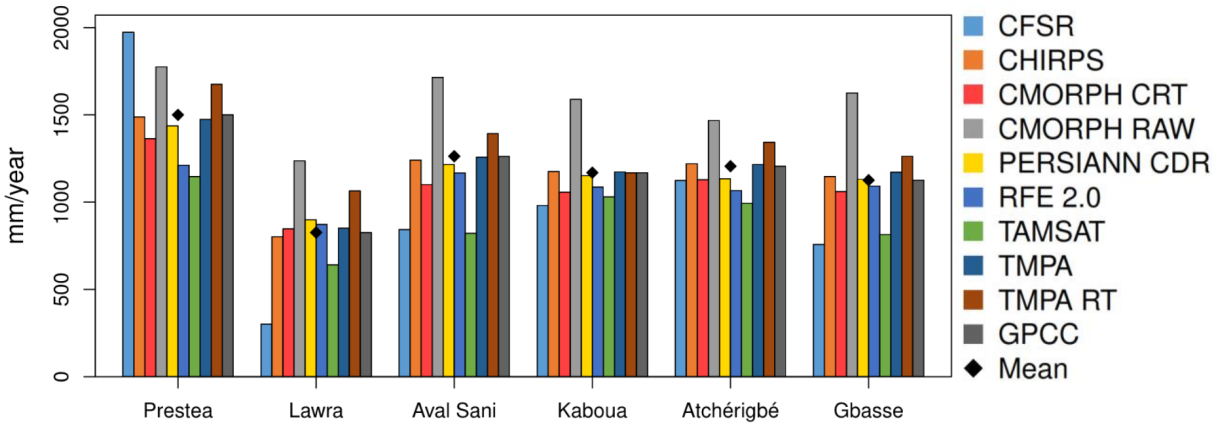

Figure 4.4: Average yearly precipitation (2003–2013) for all products and basins in mm/year.

When comparing the average monthly rainfall distribution for the same period, as done in Figure 4.5, CHIRPS, CMORPH CRT, PERSIANN-CDR, RFE 2.0 and TMPA again perform similarly. While CFSR predicts higher precipitation than the other products in the southern Prestea basin, it generally predicts lower precipitation in the other basins located further to the north, the extreme case being the Lawra basin, as well as a distinctly late onset of the rainy seasons. In the Kaboua basin, CFSR rainfalls are very high during the beginning of the

rainy season between July and August and in the Atchérigbé basin, a peak can be observed during August. CMORPH RAW estimates very high peaks for the first rainy season in the southern Prestea and Atchérigbé basin as well as the central Kaboua basin, while the rainfall distribution of the Lawra, Aval Sani and Gbasse basins follows the average, although with a clear overestimation.

In the Kaboua and Atchérigbé basins, CMORPH RAW predicts higher than average precipitation from March to May. TAMSAT rainfall estimates follow the mean distribution but exhibit lower than average peaks in all basins. TMPA RT precipitation is close to the average in the Prestea, Kaboua and Gbasse basins but overestimates the peaks in the Lawra and Aval Sani basins. In the Atchérigbé basin, TMPA RT predicts a peak in May, which is not observed by other products with the exception of CMORPH RAW. GPCC generally performs very close to the mean values of all products. With the exception of GPCC, products relying on only two out of the three possible input data sources (IR, MW, RG) display the highest deviations from the mean, notably CFSR (RG, reanalysed), CMORPH RAW (MW, IR), TAMSAT (IR, RG) and TRMM RT (IR, MW).

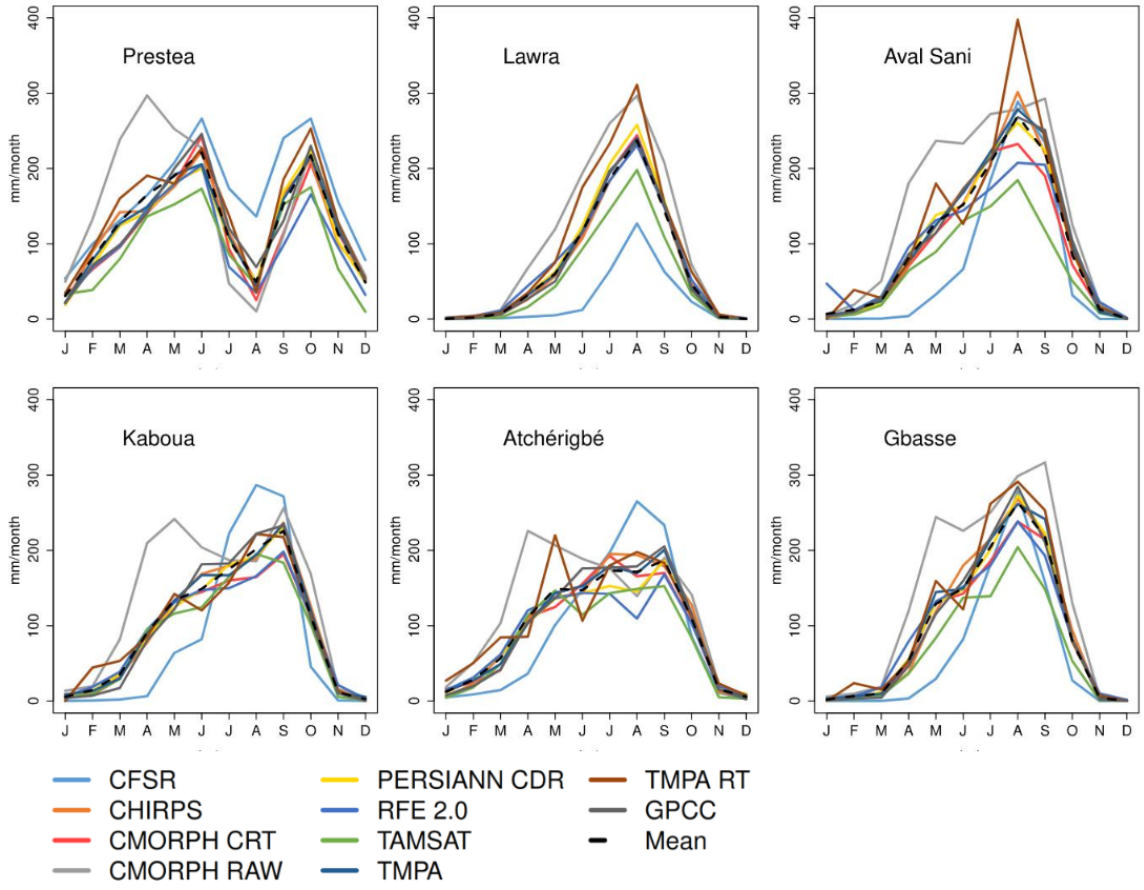

Figure 4.5: Average monthly precipitation (2003–2013) for all products and basins in mm/month.

4.3.2 Results of the Hydrological Evaluation

The influence of the model calibration strategy on the SRFEs products' performances for the Lawra basin is shown in Figure 4.6. Simulations using GPCC-based model optimum parame-

ters (initial calibration) are compared to simulations using default model parameter ranges for individual SRFE products (each product). The results show that if the model is calibrated using GPCC data, the performance of products using GPCC and/or independent RG data for bias-correction (namely CHIRPS, CMORPH CRT, PERSIANN CDR, RFE 2.0, TAMSAT and TMPA) only changes slightly as opposed to calibrations being performed for each product. However, model efficiencies were observed to be significantly lower for non-corrected precipitation products (CFSR, CMORPH RAW, TRMM RT) compared to calibrations for each product individually. It can therefore be argued that an initial calibration using a dataset that has already been used to bias-correct the evaluated datasets introduces further bias into the analysis by discriminating non-corrected products. Therefore, the initial calibration was rejected for this study.

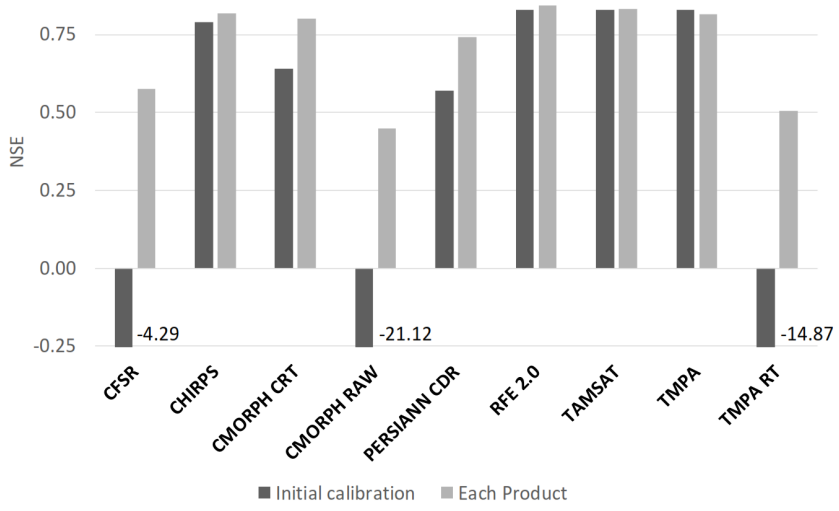

Figure 4.6: Initial calibration results compared to calibrations for each product in the Lawra basin.

The results of the hydrological evaluation are diverse, as is visible in the NSE score diagrams (Figure 4.7). Trends between the products and differences between the basins are evident. For the southern Prestea basin, none of the products yielded good simulations, the only satisfactorily performing being PERSIANN CDR (NSE 0.52) and TMPA RT (NSE 0.50). All other products scored unsatisfactorily, with the worst scoring being CMORPH RAW (NSE -0.53) and CHIRPS (NSE 0.26). For the validation phase, no product performed satisfactorily.

In the northernmost and largest Lawra basin, the overall best calibrations were achieved. Very good NSE values of between 0.80 and 0.84 were attained by CHIRPS, CMORPH CRT, GPCC, RFE 2.0, TAMSAT and TMPA. TMPA RT and CFSR achieved only satisfactory NSE values and CMORPH RAW was identified as performing poorly. For the validation phase, very good simulations were achieved by all products except CFSR and CMORPH RAW, which did not manage to simulate the streamflow accurately. For the Aval Sani basin only CMORPH CRT achieved a very good simulation (NSE = 0.79). CHIRPS, PERSIANN CDR and TAMSAT produced good simulations (NSE between 0.70 and 0.74), while CMORPH RAW, GPCC, RFE 2.0 and TMPA only performed satisfactory and CFSR as well as TMPA RT unsatisfactory. During the validation phase, very good results were observed for CHIRPS and CMORPH CRT, while

CMORPH RAW, GPCC, TMPA and TMPA RT produced good results. For the Kaboua basin calibration, CHIRPS, CMORPH CRT, RFE 2.0 and TAMSAT produced good simulations with NSE values between 0.71 and 0.73. CMORPH RAW, PERSIANN CDR and TMPA performed satisfactorily, while CFSR, GPCC and TMPA RT calibrated unsatisfactorily. Efficiencies are better for the validation period with all products except CFSR performing good or very good simulations. For the calibration of the Atchérigbé basin, good simulations were produced by CHIRPS, TMPA and TMPA RT (NSE 0.65–0.72). While CFSR and CMORPH RAW performed unsatisfactory, all other products performed satisfactory calibrations. For the validation phase, results differ. Here, TMPA produced a very good (NSE = 0.79) and CMORPH CRT, GPCC and PERSIANN CDR good simulations, while CMORPH RAW, RFE 2.0 and TAMSAT only produced satisfactory results. CFSR, CHIRPS and TMPA RT performed unsatisfactorily. For the calibration of the Gbasse basin, TAMSAT produced a very good simulation (NSE 0.76), while CHIRPS, CMORPH CRT, GPCC, PERSIANN CDR, RFE 2.0 and TMPA RT produced good simulations. CMORPH RAW and TMPA performed satisfactorily and CFSR delivered an unsatisfactory result. During the validation, almost all products performed unsatisfactorily with sometimes highly negative NSE values (-7.89 on average), with the exception of CMORPH RAW and TMPA RT, which produced satisfactory simulations.

Considering the performances for all basins during the calibration and validation it becomes apparent that CMORPH CRT performed best, followed by CHIRPS, TAMSAT and TMPA. PERSIANN CDR and RFE 2.0 performed averagely while GPCC, TMPA RT, CMORPH RAW and CFSR performed poorly.

Interestingly, discharges simulated by CMORPH CRT, TMPA, PERSIANN CDR and TMPA RT show similar correlations and standard deviations to observed data in the model results. This might be due to the same resolution of is 0.25° for each product on the one hand and the input data on the other, as some of these datasets share input data from the same satellites. Where products are bias- corrected, the limited number of available gauge products makes it likely that the same data is used for multiple products.

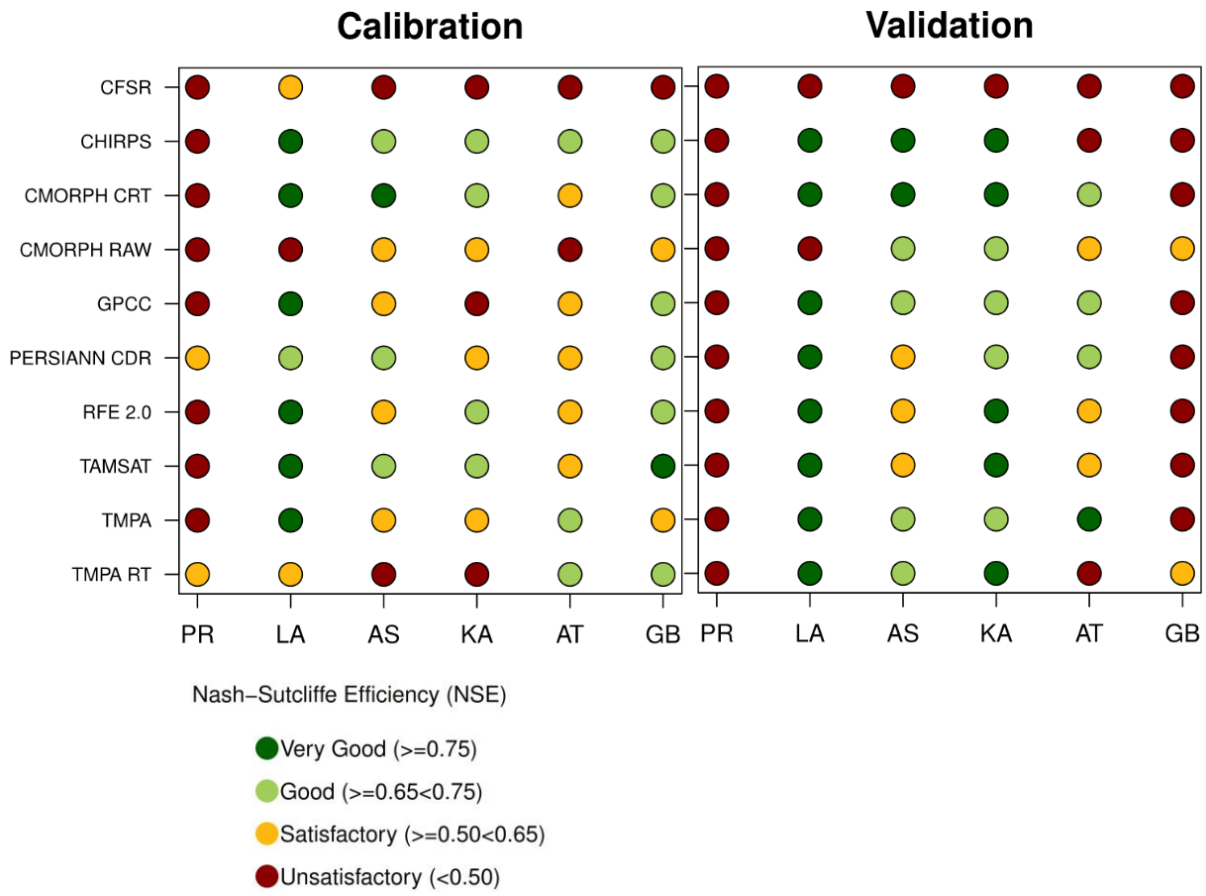

Figure 4.7: NSE score for model calibration and validation period. PR: Prestea, LA: Lawra, AS: Aval Sani, KA: Kaboua, AT: Atchérigbé, GB: Gbasse.

The evaluation of the cumulative discharges and biases for the calibration periods are depicted in Figures 4.8 and 4.9. The cumulative discharges immediately reveal inconsistencies in some of the models. Both CMORPH CRT and RAW models underestimate discharges during certain periods in the Prestea basin but overestimate in mid-2003 in Atchérigbé. During certain points in time CFSR underestimates streamflow in Lawra, Aval Sani, Kaboua, Atchérigbé and Gbasse and TMPA RT underestimates periods in Aval Sani. Interestingly, all models underestimate streamflow in Lawra. Biases for the calibration period are generally low at below 15%. Only in the Lawra basin we observed an underestimation of streamflow by every model, the strongest bias being exhibited by CFSR and TMPA RT (both $> 30\%$). The only product that performed well in this regard is CMORPH RAW ($< 15\%$). While RFE 2.0 and TMPA overestimated in the Aval Sani basin, CHIRPS overestimated in Prestea. For the validation phase, higher biases were exhibited by all models. Especially in the Gbasse basin, all models highly overestimated discharge by over 30% with the exception of CMORPH CRT and TMPA RT. Best performances were achieved in the Lawra, Aval Sani and Atchérigbé basins with only two models each being highly biased (CFSR in Lawra and Aval Sani, TMPA RT in Lawra and Atchérigbé, PERSIANN CDR in Aval Sani and CHIRPS in Atchérigbé). While most models exhibited less than 30% bias in Prestea and Kaboua, high biases were observed for CMORPH RAW, GPCC, TAMSAT and TMPA in Prestea and CFSR, GPCC, PERSIANN CDR and TAMSAT in Kaboua. For both the calibration and validation phase, the models exhibiting the least bias are CMORPH

CRT, TMPA, CMORPH RAW, TAMSAT and TMPA RT.

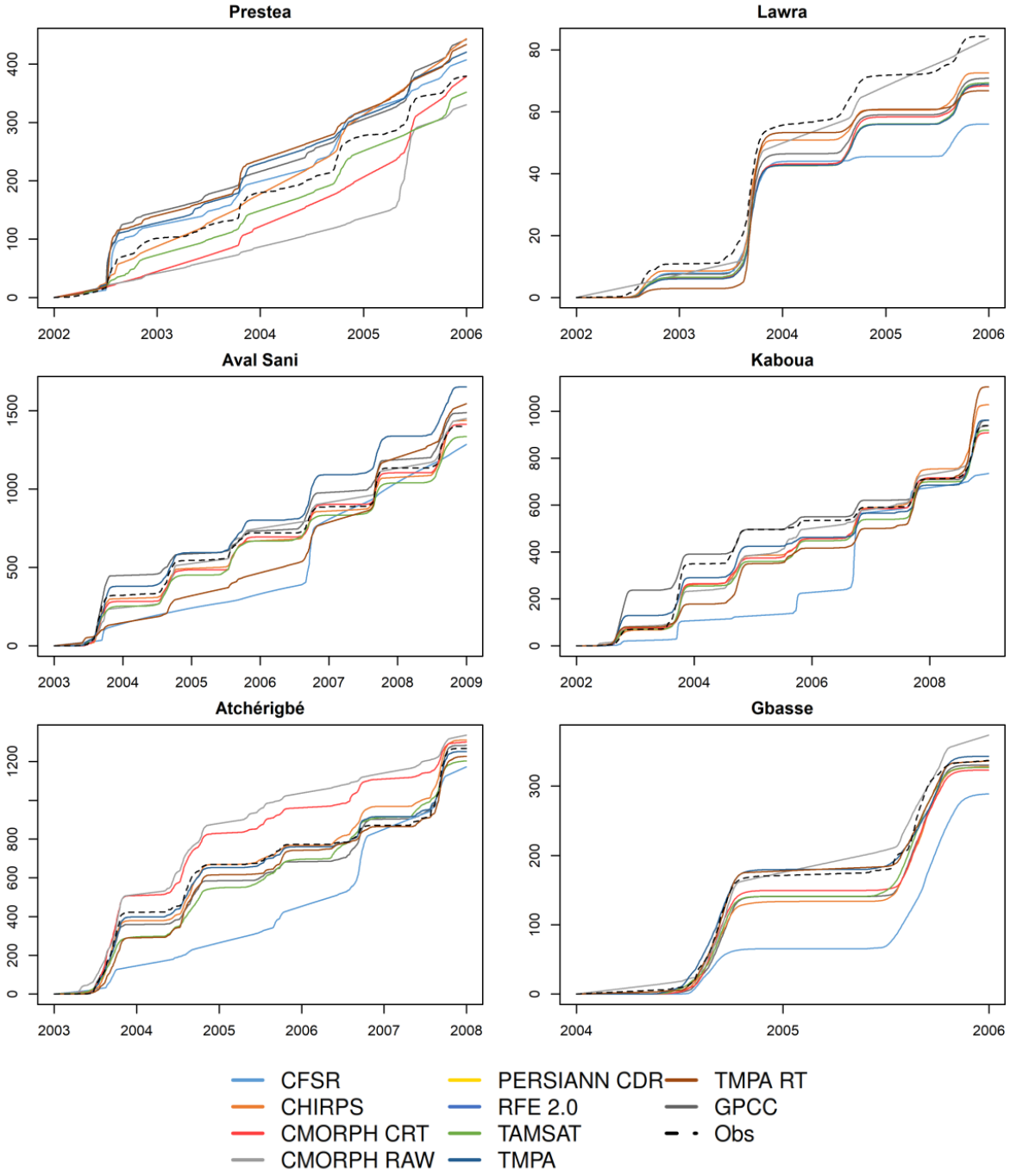

Figure 4.8: Cumulative yearly discharges for the calibration period in mm.

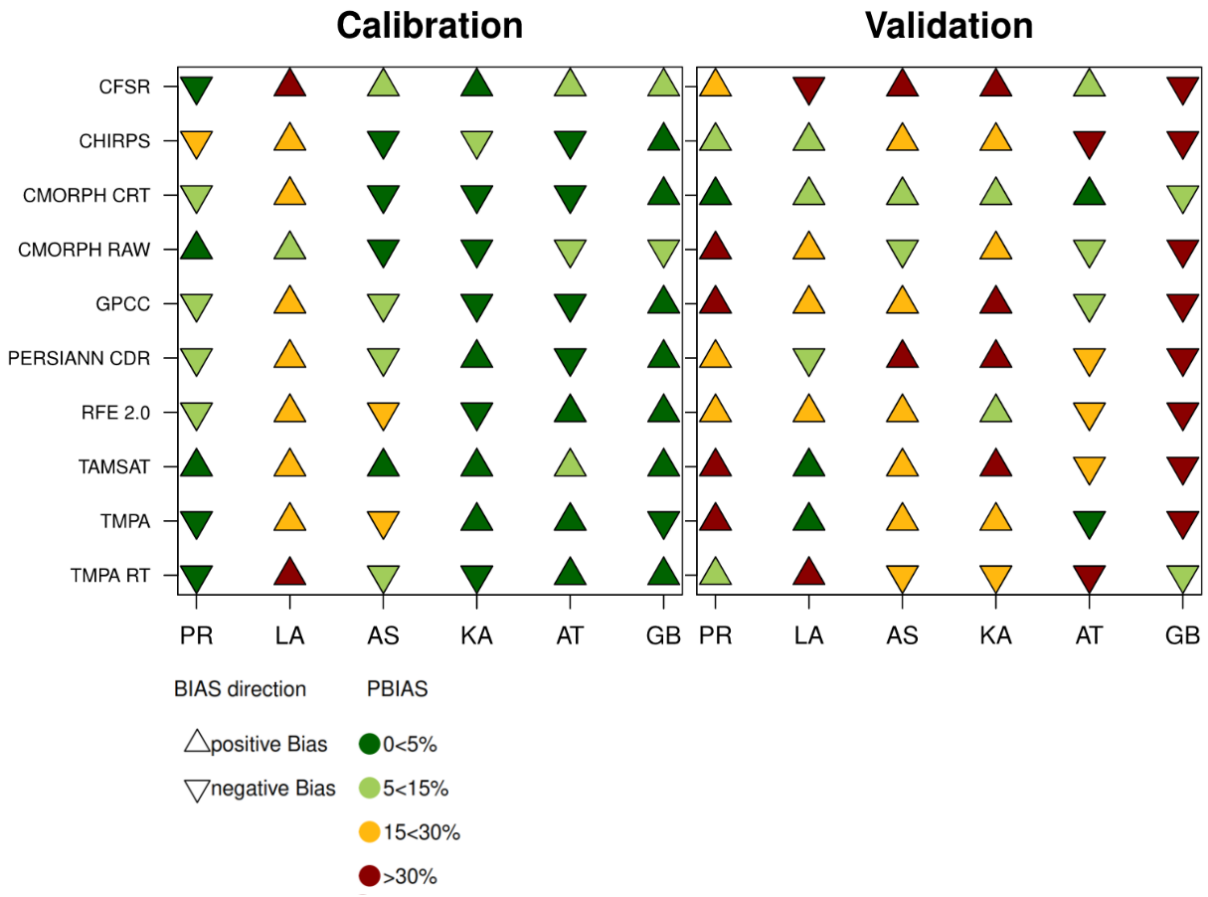

Figure 4.9: Percent bias for model calibration and validation period. PR: Prestea, LA: Lawra, AS: Aval Sani, KA: Kaboua, AT: Atchérigbé, GB: Gbasse.

When we calculated the scores for each product based on NSE and PBIAS during the calibration and validation period as shown in Figure 4.10, best results were achieved by the CMORPH CRT product, followed by TAMSAT, TMPA and CHIRPS. RFE 2.0, PERSIANN CDR, TMPA RT and GPCC which scored well to average and CMORPH RAW as well as TMPA RT which scored unsatisfactory. During calibration, best performances were reached in the Gbasse, Kaboua and Atchérigbé basins, followed by Aval Sani, Presta and Lawra. Interestingly, if NSE and PBIAS values are weighted equally, no product performed unsatisfactorily during the calibration. It follows that products with unsatisfactory NSE values show low biases and vice versa. For the validation period, results differ. Here the worst scores are exhibited for the Gbasse and Prestea basins. This may be due to the short validation periods and missing discharge data. Nevertheless, the CMORPH CRT product scored well in Prestea and satisfactory in Gbasse. The best validations were achieved for Lawra, Kaboua, Aval Sani and Atchérigbé with CMORPH CRT being the only product consistently scoring very well, followed by TMPA with only good or very good scores. CFSR performed worst with only satisfactory scores in Prestea and Atchérigbé and unsatisfactory performances in all other basins.

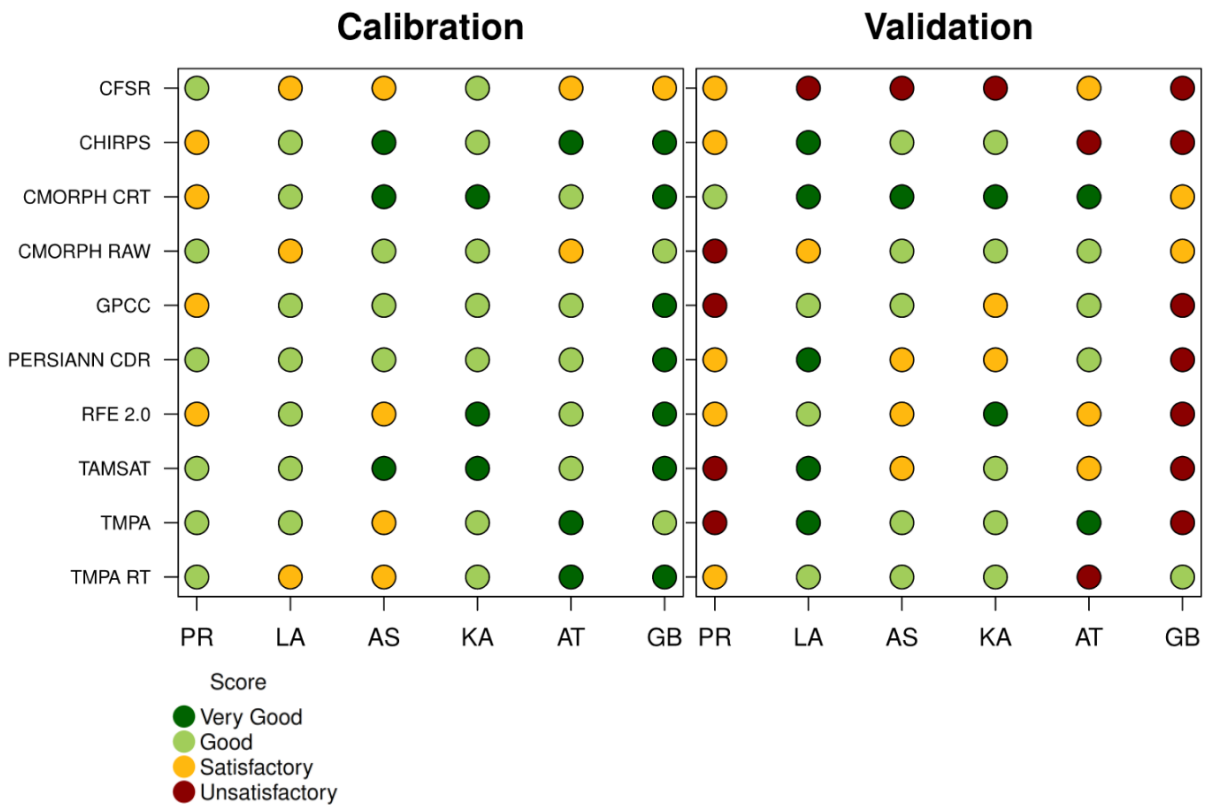

Figure 4.10: Calibration and validation score calculated based on performance of NSE and PBIAS criteria. PR: Prestea, LA: Lawra, AS: Aval Sani, KA: Kaboua, AT: Atchérigbé, GB: Gbasse.

Similar results were achieved by Thiemig et al. (2013), with RFE 2.0 and TMPA performing best over the Volta basin. While the lower performance of their PERSIANN product might be due to the fact that the PERSIANN CDR product was created using different input data, it is interesting to note that Thiemig et al. observed high biases for the CMORPH product which we were also able to reproduce with the RAW product. Gosset et al. (2013) also produced comparable results with RFE 2.0 and TMPA scoring best in Benin and Niger with good model performances and low biases, followed by CMORPH RAW which was biased and performed lower over the Volta catchment. Dile and Srinivasan (2014) discovered high over-and underestimations of CFSR for several subbasins of the Blue Nile Basin. Interestingly, the two products which rely largely on ground data, CFSR and GPCC, produce among the worst results. This may be due to the scarcity of rain gauges in West Africa and the subsequent large scale interpolation of the data. Since the resolution of these two products is the coarsest, this may also influence the model, e.g. in regions where multiple subbasins fall within the same pixel. Products relying on a combination of infrared and microwave satellite instruments as well as rain gauges for bias-correction, such as CMORPH CRT, TMPA, CHIRPS, TAMSAT (uses only infrared and rain gauge data), RFE 2.0 and PERSIANN CDR score the highest. However, neither satellite only- nor ground only products, such as TMPA RT and CMORPH RAW (both using only infrared and microwave satellite instruments), as well as GPCC (ground-only), are able to produce results of similar quality. The CFSR reanalysis product scored worst of all considered products. While the analyzed products feature different spatial resolutions, we were unable to remark any impact

on the products' performance. Although CFSR and GPCC have the largest pixel sizes and worst results, especially if considering the rain gauge density in West Africa, we conclude that the lack of data and subsequent large-scale interpolation will have a greater influence than the size of the pixel. Overall the results show that especially in the West African region, which is characterized by few reporting rain gauges but also a relatively flat topography, combined satellite-gauge precipitation products prove advantageous over single-source products.

Uncertainties may be introduced into an analysis by the choice of the hydrological model and calibrating that model for discharge. Especially parameters governing infiltration and evaporation may dampen the effects of over- or underestimated precipitation. It is therefore important to select realistic ranges for these parameters before the calibration. These parameters can only influence results in streamflow sums but not in daily distribution. This can explain the fact that CFSR, CMORPH RAW and TMPA RT show low biases in streamflow although the products either over- or underestimate precipitation. When analyzing the NSE however it can be seen that the products deliver weaker results. The model may also be influenced by the potential evapotranspiration inputs, where CFSR estimates were used, but since the same product was used for every model, it should not influence the comparison between the precipitation datasets. The length of the time series also has a certain influence on the performance of the satellite estimates using the hydrological evaluation approach, especially if it is too short for proper calibration and validation of the model. Thiemig et al. (2013) emphasized that longer time periods would obtain results that are more robust, as they only analyzed data over five years. This could be of concern for some of the basins included in this study, e.g. Gbasse, as the time series records are relatively short to sufficiently calibrate and validate the HBV model.

4.4 Conclusion

The evaluation of remotely sensed and reanalysed precipitation datasets is invaluable in undergauged regions. Especially if considering the challenges for Africa that arise due to climate change and intensifying rainfall variabilities, these datasets can provide valuable up to date information beneficial to decision makers. In search of a method to validate precipitation products in the notoriously undergauged West African subcontinent, ten datasets were evaluated, including satellite estimates, reanalysis data and one gauge product. We furthermore placed emphasis on comparing real time, post processed, global and regional products in six basins of different location and size as well as using two sampling strategies. Also, it was evaluated whether results change significantly if the HBV model is initially calibrated or calibrated for each rainfall product separately. The calibration of the model for each product seems the most sensible approach, since the possible introduction of bias during an initial calibration influences the results at the expense of unrealistic evapotranspiration values. Results showed that satellite products which use MW, IR and RG data generally produce good to very good simulations. Best calibrations in terms of NSE were achieved by CMORPH CRT, PERSIANN CDR, TAMSAT, CHIRPS and TMPA, all of which show significant potential as an alternative to RG data. RFE 2.0, GPCC and TMPA RT performed satisfactorily, while CMORPH RAW and CFSR deliver the least robust simulations. Interestingly, CFSR predicted very low precipitation rates in the northern basins and high rates in the south. This was not observed by any other product and

leads to questioning the intrinsic data quality of the set for this region. The results show that the best results can be achieved using bias-corrected satellite products, while satellite only or gauge only products deliver less robust simulations. Many hydrological evaluations of satellite and reanalysis precipitation data have employed the more complex, physically based SWAT model (Dile and Srinivasan 2014; Tobin and Bennett 2014). In this study, the easy to learn conceptual HBV light model has proven to be able to generate robust results. Nevertheless, uncertainties remain due to possible inaccuracies in the discharge data used and uncertainties in the model structure. It can be expected that, with longer and more complete discharge time series, calibration and validation results can be improved.

Acknowledgements: This study is part of the COAST project (Studying changes of sea level and water storage for coastal regions in West-Africa using satellite and terrestrial data sets) of the University of Bonn, supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) under Grant No. DI 443/6-1. We are grateful to Dr. Christophe Peugeot and the AMMA-Catch project as well as to the Global Runoff Data Centre in 56068 Koblenz, Germany for providing discharge data. The AMMA-CATCH regional observing system was set up thanks to an incentive funding of the French Ministry of Research that allowed pooling together various preexisting small scale observing setups. The continuity and long term perenity of the measurements are made possible by an undisrupted IRD funding since 1990 and by a continuous CNRS-INSU funding since 2005. The Precipitation – PERSIANN CDR used in this study was acquired from NOAA's National Centers for Environmental Information http://www.ncdc.noaa.gov. This CDR was originally developed by Soroosh Sorooshian and colleagues for NOAA's CDR Program.

Author Contributions: Bernd Diekkrüger and Thomas Poméon designed the framework of the study. Thomas Poméon and Dominik Jackisch performed data preparation and HBV modeling. Thomas Poméon prepared the original draft and visualizations. Bernd Diekkrüger secured additional discharge data and consulted during the modeling process. All authors contributed to the interpretation of the results and proofreading.

5 Multi-Objective Validation of SWAT for Sparsely-Gauged West African River Basins-A Remote Sensing Approach

This chapter has been published as: Poméon, T., Diekkrüger, B., Springer, A., Kusche, J., Eicker, A., 2018. Multi-Objective Validation of SWAT for Sparsely-Gauged West African River Basins—A Remote Sensing Approach. Water 10, 22. doi:10.3390/w10040451.
http://www.mdpi.com/2073-4441/10/4/451

Abstract: Predicting freshwater resources is a major concern in West Africa, where large parts of the population depend on rain-fed subsistence agriculture. However, a steady decline in the availability of *in situ* measurements of climatic and hydrologic variables makes it difficult to simulate water resource availability with hydrological models. In this study, a modeling framework was set up for sparsely-gauged catchments in West Africa using the Soil and Water Assessment Tool (SWAT), whilst largely relying on remote sensing and reanalysis inputs. The model was calibrated using two different strategies and validated using discharge measurements. New in this study is the use of a multi-objective validation conducted to further investigate the performance of the model, where simulated actual evapotranspiration, soil moisture, and total water storage were evaluated using remote sensing data. Results show that the model performs well (R^2 calibration: 0.52 and 0.51; R^2 validation: 0.63 and 0.61) and the multi-objective validation reveals good agreement between predictions and observations. The study reveals the potential of using remote sensing data in sparsely-gauged catchments, resulting in good performance and providing data for evaluating water balance components that are not usually validated. The modeling framework presented in this study is the basis for future studies, which will address model response to extreme drought and flood events and further examine the coincidence with Gravity Recovery and Climate Experiment (GRACE) total water storage retrievals.

Keywords: SWAT hydrological model; GRACE total water storage; MODIS evapotranspiration; ESA-CCI soil moisture; modeling framework

5.1 Introduction

The availability of freshwater is a major concern in West Africa, directly influencing food security, human health, and economic development (Schuol et al. 2008a). In the region, approximately 60% of the active labor force is employed in agriculture. However, this only contributes 35%

to the gross domestic product (Hollinger and Staatz 2015; Jalloh et al. 2013). Many West African farmers are poor and only able to produce close to subsistence levels, rendering them especially vulnerable to water stress (Jalloh et al. 2013). Therefore, knowledge of the available water resources is essential and modeling the water balance to estimate available resources can be an important tool in this respect. Several meso-scale models have been applied to the area, among others by Andersen et al., who used the physically-based MIKE SHE model to model the Senegal river basin in 2001 (Andersen et al. 2001). In 2005, Bormann studied the results of two process-based (SIMULAT-H and TOPLATS) and one conceptual hydrological model (UHP) applied to the upper Ouémé basin (Bormann 2005). Another approach was employed by Wagner et al. (2009), who downscaled meteorological fields and remote sensing data over the White Volta basin for water balance estimation. Due to the lack of *in situ* data in the region, remote sensing data is starting to be used more often for water resource estimation, e.g., by Fujihara et al., who used reanalysis and global precipitation data in 2014 to model the White Volta basin. The estimation of available water resources has furthermore been a topic of research for the whole of West Africa in several studies, e.g., by Döll et al. (2003), who used the WaterGAP Global Hydrology Model to estimate water resources worldwide or Li et al. (2005), who used the IBIS land surface model in combination with the hydrological HYDRA model for the Niger and Lake Chad basins. However, calibrating global hydrological models against long-term annual discharge may lead to poor temporal performance (Schuol et al. 2008a).

Considering these constraints, Schuol et al. employed the semi-distributed, physically-based hydrological SWAT (Soil & Water Assessment Tool) model to estimate freshwater resources of West Africa in 1998 (Schuol et al. 2008a; Schuol and Abbaspour 2006a). The model was later expanded to cover the whole of Africa (Schuol et al. 2008b) and a calibration/validation approach of this model using GRACE gravity recovery data was performed by Xie et al. (2012), who found SWAT and GRACE agreeing less in arid and humid regions, albeit using historic discharge data for the modern period simulated. In their approach, Schuol et al. covered a four mio. km^2 research area in West Africa, which was partitioned into 292 subbasins. Inaccurate or missing input data have been shown to increase uncertainties of distributed hydrological models (Schuol and Abbaspour 2006b; Arnold et al. 1998). Especially in developing countries, observation networks are sparse, and data regularly includes gaps and gross errors (Behrangi et al. 2011; Bitew and Gebremichael 2011; Koutsouris et al. 2016). A continuous decline in data availability has been observed in recent years due to political unrest, financial issues, and maintenance problems (Adjei et al. 2012; Hughes 2006). Schuol et al. circumvented this data availability problem by training a weather generator to produce daily data using 0.5° Climate Research Unit monthly data (Schuol and Abbaspour 2006b). While some authors state that hydrological models may perform poorly due to the limited data availability in the region (Forootan et al. 2014), Schuol et al. demonstrated the SWAT model to deliver robust results (Schuol et al. 2008a; Schuol and Abbaspour 2006a; Schuol et al. 2008b).

In recent years, many studies have explored the performance of remotely sensed climate data in Africa, with the focus being on the evaluation of precipitation estimates. Investigations have been conducted, among others, for West Africa (Fujihara et al. 2014; Thiemig et al. 2012; Gosset et al. 2013; Poméon et al. 2017), East Africa (Koutsouris et al. 2016; Thiemig et al. 2012; Bitew

et al. 2012; Worqlul et al. 2017; Worqlul et al. 2014), southern Africa (Pombo et al. 2015), and the entire continent (Awange et al. 2015). Most studies agree that certain precipitation products perform well and may be used in substituting for *in situ* gauge data. However, in regions with complex topography and strong altitude variations over short distances, as e.g., in Ethiopia, the performance of satellite precipitation estimates declines sharply (Bitew et al. 2012; Worqlul et al. 2014). We hypothesize that the increasing number of high resolution remote sensing data products will enable substituting data from gauge networks and render weather generators, as used in the study by Schuol and Abbaspour (2006b), obsolete. In this study, we set up a modeling framework by calibrating and validating a SWAT model covering the major West African river basins for the period of 1998–2013, using freely available remote sensing and reanalysis products including new climatological, land use, and soil datasets. The starting time is restricted due to the availability of high resolution satellite data and the end of the modeling period is determined by the lack of discharge data for model validation. While we have chosen to use only freely available data to guarantee the best-possible applicability of the framework, we realize that finding the appropriate discharge data to validate the model is a challenge. To overcome this problem, a multi-objective validation was conducted, where in addition to streamflow, the variables actual evapotranspiration, soil moisture, and total water storage were used to test the model results against available satellite datasets. Therefore, the objectives of this study are (i) to set up a hydrological modeling framework for West Africa using freely available data, (ii) to assess which simulation quality can be obtained using these datasets, (iii) to perform a multi-objective validation, and (iv) to evaluate the potential and limitations of this approach for assessing water availability at the sub-continental scale.

5.2 Materials and Methods

5.2.1 Research Area

The research area is located between 3° and 24° latitude and -18° and 16° longitude and includes, among others, the basins of the Niger, Volta, Senegal and Ouémé rivers stretching over 18 countries, as shown in Figure 5.1. The relief in West Africa is low and flat (CILSS 2016) and rainfall is strongly seasonal, with a unimodal rainy season in the northern part and bimodal rainy seasons in the south (Fink et al. 2010). Rainfall amounts show a distinct south-north gradient. Annual average precipitation in the humid Guinea-Congolian region in the south ranges from 2200 to 5000 mm, while in the arid Sahara region, annual rainfall ranges from 0 to 150 mm (CILSS 2016; Sebastian and Kate 2009; Gessner et al. 2013).

River basins were selected based on the availability of discharge data for calibration purposes. The total area of the basins selected for the model is 3.4 mio. km^2. Due to computational constraints, three different models were built: South (Volta, Ouémé, Comoé, Mono, Pra, Ankobra and Ayensu river basins, 633,000 km^2, 41 stream gauges), West (Senegal and Gambia river basins, 558,600 km^2, 9 stream gauges) and Niger (Niger river basin, 2,250,000 km^2, 12 stream gauges).

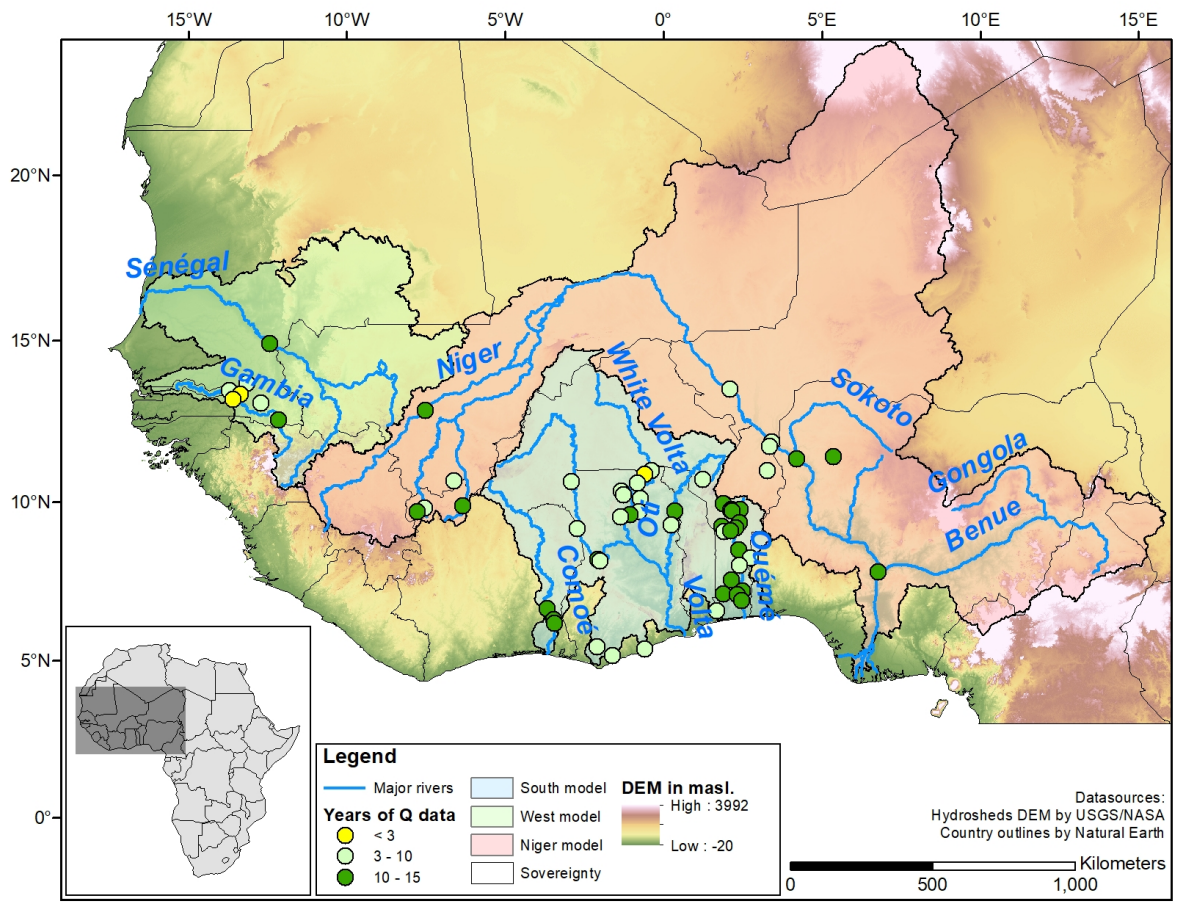

Figure 5.1: Research area, Soil and Water Assessment Tool (SWAT) models and available discharge stations.

5.2.2 The SWAT Hydrological Model

The Soil and Water Assessment Tool (SWAT) represents a continuous-time, semi-distributed, process-based river basin model. SWAT runs at a daily time step but may be calibrated using monthly or yearly observed data (Arnold et al. 2012b; Srinivasan et al. 1998). The model is comprised of eight major components: hydrology, weather, sedimentation, soil temperature and properties, crop growth, nutrients, pesticides and agricultural management. The hydrological component of SWAT is based on the water balance equation (Arnold et al. 1998). SWAT has been proven to be competitive at a number of scales from local to continental, having been employed for the modeling of water resources in Africa and Europe (Schuol et al. 2008b; Abbaspour et al. 2015), among others. In this study, SWAT 2012 was used. The major model inputs and data preparation will be described in detail below.

5.2.3 Input Datasets

- Digital Elevation Model (DEM): The hydrologically conditioned HydroSHEDS (Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales) digital elevation model (DEM) developed by the World Wildlife Fund (WWF) and the United States Geological Survey (USGS) based on the NASA SRTM (Shuttle Radar Topographic

Mission) was used for streamflow delineation. HydroSHEDS is available in 3 and 15 arc-second resolutions (approximately 90 and 500 m) (Lehner et al. 2013; Lehner et al. 2008). In this study, subbasins were generated using the 500 m version.

- Land use and land cover: The Comité permanent Inter-Etats de Lutte contre la Sécheresse dans le Sahel (CILSS) Landscapes of West Africa land use and land cover raster dataset of the year 2013 was used as a basis for developing the land use layer required by SWAT. Maps are also available for the years 1975 and 2000 at a resolution of 2 km. The maps were created using local information and remote sensing data in cooperation with US Aid and USGS (CILSS 2016). Since no data is included for the country of Cameroon, nor north of the 18th parallel in Mauritania and Mali and north of the 15.5th parallel in Niger, missing data was replaced using the European Space Agency (ESA) Globcover 2.3 dataset depicting the land use of the year 2009 in a 300 m resolution (Bontemps et al. 2011). Land use classes were converted to default SWAT classes. It is unclear whether SWAT allows to realistically simulate plant growth under tropical conditions due to its implemented heat unit growth model (Neitsch et al. 2011; Alemayehu et al. 2017; Strauch and Volk 2013). In our study, the management database was adapted by setting fixed plant and harvest dates corresponding to onset and end of rainy season. When compared to MODerate-resolution Imaging Spectroradiometer (MODIS) MOD 15A2 leaf area index (LAI) estimates produced by NASA (Myneni et al. 2015b), SWAT LAI reaches a Pearson's r of 0.62, whereas without management modifications this value drops to -0.47.

- Soil: The Harmonized World Soil Database (HWSD) version 1.2 produced by the Food and Agriculture Organization of the United Nations, the International Institute for Applied Systems Analysis, ISRIC World Soil Information, the Institute of Soil Science-Chinese Academy of Sciences and the European Commission's Joint Research Centre (JRC) in 2012 was used to generate the soil data needed in SWAT. The HWSD supplies a raster map and database containing several soil physical and chemical parameters for a top- and subsoil layer (FAO/IIASA/ISRIC/ISS-CAS/JRC 2012). Missing parameters were estimated from soil texture using pedotransfer functions (Wösten et al. 2001). The HWSD and its predecessors have been used for SWAT simulations in Africa, the Middle East, and Europe, among others (Schuol et al. 2008a; Schuol et al. 2008b; Xie et al. 2012; Abbaspour et al. 2015; Faramarzi et al. 2013; Faramarzi et al. 2010; Malagò et al. 2016).

- Climate: In a previous study, ten precipitation datasets were analyzed for six subbasins in the study area (Poméon et al. 2017). It was concluded that the Climate Prediction Center Morphing Technique (CMORPH) version 1 CRT produced by the National Oceanic and Atmospheric Administration Climate Prediction Centre (NOAA-CPC) performed best. CMORPHv1 CRT is a global precipitation analysis algorithm, including satellite infrared and microwave precipitation estimates as well as rain gauge information for bias correction. Precipitation estimates are available from 1998 onwards at a resolution of 0.25° (Joyce et al. 2004; Xie et al. 2011). Minimum and maximum 2 m daily temperature data were compiled from the NASA MERRA 2 reanalysis dataset. Inputs from both satellite and ground data are included at a resolution of 0.625° $*$ 0.5° (Bosilovich et al. 2016).

While SWAT-ready climate input files based on the National Centers for Environmental Prediction (NCEP) climate forecast system reanalysis data (CFSR) (Saha et al. 2010) are readily available, as discovered in Poméon et al. (2017), CFSR precipitation information compares worse to other products in the region. No other climate data were necessary as the authors selected Hargreaves as the potential evapotranspiration method.

- Discharge and reservoirs: Discharge data used in this study was obtained from the German Global Runoff Data Center (GRDC) in Koblenz, the French AMMA-CATCH regional observing system, as well as through personal communication with local agencies. Discharge stations and their temporal coverage (without gaps) are depicted in Figure 5.1 and summarized in Table 5.1. The 12 largest reservoirs in the study area where downstream discharge observations are available were included in the model. Reservoir information was provided by the Global Water System Project (GWSP) Global Reservoir and Dam (GRanD) database version 1.1 created by Lehner et al. (2011). Missing storage volumes information was approximated as proposed by Schuol et al. (2008a). Lake Volta was not modeled due to insufficient data being available.

Table 5.1: Selected river basins and discharge gauges in the study area.

River Basin	Area in km^2	Gauges	SWAT Model
Niger	2,246,220	12	Niger
Senegal	480,289	1	West
Volta	425,133	16	South
Comoé	84,533	3	South
Gambia	78,321	8	West
Ouémé	61,057	17	South
Mono	24,310	1	South
Pra	23,345	1	South
Ankobra	8773	1	South
Kouffo	4122	1	South
Ayensu	1753	1	South
TOTAL	3,437,856	62	All

5.2.4 Multi-Objective Validation Datasets

We decided to validate model simulations using actual evapotranspiration, soil moisture, and total water storage, in order to evaluate the model performance of processes not reflected in streamflow. This section gives an overview of the remote sensing data used in the multi-objective validation.

- Actual evapotranspiration (ETA): Data was extracted from the MODIS MOD 16 dataset supplied by NASA, available at a 1 km^2 spatial- and 8-day or monthly temporal resolution. ETA is calculated based on the Penman-Monteith equation using ground-based and remote

sensing datasets. The algorithm includes vapor pressure deficit, leaf area index, enhanced vegetation index and soil evaporation (Mu et al. 2011; Mu et al. 2007).

- Soil moisture: ESA Climate Change Initiative (CCI) 3.2 soil moisture (SM) retrievals were used to validate the soil moisture dynamics simulated by SWAT. The product is generated by blending passive and active microwave soil moisture retrievals generated by C-band scatterometers and multi-frequency radiometers on multiple spacecraft. Daily data is available at a resolution of 0.25 degrees but covering only the upper few cm of the soil (Liu et al. 2012; Liu et al. 2011; Wagner et al. 2012).

- Total water storage (TWS): Gravity Recovery And Climate Experiment (GRACE) TWS retrievals were used for further model validation. The twin satellite GRACE mission has been measuring temporal and spatial variations in the Earth's gravity field since 2002. GRACE consists of two identical satellites on the same near-circular orbit. The dual one-way K-band microwave ranging system observes the distance between the two satellites. Changes in the distance in conjunction with complementary tracking data are used to derive monthly gravity fields, which, subsequently, are converted to mass changes in terms of equivalent water height according to Wahr et al. (1998). In this study, we used the ITSG-Grace2016 time series provided by the Institute of Geodesy (IfG) at Technical University (TU) Graz as sets of spherical harmonic coefficients up to degree and order 90. As GRACE does not measure geocenter variations, degree 1 coefficients were replaced by the time series provided by Rietbroek et al. (2012a) and Rietbroek et al. (2012b). The c_{20} coefficient, which is corrupted by aliasing effects, was replaced by results from satellite laser ranging (Cheng et al. 2013). GRACE observes the integral sum of all mass variations in hydrosphere, atmosphere, biosphere, oceans and mass variations inside of the earth. Gravity field solutions from ITSG-Grace2016 are already corrected for tides (ocean, earth and pole tides) and non-tidal atmospheric and oceanic effects. Trends from glacial isostatic adjustment are about zero in the study region. Therefore, the spherical harmonic coefficients from ITSG-Grace2016 primarily reflect variations in the terrestrial water storage. As GRACE-derived gravity solutions are contaminated with correlated noise leading to the characteristic striping patterns in the north-south direction, the monthly fields were smoothed using the anisotropic DDK3 filter (Kusche 2007). Filtering implies attenuation of the signal and further distortion, known as leakage effect. Therefore, TWS time series derived for the three target areas *via* spatial averaging were rescaled using the scaling factor approach (Long et al. 2015). Here, scaling factors were derived from five global hydrological models for each target area separately (Longuevergne et al. 2010). All computations are accompanied by a thorough error propagation, which starts from the full error covariance matrices of the spherical harmonic coefficients and results into errors for the rescaled TWS time series. Since Lake Volta was not modeled in SWAT, the lake signal was computed using lake height variations and an area varying between 4450 km^2 and 9970 km^2 (Tanaka et al. 2002; Uebbing et al. 2015) and subsequently subtracted from the GRACE estimates.

5.2.5 Model Setup and Calibration/Validation

The model parametrization was conducted using the ArcSWAT 2012 interface (Winchell et al. 2010). The research areas were divided into subbasins based on the DEM and derived stream network. We used a streamflow delineation threshold of at least 500 km^2 for the southern and western models (1500 km^2 for the Niger model) and manually added outlets where data from gauging stations was available, generating 2153 subbasins (South: 712; West: 630; Niger: 811). Next, the subbasins were overlaid with land use and soil maps to derive Hydrological Response Units (HRUs), units with the same land use, soil and slope characteristics (Neitsch et al. 2011). In view of computational efficiency, we opted to derive one HRU per subbasin by considering dominant land use, soil and slope (Schuol et al. 2008b) (divided into 0-1; > 1-5 and > 5% slope).

The dominant land use distribution for each model is displayed in Figure 5.2. In the South model, range-brush is the dominant land use type (53.6%), followed by agriculture (31.4%) and forest (7.8%). Both forest and agricultural areas are mostly located in the more humid south, while rangeland dominates in the arid north. In the western model, rangeland (brush and grasses) dominates with 40.4 and 47.8%, respectively. 8.6% of the area is barren and 2.9% under agricultural use. Land use in the Niger model is to almost equal parts range grasses, barren, agriculture and range brush (26.4, 24.6, 23.3 and 22.2%). The high prevalence of barren areas can be explained by the hydrologically inactive part of the basin, located in the north-east (Itiveh and Bigg 2008). Only 2.6% of the area is predominantly forested.

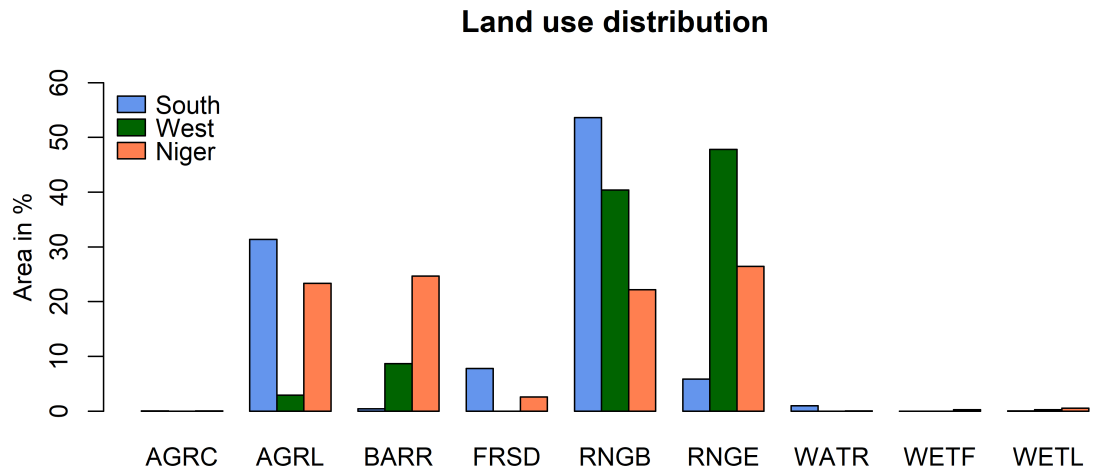

Figure 5.2: Final SWAT land use distribution for each model, where AGRC: close grown agriculture; AGRL: agriculture; BARR: barren; FRSD: forest deciduous; RNGB: range brush; RNGE: range grasses; WATR: water; WETF: wetlands forested; WETL: wetlands.

After generating the HRUs, reservoirs were included as described in Section 5.2.3. Due to uncertainties in the climate data, the potential evapotranspiration was calculated using the 1985 Hargreaves equation, which requires only temperature and extraterrestrial radiation inputs (Neitsch et al. 2011; Hargreaves and Samani 1985). In SWAT, extraterrestrial radiation is

calculated as a function of location and time of year (Neitsch et al. 2011). The Hargreaves method has been suggested if the input data quality is in doubt (Droogers and Allen 2002). Since the evapotranspiration processes in the study region are water-limited, more emphasis should be placed on the actual evapotranspiration, as it directly influences runoff generation (Weiß and Menzel 2008). SWAT calculates the Hargreaves (1985) equation as follows (Equation 5.1):

$$\lambda ETP = 0.0023 * H_0 * (T_{mx} - T_{mn})^{0.5} * (\bar{T}_{av} + 17.8) \tag{5.1}$$

where λ is the latent heat of vaporization in MJ/kg, ETP is the potential evapotranspiration in mm, H_0 is the extraterrestrial radiation in MJ/m², T_{mx} is the maximum air temperature in °C, T_{mn} is the minimum air temperature in °C, and $\bar{T}_{av}$ is the average air temperature in °C (Neitsch et al. 2011).

The simulation covered the period of 1998 to 2013 with a warm-up period from 1996 to 1997. Since CMORPH precipitation data was only available from 1998 onwards, data from 1998 to 1999 was used as fictitious data for the warm-up period in order to maximize the simulation period.

The calibration of a semi-distributed watershed model such as SWAT is challenging due to input uncertainties, model uncertainties and parameter non-uniqueness (Schuol and Abbaspour 2006a; Abbaspour et al. 2007). For the calibration of our models, the Sequential Uncertainty Fitting version 2 (SUFI-2) procedure of SWAT-CUP (Calibration and Uncertainty Programs, developed by Karim Abbaspour of the Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Dübendorf, Switzerland) (Abbaspour et al. 2007) was used. In SUFI-2, all uncertainty (parameter-, model-, and input-uncertainty) is accounted for by the respective parameter uncertainty. Uncertainties are quantified by the p-factor, which measures the percentage of the observed data falling into the 95% prediction uncertainty (95PPU) band. A further parameter, the r-factor, describes the range of the 95PPU. Ideally, one wants the p-factor to be as large as possible and the r-factor to be as small as possible (Abbaspour et al. 2007).

The model was calibrated using discharge data from 62 gauging stations. Available daily data was aggregated to monthly data by interpolation whenever seven days in a row were missing and deleting the month for longer gaps. Due to large data gaps and different lengths of the discharge time series which did not allow for fixed calibration and validation periods, the first two thirds of the discharge data were used for calibration and the last third for validation (Abbaspour et al. 2015).

Two different calibration approaches were used. In the first approach (v1), the model parameters were globally calibrated, while in the second approach (v2), upstream subbasins were individually calibrated apart from downstream subbasins in order not to influence results if discharge gauges are unevenly distributed (Schuol et al. 2008a; Schuol et al. 2008b; Abbaspour et al. 2015).

A wide variety of potential parameters and ranges for calibration were identified using available literature and the SWAT manuals (Arnold et al. 2012a; Neitsch et al. 2002). In a second step, the effects of the parameter ranges on model results were identified through a one at a time sensitivity analysis coupled with a custom R script graphically representing the reaction of

SWAT storages and flows. This way, realistic parameter ranges were defined for the research area. SWAT-CUP allows for certain parameters to be calibrated separately by soil texture or land use types. This again increases the number of parameters. In our approach, we included all potential parameters in an initial iteration with 500 (v1) and 1000 (v2) model runs and used the SWAT-CUP sensitivity analysis tool to assess the global sensitivity of each parameter (Abbaspour 2015). SWAT-CUP determines the parameter sensitivity by multiple regression of the Latin Hypercube generated parameter values against the objective function and performing a t-test. Parameters with a p-value of < 0.05 are assumed to be sensitive (Abbaspour 2015). To reach an acceptable calibration, three iterations with 500 model runs each were performed with the sensitive parameters. Parameter ranges are updated automatically after each iteration. If an acceptable calibration is reached, the validation is performed using the same parameter ranges and number of simulations. An overview of the included parameters is given in Table 5.2.

Table 5.2: Parameters included in SWAT model and initial ranges.

SWAT Parameter	Differs By	min	max
CN2*	Land use	-0.5	0.1
SOL AWC*	Soil Texture	-0.1	0.5
SOL K*	Soil Texture	-0.5	0.5
SOL BD*	Soil Texture	-0.5	0.1
EPCO*		-0.3	0.3
ESCO*	Land use	-0.3	0.3
GW DELAY		0	100
GWQMN		0	1000
RCHRG DP		0	1
GW REVAP		0.02	0.2
REVAPMN		0	500
SURLAG		0	10

CN2: runoff curve number; SOL AWC: available water capacity (mm H_2O/mm soil); SOL K: saturated hydraulic conductivity (mm/h); SOL BD: moist bulk density (g/cm^3); EPCO: plant uptake compensation factor; ESCO: soil evaporation compensation factor; GW DELAY: groundwater delay time (days); GWQMN: threshold depth for return flow to occur (mm H_2O); RCHRG DP: deep aquifer percolation fraction; GW REVAP: groundwater "revap" coefficient; REVAPMN: threshold depth for "revap" or percolation to occur (mm H_2O); SURLAG: surface runoff lag coefficient; *: relative change.

The Kling-Gupta Efficiency (KGE) was chosen as the objective function, as it can be decomposed into correlation, bias and relative variability between simulated and observed variables (Gupta et al. 2009). SWAT-CUP implements the 2009 equation (Gupta et al. 2009; Abbaspour 2015). KGE can take values from -∞ to 1 and is calculated as follows (Equation 5.2) (Gupta et al. 2009):

$$KGE = 1 - \sqrt{(r-1)^2 + \left[\left(\frac{\sigma_s}{\sigma_m}\right) - 1\right]^2 + \left[\left(\frac{\mu_s}{\mu_m}\right) - 1\right]^2} \tag{5.2}$$

where KGE is the Kling-Gupta Efficiency, r is the regression coefficient between simulated and

measured variables, σ is the standard deviation, μ is the mean value and s and m are simulated and measured values, respectively. In this study, we consider KGE values of ≥ 0.5 to be good and values ≥ 0.7 to be very good.

A further efficiency criterion used in this study is the Nash-Sutcliffe Efficiency (Equation 5.3) (Nash and Sutcliffe 1970; Moriasi et al. 2007):

$$NSE = 1 - \left[\frac{\sum_{i=1}^{n} (Y_i^{obs}) - (Y_i^{sim})^2}{\sum_{i=1}^{n} (Y_i^{obs}) - (Y^{mean})^2} \right] \tag{5.3}$$

where Y_i^{obs} is the i^{th} observation of the variable to be evaluated, Y_i^{sim} is the i^{th} simulation of the variable to be evaluated, Y^{mean} is the mean of the observed variables and n is the number of observations. Similar to KGE, NSE can range from -∞ to 1, where values ≥ 0.5 are acceptable and values ≥ 0.7 very good (Moriasi et al. 2007).

Finally, percent of model bias or PBIAS is calculated as follows (Equation 5.4) (Moriasi et al. 2007):

$$PBIAS = \frac{\sum_{i=1}^{n} (Y_i^{obs} - Y_i^{sim}) * 100}{\sum_{i=1}^{n} (Y_i^{obs})} \tag{5.4}$$

where Y_i^{obs} is the i^{th} observation of the variable to be evaluated and Y_i^{sim} is the i^{th} simulation of the variable to be evaluated. Positive values represent an underestimation and negative values an overestimation by the model.

5.2.6 Multi-Objective Validation

Calibration and validation of hydrological models is often done using observed discharge alone, whereby aspects of the water balance are being neglected (Qiao et al. 2013). In this study, we perform an additional validation of the model results by comparing ETA, SM and TWS to remote sensing data. ETA was evaluated using the MODIS MOD16 satellite product (Mu et al. 2007; Mu et al. 2011). We chose ETA to evaluate the model performance under uncertain precipitation and land use inputs, as well as to validate the Hargreaves evapotranspiration calculations. The modeled soil moisture was validated against the ESA CCI SM product (Liu et al. 2011). We chose to validate the soil moisture, as its inter-annual variability is very high in West Africa and it is an important factor for crop production. The CCI product was used, as it optimally fits our period of interest. The evaluation of the soil moisture performance of SWAT proved problematic, as outputs produced by the model provide soil moisture in mm for the whole profile or soil layers, while the CCI SM is given in percent over the upper few centimeters of the soil profile. Furthermore, SWAT calculates plant-available soil moisture rather than absolute soil moisture as given for the observation (DeLiberty and Legates 2003; Milzow et al. 2011; Rajib et al. 2016). Therefore, we decided to focus on comparing the dynamics of simulations and observations instead of absolute values. Finally, the calculated total water storage was validated using GRACE TWS data. The SWAT total water storage change was estimated from the water storages by calculating the deviation from the mean water storage during the period of GRACE

data availability (2003-2013) according to the following formula (Equation 5.5):

$$\Delta TWS^t = (SW^t + SA^t + DA^t) - \overline{SW^t + SA^t + DA^t} \tag{5.5}$$

where ΔTWS^t is the total water storage change at time step t, SW^t is the soil water storage, SA^t is the shallow aquifer storage and DA^t is the deep aquifer storage. All units are in mm.

5.3 Results

5.3.1 Calibration and Validation Results

Results for the three models and two calibration approaches are listed in Table 5.3 and will be described in detail.

Table 5.3: Calibration and validation results for v1 and v2 models.

Model	Objective Function						% of Discharge Stations		
	p	r	R^2	PBIAS	KGE	KGE $\geq$ 0	KGE $\geq$ 0	KGE $\geq$ 0.5	KGE $\geq$ 0.7
					Calibration				
South v1	0.37	0.36	0.53	5.54	0.23	0.48	85	54	20
South v2	0.31	0.71	0.51	-29.71	0.15	0.47	66	39	5
West v1	0.73	1.42	0.57	7.01	0.40	0.54	90	50	20
West v2	0.33	0.79	0.61	-53.50	-0.13	0.38	78	33	0
Niger v1	0.30	0.65	0.46	30.53	0.14	0.38	58	17	0
Niger v2	0.29	0.62	0.41	11.98	0.08	0.35	75	25	25
Average v1	0.47	0.81	0.52	14.36	0.26	0.47	78	40	13
Average v2	0.31	0.71	0.51	-23.74	0.03	0.40	73	32	10
					Validation				
South v1	0.36	0.44	0.61	-1.39	0.03	0.48	78	37	17
South v2	0.30	0.80	0.60	-62.67	-0.21	0.54	73	49	24
West v1	0.72	12.22	0.74	20.43	0.17	0.47	67	44	11
West v2	0.30	0.97	0.71	-825.36	-8.72	0.32	67	0	0
Niger v1	0.30	0.57	0.52	33.73	0.03	0.30	67	17	8
Niger v2	0.36	0.59	0.53	30.18	0.10	0.48	50	25	8
Average v1	0.46	4.41	0.63	17.59	0.07	0.42	70	33	12
Average v2	0.32	0.78	0.61	-285.95	-2.94	0.45	63	25	11

p: fraction of data bracketed by the 95PPU; r: 95PPU range (dimensionless); R^2: coefficient of determination; PBIAS: percent model bias; KGE: Kling-Gupta Efficiency, KGE $\geq$ 0: mean KGE of stations scoring higher than 0.

For the v1 calibration (Figure 5.3), 78% of gauging stations reach a KGE of higher than zero, meaning that the model performs better than if using observed mean values as predictors (Gupta et al. 2009). On average, 40% of the gauging stations reach a KGE of more than 0.5,

while 13% are above 0.7, with the highest average KGE of 0.40 in the West (the overall best result) and the lowest average KGE of 0.14 found in the Niger region. The average bias is 14.36% and R^2 amounts to 0.52, with the highest R^2 of 0.57 reached in the West model and the lowest value of 0.46 in the Niger model. While the range of the model uncertainty (r-Factor) is 0.82, the percentage of data bracketed by the 95 PPU (p-Factor) is 47%. Calibrations of the southern model perform best in the Ouémé and White Volta basins and worst in the Black Volta basin. For the western model, best performances can be observed for the downstream Gambia tributary rivers, while some upstream stations perform less well. For the Niger model, the best performance is reached downstream of the confluence of the Benue and the Niger in Lokoja, while it performs worst in most of the most upstream subbasins.

Concerning validation, 33% of v1 stations reach a KGE above 0.5 and 12% above 0.7. However, the average KGE is 0.07 due to some poorly-performing stations. If we removed these stations, KGE would increase to 0.42. R^2 is the only factor performing better in the validation than the calibration (0.63 as opposed to 0.52). While the p-factor is similar to the calibration, the r-factor is influenced by the large uncertainty band of the West model and reaches 4.41. PBIAS has also increased to 17.59%. Best performances are reached in the West model.

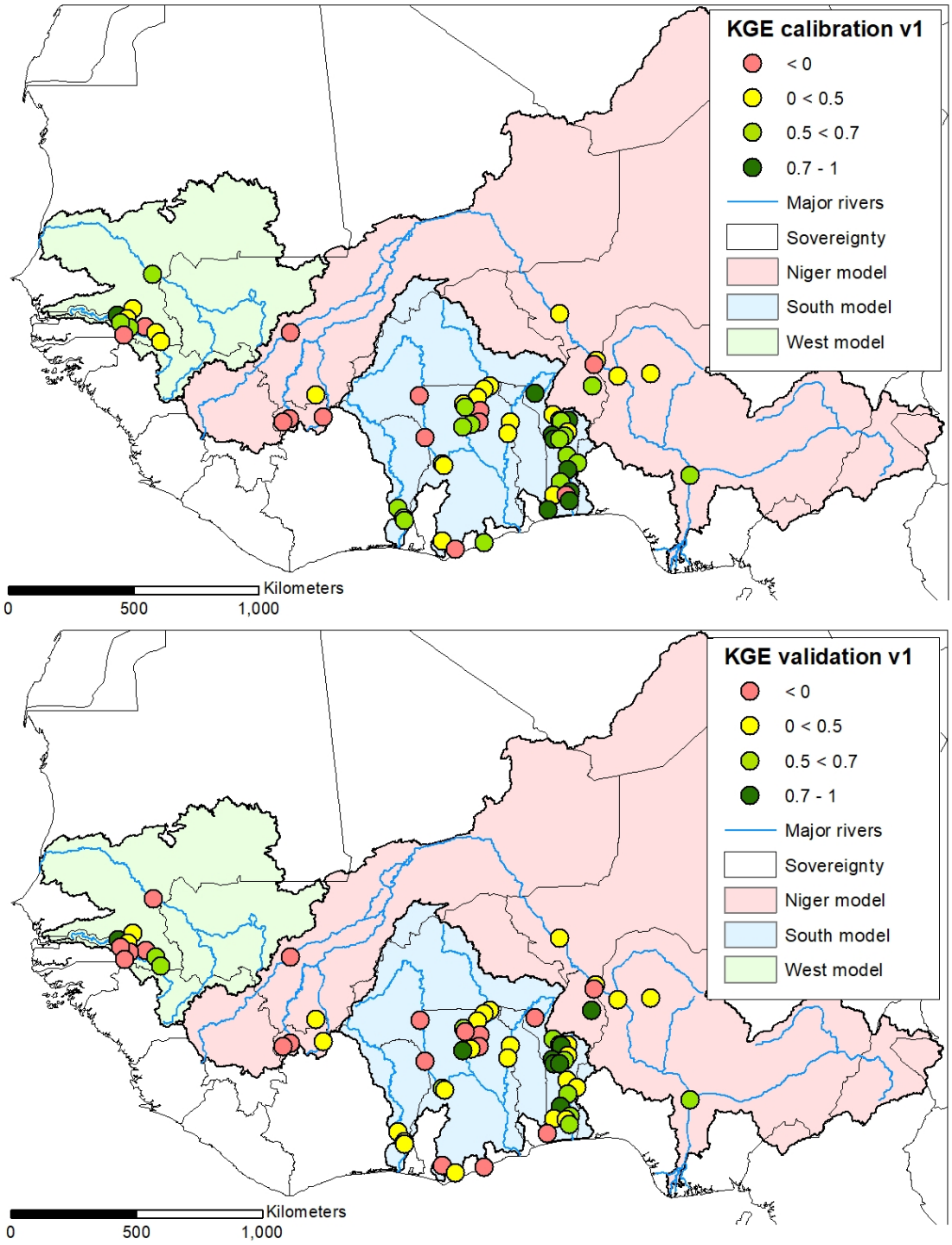

Figure 5.3: Calibration and validation results of the v1 (global calibration) models.

While for the v2 approach (Figure 5.4), about the same amount of stations score a KGE of higher than 0.5 (73%), it generally delivers less robust solutions, with only 32% of discharge stations reaching a KGE of 0.5 or higher and 10% reaching above 0.7, as opposed to 13% in v1.

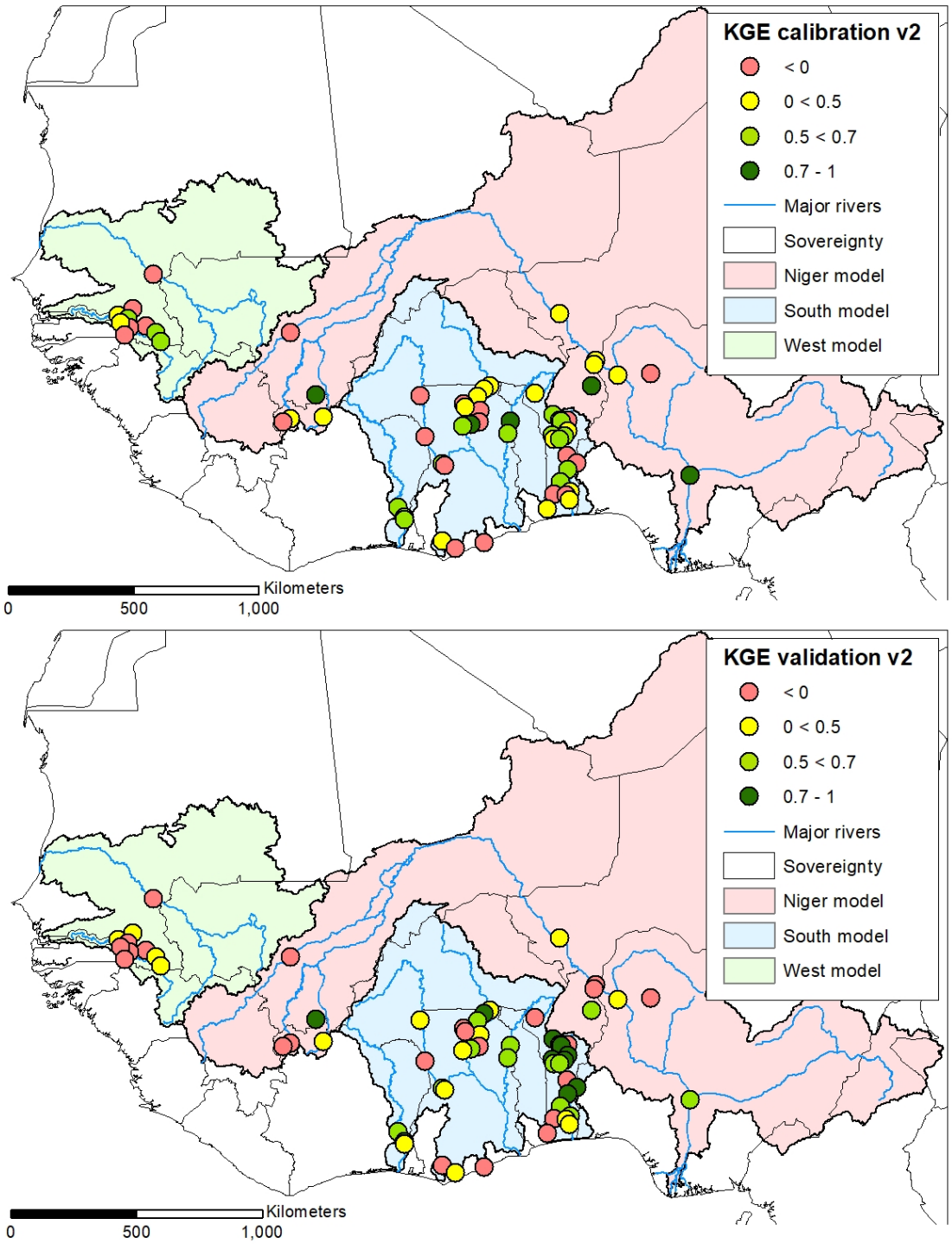

Figure 5.4: Calibration and validation results of the v2 (local calibration) models.

The average KGE value of simulation v2 is worse in the South (0.15) and Niger (0.08) models and decidedly worse in the West (-0.13). While p and r perform slightly worse in the v2 approach, R^2 remains almost constant at 0.51. PBIAS is worse in v2, dropping from 14.36% underestimation to -23.74% overestimation of streamflow. The Comoé, as well as certain up-

stream Niger basins, perform better. The v2 validation performs worse than the v1 validation, with 25% of stations reaching KGE values above 0.5 and 1% above 0.7. The average KGE is low with -2.94, due to bad performance in the West model. When only taking stations performing above zero into account, KGE is higher with 0.45 than in the v1 validations (0.42). The highest KGE is reached in the Niger (0.10). While the r value for v2 is decidedly better (0.78 as opposed to 4.41), p is slightly worse (0.32 vs. 0.46). v2 strongly overestimates streamflow (PBIAS: -285.95), again mainly due to the performance of the West model. R^2 performs similar to v1. Best validation results are reached in the White Volta, Oti and Ouémé.

An example of monthly calibration and validation results for four selected discharge stations is given in Figure 5.5.

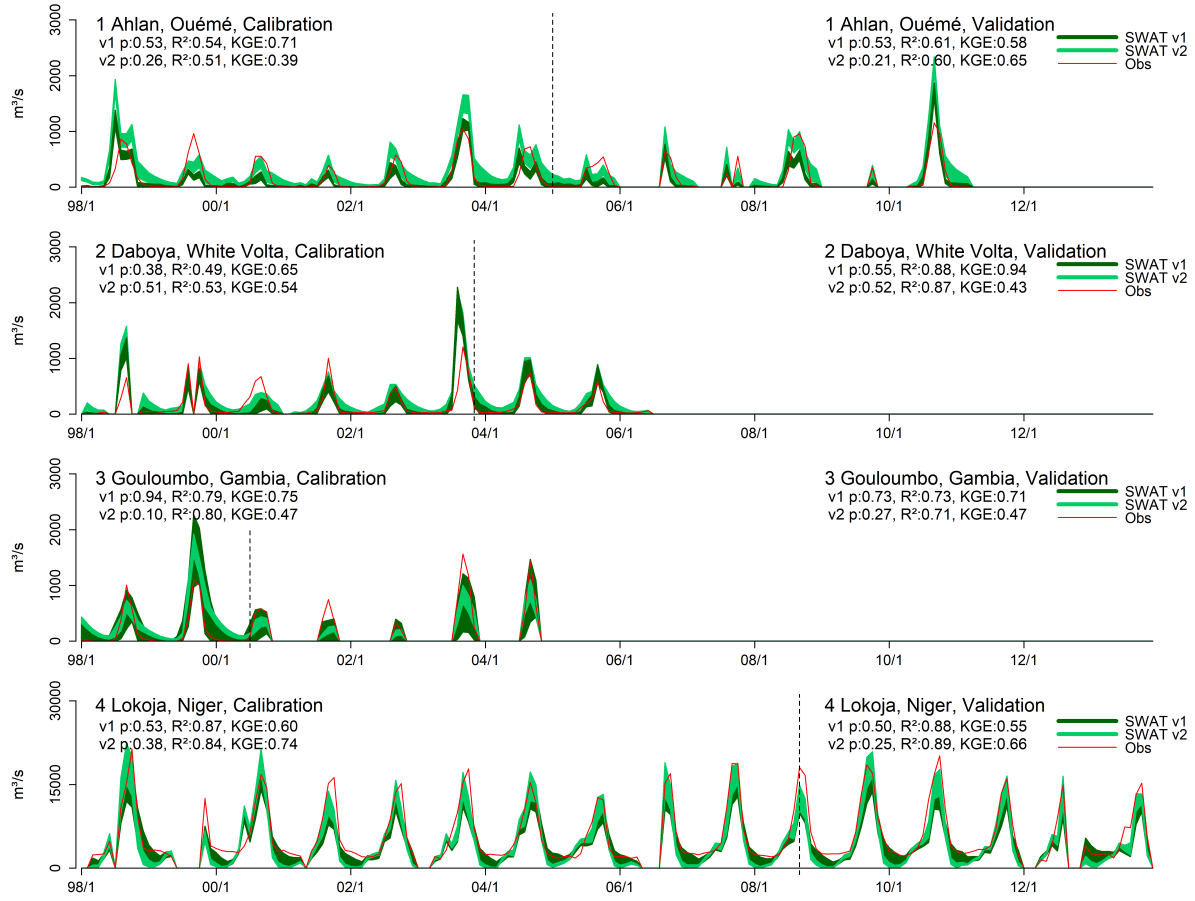

Figure 5.5: Example discharge results for the South (1&2), West (3) and Niger (4) models.

Displayed are the 95PPU ranges of both v1 and v2 calibration and validation, the observed data, as well as the key efficiency criteria p, R^2 and KGE. The stations are located in the Ouémé (1 Ahlan), White Volta (2 Daboya), Gambia (3 Gouloumbo) and Niger (4 Lokoja) river basins. For v1, KGE values for all stations are between 0.65 and 0.75. On average, validations perform less well than calibrations except in Daboya (Validation: 0.94). While performances increase during v2 in Lokoja, decreases can be observed for the other stations. Only during validation do Ahlan and Lokoja perform better than v1. At this point, we conclude that for the calibration and validation of the model with discharge alone, the global calibration (v1) performed slightly better than the local calibration (v2).

5.3.2 Multi-Objective Validation Results

During the multi-objective validation, several output variables which were not used for calibration were kept for further validation by comparing to MODIS ETA, ESA CCI SM and GRACE data. Concerning actual evapotranspiration, validation results reach good scores, as shown in Table 5.4 and Figure 5.6.

Table 5.4: Actual evapotranspiration validation against MODIS MOD 16 Data.

Model	R^2	sig.	PBIAS	NSE	KGE
South v1	0.93	< 0.001	5.0	0.81	0.71
South v2	0.92	< 0.001	6.6	0.73	0.63
West v1	0.92	< 0.001	-15.8	0.88	0.80
West v2	0.91	< 0.001	-15.1	0.87	0.82
Niger v1	0.94	< 0.001	2.8	0.81	0.67
Niger v2	0.94	< 0.001	4.4	0.82	0.70
Average v1	0.93		2.67	0.83	0.73
Average v2	0.92		1.37	0.81	0.72

R^2: coefficient of determination; sig: significance level; PBIAS: percent model bias; NSE: Nash-Sutcliffe Efficiency; KGE: Kling-Gupta Efficiency.

Both the v1 and v2 calibrations of all models fit the observed data very well, with all criteria being almost identical. Best R^2 performances are observed in the Niger model with 0.94. No model R^2 performs below 0.91. In both the West and the Niger models, SWAT overestimates ETA during the wet seasons. In the South and Niger models, underestimations during the dry season also occur. On average, the South and Niger model underestimate ETA (PBIAS: -2.8 - -6.6%), while the West model overestimates by around 15%. In the West model, no underestimations can be observed due to observed ETA being very low during the dry season.

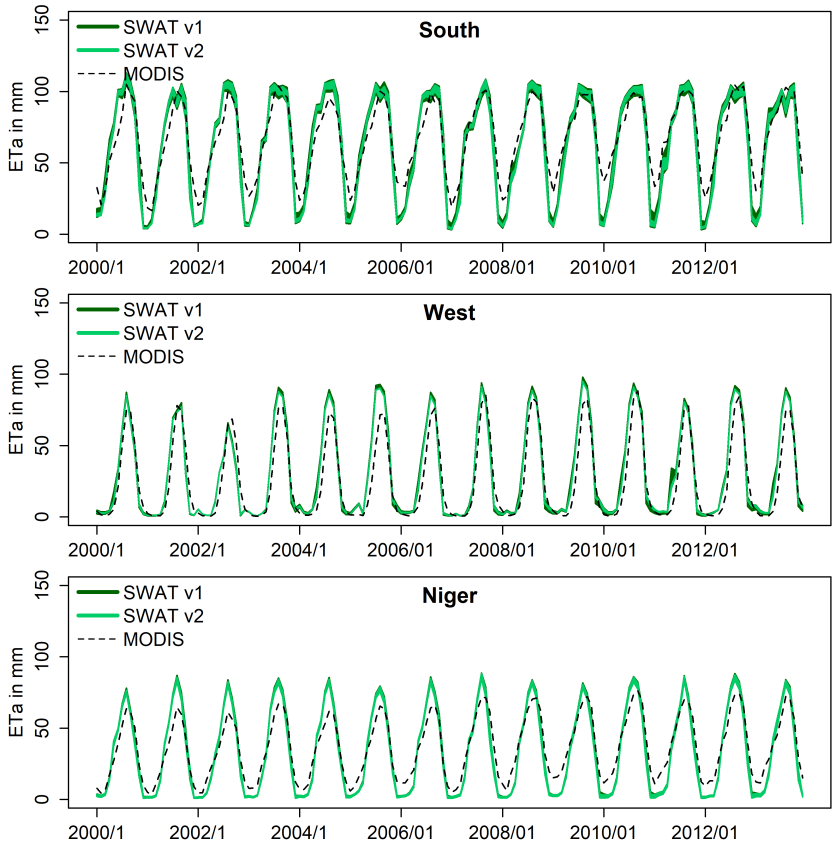

Figure 5.6: Monthly simulated actual evapotranspiration validation against MODIS MOD 16 data, where SWAT v1 is the global and v2 the local calibration.

Performance metrics for the soil moisture validation against ESA CCI data are given in Table 5.5 and graphical representations in Figure 5.7. Overall, the dynamics fit very well. Years with lower soil moisture content such as 2002-2003 in the West, or 2001-2003 in the South model regions are also well represented in the simulations. R^2 values are above 0.69 for all models with best performances reached in South v2 and Niger v1/v2 (0.82 and 0.80). West v1 performed least well with an R^2 of 0.69. v1 and v2 again perform very similarly (v1: 0.75, v2: 0.78).

Table 5.5: Soil moisture validation against ESA CCI data.

Model	R^2	sig.
South v1	0.77	< 0.001
South v2	0.82	< 0.001
West v1	0.69	< 0.001
West v2	0.73	< 0.001
Niger v1	0.80	< 0.001
Niger v2	0.80	< 0.001
Average v1	0.75	
Average v2	0.78	

R^2: coefficient of determination; sig: significance level.

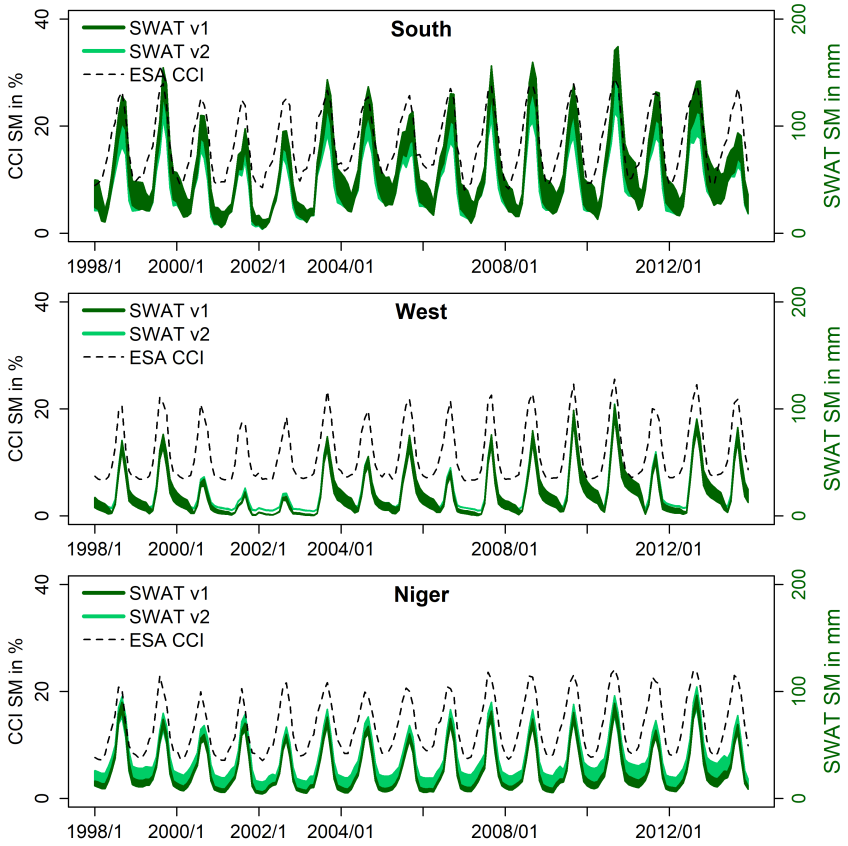

Figure 5.7: Monthly simulated soil moisture validation against ESA CCI data, where SWAT v1 is the global and v2 the local calibration.

Finally, total water storage was calculated from SWAT outputs and validated using GRACE data (see Table 5.6 and Figure 5.8). Again, results show a good fit. Nonetheless, an overestimation of TWS during the dry seasons is apparent in all models, as well as a slight underestimation during the wet seasons in the Niger and South models. Also apparent is a phase shift in the model results by approximately half a month. Some fast changes, e.g., the sharp drop and rise in TWS during the wet season 2012, are not visible in the simulation results at all. Performances vary, and a very high uncertainty in the West v2 model is immediately apparent. Otherwise, the dynamics of both calibrations perform similarly with best R^2 and NSE results reached in the globally calibrated models (0.82 and 0.79 as opposed to 0.61 and 0.56 in the locally calibrated models, respectively). All models except West v2 reach between acceptable and very good R^2 and NSE values with the West v1 model performing best and the West v2 model performing worst.

Table 5.6: Total water storage validation against GRACE data.

Model	R^2	sig.	NSE
South v1	0.75	< 0.001	0.75
South v2	0.70	< 0.001	0.68
West v1	0.90	< 0.001	0.87
West v2	0.35	< 0.001	0.27
Niger v1	0.82	< 0.001	0.76
Niger v2	0.79	< 0.001	0.72
Average v1	0.82		0.79
Average v2	0.61		0.56

R^2: coefficient of determination; sig: significance level; NSE: Nash-Sutcliffe Efficiency.

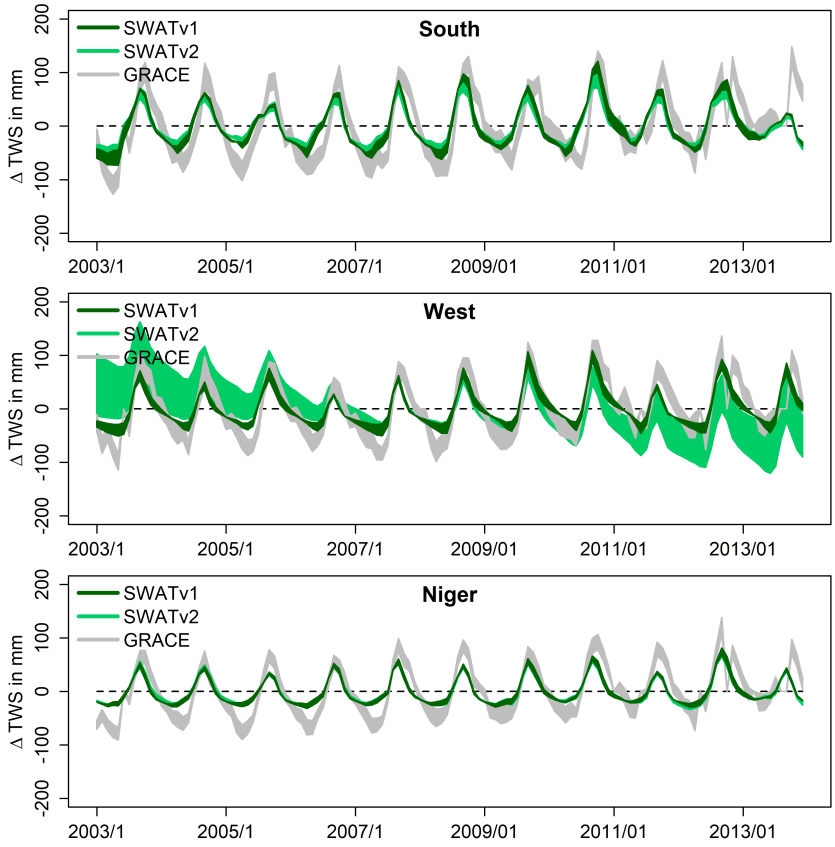

Figure 5.8: Monthly simulated total water storage validation against GRACE data, where SWAT v1 is the global and v2 the local calibration.

5.4 Discussion

5.4.1 Model Calibration/Validation Discussion

Results show that satellite and remote sensing data can be used to substitute missing observations and boundary conditions in a SWAT simulation. Results are promising with especially

successful calibrations and validations generated for the Ouémé, Gambia and lower Niger basins. It may be argued that during global calibration (v1), a prevalence of stations in a certain region may unduly influence the model if the weights of the stations remain the same (Abbaspour et al. 2015). In our case, this can be observed in the Niger basin, where two of the most upstream stations perform poorly due to the calibration being influenced by the downstream gauging stations but performing better when separately calibrated in v2. However, this effect does not explain the poor performance along the Black Volta river, as similarly poor results are observed in the v2 simulation. This was also reported by Schuol et al. (2008a). Some of the discharge stations are highly influenced by upstream reservoirs for which no outflow data is available. Even when including reservoirs in the SWAT model, we noticed downstream stations often performed poorly due to the limited amount of data available for proper reservoir setup. Also problematic is the decline in the availability of discharge measurements and uncertainty as to their quality, coupled with data gaps. In contrast to Schuol et al. (2008a), we did not include the Inner Niger Delta in the model. While they set up the delta as an artificial reservoir and defined the outflow as according to a close downstream station, the closest station for our timeframe is located almost 500 km downstream.

Also, the Akosombo dam in southern Ghana, which creates Lake Volta, could not be included due to missing information about in- and outflows. While the lake was removed from GRACE-derived water storage change (by deriving mass variations using altimeter measurements and information on the lake area) to correspond to the simulations, the missing lake might lead to lower actual evapotranspiration simulations in this area. If comparing the amount of discharge data available for the period modeled by Schuol et al. (calibration from 1970 to 1995) and this study (1998–2013), the decline in available discharge measurements becomes apparent, with the exception of the Ouémé basin, where we were able to secure additional stations. Interestingly, the distribution of well-performing stations is very similar in results from both studies, except for the upstream Niger stations, which performed less well in our approach. We observed v1 performing stronger in the calibration and validation periods. We attribute this to the global sensitivity evaluation used in this study. While the 500 runs used to evaluate the sensitivity of the global calibration seem appropriate, we believe 1000 runs for the local sensitivity analysis might have been too low, especially considering the large number of parameters used, which influences the relative sensitivity of each parameter (Abbaspour et al. 2017). While Abbaspour et al. suggest between 500 and 1000 runs suffice, we believe the effects of more runs especially when using many parameters should be studied. Opening the parameter ranges further might lead to increased p-values and better calibrations/validations. However, effects of the parameters on hydrological processes not represented in the streamflow must be carefully assessed. We encountered several difficulties with unrealistic soil moisture and aquifer behavior using less restricted ranges, which led to bad multi-objective validation results. Furthermore, it can be assumed that if more stations with longer and more complete time series are available, better and more accurate results can be generated.

5.4.2 Multi-Objective Validation Discussion

It seems unrealistic to expect more discharge observations becoming available in the near future. So far, discharge measurements based on satellite-derived water levels have been limited to rivers wider than about 100 m, their spatial coverage is limited by orbit patterns, and they rely on assumptions inherent to rating-curve approaches or river hydraulic modeling which are difficult to verify. Therefore, alternative methods for verifying the accuracy of hydrological model outputs must be explored (Milzow et al. 2011). The multi-objective validation allows us to assess the performance of the model for multiple aspects of the water balance. In terms of actual evapotranspiration, the remote sensing and reanalysis climate forcings allowed for a very good performance at the basin scale.

When looking at single subbasins, however, the model tends to overpredict ETA in extremely arid areas. In some very humid subbasins, ETA may likewise be underpredicted. When validating MOD 16 ETA for South Africa, underpredictions of between 13% and 35% have been found (Jovanovic et al. 2015; Sun et al. 2012), leading us to assume that the apparent overestimation in arid areas in our model may in part be due to inaccuracies of the MODIS validation data, while the underestimation of ETA during the dry seasons in the southern model could be explained due to Lake Volta not being simulated.

Dynamics of the modeled soil moisture fit the observations very well. The SWAT uncertainties increase markedly during the wet seasons due to a higher availability of water and thus greater influence of the governing parameters. SWAT SM outputs do not allow for a direct comparison, due to the lack of residual water content included in the results.

The validation of the simulated total water storage with GRACE showed good agreement with some peculiarities. The phase shift of one-half month that we identify, especially in the South and West models, has also been observed by Grippa et al. (2011) and Ndehedehe et al. (2016) when comparing multi-model results with GRACE solutions for West Africa. The most noticeable difference between model and GRACE solutions is the very pronounced decline in TWS during the dry season retrieved by GRACE, which is not always captured by SWAT. This discrepancy is very strong in the West and Niger models, and while it may in part be due to our calculation of the TWS change in SWAT, similar observations have been reported in other studies. Grippa et al. (2011) compared water storage anomalies derived from nine land surface models to GRACE, both for the Sahel and West Africa. Their findings are very similar to ours, with SWAT TWS change estimations of our West model comparing well to the Sahel zone and the South model to the West Africa zone, while the Niger model lies in between the two. They assume that incorrectly modeled evapotranspiration during the dry season led to these results. Boone et al. (2009) also compared land-surface models (LSMs) with GRACE over West Africa and came to the conclusion that the difference in amplitudes might either be due to deficits in the precipitation forcing of the LSMs, their insufficient soil depth (where water percolating past a certain depth is lost, similar to SWAT) or the overestimation of the storage anomalies by GRACE during the dry season. Similar observations were made by Ndehedehe et al. (2016), who speculate that differences might be due to anthropogenic influences intensifying land surface processes which the models cannot capture, or the lack of observed data for model calibration leading to improper soil moisture outputs and thus wrong TWS solutions. Werth et al. (2017)

observed an increase in total water storage over the Niger river basin of seven mm/year and conclude this to be mainly due to an accumulation of groundwater in the Sahel Zone. While we observe positive trends of the total water storage for all models except West v2, which is influenced by high uncertainties, trends in SWAT are generally lower than the GRACE solutions. Furthermore, several studies (Forootan et al. 2014; Werth et al. 2017; Rateb et al. 2017; Hassan and Jin 2016) report a clear positive trend toward a higher total water storage over the Volta basin since 2007 due to increased precipitation. We have seen a similar effect before removing the Lake Volta signal from the GRACE solution, where we observed a trend of 25 mm/year from January 2007 to December 2010. Afterwards, a positive trend is much less evident, and we conclude that their results were masked by the strong signal of the lake.

5.5 Conclusion

For the first time, to the authors' knowledge, has a SWAT model been calibrated using remote sensing and reanalysis inputs and validated for streamflow, actual evapotranspiration, soil moisture dynamics and total water storage simultaneously, proving its robustness and predictive capability. Results show that SWAT simulations for different sparsely-gauged regions of West Africa using freely available remote sensing and reanalysis datasets as input perform surprisingly well. This framework significantly eases the modeler's task of acquiring the necessary climatological, land use and soil data to parameterize a physically-based model. Especially considering the lack of measurements conducted *in situ*, the use of remote sensing is essential to produce meaningful assumptions of the water resources in West Africa. While the models perform well using two different calibration and validation schemes, it is necessary to further validate parameters apart from streamflow, otherwise errors in other parts of the water balance might be overlooked. Worqlul et al. (2017) have e.g., shown that streamflow may be well simulated even if input precipitation data has large errors. We therefore chose to additionally validate actual evapotranspiration, soil moisture and total water storage outputs. The multi-objective validation produced very good results and confirmed that the model performs well in the study area. While our approach delivers good results at the regional, sub-continental scale, we realize that it might not be appropriate to model smaller catchments. The model framework could be further improved if data becomes available to accurately model the Niger Inland Delta and Lake Volta. Also, the sensitivity analysis procedure should be improved if using a large number of potential parameters, as in our v2 approach. Furthermore, parameters such as actual evapotranspiration or leaf area index could be included in a multi-objective calibration using SWAT-CUP. Our framework offers possibilities for further evaluation of the water cycle in West Africa. In ongoing work, we plan to evaluate the model performance against global hydrological models to investigate capabilities and limitations of these models and investigate the model response to extreme drought and flood events. Also, the performance of SWAT with different remote sensing inputs can be evaluated for the region. Nonetheless, it is the authors' opinion that remote sensing data should only be used to complement and not replace discharge and other *in situ* measurements for model calibration and validation. Despite the availability of satellite measurements, we believe countries should still invest in *in situ* measurement networks.

Acknowledgements: This study is part of the COAST project (Studying changes of sea level and water storage for coastal regions in West-Africa using satellite and terrestrial data sets) of the University of Bonn, supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) under Grants No. DI443/6-1 and KU1207/20-1. We are grateful to Christophe Peugeot and the AMMA-Catch project as well as to the Global Runoff Data Centre in 56068 Koblenz, Germany for providing discharge data. The AMMA-CATCH regional observing system was set up thanks to an incentive funding of the French Ministry of Research that allowed pooling together various preexisting small scale observing setups. The continuity and long term perenity of the measurements are made possible by an undisrupted IRD funding since 1990 and by a continuous CNRS-INSU funding since 2005.

Author Contributions: Bernd Diekkrüger and Thomas Poméon designed the framework of the study. Thomas Poméon performed data preparation and SWAT modeling. Bernd Diekkrüger secured additional discharge data and consulted during the modeling process. Anne Springer performed the computations related to the GRACE data. Annette Eicker and Jürgen Kusche contributed their knowledge about GRACE processing and analysis. All authors contributed to the interpretation of the results and proofreading.

6 Computationally Efficient Multi-Parameter Calibration and Validation of a Grid-Based Hydrological Model in Sparsely-Gauged West African River Basins

This chapter has been published as: Poméon, T., Diekkrüger, B., Kumar, R. 2018. Computationally Efficient Multi-Parameter Calibration and Validation of a Grid-Based Hydrological Model in Sparsely-Gauged West African River Basins. Water. 10, 26. doi:10.3390/w10101418. https://www.mdpi.com/2073-4441/10/10/1418

Abstract: The prediction of freshwater resources remains a challenging task in West Africa, where the decline of *in situ* measurements has a detrimental effect on the quality of estimates. In this study, we establish a series of modeling routines for the grid-based mesoscale Hydrologic Model (mHM) using Multiscale Parameter Regionalization (MPR). We provide a computationally efficient application of mHM-MPR across a diverse range of data-scarce basins using *in situ* observations, remote sensing, and reanalysis inputs. Model performance was first screened for four precipitation datasets and three evapotranspiration calculation methods. Subsequently, we developed a modeling framework in which the pre-screened model is first calibrated using discharge as the observed variable (mHM Q), and next calibrated using a combination of discharge and actual evapotranspiration data (mHM Q/ET). Both model setups were validated in a multi-variable evaluation framework using discharge, actual evapotranspiration, soil moisture and total water storage data. The model performed reasonably well, with mean discharge KGE values of 0.53 (mHM Q) and 0.49 (mHM Q/ET) for the calibration; and 0.23 (mHM Q) and 0.13 (mHM Q/ET) for the validation. Other tested variables were also within a good predictive range. This further confirmed the robustness and well-represented spatial distribution of the hydrologic predictions. Using MPR, the calibrated model can then be scaled to produce outputs at much smaller resolutions. Overall, our analysis highlights the worth of utilizing additional hydrologic variables (together with discharge) for the reliable application of a distributed hydrologic model in sparsely gauged West African river basins.

Keywords: mHM; remote sensing; GRACE total water storage; MODIS/GLEAM evapotranspiration; modeling framework

6.1 Introduction

Due to economic conditions in many West African countries, between 50 and 60% of the labor force works in the agricultural sector, mostly as self-sustaining farmers generating income by selling surpluses and cash-crops not intended for local consumption (Hollinger and Staatz 2015; Jalloh et al. 2013; African Development Bank 2018). Therefore, water availability not only directly influences the food security of large parts of the population, but also economic development (Schuol et al. 2008a). The estimation of available water resources using hydrologic modeling provides important information for planners and policy makers to mitigate problems arising due to water shortages. The subject of performing hydrologic predictions in sparsely gauged West African river basins has been well covered in recent years (Schuol et al. 2008a; Bormann 2005; Wagner et al. 2009; Fujihara et al. 2014; Schuol and Abbaspour 2006a; Schuol et al. 2008b; Xie et al. 2012; Poméon et al. 2018). Due to a continuous decline in ground-based observation networks as a consequence of political unrest and financial instability (Adjei et al. 2012; Hughes 2006), the authors explored the possibility of setting up the semi-distributed Soil and Water Assessment Tool (SWAT) model (Arnold et al. 1998; Arnold et al. 2012b; Srinivasan et al. 1998) for several West African river basins in a previous study (Poméon et al. 2018). While multi-objective validation of streamflow, actual evapotranspiration, soil moisture dynamics, and total water storage revealed the model to provide robust results, the applied scale was coarse due to computational constrains, with the smallest areas of fully distributed estimations, called subbasins in SWAT, being at least 500 km^2 large.

Hydrologic models are generally calibrated by optimizing multiple free parameters to fit a model output against an observation using objective functions to measure the performance of the solution (Beven 2012). The continuous increase in computational power and availability of remotely sensed data for model parametrization has given rise to complex spatially distributed and semi-distributed hydrologic models in recent years (Samaniego et al. 2010). Well-known examples for this are the spatially distributed MIKE-SHE model (Schulla and Jasper 2007), and the semi-distributed SWAT model (Arnold et al. 1998; Arnold et al. 2012b; Srinivasan et al. 1998). However, especially at the mesoscale, the main problems of contemporary hydrology-nonlinearity, scale, uniqueness, equifinality and uncertainty-remain (Beven 2001). It has been suggested that these problems are not adequately addressed by the continuous increase in model complexity, which does not necessarily relate to an improved performance. Especially, overparameterization and subsequent parameter equifinality are cited as key problems (Samaniego et al. 2010; Kumar et al. 2013b; Orth et al. 2015). In spatially distributed models, free parameters have to be inferred through calibration for each modeling unit, thus increasing the number of parameters if the model resolution is increased (Pokhrel et al. 2008; Beven 1993). One technique for reducing the number of free parameters is the Hydrological Response Unit (HRU) approach, applied, among others, by SWAT, where modeling units of the same physical characteristics (soil, land use, slope, etc.) are first grouped and then calibrated together (Arnold et al. 2012b; Becker and Braun 1999; Kumar et al. 2010). Another method of parameter reduction, which is applied by the mesoscale Hydrologic Model (mHM) (Samaniego et al. 2010; Kumar et al. 2013b), is multiscale parameter regionalization (MPR) (Samaniego et al. 2010). In MPR, model parameters and physical basin parameters are connected *via a priori* defined

relationships, e.g., through pedotransfer functions, and only the global parameters that define these relationships are optimized by calibration (Samaniego et al. 2010; Kumar et al. 2013b).

Multiple studies have explored the performance of the mHM model, mostly for the European continent (Samaniego et al. 2010; Kumar et al. 2013b; Kumar et al. 2010; Kumar et al. 2013a; Rakovec et al. 2016b; Rakovec et al. 2016a; Zink et al. 2017; Samaniego et al. 2013; Thober et al. 2015). Kauffeldt et al. (2016) give an overview of 24 large-scale hydrologic models and their suitability for the European Flood Awareness System. They state that mHM fulfills almost all requirements and needs little modification to be adopted. More recently, mHM has also been included in the Inter-Sectoral Impact Model Intercomparison Project-regional water sector (ISI-MIP; `www.isimip.org`), which analyzes forcing and model uncertainties using an ensemble of hydrologic models and climate scenarios for multiple regions on a century-long timescale, including the Blue Nile and Niger river basins (Hattermann et al. 2017; Samaniego et al. 2017a; Krysanova et al. 2017; Huang et al. 2017).

It has been shown by Rakovec et al. (2016a) that the implementation of Gravity Recovery and Climate Experiment (GRACE)-derived total water storage data significantly improves the partitioning of rainfall into the runoff components without impairing discharge simulations in 83 European river basins. Furthermore, mHM was used in the framework of the German drought monitor, where a drought map for the whole of Germany was generated using mHM soil moisture estimations (Zink et al. 2016). mHM has also been used to produce a high-resolution dataset of water fluxes for Germany by running the model for seven large German river basins and validating the results using 222 additional streamflow stations, as well as evapotranspiration, soil moisture and groundwater recharge data. Results have shown the model to capture the daily streamflow dynamics at the calibrated stations very well and to also sufficiently estimate stations not included in the calibration. Daily actual evapotranspiration evaluated against multiple eddy covariance stations showed little bias and no systematic over- or underestimation. Soil moisture anomalies also proved to be in good agreement with the observations. Spatial patterns of actual evapotranspiration and groundwater recharge evaluated against the observations again showed good agreement (Zink et al. 2017). The application of the model in a data-scarce basin in India using remote sensing data for model parametrization also produced promising results (Samaniego et al. 2011).

While calibrating and validating a hydrologic model using discharge observations allows the modeler to confidently predict runoff, the same cannot be said for the representation of other hydrologic processes due to parameter equifinality (Poméon et al. 2018; Rakovec et al. 2016b). If, e.g., modeled discharge is to be reduced to fit observations during calibration, the same effect can be achieved by increasing either actual evapotranspiration or percolation. Since the prediction of spatially distributed hydrologic components such as evapotranspiration, soil moisture or total water storage is increasingly desired, multivariate calibration and validation is of immense value in producing realistic results other than streamflow (Rakovec et al. 2016b; Rakovec et al. 2016a; Zink et al. 2018). This holds especially true for data-scarce regions.

Earlier studies have shown the scaling capabilities of mHM-MPR in Europe and the United States (Kumar et al. 2013a; Rakovec et al. 2016b). This approach will be tested in the data-scarce West African domain for the first time. In this study, we assess the scalability of mHM as argued

by Kumar et al. (2013a) by calibrating the model at a computationally efficient coarse scale and running the calibrated model in finer resolutions. If the scaling capability of the model has been proven to perform well, it can then be used to generate hydrologic outputs at much smaller, local scales for various research and policy applications. Furthermore, this approach could be utilized to include other hydrologic variables at their native resolutions without introducing bias through data aggregation and disaggregation (Rakovec et al. 2016b).

The objectives for this study are therefore (1) to investigate the performance of the mHM model using a multiscale parameter regionalization approach for selected sparsely gauged West African river basins under different precipitation inputs, (2) to assess the multivariate calibration options of mHM by calibrating the model for streamflow and actual evapotranspiration and validating against these, as well as against soil moisture and total water storage anomaly datasets, and (3) to assess the transferability of model parameters inferred through MPR between different spatial resolutions.

This study is part of the German Research Foundation-funded COAST (studying changes of sea level and water storage for coastal regions in West Africa using satellite and terrestrial data sets) project. The work builds on previously published studies (Poméon et al. 2017; Poméon et al. 2018), with the main focus being the assessment of the contribution of remote sensing evapotranspiration, soil moisture and total water storage estimates in hydrologic simulations of data-scarce basins.

6.2 Materials and Methods

6.2.1 Study Area

The study area in southern West Africa, located between 4.5 and 15° latitude and -6 and 3.5° longitude, is presented in Figure 6.1a. The topography of the study area is flat, with the highest mountain range stretching from southern Ghana (Akwapim Hills) through Togo (Togo Mountains) and into northern Benin (Atakora Mountains) (CILSS 2016). The highest peaks reach almost 1000 m according to DEM data (Lehner et al. 2008; Lehner et al. 2013). Rainfalls in the region are strongly seasonal; bimodal in the south (e.g., Cotonou, see Figure 6.1c) and unimodal in the north (e.g., Ouagadougou, see Figure 6.1b), with a distinct south-north gradient of reducing rainfall sums (Fink et al. 2010; Sebastian and Kate 2009; Gessner et al. 2013).

Ten river basins were chosen to be modeled based on the availability of streamflow records for calibration and validation: Comoé (Ivory Coast, Ghana and Burkina Faso, 74,604 km^2, three gauges), Ouémé (Benin, Nigeria and Togo, 47,562 km^2, 13 gauges), Couffo (Benin and Togo, 1672 km^2, one gauge), Mono (Togo and Benin, 22,711 km^2, two gauges), Pra (Ghana, 22,733 km^2, one gauge), Ayensu (Ghana, 1699 km^2, one gauge), Ankobra (Ghana, 4246 km^2, one gauge), Black Volta (Burkina Faso, Ivory Coast, Ghana and Mali, 138,440 km^2, four gauges), White Volta (Burkina Faso, Ghana and Togo, 100,036 km^2, seven gauges), and Oti (Burkina Faso, Togo, Benin and Ghana, 60,015 km^2, three gauges). Basin areas are relative to the last modeled gauging station and not the entire river basin. In total, 473,718 km^2 and 36 streamflow gauging stations were modeled.

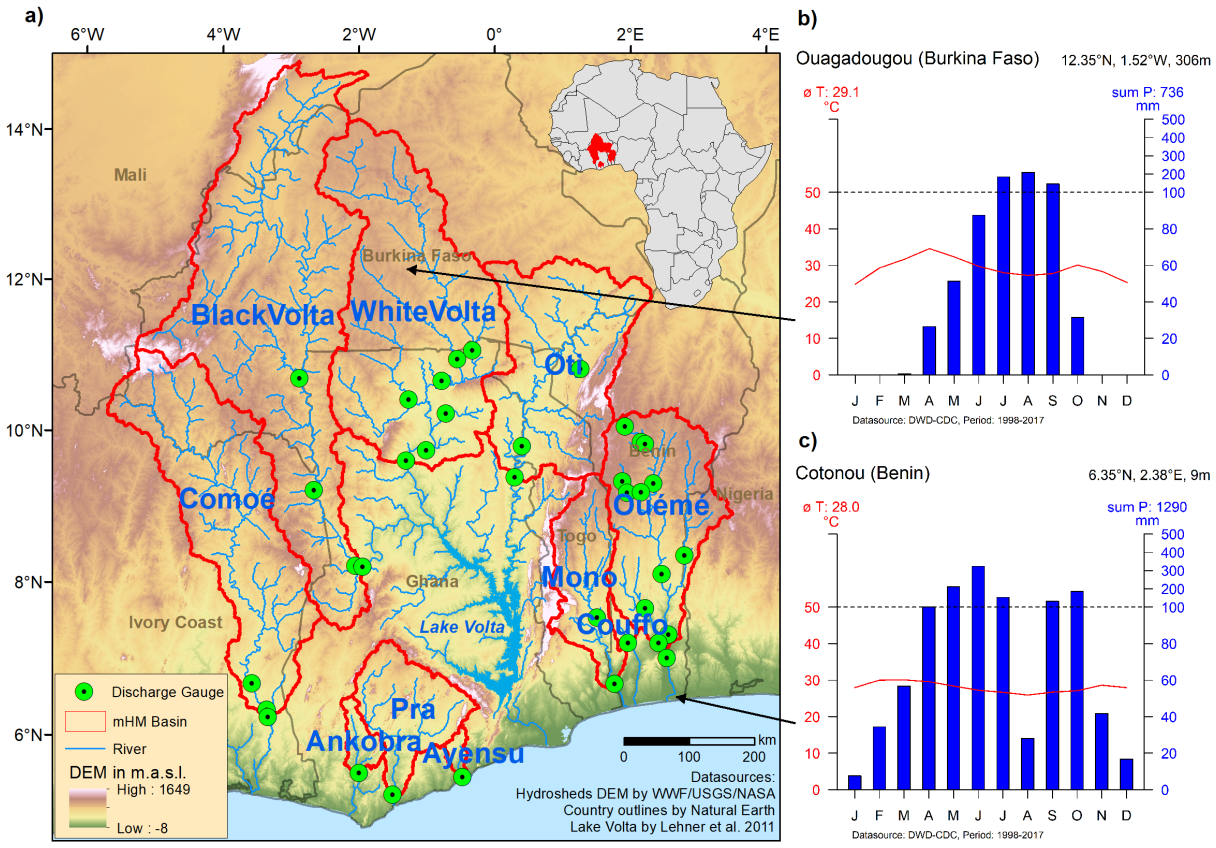

Figure 6.1: (a) Study area with basins modeled in mHM; (b) Climate diagram Ouagadougou; (c) Climate diagram Cotonou. Dashed line indicates break in precipitation axis. Station data was compiled from the German Meteorological Agency Climate Data Center (DWD-CDC).

6.2.2 The Mesoscale Hydrologic Model (mHM)

The mesoscale Hydrologic Model (mHM) (Samaniego et al. 2010; Kumar et al. 2013b) is a spatially explicit, grid-based hydrologic model created specifically for providing distributed predictions of hydrologic variables like runoff, evapotranspiration, soil moisture and discharge along the river network. The main components of mHM are based on well-established hydrologic process descriptions of large-scale models like HBV (Bergström 1976, 1992) and VIC (Liang et al. 1994). mHM accounts for the following processes: canopy interception, snow accumulation and melting, soil moisture dynamics, infiltration, surface runoff, discharge generation, evapotranspiration, subsurface storage, deep percolation, baseflow, and flood routing (Samaniego et al. 2010).

In mHM, different (spatial) levels of modeling components are considered to better account for the spatial variability of inputs and hydrologic processes. A schematic overview of mHM is given in Figure 6.2. The model source code is freely available, and further details can be found at `www.ufz.de/mhm`. The finest-resolution category is the small-scale morphological input class Level 0 (L0). This class contains variables such as elevation, slope, soil, and land use. The next highest resolution, Level 1 (L1) represents the resolution of the hydrologic model routines and output and only requires the regionalized fields of model parameters (see below). Level 2 (L2)

is the coarsest-resolution category, containing the large-scale meteorological forcing inputs such as precipitation and temperature (Samaniego et al. 2010; Samaniego et al. 2017b). mHM uses a novel Multiscale Parameter Regionalization (MPR) approach to estimate the regionalized fields of model parameters first at the L0 resolution. Then, an upscaling operator is used to generate effective parameters at the L1 modeling resolution. Regionalization at the L0 scale is performed by linking the model parameters to available catchment attributes (terrain, slope and aspect, soil textural and land cover properties) *via* a set of pedotransfer functions and free calibration parameters (see Samaniego et al. (2010) and Kumar et al. (2013b) for further details). The MPR approach significantly reduces the number of free (calibration) parameters while accounting for the spatial variability of model parameters required for distributed hydrologic model applications (see Livneh et al. (2015) for details on estimation of key soil hydrologic parameters using MPR). Kumar et al. (2013b) found both MPR and HRU approaches to perform similarly regarding daily streamflow predictions in 45 calibrated southern German river basins. However, MPR proved to be superior in the preservation of spatiotemporal patterns of water fluxes and, since free parameters of MPR are not scale-specific, they can be transferred to different modeling scales without time-intensive recalibration and without introducing significant bias. This effectively means that mHM can be calibrated at a computationally efficient, coarser L1 resolution and that the calibrated (free) parameters can then be transferred to a finer resolution. Further studies also demonstrated the scaling capabilities of mHM on comparatively large domains of European basins and the Mississippi river basin (Kumar et al. 2013a; Rakovec et al. 2016b).

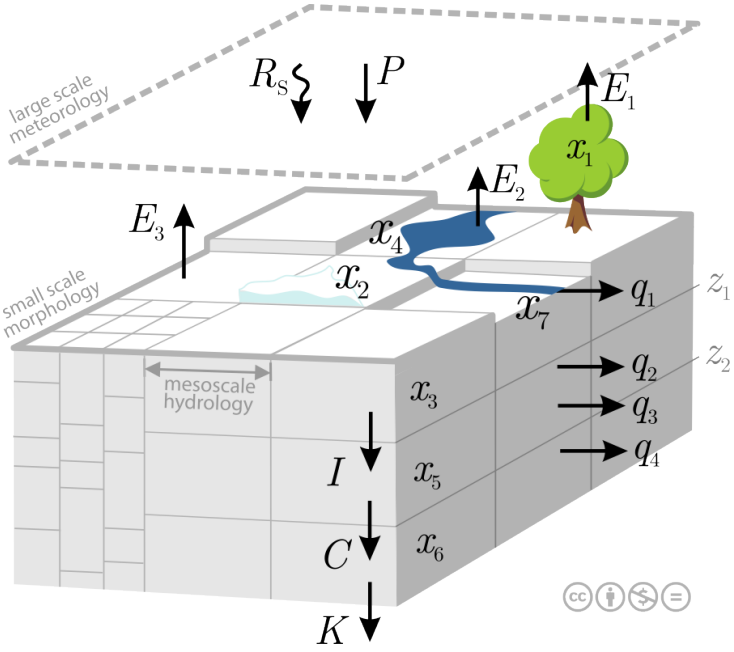

Figure 6.2: Schematic overview of the mesoscale Hydrologic Model, mHM, developed at the Helmholtz Centre for Environmental Research—UFZ (Samaniego et al. 2017b).

mHM offers multiple optimization algorithms, such as Monte Carlo Markov Chain (MCMC), Dynamically Dimensioned Search (DDS), Simulated Annealing (SA) and Shuffled Complex Evolution (SCE) (Samaniego et al. 2017b). In this study, the Dynamically Dimensioned Search algorithm was used. DDS has been shown to perform well when compared to SCE, converging

on a good solution in only 15 to 20% of the model runs required by SCE while also avoiding local optima (Tolson and Shoemaker 2007). Behrangi et al. (2008) noted that DDS finds a solution faster than SCE because it has been developed especially for computationally demanding models, while SCE was originally developed for lumped models, and go on to show that SCE may be adapted to decrease the number of iterations needed to find a good solution. The authors suggest a large number of model iterations to ensure good solutions are found by DDS.

6.2.3 Input Data

Morphological Inputs

The HydroSHEDS (Hydrological data and maps based on SHuttle Elevation Derivatives at multiple Scales) hydrologically conditioned digital elevation model (DEM) (Lehner et al. 2008; Lehner et al. 2013) was used as a basis to create the morphological input data. The DEM is based on Shuttle Radar Topographic Mission (SRTM) data and was developed by the World Wildlife Fund (WWF) and the United States Geological Survey (USGS). While available in multiple resolutions, in this study the 500 m resolution was used. The required slope, aspect, flow direction and flow accumulation data were derived from this dataset.

Soil Inputs

Concerning soil input data, a raster dataset with the location of soil classes and a corresponding lookup table with the attributes' depth, texture and bulk density are required. This information was compiled from the Harmonized World Soil Database (HWSD) version 1.2 (FAO/IIASA/ISRIC/ISS-CAS/JRC 2012). The dataset contains physical and chemical parameters for a topsoil (0-30 cm) and subsoil (>30-100 cm) layer at a 1 km resolution.

Land Use Inputs

Two land use inputs are necessary, a Land Use and Land Cover (LULC) map and a Leaf Area Index (LAI) class map with a corresponding lookup table. While mHM recognizes only three LULC classes (forest, pervious, impervious), an arbitrary number of LAI classes can be defined. Both maps were generated using the Globcover 2.3 product developed by the European Space Agency (ESA), representing the global land cover of the year 2009, available at a 300 m resolution (Bontemps et al. 2011). To generate the mHM LULC map, Globcover classes were reclassified to the three mHM classes. The mHM LAI class map was produced using the original Globcover classes and calculating long-term monthly LAI values (2003-2013) for each class from MODerate-resolution Imaging Spectroradiometer (MODIS) MCD15A2v5 LAI data, produced by the National Aeronautics and Space Administration (NASA) (Myneni et al. 2015a; Knyazikhin et al. 1999).

Meteorological Inputs

In Poméon et al. (2017), ten remotely sensed and reanalyzed precipitation datasets were evaluated for the study region. Products relying on a combination of satellite microwave and infrared measurements as well as bias correction using ground-based data generally performed best. Four

of the best-performing datasets were chosen for exploratory mHM runs. The combination that performed best in the exploratory runs was used for model calibration. Firstly, the Climate Prediction Center Morphing Technique (CMORPH) v1 CRT was chosen, which is globally available at a 0.25° resolution from 1998 onwards (Joyce et al. 2004; Xie et al. 2011). CMORPHv1 CRT, hereafter named CMORPH, was also used in SWAT simulations of sparsely gauged river basins in West Africa (Poméon et al. 2018). Secondly, the Tropical Rainfall Measuring Mission-Multi-satellite Precipitation Analysis (TRMM-MPA) product version 3B42v7, hereafter named TMPA, was evaluated. It is available from 50° north to 50° south at a 0.25° resolution from 1998 onwards (Huffman et al. 2007). Further evaluated was PERSIANN CDR version 1, available from 60° north to 60° south at a 0.25° resolution (Ashouri et al. 2015). All three products include bias-corrected satellite microwave and infrared measurements (Joyce et al. 2004; Xie et al. 2011; Huffman et al. 2007; Ashouri et al. 2015). Lastly, the Global Precipitation Climatology Centre Full Data Daily (GPCC FDD) version 1 global rain gauge product, developed by the German Meteorological Agency at a 1° resolution, was evaluated (Schamm et al. 2015). For more information on the used precipitation datasets, please refer to Poméon et al. (2017) (Chapter 4). While rainfall dynamics are very similar for all datasets, CMORPH estimates on average slightly less precipitation during the peak of the wet season from June to September (Figure 6.3). For the 1998 to 2013 long-term monthly average, CMORPH predicts 77 mm monthly precipitation, TRMM 87 mm, PERSIANN 92 mm and GPCC 86 mm.

Depending on the method chosen to model the potential evapotranspiration (ETP), further meteorological inputs are needed to run mHM (Samaniego et al. 2017b). Three methods were evaluated in this study, namely: the estimation based on air temperature data using the Hargreaves-Samani method (Hargreaves and Samani 1985) and a direct in-built read-in of potential evapotranspiration estimates with either LAI or aspect correction. Minimum and maximum 2 m daily temperature was extracted from NASA Modern-Era Retrospective analysis for Research and Applications 2 (MERRA 2) reanalysis data, available at a global grid of 0.625° * 0.5° from 1980 onwards (Bosilovich et al. 2016). Potential evapotranspiration data was obtained from the Global Land Evaporation Amsterdam Model (GLEAM) v3.1a (Martens et al. 2017; Miralles et al. 2011). GLEAM data was chosen since it is available for the studied period (model warm-up starting in 1998) at a daily timestep.

Discharge Inputs

The discharge data used for calibrating and validating the mHM model was obtained from the German Global Runoff Data Center (GRDC), the French AMMA-CATCH (Analyse Multidisciplinaire de la Mousson Africaine-Couplage de l'Atmosphère Tropicale et du Cycle Hydrologique) regional observing system, and through personal communication with local agencies and partners. Due to different observation periods and gaps in the discharge data that did not allow for fixed calibration and validation periods, two thirds of the data were used for calibration and one third for validation (Poméon et al. 2018; Abbaspour et al. 2015).

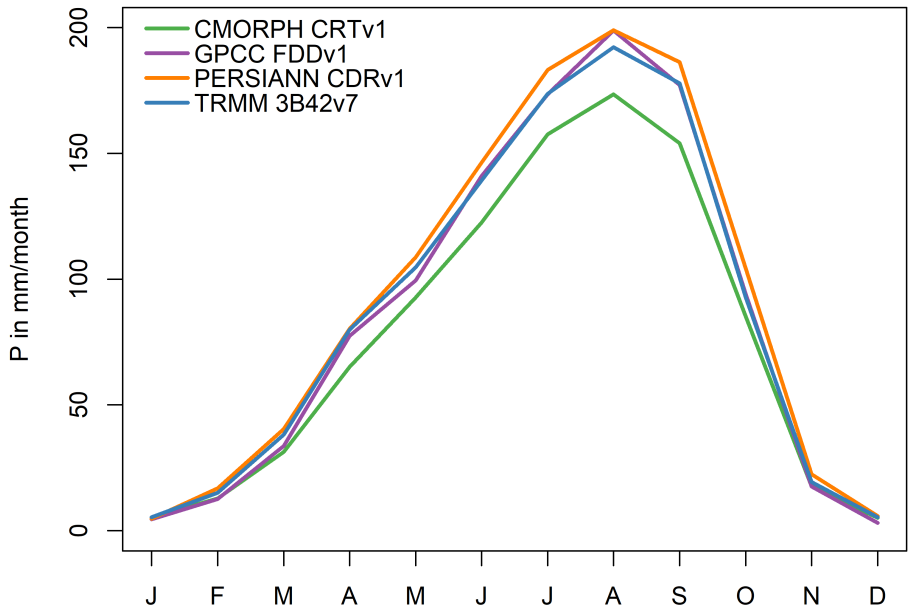

Figure 6.3: Long-term (1999–2013) monthly sums of chosen precipitation inputs. Products were projected in the mHM L2 resolution and averaged over the study area.

Validation Data

Multivariate validation of the modeled soil moisture, actual evapotranspiration and total water storage anomalies was furthermore performed. For the soil moisture validation, the ESA Climate Change Initiative (CCI) soil moisture (v4.2; www.esa-soilmoisture-cci.org) combined product was used. The combined product was created by merging active scatterometer and passive soil moisture retrievals from multiple satellites. Daily retrievals, although only in the satellite swaths, are available at a resolution of 0.25° (Liu et al. 2012; Liu et al. 2011; Wagner et al. 2012; Gruber et al. 2017; Dorigo et al. 2017). Three different datasets were used for the validation of the actual evapotranspiration, namely MODIS MOD16A2 (Mu et al. 2007; Mu et al. 2011), GLEAM 3.2a and GLEAM 3.2b (Martens et al. 2017; Miralles et al. 2011). MOD16A2 ETP estimates were further used to validate model results during the exploratory model runs. For the validation of total water storage anomaly (ΔTWS), remotely sensed GRACE retrievals from ITSG-Grace 2016 time series provided by the Institute of Geodesy at Technical University Graz (Mayer-Gürr et al. 2016) and adapted for the study area (Poméon et al. 2018) were used. GRACE estimates are available at a minimum spatial resolution of 400 * 400 km and a temporal resolution of 30 days (Tapley et al. 2004). While the model can be calibrated using TWS, the processed timeseries covers all basins and cannot be further spatially disaggregated into single basins.

6.2.4 Framework of the Modeling Experiment

To be able to place confidence in the model output, the following framework was developed, where the input data and modeling routines are examined in twelve default simulation runs before the model is optimized. In the first step after preparing the input data and setting up

the model, a simulation using the default parameter values was performed. For each subsequent run, the precipitation inputs were varied between the four datasets (CMORPH, GPCC, PERSIANN and TMPA). Likewise, the potential evapotranspiration calculation method was changed between three options available in mHM, namely the Hargreaves-Samani temperature-based method (HAR), read-in with automatic LAI correction (LAI) and read-in with automatic aspect-driven correction (ASP) (Samaniego et al. 2017b) for each precipitation dataset. A graphical overview of the setup and associated combination numbers is presented in Figure 6.4.

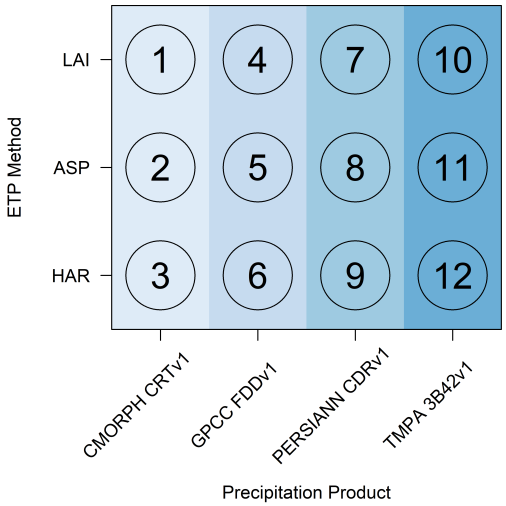

Figure 6.4: Initial model evaluation simulation numbers. ETP method: chosen potential evapotranspiration method; HAR: Hargreaves-Samani; ASP: ETP read-in with aspect correction; LAI: ETP read-in with LAI correction.

In our case, the model is run at a daily timestep from 1998 to 2013, with the year 1998 considered as the warm-up period. While daily streamflow outputs are produced, other variables are output at a monthly resolution for better comparison to monthly remote sensing estimates.

It has been shown that precipitation estimates have a strong impact on model performance (Poméon et al. 2017; López et al. 2017). The performance of the simulated streamflow, potential and actual evapotranspiration and total water storage anomaly were evaluated for each combination. Once the best initial combination was determined, the most efficient model resolution was chosen by running the previously defined simulation at different resolutions and evaluating the required runtime for each resolution. After establishing the best combination of precipitation input data, evapotranspiration calculation method, and modeling resolution, the model was calibrated for each river basin separately using discharge records as the observed variable. The optimization method based on the Kling-Gupta Efficiency (KGE) measure was chosen in this study. KGE offers the advantage of including correlation, bias and relative variability in its composition. The dimensionless result ranges from $-\infty$ to 1 (Gupta et al. 2009) and will be regarded as acceptable from 0.5 upwards. For the chosen optimization method, mHM calculates 1 minus the average KGE, where all stations within the specific basin are weighted equally.

As a separate setup, the model was calibrated using an optimization method which includes both discharge and basin-wide mean actual evapotranspiration (where we used GLEAM 3.2a), as shown in Equation 6.1.

$$SO_{30} = [1 - KGE(Q)] * RMSE\left(\overline{basinETA}\right) \tag{6.1}$$

where SO_{30} is the mHM objective function 30, $KGE(Q)$ is the average Kling-Gupta Efficiency of the discharge simulation (all stations are again weighted equally) and $RMSE\left(\overline{basinETA}\right)$ is the root mean squared error of the basin average actual evapotranspiration (ETA) simulation.

The second calibration was chosen to assess the gains of calibrating the model using an additional dataset (complimenting discharge observations). This approach can be especially useful in sparsely gauged regions, such as West Africa, where only few discharge observations exist, but remote sensing evapotranspiration estimates are easily available. Also, more realistic model results may be obtained if using multiple observed variables in the calibration (Rakovec et al. 2016a). mHM may also be calibrated using SM or TWS inputs. We chose to use ETA as an additional variable for two reasons: (1) Several remote sensing ETA products are readily available and only limited preprocessing is required to adopt the data to mHM. The use of GRACE TWS fields requires extensive preprocessing and SM retrievals are for the most part only available for the top few cm of the soil, for which no information is available in the study area. (2) If SM and TWS information is excluded from the calibration, they remain as independent datasets for validation (Rakovec et al. 2016a).

Both model set-ups were validated using discharge, actual evapotranspiration, soil moisture anomaly, and total water storage anomaly data. Soil moisture anomaly was calculated according to Equation 6.2. This was necessary since absolute comparisons could not be performed using the model outputs and satellite data, as the soil horizons of the variables differ in depth.

$$\%SM_{anomaly} = \frac{\left(SWC_{layer1} - \overline{SWC_{layer1}}\right)}{\overline{SWC_{layer1}}} * 100 \tag{6.2}$$

where $\%SM_{anomaly}$ is the soil moisture anomaly of the first soil horizon in percent and SWC_{layer1} is the soil water content of the first soil horizon in mm.

Finally, total water storage anomaly was calculated according to Equation 6.3. While GRACE provides TWS in mm, the model outputs only allow estimates of the total water storage change due to the initial conditions being unknown.

$$\begin{gathered} TWS = SAT_{stw} + UNSAT_{stw} + SEALEDstw + SWC_{layer1} + SWC_{layer2} \\ \Delta TWS = TWS - \overline{TWS} \end{gathered} \tag{6.3}$$

where TWS is the total water storage, SAT_{stw} is the storage in the saturated groundwater reservoir, $UNSAT_{stw}$ is the storage in the unsaturated reservoir, $SEALED_{stw}$ is the reservoir of sealed areas, SWC_{layerX} is the soil water content of the respective soil layer and ΔTWS is the total water storage deviation from the mean.

After an initial trial model run, it was found that actual evapotranspiration may be over-estimated by mHM using the default parameter ranges for the Hargreaves-Samani coefficient if calibrating for discharge only, and ranges were restricted to 0.0021-0.0023, as according to Gavilán et al. (2006), who examined the adjustment of the coefficient in semi-arid Spain. The minimum and maximum ETP correction factors related to different vegetative covers were also

restricted to generate reasonable ETA estimations (minimum: 0.90-0.96; maximum: 0.17-0.20). Other adjustments include the maximum ranges for the recharge coefficient and the geological recession coefficient parameters, which increased to 200 and 1500 days, to reasonably match groundwater-dominated low-flow periods. When calibrating for discharge in combination with ETA during the second calibration, default ranges were used. In total, parameters regulating interception, soil moisture, direct sealed area runoff, potential evapotranspiration, interflow, percolation and geology were calibrated (see also Appendix A).

6.3 Results

6.3.1 Initial Model Setup Results

During exploratory screening, the model was run with default parameters while varying the precipitation inputs and potential evapotranspiration estimation methods. The runs were validated against discharge, MODIS derived potential (ETP) and actual (ETA) evapotranspiration, as well as GRACE ΔTWS without calibration of free parameters. Results of these analyses are displayed in Figure 6.5.

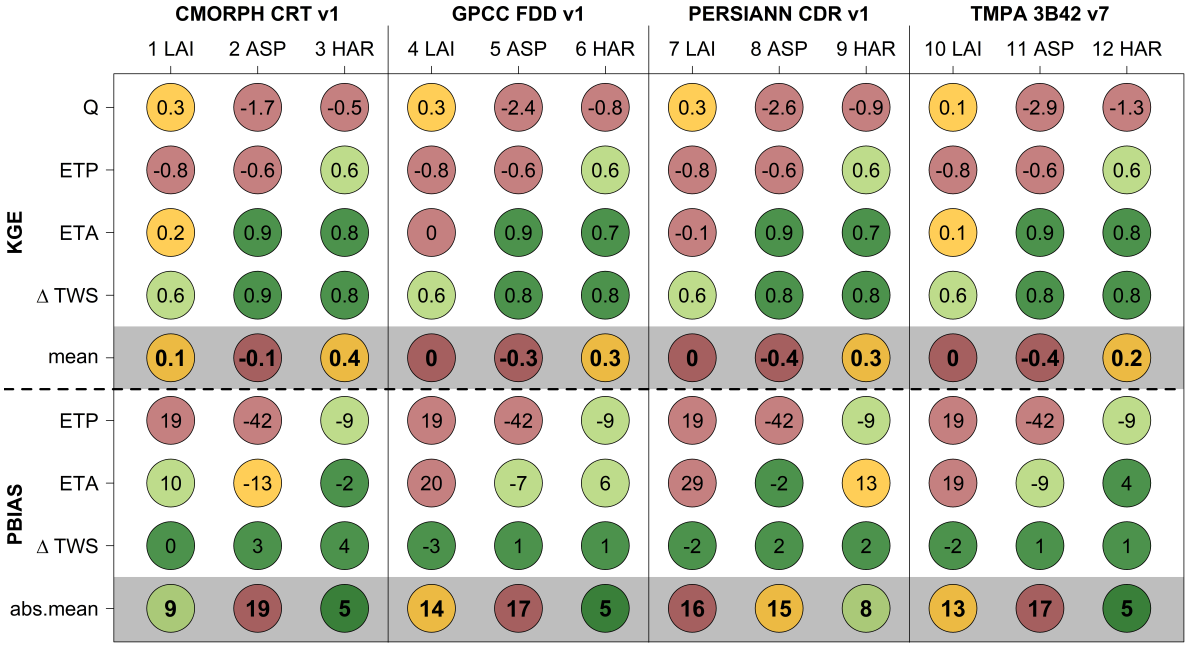

Figure 6.5: Exploratory model results. Where CMORPH CRTv1, GPCC FDDv1, PERSIANN CDRv1 and TMPA 3B42v7 represent precipitation inputs; LAI, ASP and HAR are evapotranspiration calculation methods: HAR: Hargreaves-Samani; ASP: ETP read-in with aspect correction; LAI: ETP read-in with LAI correction; KGE: Kling-Gupta Efficiency; PBIAS: Percent bias; Q: Discharge; ETP: Potential evapotranspiration; ETA: Actual evapotranspiration; ΔTWS: Total water storage deviation from the mean.

It is apparent that the ASP method (models 2, 5, 8, 11) produces the worst KGE results regarding Q and ETP, as well as strong negative bias in ETP and between -2% and -13% bias in ETA validations. It was therefore rejected outright. While the best uncalibrated streamflow

KGE was achieved with the LAI method (models 1, 4, 7, 10), it proved to perform poorly regarding ETP and ETA, with either negative or low KGE values and strong biases with the lowest being 10% (CMORPH, model 1) and the highest 29% (PERSIANN, model 7). If considering mean KGE and absolute mean bias, the HAR method (models 3, 6, 9, 12) performs best for all precipitation products with high KGE values for ETP, ETA and ΔTWS, and the second-best streamflow results after the LAI method. Biases are also the lowest for all precipitation methods. CMORPH precipitation and HAR evapotranspiration (model 3) produce the best results out of all exploratory model runs. This may be due to CMORPH estimating lower precipitation rates for the region than other products (Poméon et al. 2017). CMORPH precipitation in combination with the Hargreaves-Samani ETP method has furthermore been proven to produce good results in SWAT simulations of the region (Poméon et al. 2018). Due to the good performance, this combination was retained for further analysis.

To choose the most efficient modeling (L1) resolution, the model was set up and run with three different resolutions. The only constraint was that these resolutions are divisible by the input data (L0) resolution, and the meteorological forcing (L2) resolution being divisible by the modeling resolution (L1) - so that the different spatial resolutions are compatible with the grid-based structure of mHM facilitating the smooth operation of upscaling and downscaling procedures.

The chosen resolutions were 6.5, 13 and 26 (equal to L2) km. On our intel i5-2400 system with eight gigabyte RAM, one model run (without loading data) took 4.0 min (6.5 km), 1.6 min (13 km) and 0.9 min (26 km), respectively. Smaller resolutions caused the system to run out of memory. If considering the 5000 model iterations chosen for the DDS optimization, this translates to runtimes of 13.9 days, 5.6 days and 3.1 days, respectively. It was therefore decided to run the model in the 26 km resolution, where L1 is equal to L2, and to test the MPR performance by running the final calibrated model parameterization at finer resolutions afterwards.

6.3.2 Calibration and Discharge Validation Results

mHM discharge simulations perform reasonably well for the study region. Results of the two calibration approaches discharge only (Q) and discharge/actual evapotranspiration (Q/ET) are listed in Table 6.1 and will be described in detail.

Concerning the Q calibration and validation (Table 6.1 and Figure 6.6), calibration results are acceptable with an average KGE of 0.53 and R^2 of 0.61. 83% of the 36 stations over 10 basins achieve an R^2 of greater than 0.5 and 75% reach a KGE of above 0.5. KGE results in the southern basins tend to reach from acceptable to good except for two stations in the Ouémé basin, which perform poorly. The model has a comparatively lower accuracy when applied in the larger northern basins, with two out of seven stations in the White Volta basin not reaching a positive KGE; and only one station out of four in the Black Volta basin showing reasonable skill in capturing observed streamflow dynamics. For the validation, 53% of the stations reach an at least acceptable KGE, resulting in a mean KGE of 0.23. Mean R^2 is slightly higher than during the calibration at 0.65, and 75% of the stations reach an acceptable R^2. While streamflow was underestimated during the calibration by -8.8%, it is overestimated during the validation by

12.9%. Interestingly, southern stations in the Ankobra, Pra, Ayensu, Mono, and Ouémé river basins do not perform well during the validation. Reasons may include higher rainfall sums and different rainfall distributions (two rainy seasons in the south, one rainy season in the north), as well as data quality issues. White Volta and Oti basins also see some deterioration. Contrarily, some of the northern Ouémé stations, as well as select stations in the central White Volta and Oti catchments, outperform the calibration.

Table 6.1: mHM discharge results for both discharge and discharge and evapotranspiration calibration averaged over study basins.

	Objective Function				Discharge Stations		
	R^2	sig.	PBIAS	KGE	$R^2 \geq 0.5$	KGE ≥ 0.5	KGE ≥ 0.7
				Q Calibration			
Calibration	0.61	< 0.001	-8.8	0.53	83%	75%	39%
Validation	0.65	< 0.001	12.9	0.23	75%	53%	39%
Average	0.63		2.0	0.38	79%	64%	39%
				Q/ET Calibration			
Calibration	0.56	< 0.001	-8.1	0.49	72%	69%	33%
Validation	0.62	< 0.001	25.2	0.13	75%	47%	28%
Average	0.59		8.5	0.31	74%	58%	31%

R^2: coefficient of determination, sig.: significance level, PBIAS: percent bias, KGE: Kling-Gupta Efficiency, Q Calibration: mHM calibrated using observed discharge, Q/ET Calibration: mHM calibrated using observed discharge and actual evapotranspiration.

Streamflow predictions of the Q/ET calibration scheme perform slightly worse than the Q calibration. Results are depicted in Table 6.1 and Figure 6.7. Nonetheless, 72% of the stations reach an R^2 of over 0.5 with an average of 0.56, while 69% achieve an acceptable KGE with an average of 0.49. Bias is similar to the Q calibration at -8.1%. KGE performance distribution is almost identical to the Q calibration, except for some stations in the White Volta and Ouémé catchments, which performed less well. During the validation, 75% of the stations reach an acceptable R^2, which relates to a mean of 0.62, while 47% reach a KGE of at least 0.5, with the average KGE being 0.13. Positive bias is strong, with 25.2%. KGE distribution is again similar to the first calibration approach, albeit with some stations in the Black Volta, White Volta, Oti and Ouémé catchments performing less well. Still, southern stations in the Ankobra, Pra, Ayensu, Mono and Ouémé catchments perform similarly poor. When comparing averaged calibration and validation results for both calibration methods, average R^2 decreases from 0.63 (Q) to 0.59 (Q/ET), with 79% reaching an acceptable R^2 during the first and 74% during the second calibration. The same holds true for KGE, which decreases from 0.38 (Q) to 0.31 (Q/ET), with 64% (Q) and 58% (Q/ET) of stations performing acceptably. 39% (Q) and 31% (Q/ET) of the stations perform well. However, average bias is markedly increased in the second calibration approach due to overestimation in the validation (8.5% as opposed to 2%).

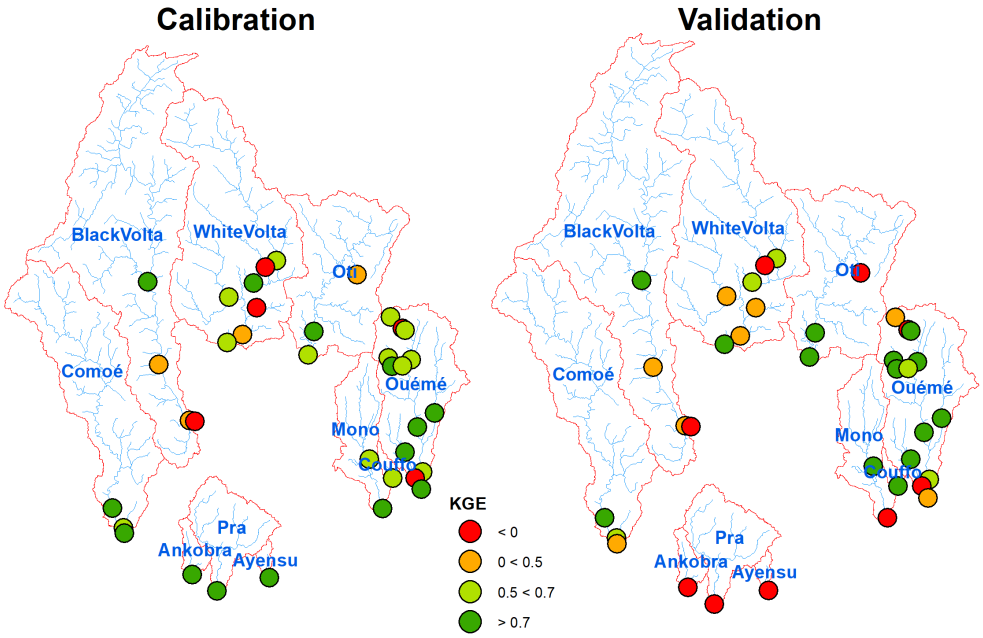

Figure 6.6: Discharge calibration and validation results, discharge (Q) calibration method.

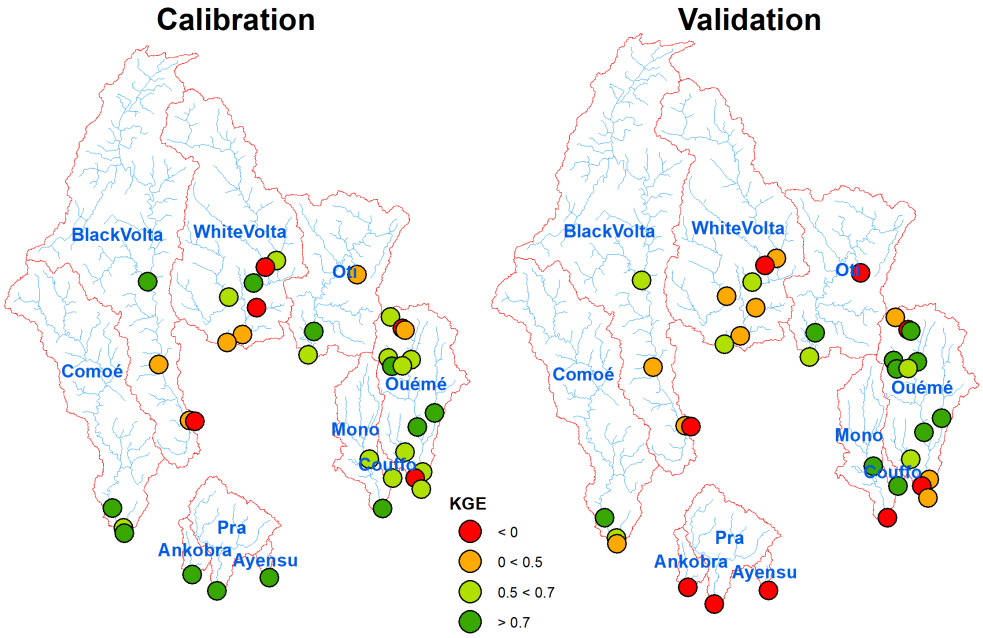

Figure 6.7: Discharge calibration and validation results, discharge and actual evapotranspiration (Q/ET) calibration method.

Two example hydrographs for the Comoé and Oti river basin calibration and validation are given in Figure 6.8. Displayed is the mean monthly observed and simulated discharge of the Q and Q/ET calibration schemes, as well as the respective key efficiency criteria (R^2 and KGE). In

both examples, Q and Q/ET model setups perform very well, with slightly higher KGE values reached by the Q model. As can be seen, performance of both models is higher during the validation period for these basins.

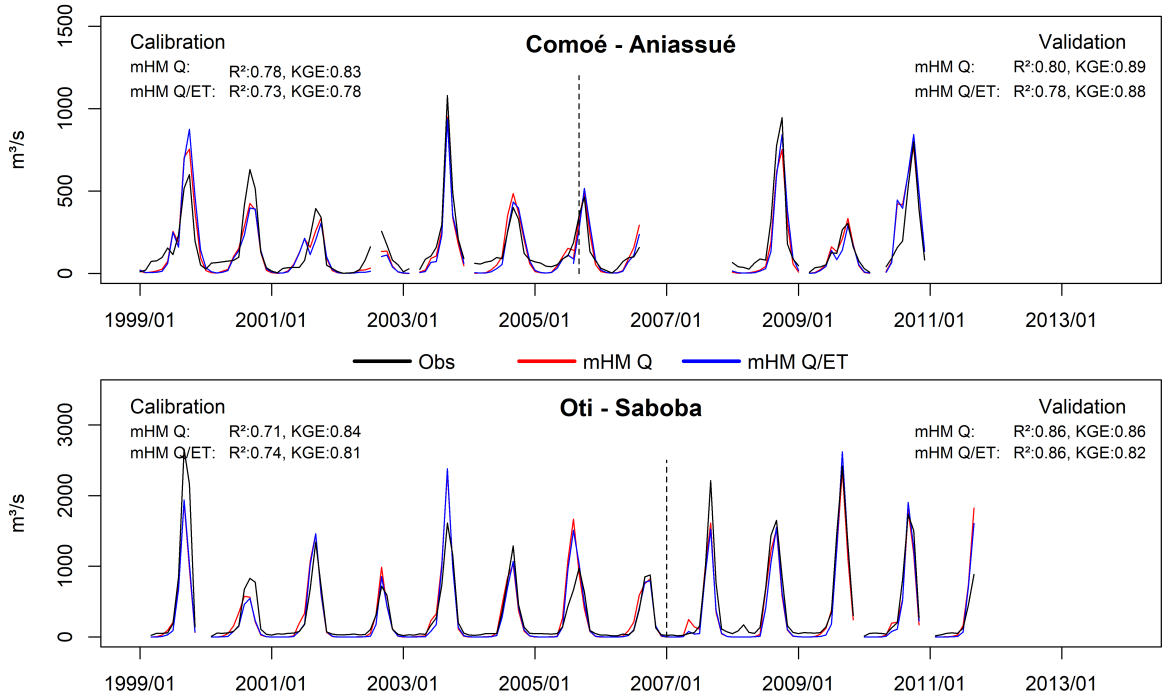

Figure 6.8: Example hydrographs for discharge (Q) and discharge and actual evapotranspiration (Q/ET) model calibrations. Daily model outputs have been aggregated to monthly data for visual clarity. Model efficiency calculations were performed on daily data.

6.3.3 Multivariate Validation Results

Results of the two calibration approaches were further validated against actual evapotranspiration, soil moisture anomaly and total water storage anomaly data. Since mHM was set up to produce monthly output data (except for streamflow, which was daily), all validations are performed with monthly data. Results of the ETA validation are shown in Table 6.2 and Figure 6.9. For the Q calibration, the model overestimated ETA during the rainy seasons, but acceptable results are still reached. Compared against the three reference datasets, mHM performs best against GLEAM 3.2b, with an R^2 of 0.93, 6.6% bias and a KGE of 0.73. Good results were also achieved when validating against GLEAM 3.2a, albeit with a slightly higher bias (8.1%) and lower, but still acceptable, KGE of 0.64. However, a marked difference could be seen compared to MODIS data, with a bias of 11.3% and a no longer acceptable KGE of 0.44. It can be observed that the amplitude of MODIS is lower than in the other products. On average, R^2 is very good at 0.93 and bias is reasonable at 8.7%. Average KGE is still acceptable at 0.60. Results are markedly improved for the Q/ET calibration, but the order of validation performance is identical, even though GLEAM 3.2a was used as calibration input. Best validation is against GLEAM 3.2b, with an R^2 of 0.97, 5.9% bias and a very good KGE of 0.93. Second highest validation occurs against GLEAM 3.2a (R^2: 0.97; bias: 7.4%, KGE: 0.88) and the validation

against MODIS, although in itself still good, performs worst of the three (R^2: 0.92, bias: 10.7%, KGE: 0.72). While the mean KGE strongly increases from 0.60 (Q) to 0.84 (Q/ET), mean RMSE decreases from 15.3 mm (Q) to 9.0 mm (Q/ET).

Table 6.2: mHM modeled actual evapotranspiration validation against remote sensing datasets averaged over study basins.

		Objective Function			
Dataset	R^2	sig.	PBIAS	KGE	RMSE
		Q Calibration			
MOD 16A2	0.93	< 0.001	11.3	0.44	17.9
GLEAM 3.2a	0.93	< 0.001	8.1	0.64	14.8
GLEAM 3.2b	0.93	< 0.001	6.6	0.73	13.1
Average	0.93		8.7	0.60	15.3
		Q/ET Calibration			
MOD 16A2	0.92	< 0.001	10.7	0.72	12.3
GLEAM 3.2a	0.97	< 0.001	7.4	0.88	7.9
GLEAM 3.2b	0.97	< 0.001	5.9	0.93	6.9
Average	0.95		8.0	0.84	9.0

R^2: coefficient of determination, sig.: significance level, PBIAS: percent bias, KGE: Kling-Gupta Efficiency, RMSE: root mean squared error in mm/month, Q Calibration: mHM calibrated using observed discharge, Q/ET Calibration: mHM calibrated using observed discharge and actual evapotranspiration.

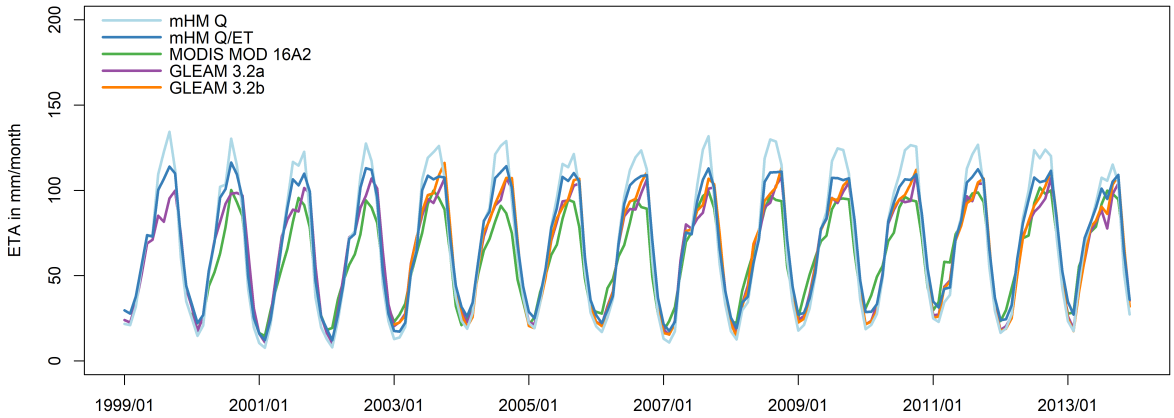

Figure 6.9: Monthly actual evapotranspiration validation averaged over study basins.

Figure 6.10 shows monthly mean raster comparison maps of model ETA against each of the validation datasets. The maps were produced by first resampling the remote sensing data to the mHM resolution, aggregating daily to monthly data and then subtracting mHM from remote sensing results for the period of data availability. Results indicate that mHM Q calibration-based ETA simulations and MODIS ETA compare well, especially in the central part of the

study area, but results towards the very north of the Black Volta and White Volta basins show some overestimation. ETA is underestimated by mHM in the south of the Comoé and Black Volta basins. In the southern Ankobra, Pra and Ayensu basins, mHM did not manage to capture the ETA amounts of the MODIS data well, showing strong underestimations. While comparison against GLEAM 3.2a also performs less well in the south, results are altogether better than for MODIS with lower over- and underestimations. Comparisons against GLEAM 3.2b also show good correspondence, except in the very south. Spatial distribution of the Q/ET calibration results is similar, albeit with slightly lower over- and underestimations.

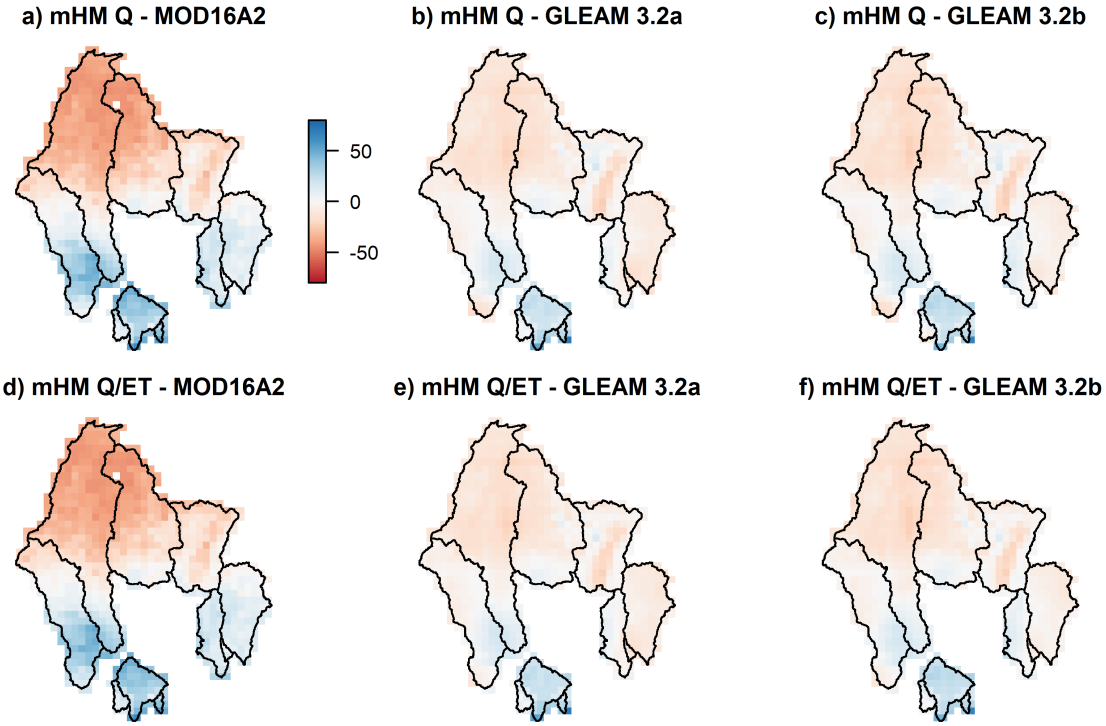

Figure 6.10: Actual evapotranspiration comparison maps. Shown is the mean monthly deviation of mHM model results from remote sensing results in mm. Positive values signify an underestimation of mHM and negative values an overestimation.

Both mHM calibration approaches capture the observed soil moisture anomalies well (Figure 6.11), reaching an R^2 of 0.96, with a slightly better KGE performance of the Q/ET calibration (0.75) than the Q calibration (0.70). In both cases, mHM results show overestimation during the dry seasons and the Q/ET calibration shows an overall greater amplitude. Spatial correlations are high, as evident in Figure 6.12, except for weaker correlations in the south.

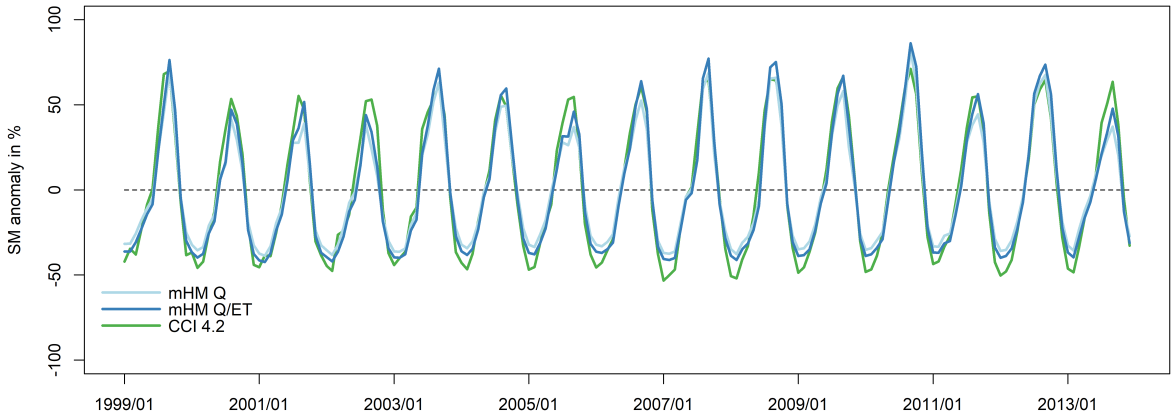

Figure 6.11: Monthly soil moisture anomaly validation averaged over study basins.

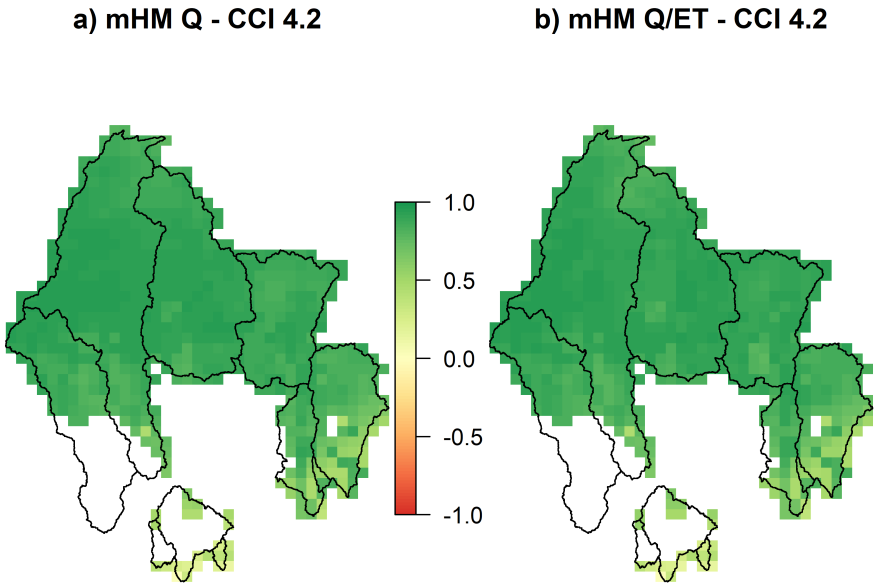

Figure 6.12: mHM soil moisture correlation against ESA-CCI 4.2. (a) mHM discharge calibration; (b) mHM discharge and actual evapotranspiration calibration. White color = no data.

The simulated total water storage anomaly was evaluated against GRACE estimates. Results are shown in Figures 6.13 and 6.14. Both model calibrations fit the ΔTWS estimation observed by GRACE well, but show a phase shift of approximately one month. The calibration approaches score a very good R^2 of 0.78 for Q and 0.87 for Q/ET against GRACE retrievals. KGE is also very good at 0.88 and 0.85, while the root mean squared error is similar at 26.2 and 24.9 mm, respectively. Starting from 2008 onwards, a slight positive trend is established.

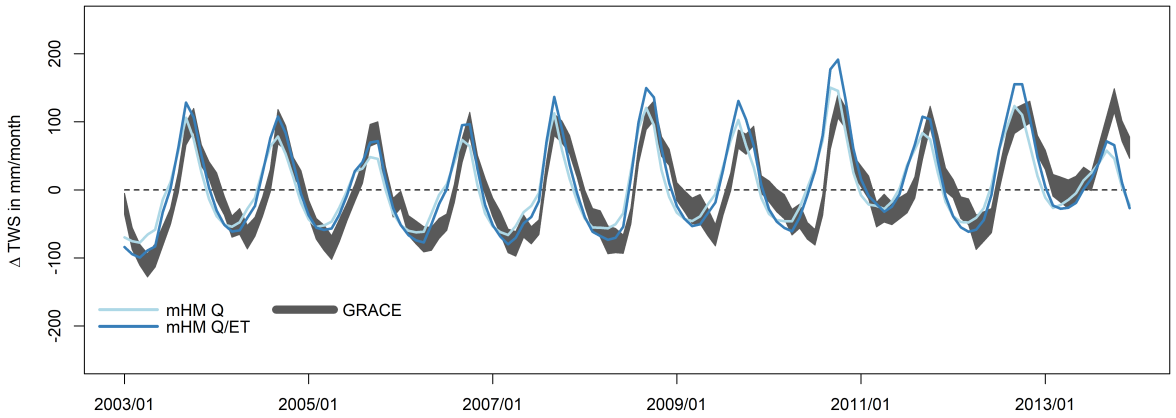

Figure 6.13: Mean monthly total water storage anomaly validation averaged over study basins.

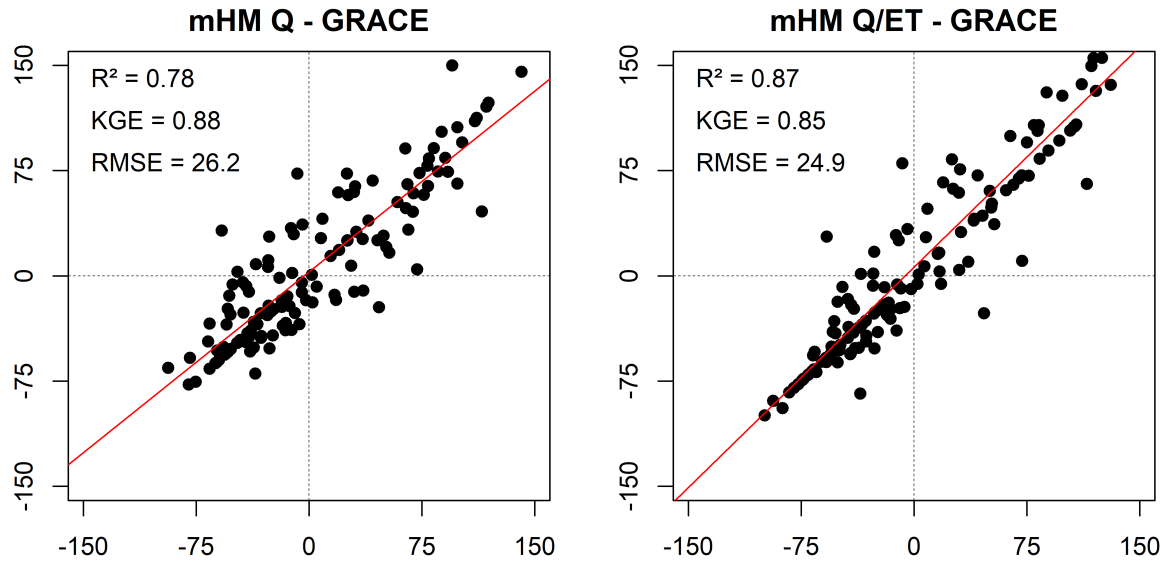

Figure 6.14: Linear regression of mean monthly mHM and GRACE total water storage anomaly averaged over study basins.

6.3.4 Evaluation of mHM-MPR for Transferability across Scales

After successful calibration of both Q and Q/ET models at a resolution of 26 km, the performance of the Multiscale Parameter Regionalization was evaluated by re-running the models under a different L1 output resolution with the calibrated optimal parameter sets. Results, depicted in Table 6.3, show the discharge simulation to remain stable during runs in finer resolutions of 13 and 6.5 km. While the results for both models remain identical for the calibrations (Q: 0.53, Q/ET: 0.49), validation results increase slightly from 0.23 to 0.28 (Q) and 0.13 to 0.17 (Q/ET). The number of modeled grid cells increases from 830 (26 km) to 3320 (13 km) and 13,280 (6.5 km), while the area per grid cell decreases from 676 to 169 and 42.3 km^2. These results highlight the computational efficiency of the model, which can be calibrated at a coarse resolution and later run at a more resource-demanding, finer, resolution without introducing further bias.

Table 6.3: mHM multiscale parameter regionalization results, pixel count, and pixel size comparison for selected resolutions.

	Spatial Resolution		
	26 km	13 km	6.5 km
	Mean KGE Discharge		
Calibration Q	0.53	0.53	0.53
Calibration Q/ET	0.49	0.49	0.49
Validation Q	0.23	0.26	0.28
Validation Q/ET	0.13	0.15	0.17
	Model Domain		
Pixel Count	830	3320	13,280
Pixel Area	676 km^2	169 km^2	42.3 km^2

Q: mHM calibrated using observed discharge, Q/ET: mHM calibrated using observed discharge and actual evapotranspiration; Mean KGE: mean Kling-Gupta Efficiency of all discharge gauges; discharge is averaged over study basins.

6.4 Discussion

6.4.1 Initial Model Runs and Discharge Calibration and Validation

Results of the twelve exploratory model runs have shown ambiguity, depending on the combination of precipitation input and actual evapotranspiration calculation method. In our case, the use of ETA data with aspect correction based on the 500 m DEM was rejected outright, as it performed quite poorly, the exception being ETA and ΔTWS dynamics. While best uncalibrated streamflow results were achieved with the LAI correction method, it did not manage to sufficiently capture either ETP, nor ETA dynamics or amounts. The distribution of LAI classes has been dismissed as a factor, since selected LAI classes accurately depict a vegetation shift in line with the changing precipitation patterns from south to north. This leads us to question the quality of the ETP product used. Due to the period modeled, only the mostly reanalysis-based GLEAM 3.1a dataset was available, which uses ERA-Interim radiation and air temperature-, as well as MSWEP precipitation inputs to calculate ETP (Martens et al. 2017). The Hargreaves-Samani method produced the best ETP simulations. This method has also been proven to produce robust ETP estimations for the region in SWAT simulations and is preferred when data availability is limited or the quality in doubt (Droogers and Allen 2002; Poméon et al. 2018). Out of all precipitation products, CMORPH performed best in combination with the Hargreaves-Samani method. Due to equifinality, when calibrated, other datasets might perform similarly well. This has also been observed by Qi et al. (2016) and Thiemig et al. (2013), who both show product-specific model calibration to improve model results under various remote sensing precipitation inputs. Therefore, a multi-parameter validation is paramount to confirming the model robustness.

Calibration of the CMORPH and Hargreaves-Samani model using both Q and Q/ET as ob-

served variables has shown the model to perform well for the region, leading to acceptable discharge simulations. However, some of the more northern stations in the Black and White Volta river basins perform poorly for both calibrations. The same is evident during the validations, where there are additional problems with several southern stations. Since the model performs well for other stations in the immediate vicinity, we attribute this to three main factors: (1) Data quality: Many stations show large gaps in the time-series or other inconsistencies. While obviously questionable data has been removed, performance might still be influenced. Especially in later years, data quality decreases while gaps increase. (2) Anthropogenic influences: mHM currently does not simulate reservoirs or water abstraction. Several reservoirs exist in the region, such as the Nangbeto reservoir in the southern Mono basin, upstream of the southernmost station, or the Bagre reservoir in the upper White Volta basin. Additionally, there are a multitude of smaller dams scattered mostly across the northern parts of the Black and White Volta basins in Burkina Faso (Lehner et al. 2011). (3) Calibration and validation periods: Due to having insufficient overlapping discharge timeseries, it was decided to run the model for the whole period and to use the first two thirds of the discharge records for calibration and the last third for validation. This approach allows the modeler to use all discharge records, even without overlap, and is opportune in basins with severely limited data availability (Poméon et al. 2018; Abbaspour et al. 2015). In this approach, it can occur that data for several stations in a given basin is available for approximately the same period, while fewer stations have data for a different period. As the stations are equally weighted during optimization, this might negatively influence the results of the stations which are in the minority, especially under climatic anomalies.

6.4.2 Multivariate Validation and Scale Transferability of the mHM-MPR Scheme

In addition to discharge, actual evapotranspiration, soil moisture and total water storage anomalies were also validated. This allows the modeler to assess the robustness of the model (Poméon et al. 2018; López et al. 2017). According to the principle of equifinality, multiple parameter combinations lead to the same model result. Multi-parameter validation allows the modeler to check for inconsistencies or unrealistic behavior in the model and to rectify problems. This furthermore increases the model's predictive capability, as confidence in the prediction of uncalibrated, but validated, variables is increased. During model calibration, it became obvious that if calibrated with default parameter ranges and discharge only, the model tends to remove surplus water into the atmosphere, leading to elevated ETA results. This was circumvented by restricting evapotranspiration parameters and opening infiltration parameters. The final result shows an overall good representation of ETA, although with some positive bias. Interestingly, when calibrated in combination with discharge and ETA, no such adjustment was necessary, and the model was able to find a sufficient solution within the default parameter ranges.

While the streamflow performance decreased slightly for this method, ETA performance increased markedly. Both models perform better against GLEAM ETA (a and b) than against MODIS data; this is especially evident in the southern Comoé, as well as Ankobra, Pra and Ayensu river basins. Generally, MODIS estimates follow a similar dynamic but are lower than GLEAM estimates. MODIS data has also been reported to underpredict ETA in parts of South

Africa (Jovanovic et al. 2015; Sun et al. 2012). GLEAM estimates show a single peak during the rainy season, while mHM and MODIS estimates tend to exhibit a longer period of maximum evapotranspiration for most years. Generally, mHM overestimates ETA in the north and underestimates in the south. This is more pronounced when compared to MODIS than to GLEAM data. Soil moisture anomalies are captured very well by both Q and Q/ET calibrations, with some slight overestimations during the wet seasons from 2006 to 2010. Other wet seasons are slightly underestimated. It is apparent that the model simulates a smaller loss of soil moisture in the dry season than CCI estimates. This might be because the upper soil layer simulated in mHM is 30 cm deep and will dry out significantly slower than the top few centimeters observed by satellites. While further horizons may be implemented in the model, it remains questionable whether model performance can be increased, since none of the soil physical parameters governing infiltration are available for the top few centimeters in the region.

Both Q and Q/ET model calibrations show a generally good agreement with GRACE total water storage estimates. mHM results appear to be slightly shifted by approximately one month. The same phenomenon has also been observed for other model simulations of the region (Poméon et al. 2018; Grippa et al. 2011; Ndehedehe et al. 2016). The ΔTWS amplitude of the Q/ET calibration is greater than the Q calibration for all years. This can be attributed to the generally lower ETA estimates, leading to higher groundwater recharge in wet years and subsequently higher water storages during the wet seasons.

From 2008 to 2010, mHM tends to produce higher estimates than GRACE. Grippa et al. (2011) come to the conclusion that incorrectly modeled evapotranspiration during the dry season may negatively influence TWS results. After comparing model and GRACE results for West Africa, Boone et al. (2009) conclude that differences in amplitudes are due to either inaccuracies in the model precipitation forcing, deficiencies in the model structure, or the overestimation of the storage anomalies by GRACE. Ndehedehe et al. (2016) hypothesize that land surface processes may be influenced anthropogenically, which is not sufficiently reflected in the model. Furthermore, they state that a lack of observed data for the region might lead to uncertain model predictions of soil moisture, and subsequently questionable TWS solutions. Some studies also indicate a clear positive trend towards higher water storage in the entire Volta basin from 2007 onwards, which is attributed mainly to an increase in precipitation (Forootan et al. 2014; Rateb et al. 2017; Hassan and Jin 2016). This is only partly observed in our study. In Poméon et al. (2018), we conclude that the strong signal of Lake Volta may mask TWS solutions. When the lake signal was removed during GRACE preprocessing, a positive trend was less obvious. Similar findings have been reported by Ndehedehe et al. (2016), who observed a decline in TWS between 2007 and 2011 after removing the signal. Since only the upper Volta basins upstream of the lake were modeled in mHM, a TWS trend is also less evident.

In the context of multivariate calibration and validation, mHM has previously been shown to perform well when calibrated with SM and TWS (Rakovec et al. 2016b; Rakovec et al. 2016a). However, applications have been limited to data-rich regions, e.g., Europe. Here we are considering the use of multivariate calibration and validation in a data-scarce domain to offset the poor availability of *in situ* data. Results show that using this approach, mHM produces reasonably realistic predictions even in these environments. It should however be considered that remote

sensing estimates may themselves contain significant bias and generally should be validated for the specific region of interest. This is especially relevant for evapotranspiration products, where models are used to generate estimates based on satellite observations. In data-scarce regions, this proves especially challenging, and was neither conducted in this study, nor are the authors familiar with validations of the specific products for the study area. Studies conducted in Germany (Zink et al. 2018) and China (Fang et al. 2017) have also shown land surface temperature inputs to further constrain hydrologic model parameters.

Multiscale parameter regionalization has proven to perform well for discharge calibration and validation results under both Q and Q/ET calibrations. Calibration results are almost identical for all three resolutions, and validation results increase slightly for the finer resolutions. MPR seems to be a promising approach for reducing the runtimes of large-scale hydrologic models by calibrating at a computationally efficient coarse resolution and running the model in a significantly finer resolution. Scaling results obtained here are consistent with previous findings for EU and US regions (Samaniego et al. 2010; Kumar et al. 2013b; Kumar et al. 2013a; Rakovec et al. 2016b).

6.5 Conclusion

In this study, we have proven mHM to perform reasonably well and to produce robust results for a (sparsely-gauged) study area in southern West Africa. When considering the lack of *in situ* measurements in the region, the use of remote sensing data becomes a necessity to produce meaningful hydrologic predictions. The model has been developed for the express purpose of integrating these datasets. Running the uncalibrated model for twelve combinations of precipitation and evapotranspiration inputs and calculation methods has allowed us to rapidly screen the performance regarding each combination. This proved to be helpful in choosing the final model combination for further investigation. The model was first calibrated using discharge as the observed variable and then, to further reduce model uncertainty, calibrated using discharge and actual evapotranspiration data. Equifinality dictates that the same result can be achieved with an infinite combination of model parameters. Since some of these combinations may be unrealistic, the model may predict variables that were not included in the calibration poorly. It therefore becomes necessary to either include in the calibration, or to separately validate, all further variables of interest. This was performed using actual evapotranspiration, soil moisture anomaly and total water storage anomaly data. It has been shown that mHM can reasonably well predict discharge in the region. Further validation has confirmed actual evapotranspiration as well as soil moisture and total water storage anomaly predictions to perform well when compared to remote sensing products. When model parameters are further constrained using a calibration based on streamflow and actual evapotranspiration, the performance of the evapotranspiration simulation is vastly improved, while soil moisture and total water storage anomaly predictions remain stable. Discharge simulations, however, decrease in performance. Using Multiscale Parameter Regionalization, the calibrated model can then be run at a finer resolution while results remain stable. This approach is computationally efficient, as a calibration at a fine scale would take significantly longer to complete.

Further work is necessary to fully investigate the model uncertainty under different remote

sensing inputs, as well as the performance of product-specific calibration for each of the considered precipitation datasets. Also, the contribution of land surface temperature to further constrain model parameters should be assessed. Since the remote sensing products have not been validated due to the lack of *in situ* data in the region, it would be interesting to assess the uncertainty between different datasets. This is especially relevant for datasets which rely on models to generate estimates. As a SWAT simulation has already been performed for the study area, a thorough model comparison will be conducted to assess the strengths of the individual models in the future. Due to the climate-data friendly nature of the model, climate scenarios can also be easily implemented.

Our modeling approach allows the modeler to build a robust hydrologic model relatively quickly by using only freely and easily available remote sensing and reanalysis data and software. This is especially interesting for predictions in sparsely gauged basins such as in West Africa, where the availability of data continues to decline. However, remote sensing should ideally be used to complement, and not replace, *in situ* measurements. We therefore firmly believe that there is a necessity for further investment into observation networks in West Africa.

Acknowledgements: We thank Jürgen Kusche and Anne Springer of the Institute of Geodesy and Geoinformation, University Bonn, for providing the GRACE solutions. We are furthermore grateful to Christophe Peugeot and the AMMA-Catch project, as well as to the Global Runoff Data Centre in 56068 Koblenz, Germany for providing discharge data. The AMMA-CATCH regional observing system was set up thanks to an incentive funding of the French Ministry of Research that allowed pooling together various preexisting small-scale observing setups. The continuity and long term perenity of the measurements have been made possible by undisrupted IRD funding since 1990 and by a continuous CNRS-INSU funding since 2005.

Author Contributions: Conceptualization, T.P., B.D and R.K.; methodology, T.P., B.D. and R.K.; software, T.P and R.K; formal analysis, T.P., B.D. and R.K.; investigation, T.P.; writing—original draft preparation, T.P.; writing—review and editing, T.P., B.D. and R.K.; visualization, T.P.; supervision, B.D.; project administration, B.D.; funding acquisition, B.D.

7 Comparison of SWAT and mHM Results

7.1 Introduction, Workflow and Methods

In this chapter, key simulation results of the mHM and SWAT model runs will be compared with each other and, where available, against remote sensing or simulation results of global model runs. Actual evapotranspiration, soil moisture anomaly, total water storage anomaly, groundwater recharge and water yield were identified as the most relevant variables. Furthermore, discharge simulations were compared for three exemplary gauges. Only the best-performing model setups will be compared. This comprises the mHM calibration using discharge and actual evapotranspiration (mHM Q/ET) at a 6.5 km resolution and the SWAT global calibration results (SWAT v1). Differing spatial domains and model output formats of the simulation experiments require results to be processed before evaluation. First, an area with model overlap was chosen for conducting the comparison. In this case, the area comprising the Comoé, Ouémé, Couffo, Mono, Pra, Ayensu, Ankobra and Volta river basins offers the greatest possible overlap. mHM produces data in an easily processable netCDF raster format. This is not the case for SWAT, where results for each subbasin are written in a datafile. SWAT results were processed by writing the results of each timestep from the datafile to a shapefile containing all subbasins. Each timestep was then rasterized and timesteps were aggregated into a netCDF raster-brick for further processing. Both model and remote sensing spatial data were projected to the same resolution and subset to an identical time period. While the entire river basins were modeled in SWAT, only the basin areas up to the most downstream discharge gauge were simulated in mHM. Therefore, SWAT results were masked using the mHM model outline, so as not to bias the results of the statistical analysis. The time periods of the comparisons were chosen based on data availability.

A special case presents itself in the actual evapotranspiration comparison, where data was extracted not only for the entire region, but also for three specific windows (south: 4.5-7.5°N; central: 7.5-10.5°N; north: 10.5-15°N) based on the observed land cover, to better account for the variability within the region (see also Figure 7.1). Actual evapotranspiration was further compared against MODIS MOD16A2 and GLEAM 3.2b data, soil water anomaly against ESA CCI 4.2 data, and total water storage anomaly against GRACE. For more information about the datasets used, see chapter 5.2.3 and chapter 6.2.3. New data introduced in this chapter are ΔTWS estimates compiled from five global models for the region. Similar to GRACE, data processing (extraction for the region, removal of lake and tidal signals) was conducted by the Institute of Geodesy of the University Bonn. The five models shall be briefly described:

Global Land Data Assimilation System Models Four land surface models of the Global Land Data Assimilation System (GLDAS, Fang et al. (2009)) were used for ΔTWS comparison, namely

the Community Land Model Version 4 (CLM 4), the MOSAIC land surface model (MOSAIC), the Variable Infiltration Capacity Macroscale Hydrological Model (VIC) and the NOAH land surface model (NOAH). CLM 4 is an improvement of the previous CLM 3.5 model. It models surface energy fluxes, as well as hydrologic and biogechemical cycles. The hydrologic model is based on the Richards equation governing soil water movement and has been improved from CLM 3.5 using a revised numerical solution. Furthermore, the ground evaporation parameterization has been improved (Lawrence et al. 2011). The MOSAIC model computes energy and water fluxes from the land surface, which are area-averaged. The name derives from the "mosaic" strategy employed by the model, where surface grid cells are divided into homogeneous mosaic tiles with a unique vegetation or soil type, thereby accounting for the subgrid heterogeneity of surface characteristics (Koster and Suarez 1996). VIC is a macroscale semi-distributed conceptual hydrologic model. Both water and energy balances are calculated within each grid cell. It includes subgrid variability in vegetation classes and soil moisture storage capacity, baseflow as a nonlinear recession and surface topography to derive realistic precipitation and temperature lapse rates (Gao et al. 2009). Lastly, NOAH is a mesoscale global land model which has been continuously improved since its inception in 1996. Included are evapotranspiration, soil, canopy and surface hydrology models as well as the parametrization of snow and sea ice (Ek et al. 2003).

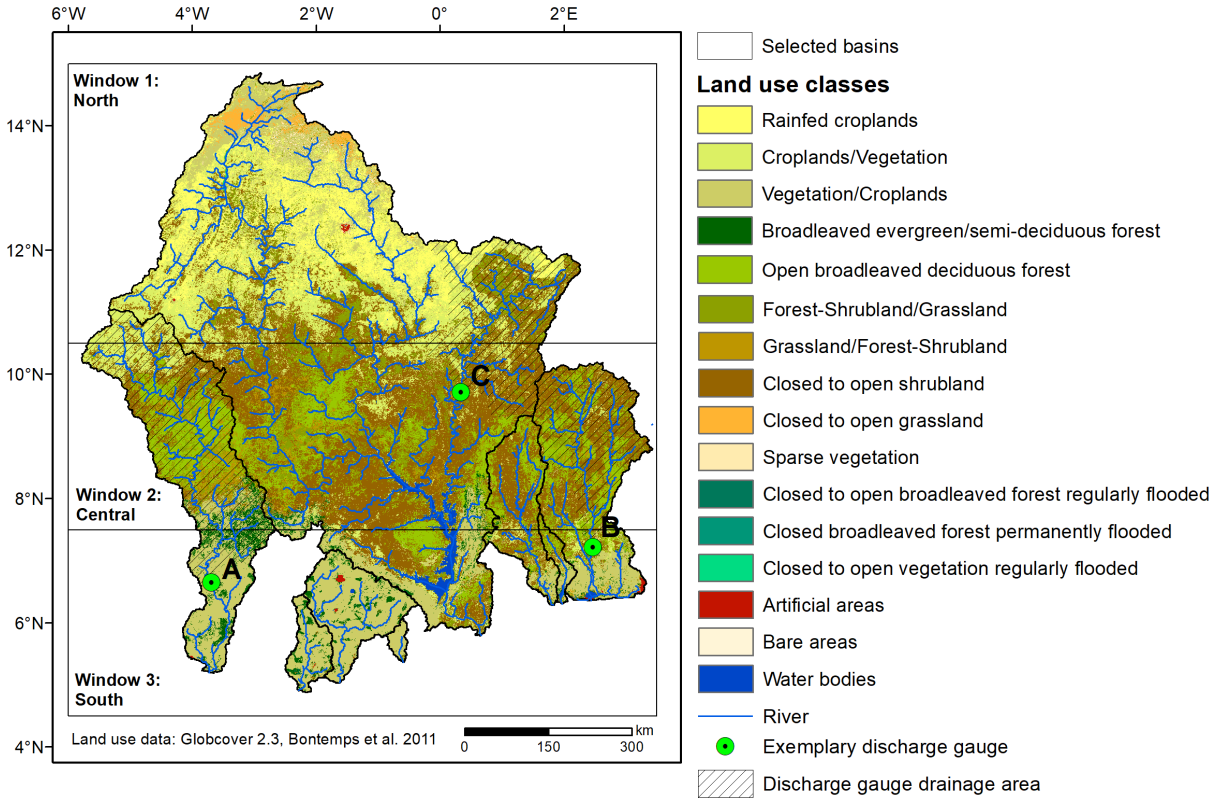

Figure 7.1: Spatial domain of model comparison, land use and extraction windows. A: Comoé - Aniassué; B: Ouémé - Ahlan; C: Oti - Saboba.

Water GAP Global Hydrology Model The Water GAP Global Hydrology Model (WGHM) is a submodel of the global WaterGAP 2 water use and availability model. WGHM is a conceptual

model simulating flows between continental water storage compartments (with the exception of glaciers) and accounts for anthropogenic influences, such as pumping and irrigation. Water is compartmentalized into canopy, snow, soil, groundwater, lake, reservoir, wetland and river storage. Between one and three free parameters are calibrated against mean annual river discharge at over 1200 gauging stations (Döll et al. 2003; Döll et al. 2012).

7.2 Results and Discussion

7.2.1 Actual Evapotranspiration

Long-term monthly actual evapotranspiration results for the month of January are presented in Figure 7.2.

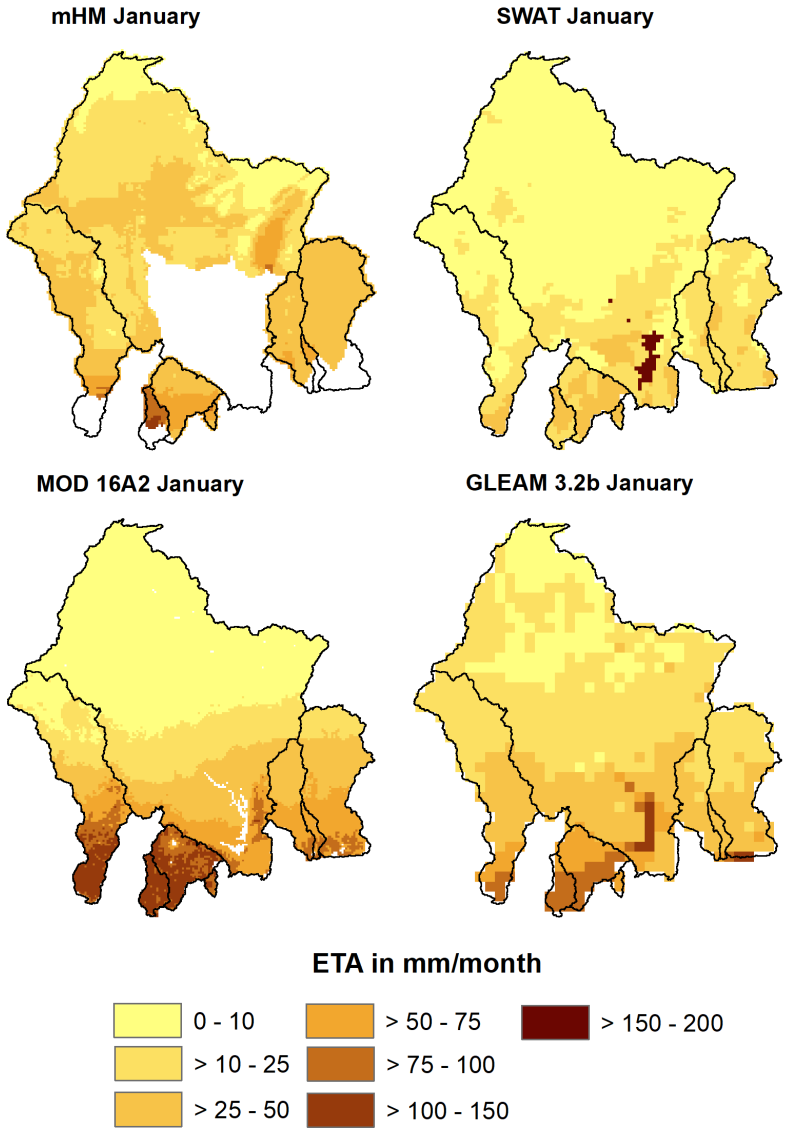

Figure 7.2: Long-term (2003-2013) January actual evapotranspiration. mHM and SWAT results compared against MODIS MOD 16A2 and GLEAM 3.2b data. White areas signify missing data.

mHM results tend to be elevated in the north compared to SWAT, MODIS and GLEAM data, while in the south, patterns of mHM and GLEAM resemble each other more closely. Contrarily, SWAT predicts low ETA in the southern areas and MODIS estimates high ETA in the southwest, which is observed by no other evaluated dataset. Clear differences can be remarked for the long-term July evapotranspiration (Figure 7.3). Here, mHM and SWAT perform very similarly both in the north, with ETA rates in the 100-150 mm range, and the south, with rates between 75 and 100 mm. MODIS and GLEAM show quite different patterns. While GLEAM predicts rather constant ETA values of between 75 and 100 mm for the entire area except the lake Volta region, MODIS predicts highest ETA rates in the south (100-150 mm), which steadily decrease towards the north.

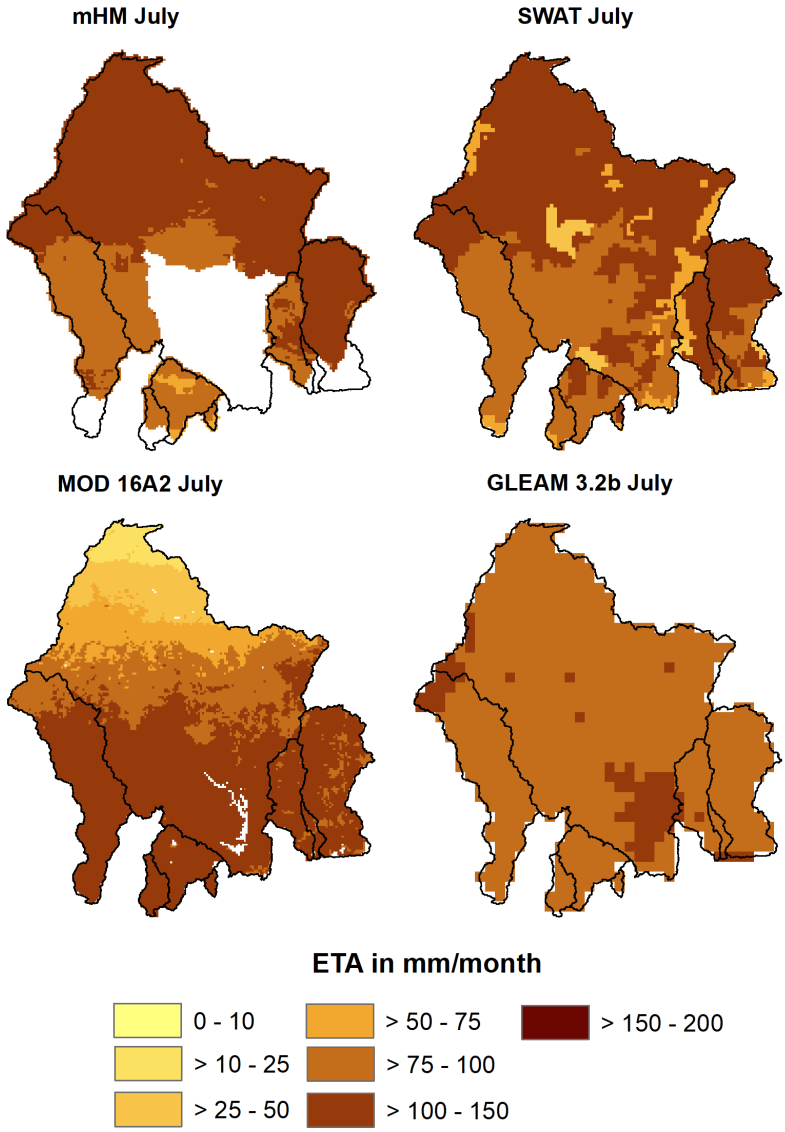

Figure 7.3: Long-term (2003-2013) July actual evapotranspiration. mHM and SWAT results compared against MODIS MOD 16A2 and GLEAM 3.2b data. White areas signify missing data.

Decreasing precipitation amounts from south to north strongly influence ETA estimates. Due to the applied calculation methods, long-term July ETA predictions differ between the applied models, MODIS and GLEAM data. Therefore, it was decided to generate monthly ETA time-series for the entire region, as well as three extraction windows (north, central and south). Results are presented in Figure 7.4 and Table 7.1.

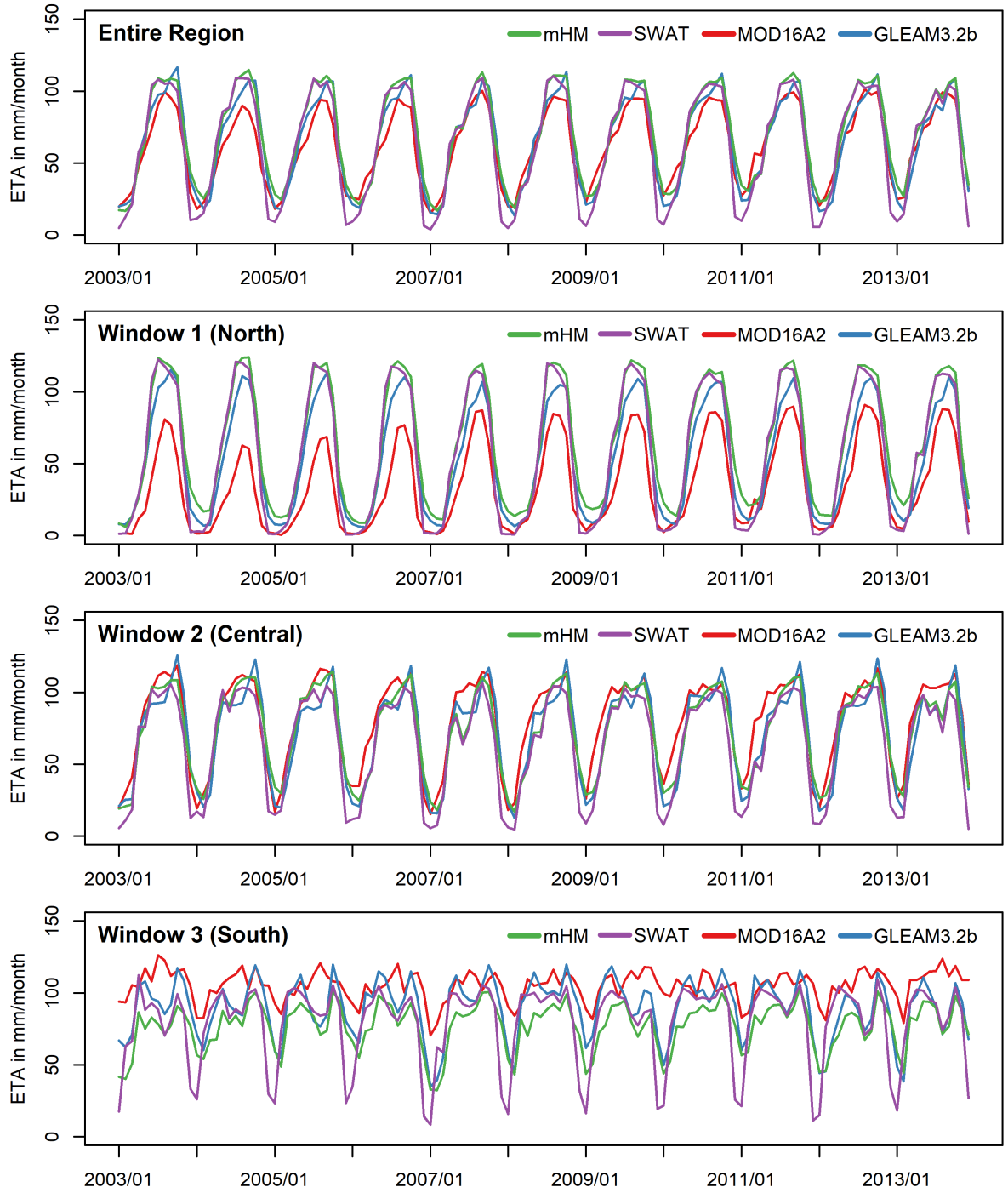

Figure 7.4: Monthly actual evapotranspiration. mHM and SWAT results compared against MODIS MOD 16A2 and GLEAM 3.2b data.

Monthly ETA patterns averaged over the entire region are similar for all considered estimates. However, MODIS peak rates are the lowest of all products, while SWAT minimum rates are

below those estimated by other products. Agreement between the datasets is generally good, with KGE values ranging from 0.74 (mHM-MODIS) to 0.93 (mHM-GLEAM), the exception being the comparison of SWAT to MODIS, which performs less well with a KGE of 0.58. RMSE ranges from 7.3 (mHM-GLEAM) to 14.2 mm (SWAT-MODIS).

Table 7.1: Monthly actual evapotranspiration statistics.

		mHM to		SWAT to	
	mHM-SWAT	MOD16A2	GLEAM 3.2b	MOD16A2	GLEAM 3.2b
			Entire Region		
R^2	0.93	0.92	0.97	0.92	0.91
sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
KGE	0.82	0.74	0.93	0.58	0.82
PBIAS	11.0	13.7	6.9	2.5	-3.6
RMSE	12.6	13.2	7.3	14.3	12.2
			Window 1 (North)		
R^2	0.96	0.88	0.98	0.86	0.95
sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
KGE	0.84	0.06	0.78	0.16	0.79
PBIAS	12.8	86.0	20.9	64.8	7.1
RMSE	12.2	34.3	12.9	31.3	13.1
			Window 2 (Central)		
R^2	0.92	0.88	0.93	0.91	0.85
sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
KGE	0.81	0.90	0.93	0.77	0.85
PBIAS	12.9	-7.6	1.8	-18.2	-9.8
RMSE	13.2	12.4	8.8	18.0	15.2
			Window 3 (South)		
R^2	0.47	0.44	0.82	0.20	0.53
sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
KGE	0.48	0.30	0.75	-0.74	0.51
PBIAS	-2.4	-26.4	-14.0	-24.6	-11.9
RMSE	20.4	30.3	15.2	35.8	21.8

R^2: coefficient of determination; sig: significance level; KGE: Kling-Gupta Efficiency; PBIAS: Percent bias; RMSE: root mean squared error in mm/month.

When data is compared for the three extraction windows, results differ. While mHM, SWAT and GLEAM predict similar ETA in the north (Window 1), MODIS predictions fall considerably short, with an RMSE of 34.3 mm against mHM and 31.3 against SWAT simulations. In the central part of the area (Window 2), results are again comparable, but variability between the estimates is high during the peaks of the rainy and dry seasons. GLEAM predicts a peak, which is less pronounced in mHM, SWAT and MODIS results. SWAT again predicts low dry season ETA. In the south (Window 3), differences between the estimates are more pronounced. Here, not only intensities, but also patterns diverge. While mHM and GLEAM display similar patterns, SWAT simulates very low ETA during the dry season, leading to poor agreement

between SWAT and mHM (KGE: 0.48, RMSE: 20.4 mm). MODIS produces a different pattern with high ETA during the dryer periods and does not compare well to mHM (KGE: 0.48, RMSE: 20.4 mm) and SWAT (KGE: -0.74, RMSE: 35.8).

The potential evapotranspiration for both mHM and SWAT was calculated using the Hargreaves equation based on MERRA 2 temperature estimates. As the modeled system is water-, and not energy-limited, actual evapotranspiration is calculated based on the amount of water available for evaporation in the soil (as well as in reservoirs and lakes) and transpiration in the canopy. One major drawback of the SWAT model in tropical regions is a general inability to accurately simulate plant growth, as its plant routine assumes temperature to be the leading driver of plant growth. In a water-limited system, however, water availability is the key growth criterion (Neitsch et al. 2011; Alemayehu et al. 2017; Strauch and Volk 2013). Although a manual crop management routine based on LAI observations was established in SWAT to more closely resemble plant development, simulated plant growth remained underdeveloped during dry seasons, leading to very low LAI values and, consequentially, low estimated transpiration throughout the region. As plants are modeled to grow more realistically during the rainy seasons, no clear underestimations are observed. In mHM, long-term monthly LAI timeseries can be defined for an arbitrary number of classes, circumventing this problem.

While MODIS and GLEAM estimates compare more or less well when averaged over the whole region, they differ strongly for the northern and southern extraction windows. This can be traced back to the calculation method and data used for each product. As evapotranspiration cannot be measured from space, proxies are used. MODIS calculates the potential evapotranspiration by means of a modified Penman-Monteith equation, including soil heat flux calculations and LAI observations. However, climatological data to calculate the equation is compiled from reanalysis datasets and reanalyzed precipitation data is further used to drive an actual evapotranspiration model (Mu et al. 2011). GLEAM, on the other hand, calculates potential evapotranspiration based on the Priestley and Taylor equation using mainly satellite data, with precipitation data derived from a mixture of satellite, gauge and reanalysis products (Martens et al. 2017).

Concerning mHM and SWAT ETA estimates, it has to be stressed again that while both models calculate potential evapotranspiration based on the same equation and input data, mHM has been calibrated using discharge data and GLEAM 3.2a actual evapotranspiration data, while ETA was not calibrated in SWAT.

7.2.2 Soil Moisture Anomaly

Long-term soil moisture anomaly estimates for the month of January are shown in Figure 7.5. Anomalies are calculated locally for each pixel. Patterns of mHM and CCI are comparable, albeit with some differences. For example, CCI estimates soil moisture to reduce most (-50 to -75%) in the central parts of the area, while mHM results show SM anomaly to be strongest towards the north. SWAT shows a very strong SM anomaly of over -75% for the central and northern areas, while results in the very south are comparable to CCI.

Concerning the long-term July SM anomaly presented in Figure 7.6, patterns of all three solutions differ. mHM predicts higher values (25 to 50%) towards the east, while CCI data shows highest anomalies to occur in the north. In the very west, mHM and CCI are comparable.

While SWAT also simulates higher values in the east, the anomaly is stronger than observed by CCI and simulated by mHM. Some regions show negative values, while in the south, the anomaly is higher than indicated by CCI data.

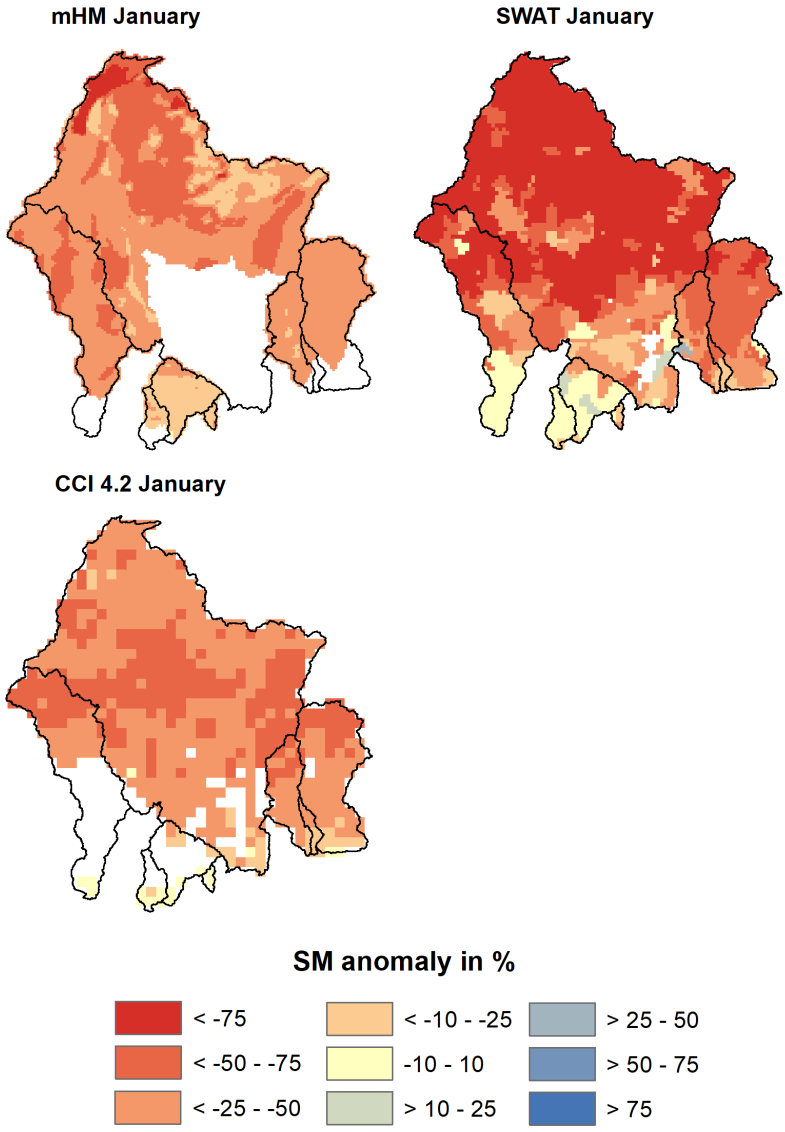

Figure 7.5: Long-term (1999-2013) January soil moisture anomaly in %. mHM and SWAT results compared against ESA CCI 4.2 data. White areas signify missing data.

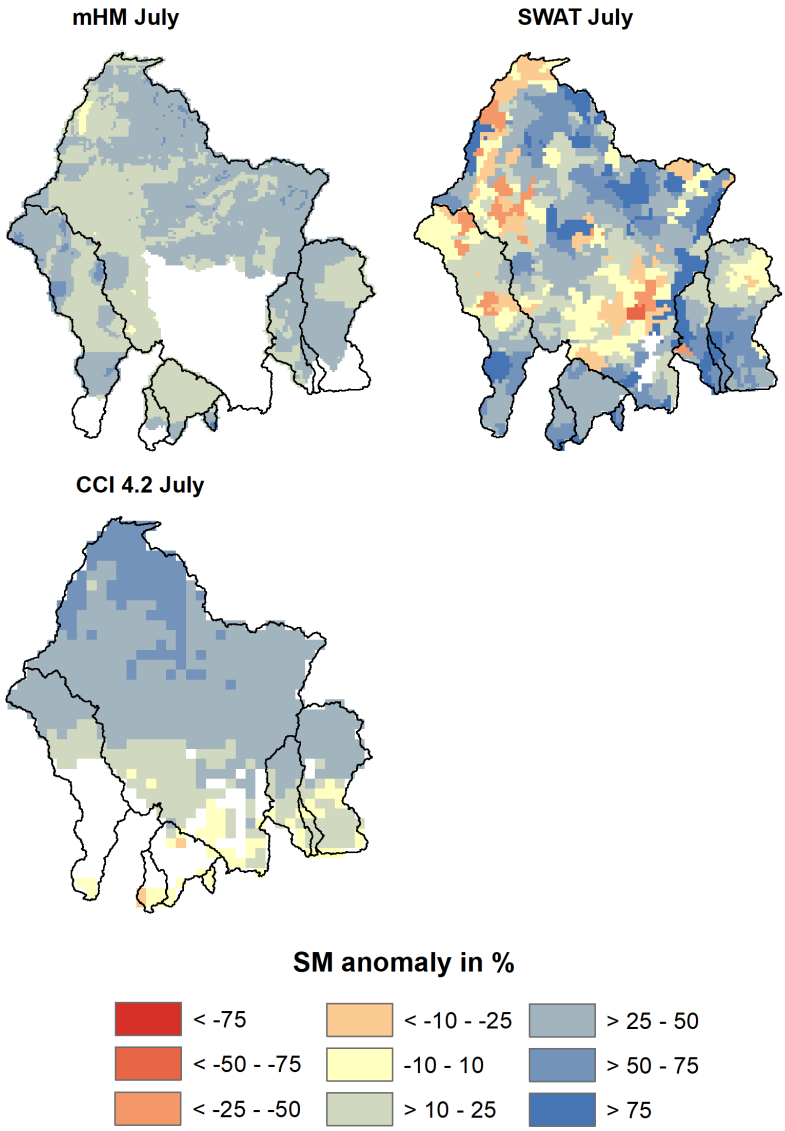

Figure 7.6: Long-term (1999-2013) July soil moisture anomaly in %. mHM and SWAT results compared against ESA CCI 4.2 data. White areas signify missing data.

The timeseries of the monthly soil moisture anomaly averaged over the area is presented in Figure 7.7. Accompanying statistics can be found in Table 7.2. mHM and CCI perform quite similarly for most years, albeit with some overestimation of SM increase in rainy seasons and underestimation of its decrease during the dry seasons by mHM for the years 2003-2004 and 2006-2012. Contrarily, SWAT results show a far stronger amplitude with both a stronger SM depletion during the dry seasons and increased storage during the rainy seasons. Consequentially, while the averaged dynamics of mHM and SWAT simulations are comparable with an R^2 of 0.87, the overall comparison is acceptable with a KGE of 0.60 and RMSE of 31.7%. Compared to CCI data, SWAT simulations also do not perform very well, with a KGE of 0.55 and RMSE of 39.3%. When comparing mHM estimates to CCI data, performance is good, with a KGE of 0.93 and RMSE of 13.2%.

The difference in soil moisture anomaly estimates between the two models cannot be explained by the soil input data used, as HWSD inputs were used in both cases. However, while mHM requires only information on the soil texture and bulk density and calibrates all further parameters, necessary parameters must first be defined in SWAT. As not all parameters are known and need to be derived using literature data or transfer functions, this introduces further uncertainty into the model. Also, the depth of the soil profiles used to calculate SM anomaly varies between mHM (top layer: 0-30 cm), SWAT (entire profile: 0-30 or 100 cm) and CCI (only top 2-4 cm), making a direct comparison difficult. Another factor is the actual SWAT output data, as residual water content of the soil is not taken into account (Milzow et al. 2011), while it is accounted for by both mHM and CCI.

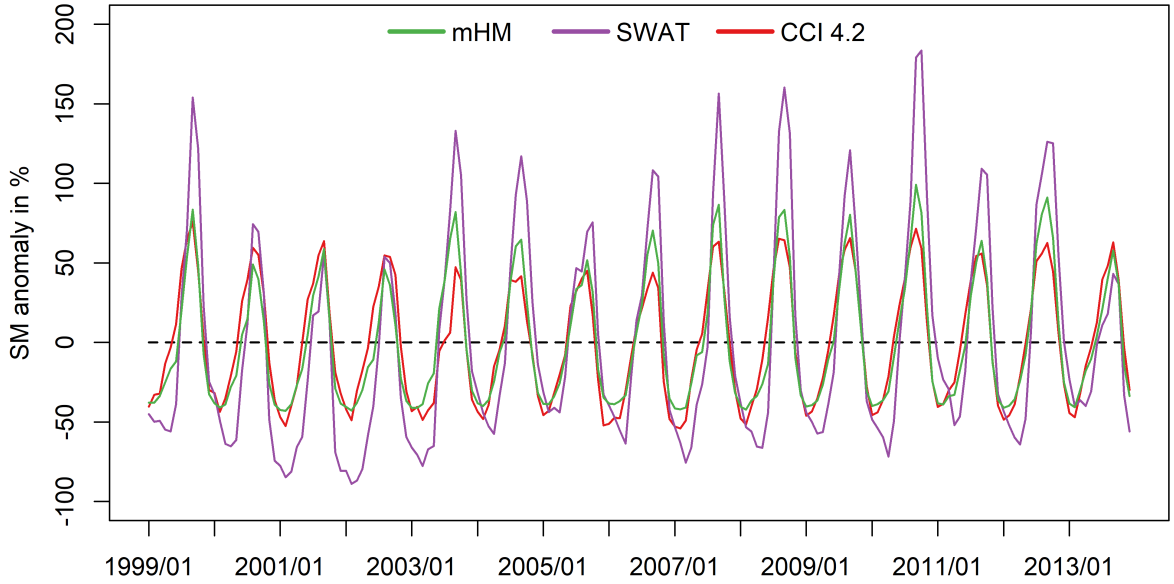

Figure 7.7: Monthly soil moisture anomaly in %. mHM and SWAT results are compared against ESA CCI 4.2 data.

Table 7.2: Monthly soil moisture anomaly statistics.

		CCI 4.2 to	
	mHM-SWAT	mHM	SWAT
R^2	0.87	0.89	0.70
sig.	< 0.001	< 0.001	< 0.001
KGE	0.60	0.93	0.55
RMSE	31.7	13.2	39.3

R^2: coefficient of determination; sig: significance level; KGE: Kling-Gupta Efficiency; RMSE: root mean squared error in %.

7.2.3 Total Water Storage Anomaly

Spatial patterns of total water storage anomaly simulated by both mHM and SWAT compare well to each other, as is shown in Figure 7.8 for long-term January and July ΔTWS. For the

long-term January estimates, some slight disagreements are apparent in the west, where mHM suggests a higher depletion, and the east, where the anomaly is stronger in SWAT simulations. Also, the break between lower and higher depletion in the Volta basin simulated by mHM is less represented in SWAT. When comparing long-term July ΔTWS, spatial patterns resemble each other more closely, the exception being mHM, which simulates higher TWS increases in the south.

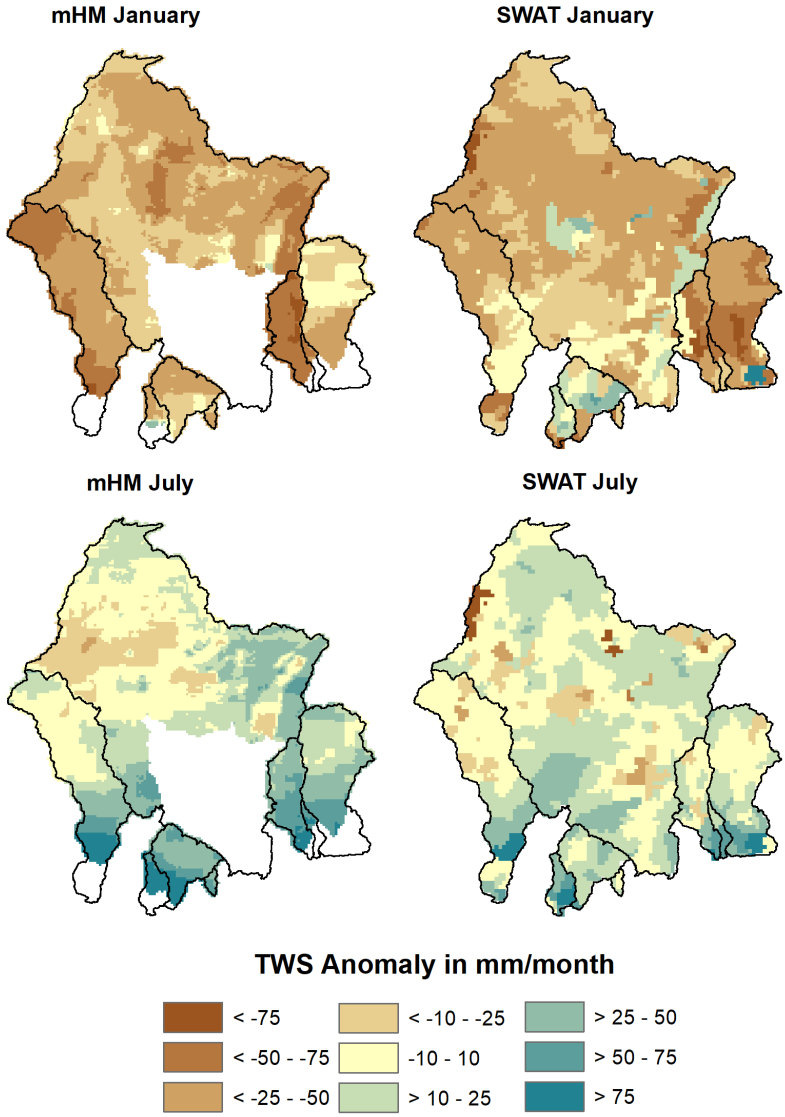

Figure 7.8: Long-term (2003-2013) January and July mHM and SWAT ΔTWS estimates. White areas signify missing data.

Spatially averaged mHM and SWAT ΔTWS results were further compared against data derived from GRACE and five global models. Results are depicted in a Taylor diagram in Figure 7.9. The diagram shows the (normalized) standard deviation, Pearson's r correlation and KGE of the comparison. Concerning the individual performances, circles represent comparisons against mHM and squares comparisons against SWAT. Numbers signify the product used in the statistical comparison (1: GRACE; 2: CLM 4; 3: MOSAIC; 4: NOAH 2.1; 5: VIC; 6: WGHM).

The diamond (⋄) sign stands for the comparison of mHM and SWAT simulation results. Results show that all products are highly correlated with Pearson's r values between 0.85 and 0.95. KGE values are for the most part above acceptable (≥0.5), except for comparisons of mHM and SWAT to CLM 4 model results and the direct comparison of mHM to SWAT simulations.

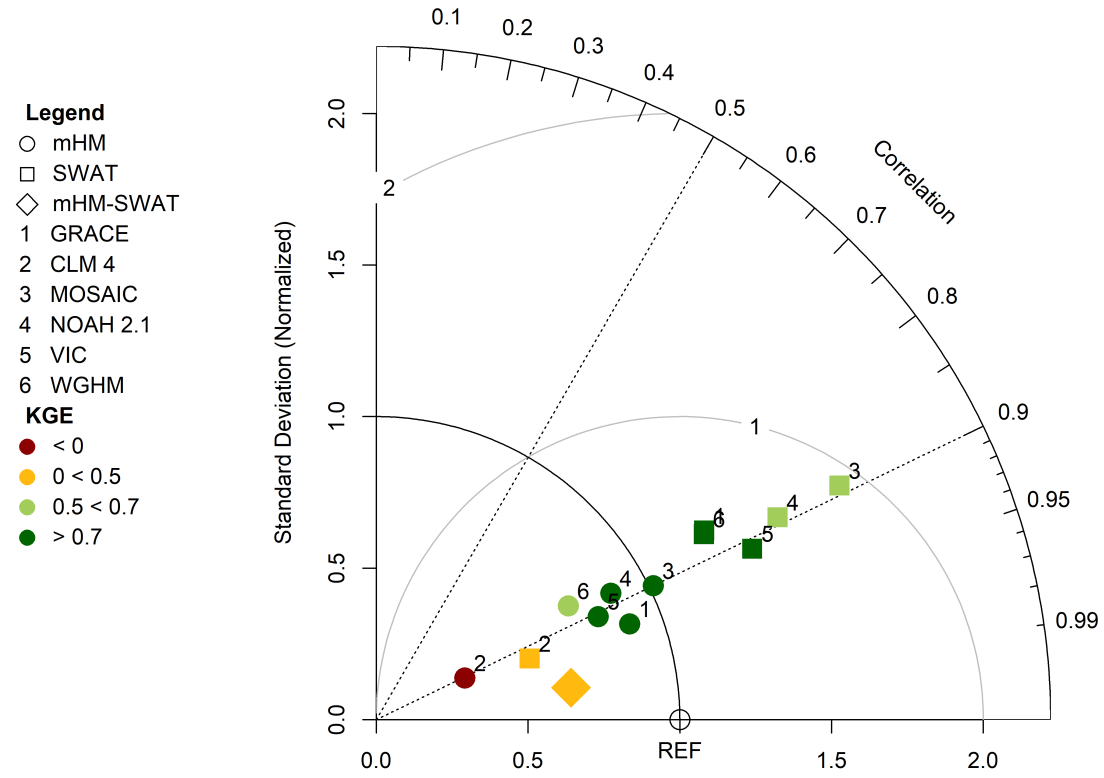

Figure 7.9: mHM and SWAT ΔTWS estimates compared against each other and GRACE, as well as CLM 4, MOSAIC, NOAH 2.1, VIC and WGHM model estimates.

Highest KGE and r results as well as lowest difference in standard deviation compared to mHM are shown for GRACE, MOSAIC, NOAH 2.1 and VIC estimates. GRACE, VIC and WGHM compare best to SWAT.

Area-averaged monthly ΔTWS timeseries of mHM, SWAT, GRACE and an ensemble of global model results are shown in Figure 7.10, with accompanying statistics presented in Table 7.3. It can be observed that global model, mHM and SWAT results appear to be slightly shifted by one half to one month against GRACE. Also, for some years, some of the global models and also mHM predict very high TWS storage anomalies during the wet seasons, which are not observed by GRACE or simulated in SWAT. GRACE furthermore predicts the highest depletion in the dry season out of all products for most years. The amplitude shown by SWAT is considerably lower than simulated by mHM and observed by GRACE. While the dynamics of mHM and SWAT compare well (R^2: 0.97), KGE performance is below acceptable at 0.45. It should be stated, however, that SWAT estimates consistently predict ΔTWS within the uncertainty band of the global model ensemble. As has been shown in Figure 7.9, mHM compares best to GRACE, with an R^2 of 0.87, KGE of 0.85 and RMSE of 24.9 mm/month. Good performance is also observed when compared against MOSAIC (R^2: 0.81, KGE: 0.86), NOAH 2.1 (R^2: 0.77, KGE: 0.79),

and VIC (R^2: 0.82, KGE: 0.73) model results. Agreements of SWAT model results compared to external data are generally lower, with best results reached when compared against WGHM (R^2: 0.76, KGE: 0.77), GRACE (R^2: 0.75, KGE: 0.76) and VIC (R^2: 0.83, KGE: 0.72).

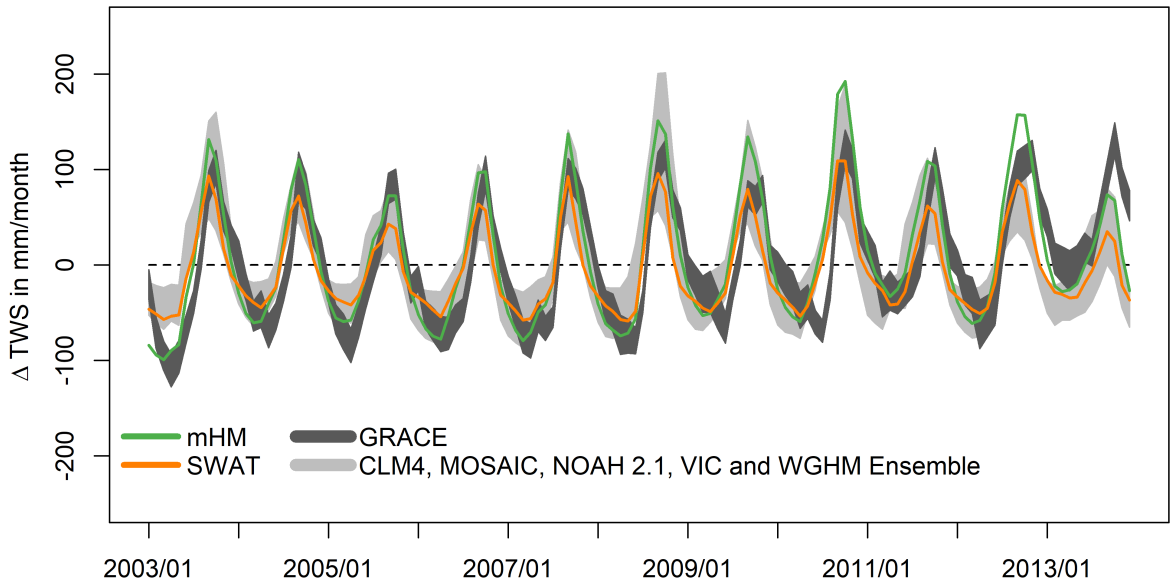

Figure 7.10: Monthly ΔTWS estimates. mHM and SWAT results are compared against GRACE data and against an ensemble of CLM 4, MOSAIC, NOAH 2.1, VIC and WGHM model estimates.

Table 7.3: mHM and SWAT ΔTWS estimates compared against each other and against global datasets.

				mHM to			
	GRACE	CLM 4	MOSAIC	NOAH 2.1	VIC	WGHM	**SWAT**
R^2	0.87	0.82	0.81	0.77	0.82	0.74	0.97
sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
KGE	0.85	-1.10	0.86	0.79	0.73	0.60	0.45
RMSE	24.9	49.6	31.7	33.3	30.6	36.6	26.8
				SWAT to			
	GRACE	CLM 4	MOSAIC	NOAH 2.1	VIC	WGHM	**mHM**
R^2	0.75	0.86	0.80	0.80	0.83	0.76	0.97
sig.	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
KGE	0.76	0.16	0.57	0.66	0.72	0.77	0.64
RMSE	25.2	21.4	37.4	29.7	24.5	24.7	26.8

mHM and SWAT results are compared against GRACE remote sensing and against global model (CLM 4, MOSAIC, NOAH 2.1, VIC and WGHM) ΔTWS estimates. R^2: coefficient of determination; sig: significance level; KGE: Kling-Gupta Efficiency; RMSE: root mean squared error in mm/month.

7.2.4 Groundwater Recharge

As no external groundwater recharge data was available, mHM and SWAT results will be compared only to each other. Figure 7.11 shows long-term January and July recharge rates. Both mHM and SWAT simulate low recharge during the dry seasons in January, with mHM results not exceeding 10 mm/month. SWAT simulations are similar, albeit with higher rates of up to 30 mm/month in the south and in some central areas. Spatial patterns are similar during the rainy season in July, but ranges differ. Both models predict higher recharge in the south and in a band along the eastern Volta basin through the Mono basin in the south. Here, SWAT simulates higher recharge rates than mHM. Also, high rates in the southern Ouémé basin are not simulated by mHM. While SWAT simulates some isolated areas of high recharge in the Volta basin, mHM results show no such anomalies.

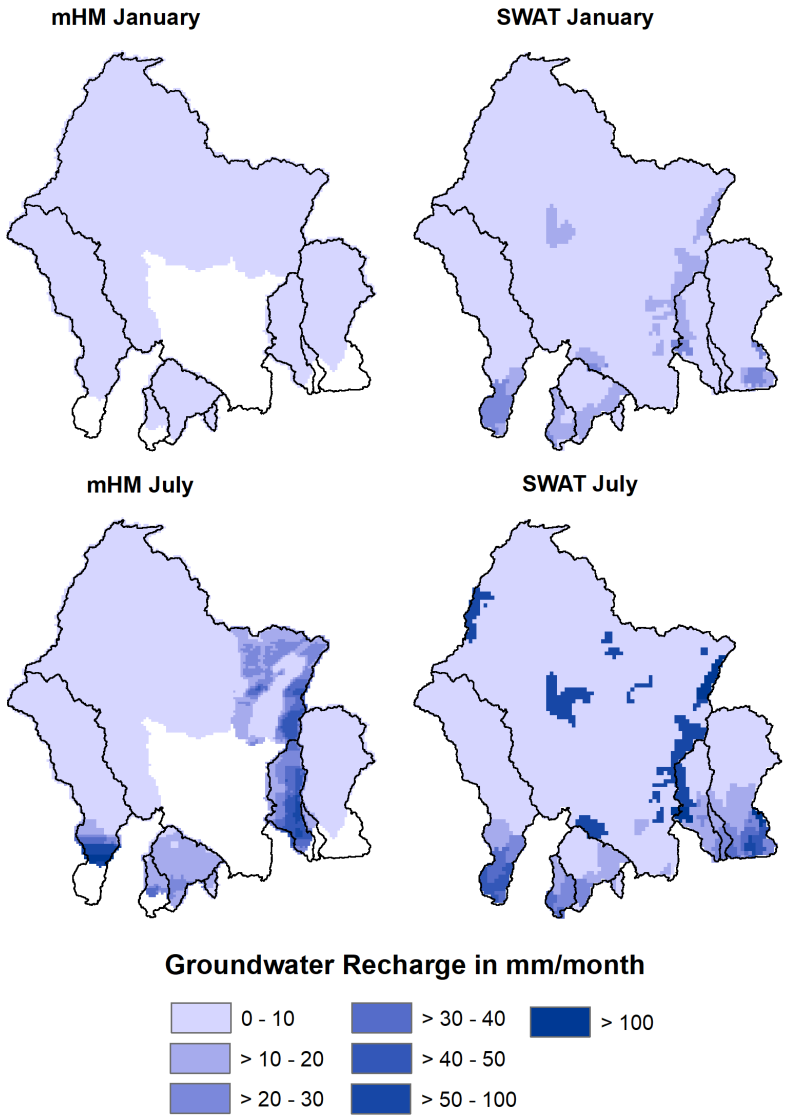

Figure 7.11: Long-term (1999-2013) January and July mHM and SWAT groundwater recharge estimates.

Monthly area-averaged groundwater recharge simulated by mHM and SWAT is displayed in Figure 7.12. Dynamics are similar, albeit with some peculiarities. SWAT, e.g., shows higher and prolonged recharge rates during almost all rainy seasons except in 2010. Higher recharge is also simulated during the dry seasons and results appear shifted by approximately one month in relation to mHM. Consequentially, performance is mediocre with an R^2 of 0.58 and KGE of 0.40. PBIAS of mHM compared to SWAT lies at 50.8% and RMSE is 7.3 mm. Since SWAT simulates generally lower actual evapotranspiration than mHM, but water yield (see section 7.2.5) remains remarkably similar, excess water could contribute to groundwater recharge in this case.

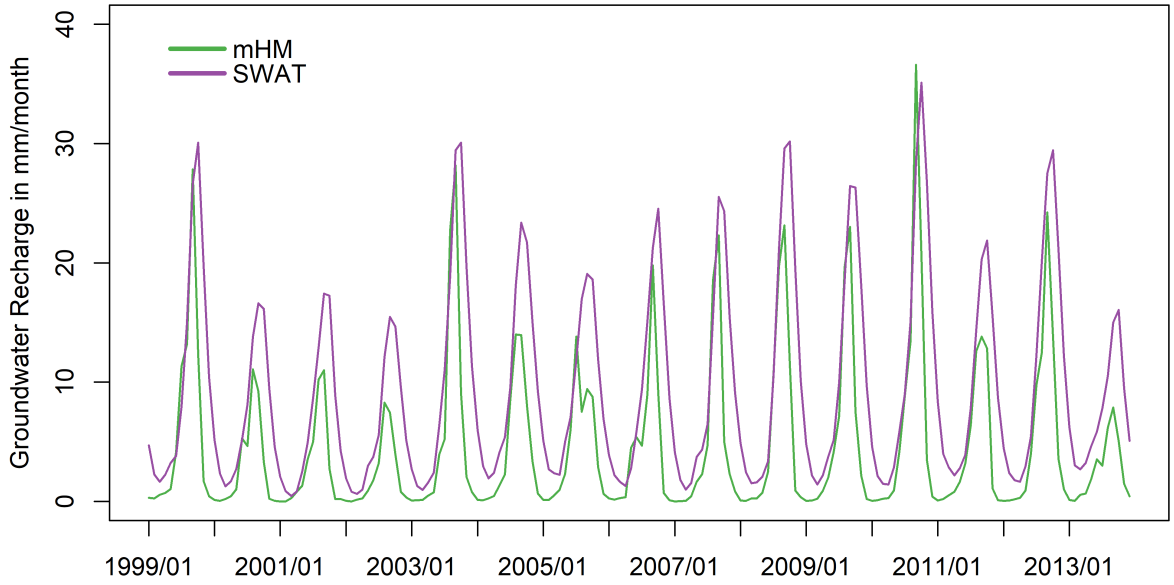

Figure 7.12: Monthly mean mHM and SWAT groundwater recharge estimates.

7.2.5 Water Yield

Lastly, simulations of water yield (water entering the stream from each modeling unit) by mHM and SWAT were compared with long-term January and July results presented in Figure 7.13. Here, mHM and SWAT perform quite similarly. During the dry season in January, both models predict low water yield of mainly between 0 and 10 mm/month. But while mHM simulates slightly elevated water yield in the south, SWAT results show higher rates for isolated regions, similar to the January groundwater recharge results presented in Figure 7.11. Concerning water yields in July, spatial patterns again closely resemble each other. The only differences are discernible in the eastern Volta basin and the central south, where mHM simulates higher yields.

When comparing the timeseries of the area-averaged results, as presented in Figure 7.14, performances are remarkably similar. Only for certain years does mHM simulate slightly higher water yields during the peaks of the rainy seasons. Both R^2 and KGE reach 0.94, with 4.7% bias and an RMSE of 2.2 mm/month.

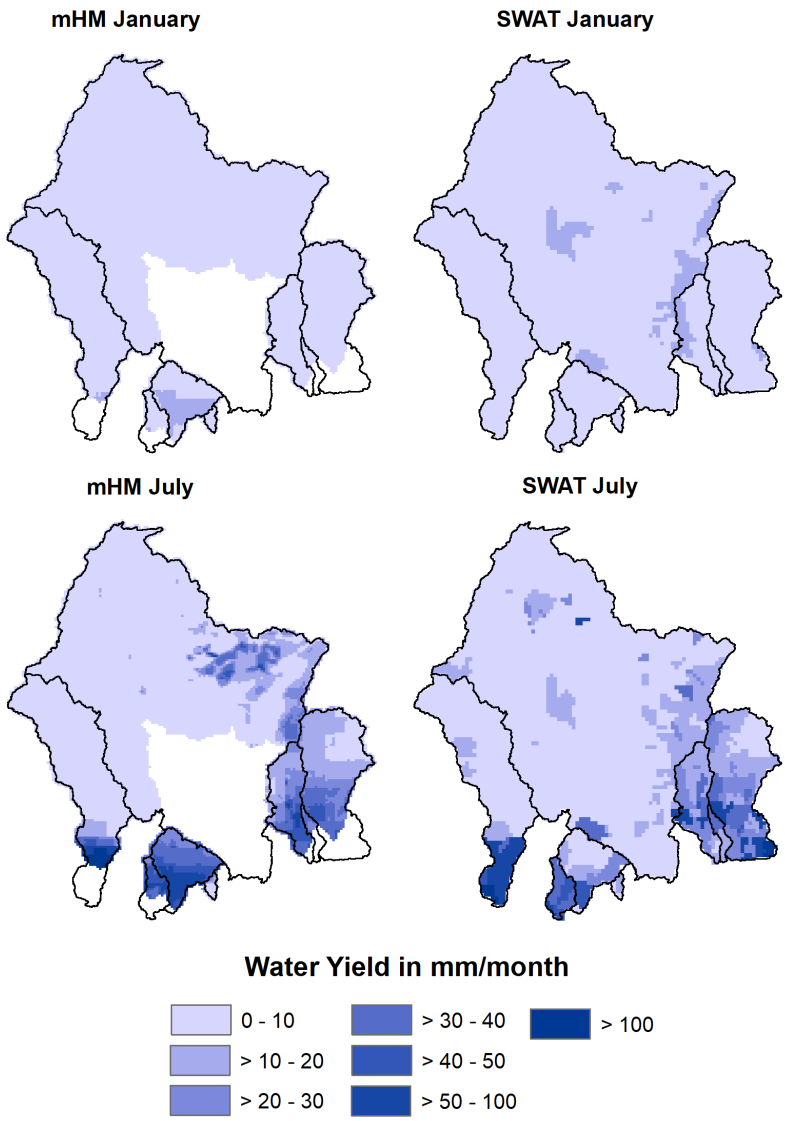

Figure 7.13: Long-term (1999-2013) January and July mHM and SWAT water yield estimates.

The performance of the mHM and SWAT discharge simulations for three exemplary gauges is shown in Figure 7.15. Gauges include Aniassué in the Comoé basin (71,252 km^2 drainage area), Ahlan in the Ouémé basin (56,556 km^2 drainage area) and Saboba in the Oti basin (39,149 km^2 drainage area), which is part of the larger Volta river basin. Since SWAT was run at a monthly timestep, daily mHM predictions were aggregated to monthly data. Statistics were calculated at the monthly timestep for both models. While the year 1998 was simulated in SWAT, mHM simulation outputs start in 1999, due to not all necessary data being available previously.

As can be seen, mHM streamflow simulations outperform SWAT results for almost all gauges with higher R^2 and KGE values and lower biases. Only during the validation at the Ahlan gauge does SWAT reach a higher KGE than mHM. While streamflow dynamics are acceptably to well matched by SWAT, biases are generally strong with underestimations during the calibration of between -13 and -43% and during the validation of between -35 and -51%. Contrarily, mHM biases range between -13 and 12% for the calibration and -21 and 23% for the validation.

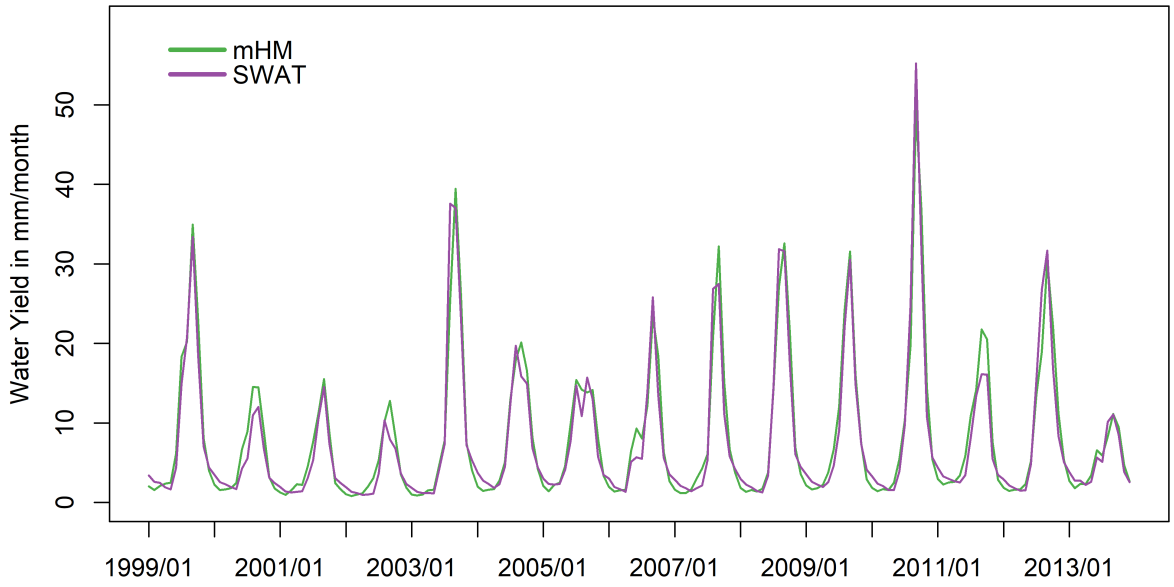

Figure 7.14: Monthly mean mHM and SWAT water yield estimates.

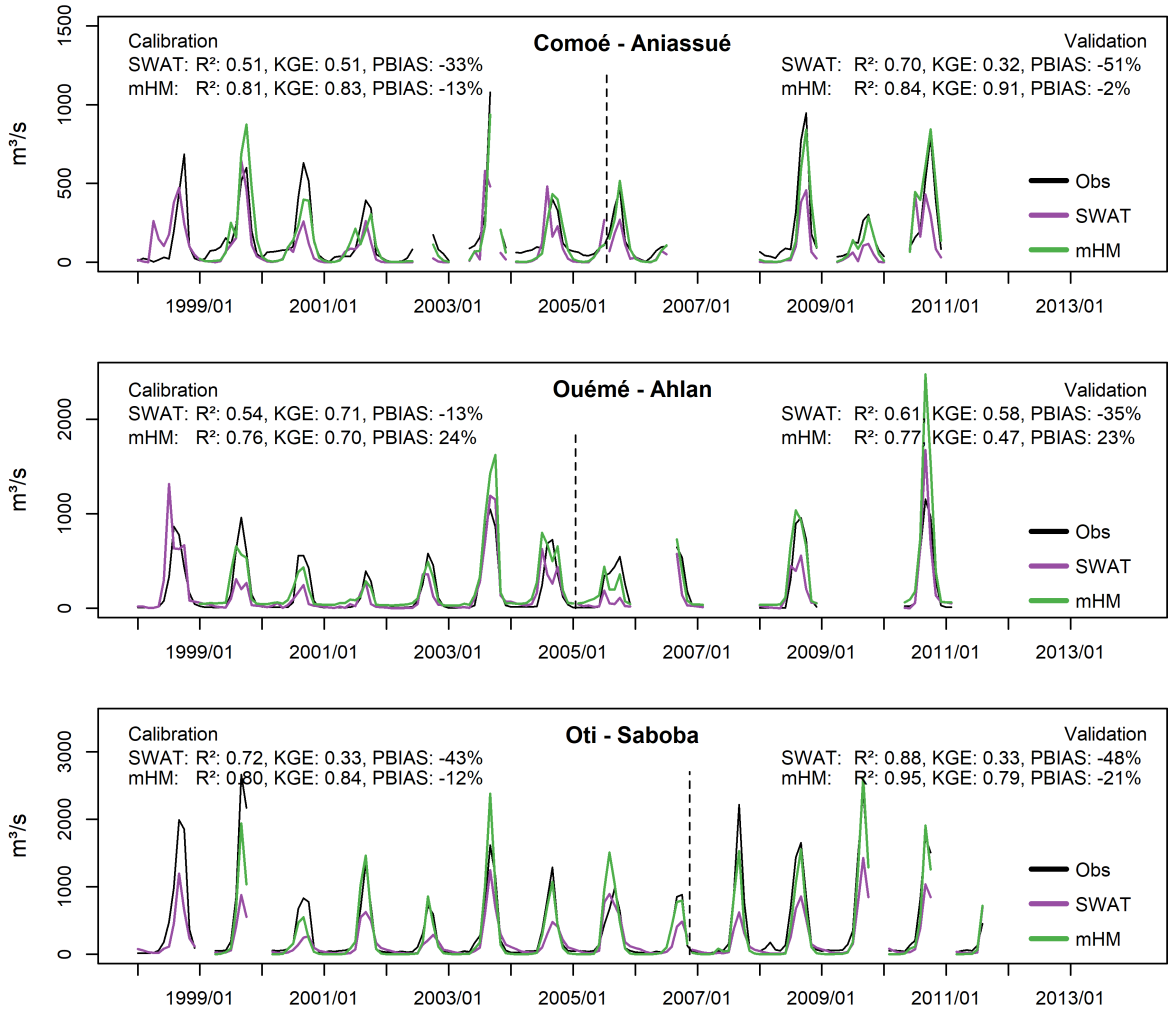

Figure 7.15: Exemplary mHM and SWAT discharge simulations.

As both models rely on the same precipitation input data and evapotranspiration calculation method, these datasets alone cannot account for the observed variability. However, the semi-distributed approach employed by SWAT only considers climatological information from the station or pixel located closest to the center of the subbasin polygon. All other points located within the subbasin are disregarded. In mHM, all spatial input information is accounted for in the model. Furthermore, while the SWAT model was globally calibrated for all basins together, river basins were calibrated separately in mHM.

7.3 Conclusion

In this chapter, the best simulation results of the SWAT (Chapter 5) and mHM (Chapter 6) models have been compared against each other and, where available, against remote sensing data and global model simulations. Key variables chosen for the comparison include actual evapotranspiration, soil moisture anomaly, total water storage anomaly, groundwater recharge and water yield. Long-term spatial comparison of the months of January and July have allowed a direct comparison of model results during the height of the dry and rainy seasons. Further spatially-averaged timeseries analysis permits the statistical evaluation of model outputs and global model and remote sensing data.

Globally, mHM has been shown to compare better than SWAT to external data. While SWAT offers a solid performance in most cases, this is not the case for each variable in question. Especially soil moisture anomaly and groundwater recharge outputs remain questionable, as do simulations of actual evapotranspiration in the south. However, performance of actual evapotranspiration for the central and northern region, as well as total water storage anomaly, compare very well to external datasets and water yield estimates perform very well when compared to mHM simulations. Interestingly, SWAT simulates a smaller amplitude of ΔTWS, albeit well within the global model ensemble uncertainty, while mHM simulates a larger amplitude, fitting better to GRACE remote sensing observations.

A clear advantage of using the mHM model is not only the reduced input data requirements, but especially the spatial representation of input data. mHM requires spatial data to be prepared in especially formatted netCDF input files. While this requires a not insignificant amount of know-how in preparing such datasets, several advantages become apparent. Since most remote sensing data is supplied in raster format, treatment of inputs remains largely similar and software developed for data treatment can be applied to a wide range of inputs, as was done in this study. Multidimensional raster data can also be processed and plotted easily. In SWAT, input data first has to be written to a single textfile for each raster pixel and data to be used is then chosen by the model according to the location in relation to the centeroid of each subbasin. Consequentially, only data from a single pixel is used to parameterize a subbasin. Especially when subbasins are larger, a great amount of spatial information is lost. Furthermore, mHM outputs all variables in a single netCDF file, while SWAT writes outputs in text format, requiring the data to be processed before evaluation.

8 General Conclusion

The accurate assessment of available water resources of sparsely-gauged West African river basins is a major challenge, and has been covered in many studies during the last decades. Hydrologic models can be used to aid researchers in this regard. While new hydrologic models are constantly developed and old models improved, the decline of *in situ* measurement networks since the 1990s has taken its toll on the quality of the predictions. However, we are currently at an unprecedented stage in history where remote sensing and reanalysis data products are available locally and globally for nearly all variables of interest. This change is also reflected in the literature, where global data and remote sensing products are increasingly used for model parametrization. Even though many data products are available for the West African region, less focus seems to be placed on the evaluation of the data quality. Although hydrologic models are more than ever parameterized using remote sensing data, most often, model calibration and validation are performed against observed streamflow only.

Here, the use of remote sensing data offers the unparalleled potential of further constraining the model parameter space during calibration, leading to more robust results. Multivariate validation can then be used to assess the performance of key hydrologic variables not included in the calibration. This immense opportunity is furthered by the fact that most of these data products are globally accessible and free of charge. The main goal of this study is to assess the contribution of these remote sensing and reanalysis data products in simulations of hydrologic processes in data-scarce West African river basins. It furthermore contributes to the ongoing debate on how hydrologic models can be improved using remote sensing products in sparsely-gauged basins worldwide.

In-depth conclusions have been provided at the end of each of the principal chapters (4, 5, 6 and 7). In the following, the main conclusions of this study will be summarized according to the four research questions formulated in the general introduction, with conclusions encompassing the whole study and potential for further investigation presented afterwards.

1. **How can the performance of remotely-sensed and reanalyzed precipitation data be assessed for the study area?**

The general lack of *in situ* precipitation measurements in the study area necessitates the use of remotely-sensed or reanalyzed data products. Many such products exist and the question arises how the performance of these products can be validated without ground-based observations. This question has been divided into two parts:

a) How can differences between ten precipitation datasets be quantified for West African river basins of differing locations and sizes?

Ten remotely-sensed and reanalyzed precipitation datasets were identified for comparison in the region. Data was then extracted for six subbasins of differing size and location in West Africa. For larger basins, multiple pixels were extracted and averaged, while for smaller basins, only a single datapoint was used. In a first step, precipitation estimates were directly compared against each other and against the GPCC gauge-based precipitation product. Results show a great variability between the products, with some products, such as TAMSAT (satellite infrared and gauge correction), CFSR (reanalysis), CMORPH RAW (satellite infrared and microwave measurements) and TMPA RT (satellite infrared and microwave measurements) consistently providing estimates not in line with predictions of other products. On the other hand, products relying on satellite infrared and microwave measurements and including gauge-based bias correction (such as CMORPH CRT, PERSIANN CDR and TMPA) as well as the GPCC product performed very similarly in all subbasins.

b) How robust are these products in accurately simulating streamflow during a hydrologic evaluation?

In order to assess how well these products perform in generating streamflow in a hydrologic simulation, the HBV-light model was set up and calibrated/validated for each combination of subbasin and precipitation product (60 models in total). For evaluation purposes, a model score was constructed based on NSE and PBIAS of the resulting simulations. Results largely confirm observations of the direct comparison, with satellite products utilizing infrared and microwave measurements with gauge correction performing best. A drawback of the hydrologic analysis is that during calibration, free infiltration and evaporation parameters might dampen effects of unrealistic precipitation estimates. It is therefore necessary to carefully choose realistic parameter ranges.

2. **How can a SWAT model of the study area be set up and multi-objectively validated using remote sensing observations?**

There have been some studies in the recent past focusing on setting up complex physically-based hydrologic models for the region. With the advent of easily accessible remote sensing observations, SWAT models using almost exclusively global data and remote sensing products were applied to the region. The main research question can be divided into four subquestions:

a) How can a hydrologic modeling framework for West Africa be set up using only freely available data?

It has been shown that almost all data needed to set up the SWAT model for the West African domain is readily available online. However, some soil characteristics need to be calculated. Alternatively, literature values must to be applied. The best-performing precipitation product from the HBV-light analysis (CMORPH CRT) was chosen to drive the SWAT simulations. SWAT cannot read grid-based meteorological data and thus, a textfile containing timeseries information for each pixel of the input raster needed to be created. This proved to be inconvenient. A major drawback of the SWAT model proved to be its inability to accurately simulate plant growth under the alternating rainy and dry seasons of the West African tropics. In order to generate

somewhat realistic plant growth, remote sensing LAI values were used to define growth in the plant management routine of the model.

b) Which simulation quality can be obtained using these datasets?

Three SWAT models were set up for the study area: the South model (Volta, Comoé, Ouémé, Mono, Pra, Ankobra, Couffo and Ayensu river basins), West model (Sénégal and Gambia river basins) and Niger model (Niger basin). Each model was calibrated using two different approaches. First, models were globally calibrated and in a second step, upstream subbasins were calibrated apart. Results show that the models can be successfully calibrated and validated for some regions, while results for other regions remain poor. When comparing both calibration methods, the global approach outperforms local calibration. This may be attributed to the vast number of parameters present for the local calibration and the limited number of model runs performed during sensitivity analysis. Only considering global calibrations, the South model performs best over the Ouémé basin, with particularly strong calibrations and validations. This is explained through the relatively good availability of discharge data with only few gaps in the region. Select stations in the central Volta and Comoé river basins also calibrate well. The West model performs best for the downstream Gambia basin, while in the Niger model, only some select tributary and downstream gauges perform well. The overall mediocre performance can be explained due to bias introduced when globally calibrating the model for such vast areas. Additionally, no outflow data was available for reservoirs in the region, which highly influences the quality of the downstream discharge simulation. The quality of the discharge data is an additional source of incertitude.

c) How can a multi-objective validation be performed?

While using observed discharge during model calibration allows the researcher to have appropriate confidence in the discharge simulation, uncertainty exists over the performance of additional key hydrologic variables. Therefore, it was decided to further validate actual evapotranspiration, soil moisture dynamics and total water storage using remote sensing data. In order to perform this analysis, SWAT outputs must first be reformatted into a usable format. Likewise, timeseries need to be extracted from remote sensing raster data. In this comparison, results were area-averaged for each timestep and subsequently compared. Concerning actual evapotranspiration and soil moisture dynamics, both SWAT calibration approaches compare well to remote sensing data. However, problems in the model structure relating to plant growth lead to slight deviations of actual evapotranspiration. A comparison of total water storage anomalies shows that while SWAT simulates the dynamics very well, amplitudes remain lower than estimated by the satellite product. It is also apparent that the total water storage anomaly simulation failed during the local calibration of the West model.

d) How can the potential and limitations of this approach for assessing the water availability at the regional scale be evaluated?

The good results attained by the multivariate validation highlight its merits in confirming the model robustness. The applied methodology offers the possibility to validate variables not

included in the calibration against readily available remote sensing data. However, the evaluation is limited by the fact that results are averaged over the model areas before comparison, which does not allow to determine the performance for smaller subbasins. Further limits are imposed by the extensive number of parameters which need to be estimated when running SWAT for large regions. Not only does this require considerable computing power, parameter equifinality can also pose problems. Overall, the model structure complicates the use of gridded input data and comparison of results against external observations and additional routines needed to be developed to prepare input and output data, which might impede the easy adoption of the proposed framework. The model could be further improved by including the Niger Inland Delta and Lake Volta, if observations were available.

3. **How can the grid-based mHM model be calibrated and validated in a multivariate framework for the region?**

mHM is a relatively new grid-based hydrologic model designed specifically for predictions in data-scarce mesoscale river basins. Due to its multiscale parameter regionalization approach, it can be efficiently calibrated at a coarse resolution and the calibrated model can then be scaled and run in a finer resolution. Also, as parameters are linked across different scales, less parameters need to be estimated, reducing uncertainty introduced through parameter equifinality. Here, the model was applied to a study area in southern West Africa encompassing the Comoé, Ouémé, Couffo, Mono, Pra, Ayensu, Ankobra, Black Volta, White Volta and Oti river basins. Three subquestions were formulated:

a) How can the performance of the mHM model using a multiscale parameter regionalization approach be assessed?

Precipitation and potential evapotranspiration inputs and calculation methods can have a profound influence on model performance. Therefore, it was decided to first assess the performance of four precipitation inputs and three evaporation calculation methods in twelve exploratory model runs for the region. Results show that CMORPH CRT precipitation together with the Hargreaves potential evapotranspiration method perform best. The same well-performing combination was also employed in the SWAT models and CMORPH CRT also performed best in simulating streamflow during the HBV-light hydrologic evaluation. To assess the overall model performance, mHM was calibrated using two methods, once using only discharge as the observed variable and once using discharge and remotely-sensed actual evapotranspiration. The calibrated models were evaluated against discharge and remotely-sensed actual evapotranspiration, soil moisture anomaly and total water storage anomaly.

b) How does the model perform under multivariate calibration inputs?

Results show that the mHM model is well-suited for hydrologic simulations of data-scarce West African river basins. While both calibration methods closely resemble each other when evaluated against discharge (except during the validation), the discharge and actual evapotranspiration calibration method produces vastly improved actual evapotranspiration estimates. This confirms the advantage of using a second observed variable to further constrain boundary conditions

during model calibration. However, some under- and overestimations can be observed in the humid south and arid north. Interestingly, remote sensing estimates of the GLEAM and MODIS actual evapotranspiration products differ quite strongly, which can be attributed to different data and calculation approaches. Concerning soil moisture and total water storage anomaly, both calibration methods again perform almost equally well.

c) How can model parameters be transferred between different spatial scales using multiscale parameter regionalization?

In a last step, the multiscale parameter regionalization approach of mHM was tested by rerunning the calibrated models in two finer resolutions with a strongly decreased pixel size and increased pixel count. When evaluated against streamflow observations, calibration results remained stable, while results obtained for the validation period slightly increased under both calibration methods. This confirms the good performance of the multiscale parameter regionalization approach in the West African domain. By using this approach, computational demand and subsequently calibration times can be significantly reduced by calibrating the model in a coarse resolution and only rerunning the calibrated model once in a finer resolution.

Overall, mHM proved to be well-suited for generating hydrologic predictions of sparsely-gauged basins in southern West Africa under remote sensing inputs. Multivariate validation has confirmed the model robustness in estimating variables not included in the calibration. The grid-based nature of the model eases the modelers task in preparing model in- and outputs significantly.

4. **How well do SWAT and mHM simulations of key hydrologic variables compare against each other and against remote sensing, as well as global model results?**

Lastly, to better assess the performance of both the SWAT and mHM models, results were compared against each other and against remote sensing, as well as global model results for a region in southern West Africa. In this case, only the best-performing model iterations were evaluated, namely the global SWAT calibration and the discharge and actual evapotranspiration mHM calibration. Simulations of actual evapotranspiration, soil moisture anomaly, total water storage anomaly, groundwater recharge and water yield were assessed. It has been shown that mHM estimates consistently outperform SWAT results when compared against remote sensing and global model results. It should be stressed, however, that while SWAT was calibrated globally for the entire region, mHM was calibrated locally for each river basin. Of particular interest is the comparison of actual evapotranspiration outputs, where the divergence of not only the model results, but also remote sensing estimates can be observed for three different regions of the study area. Here, the inability of SWAT to accurately simulate tropical plant growth under alternating wet and dry seasons becomes apparent, as actual evapotranspiration during the dry seasons is low. This also relates to total water storage simulations, where the SWAT total water storage anomaly amplitude is lower than mHM and GRACE results. It is interesting to note that SWAT total water storage simulations compare better to global model results, while mHM simulations reach a higher degree of agreement with GRACE satellite observations. In cases like the soil moisture anomaly comparison, the SWAT output format renders a direct comparison

difficult, as residual water content is not taken into account. Discharge observations were included in the SWAT and mHM calibration schemes and water yield estimations of both models are remarkably similar.
It can be concluded that the use of remote sensing data for model setup, calibration and validation positively contributes to achieving robust simulations. Already during the model setup phase, consistent quality checks using remote sensing and global data products help spot potential problems early on. Model uncertainty is reduced by constraining the parameter space and outputs for variables not included in the calibration can be easily validated. The methodological approaches developed in this study can be easily adapted to improve hydrologic simulations in sparsely-gauged basins in West Africa, particularly since all data and software used are freely available, albeit some knowledge in the handling of gridded data is necessary.

Some drawbacks, however, remain. First and foremost, the availability of high-resolution remote sensing estimates limits analysis to more recent periods, with few products being available before 1998. Also, satellites have a limited service life and as they are replaced with newer instruments, the questions remains whether the produced estimates are comparable. The hydrologic evaluation using HBV-light is limited by the fact that during calibration, large parameter ranges may dampen the effects of unrealistic precipitation estimations. While the SWAT model offers the physically-based estimation of many parameters also relating to plant growth and nutrient transport, vast amounts of input data are required which are not easily attainable for data-scarce basins. The crop growth module has been developed for temperate regions, where crop growth is defined as a function of temperature, rather than water availability. While a new routine suitable for tropical regions has recently been developed, it has so far not been applied in the study area. Input and output data formats are also less than convenient, relying on Fortran-type datafiles instead of spatial formats such as polygons or raster grids. While mHM offers spatial in- and outputs and generally performs well for the region, some model routines remain simplistic, such as the inability to account for more than three land use classes, rendering land use change analysis difficult.

This study contributes methodological advances to the accurate assessment of water resources in sparsely-gauged catchments in West Africa using remote sensing, reanalysis and global datasets. Still, a potential for further studies remains. Some key aspects to be further studied were thus identified:

- **Assess improvements to the SWAT model by including the Inner Niger Delta and Lake Volta in the simulations. Further investigate performance at the small scale using measured soil attributes.**

  SWAT simulations of the Niger river basin could be further improved by parameterizing the Inner Niger Delta, a wetland which, due to flooding during the rainy season, has a profound influence on downstream discharge. However, information on flooding in the delta is scarce, with studies only recently focusing of this issue, see e.g. Ibrahim et al. (2017). In SWAT simulations, the delta could be introduced as an artificial reservoir, if in- and outflow information was available. Furthermore, simulations of the Volta river basin could be improved if the Akosombo dam, which creates Lake Volta, is included. Again, little information on in- and outflows is available.

- **Investigate the improvement of SWAT actual evapotranspiration, soil moisture and total water storage estimates by using a plant routine adapted to tropical conditions.**

  Several researchers have explored the possibility of integrating a water availability-based plant growth module in SWAT, see e.g. Alemayehu et al. (2017). Setting up a SWAT model for the study area with such a routine could lead to improved actual evapotranspiration predictions and subsequently more realistic soil moisture and total water storage estimates.

- **Assess mHM model uncertainty for exploratory model runs by performing product-specific calibrations of each combination.**

  Due to time constrains, the twelve mHM exploratory model runs with different combinations of precipitation and potential evapotranspiration inputs were not calibrated, but run using default parameters. Assessing the model uncertainty using different combinations would lead to a more nuanced understanding of the effects precipitation and evapotranspiration inputs have on model results.

- **Determine the contribution of land surface temperature and total water storage estimates in mHM calibrations.**

  In this study, mHM was calibrated using a combination of discharge and actual evapotranspiration measurements. It has been suggested that land surface temperature and total water storage estimates included in the calibration also lead to very robust model results. This needs to be further investigated for the West African study area, especially as temperatures remain relatively constant over the course of the year.

- **Explore the performance of the mHM model applied to the Niger, Sénégal and Gambia river basins.**

  In this study, the mHM model was only applied to the southern West African river basins, where it performed well. Further research is necessary to confirm good model performance also for the Niger, Sénégal and Gambia river basins.

- **Investigate possibilities for climate change analysis using mHM.**

  Lastly, it would be interesting to assess how well mHM is suited to perform climate change and land use and land cover scenario analysis. While the grid-based structure and possibilities for netCDF data inputs facilitate large-scale climate data integration, its limited land cover parametrization of only three classes (previous, impervious, and forest) might reduce the informational value of the results.

While it has been shown that the quality of hydrologic simulations profits from integrating remote sensing estimates, it should be stressed that a continued need for high quality *in situ* measurements persists, especially regarding streamflow data. Several streamflow timeseries had to be excluded from this study due to large data gaps or general concerns regarding the data quality. Also, the quality of remote sensing estimates could be strongly improved if a network of ground-based measurements was available for appropriate bias correction. Some researchers are of the opinion that the need for expensive ground-based measurements has been reduced

by the availability of high resolution satellite data. In the author's opinion, however, there is a clear advantage of using a combination of gauge- and satellite-based data, with remote sensing complementing, and not replacing *in situ* measurements.

Bibliography

Abbaspour, K. C., E. Rouholahnejad, S. Vaghefi, R. Srinivasan, H. Yang, and B. Kløve. 2015. "A continental-scale hydrology and water quality model for Europe: Calibration and uncertainty of a high-resolution large-scale SWAT model". *Journal of Hydrology* 524:733–752. doi:10.1016/j.jhydrol.2015.03.027.

Abbaspour, K. C. 2015. *SWAT-CUP: SWAT Calibration and Uncertainty Programs - A User Manual.* 100. Dübendorf, Switzerland: Swiss Federal Institute of Aquatic Science / Technology (EAWAG).

Abbaspour, K. C., M. Vejdani, S. Haghighat, and J. Yang. 2007. "SWAT-CUP Calibration and Uncertainty Programs for SWAT". *The fourth International SWAT conference*: 1596–1602.

Abbaspour, K., S. Vaghefi, and R. Srinivasan. 2017. "A Guideline for Successful Calibration and Uncertainty Analysis for Soil and Water Assessment: A Review of Papers from the 2016 International SWAT Conference". *Water* 10 (1): 6. doi:10.3390/w10010006.

Adeyewa, Z. D., and K. Nakamura. 2003. "Validation of TRMM Radar Rainfall Data over Major Climatic Regions in Africa". *Journal of Applied Meteorology* 42 (2): 331–347. doi:10.1175/1520-0450(2003)042<0331:VOTRRD>2.0.CO;2.

Adjei, K. A., L. Ren, E. K. Appiah-Adjei, K. Kankam-Yeboah, and A. A. Agyapong. 2012. "Validation of TRMM Data in the Black Volta Basin of Ghana". *Journal of Hydrologic Engineering* 17 (5): 647–654. doi:10.1061/(ASCE)HE.1943-5584.0000487.

African Development Bank. 2018. *West Africa Economic Outlook 2018.* Abidjan, Côte d'Ivoire: African Development Bank Group. https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/2018AEO/African_Economic_Outlook_2018_West-Africa.pdf.

Aghakouchak, A., and E. Habib. 2010. "Application of a Conceptual Hydrologic Model in Teaching Hydrologic Processes". *International Journal of Engineering Education* 26 (4): 1–11.

Ahn, P. 1970. *West African Agriculture Volume 1 - West African Soils.* 3rd ed. London, United Kingdom: Oxford University Press.

Alemayehu, T., A. van Griensven, B. T. Woldegiorgis, and W. Bauwens. 2017. "An improved SWAT vegetation growth module and its evaluation for four tropical ecosystems". *Hydrology and Earth System Sciences* 21 (9): 4449–4467. doi:10.5194/hess-21-4449-2017.

Allen, R. G., L. Pereira, D. Raes, and M. Smith. 1998. "Crop evapotranspiration: Guidelines for computing crop requirements". *Irrigation and Drainage Paper No. 56, FAO*, no. 56: 300.

Andersen, J., J. C. Refsgaard, and K. H. Jensen. 2001. "Distributed hydrological modelling of the Senegal River Basin - Model construction and validation". *Journal of Hydrology* 247 (3-4): 200–214. doi:10.1016/S0022-1694(01)00384-5.

Arnold, J. G., J. R. Kiniry, R. Srinivasan, J. R. Williams, E. B. Haney, and S. L. Neitsch. 2012a. "Soil & Water Assessment Tool: Input/output documentation". *Texas Water Resources Institute, TR-439.*

Arnold, J. G., R. Srinivasan, R. S. Muttiah, and J. R. Williams. 1998. "Large area hydrologic modeling and assessment Part 1: model development". *Journal of the American Water Resources Association* 34 (1): 73–89. doi:10.1111/j.1752-1688.1998.tb05961.x.

Arnold, J. G., D. N. Moriasi, P. W. Gassman, K. C. Abbaspour, M. J. White, R. Srinivasan, C. Santhi, R. D. Harmel, a. V. Griensven, M. W. VanLiew, N. Kannan, and M. K. Jha. 2012b. "Swat: Model Use, Calibration, and Validation". *Asabe* 55 (4): 1491–1508.

Ashouri, H., K. L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat. 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. doi:10.1175/BAMS-D-13-00068.1.

Awange, J., V. Ferreira, E. Forootan, Khandu, S. Andam-Akorful, N. Agutu, and X. He. 2015. "Uncertainties in remotely sensed precipitation data over Africa". *International Journal of Climatology*. doi:10.1002/joc.4346.

Balek, J. 1977. *Hydrology and Water Resources in Tropical Africa*. New York, USA: Elsevier.

Bárdossy, A. 2007. "Calibration of hydrological model parameters for ungauged catchments". *Hydrology and Earth System Sciences* 11 (2): 703–710. doi:10.5194/hess-11-703-2007.

Becker, A., and P. Braun. 1999. "Disaggregation, aggregation and spatial scaling in hydrological modelling". *Journal of Hydrology* 217 (3-4): 239–252. doi:10.1016/S0022-1694(98)00291-1.

Behrangi, A., B. Khakbaz, T. C. Jaw, A. AghaKouchak, K. Hsu, and S. Sorooshian. 2011. "Hydrologic evaluation of satellite precipitation products over a mid-size basin". *Journal of Hydrology* 397 (3-4): 225–237. doi:10.1016/j.jhydrol.2010.11.043.

Behrangi, A., B. Khakbaz, J. A. Vrugt, Q. Duan, and S. Sorooshian. 2008. "Comment on "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration" by Bryan A. Tolson and Christine A. Shoemaker". *Water Resources Research* 44 (12): 2–4. doi:10.1029/2007WR006429.

Bergström, S. 1976. *Global Perspectives on Loss of Human Life Caused by Floods*, vol. RHO 7. 2. Norrköping: Swedish Meteorological / Hydrological Institute.

— . 1992. *The HBV model - its structure and applications*. 4. Norrköping: Swedish Meteorological / Hydrological Institute.

Beven, K. J. 1993. "Prophecy, reality and uncertainty in distributed hydrological modelling". *Adv. Water Resour.* 16:41–51.

Beven, K. 2012. *Rainfall-runoff modelling: the primer*. Second Edi, 1–457. Chichester, UK: Wiley-Blackwell. doi:10.1002/9781119951001.

Beven, K. 2001. "How far can we go in distributed hydrological modelling?" *Hydrology and Earth System Sciences* 5 (1): 1–12. doi:10.5194/hess-5-1-2001.

Bitew, M. M., and M. Gebremichael. 2011. "Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands". *Hydrology and Earth System Sciences* 15 (4): 1147–1155. doi:10.5194/hess-15-1147-2011.

Bitew, M. M., M. Gebremichael, L. T. Gebremichael, and Y. A. Bayissa. 2012. "Evaluation of High-Resolution Satellite Rainfall Products through Streamflow Simulation in a Hydrological Modeling of a Small Mountainous Watershed in Ethiopia". *Journal of Hydrometeorology* 13 (1): 338–350. doi:10.1175/2011JHM1292.1.

Blacutt, L. A., D. L. Herdies, L. G. G. de Gonçalves, D. A. Vila, and M. Andrade. 2015. "Precipitation Comparison for the CFSR, MERRA, TRMM3B42 and Combined Scheme datasets in Bolivia". *Atmospheric Research* 163:117–131. doi:10.1016/j.atmosres.2015.02.002.

Bodian, A., A. Dezetter, A. Deme, and L. Diop. 2016. "Hydrological Evaluation of TRMM Rainfall over the Upper Senegal River Basin". *Hydrology* 3 (15): 1–18. doi:10.3390/hydrology3020015.

Bontemps, S., P. Defourny, E. Van Bogaert, O. Arino, V. Kalogirou, and J. Perez. 2011. *GLOBCOVER 2009 Products Description and Validation Report*. 53. Louvain, Paris: Université catholique de Louvain, European Space Agency.

Boone, A., B. Decharme, F. Guichard, P. de Rosnay, G. Balsamo, A. Beljaars, F. Chopin, T. Orgeval, J. Polcher, C. Delire, A. Ducharne, S. Gascoin, M. Grippa, L. Jarlan, L. Kergoat, E. Mougin, Y. Gusev, O. Nasonova, P. Harris, C. Taylor, A. Norgaard, I. Sandholt, C. Ottlé, I. Poccard-Leclercq, S. Saux-Picart, and Y. Xue. 2009. "The AMMA Land Surface Model Intercomparison Project (ALMIP)". *Bulletin Of The American Meteorological Society* 90 (12): 1865–1880. doi:10.1175/2009BAMS2786.1.

Bormann, H., and B. Diekkrüger. 2004. "A conceptual, regional hydrological model for Benin (West Africa): Validation, uncertainty assessment and assessment of applicability for environmental change analyses". *Physics and Chemistry of the Earth* 29 (11-12 SPEC. ISS.): 759–768. doi:10.1016/j.pce.2004.05.003.

Bormann, H. 2005. "Regional hydrological modelling in Benin (West Africa): Uncertainty issues versus scenarios of expected future environmental change". *Physics and Chemistry of the Earth* 30 (8-10): 472–484. doi:10.1016/j.pce.2005.07.002.

Bosilovich, M. G., R. Lucchesi, and M. Suarez. 2016. *MERRA-2: File Specification. GMAO Office Note No. 9 (Version 1.1)*. 73. Greenbelt, Maryland: NASA Goddard Space Flight Center.

Buckle, C. 1996. *Weather and Climate in Africa*. Harlow, England: Longman.

Cheng, M., B. D. Tapley, and J. C. Ries. 2013. "Deceleration in the Earth's oblateness". *Journal of Geophysical Research: Solid Earth* 118 (2): 740–747. doi:10.1002/jgrb.50058.

Church, R. 1957. *West Africa - A Study of the Environment and of Man's Use of it*. London, New York, Toronto: Longmans, Green / Co.

CILSS. 2016. *Landscapes of West Africa - A Window on a Changing World*. 47914 252nd St, Garretson, SD 57030, UNITED STATES: U.S. Geological Survey EROS.

Cohen Liechti, T., J. P. Matos, J. L. Boillat, and A. J. Schleiss. 2012. "Comparison and evaluation of satellite derived precipitation products for hydrological modeling of the Zambezi River Basin". *Hydrology and Earth System Sciences* 16 (2): 489–500. doi:10.5194/hess-16-489-2012.

Cornelissen, T., B. Diekkrüger, and S. Giertz. 2013. "A comparison of hydrological models for assessing the impact of land use and climate change on discharge in a tropical catchment". *Journal of Hydrology* 498:221–236. doi:10.1016/j.jhydrol.2013.06.016.

Danvi, A., S. Giertz, S. J. Zwart, and B. Diekkrüger. 2017. "Comparing water quantity and quality in three inland valley watersheds with different levels of agricultural development in central Benin". *Agricultural Water Management* 192:257–270. doi:10.1016/j.agwat.2017.07.017.

Danvi, A., T. Jütten, S. Giertz, S. J. Zwart, and B. Diekkrüger. 2016. "A spatially explicit approach to assess the suitability for rice cultivation in an inland valley in central Benin". *Agricultural Water Management* 177:95–106. doi:10.1016/j.agwat.2016.07.003.

DeLiberty, T. L., and D. R. Legates. 2003. "Interannual and seasonal variability of modelled soil moisture in Oklahoma". *International Journal of Climatology* 23 (9): 1057–1086. doi:10.1002/joc.904.

Diekkrüger, B., H. Busche, S. Giertz, and G. Steup. 2010. "Hydrology". In *Impacts of Global Change on the Hydrological Cycle in West and Northwest Africa*, ed. by P. Speth, M. Christoph, and B. Diekkrüger. Heidelberg, Germany: Springer.

Dile, Y. T., and R. Srinivasan. 2014. "Evaluation of CFSR climate data for hydrologic prediction in data-scarce watersheds: an application in the Blue Nile River Basin". *JAWRA Journal of the American Water Resources Association* 77845:1–16. doi:10.1111/jawr.12182.

Döll, P., H. Hoffmann-Dobrev, F. T. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg, and B. R. Scanlon. 2012. "Impact of water withdrawals from groundwater and surface water on continental water storage variations". *Journal of Geodynamics* 59-60:143–156. doi:10.1016/j.jog.2011.05.001.

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

Dorigo, W., W. Wagner, C. Albergel, F. Albrecht, G. Balsamo, L. Brocca, D. Chung, M. Ertl, M. Forkel, A. Gruber, E. Haas, P. D. Hamer, M. Hirschi, J. Ikonen, R. de Jeu, R. Kidd, W. Lahoz, Y. Y. Liu, D. Miralles, T. Mistelbauer, N. Nicolai-Shaw, R. Parinussa, C. Pratola, C. Reimer, R. van der Schalie, S. I. Seneviratne, T. Smolander, and P. Lecomte. 2017. "ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions". *Remote Sensing of Environment* 203:185–215. doi:10.1016/j.rse.2017.07.001.

Droogers, P., and R. G. Allen. 2002. "Estimating reference evapotranspiration under inaccurate data conditions". *Irrigation and drainage systems* 16:33–45. doi:10.1023/A:1015508322413.

DWD. 2018. *Deutscher Wetterdienst - Climate Data Center*. Visited on 12/09/2018. https://www.dwd.de/DE/klimaumwelt/cdc/cdc_node.html.

Ek, M. B., K. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. Tarpley. 2003. "Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model". *Journal of Geophysical Research* 108 (D22): 8851. doi:10.1029/2002JD003296.

Encyclopaedia Britannica. 2018a. *Gambia River*. Visited on 12/09/2018. https://www.britannica.com/place/Gambia-River.

— . 2018b. *Mount Cameroon*. Visited on 12/09/2018. https://www.britannica.com/place/Mount-Cameroon.

Fafchamps, M., C. Udry, and K. Czukas. 1998. "Drought and saving in West Africa: are livestock a buffer stock?" *Journal of Development Economics* 55:273–305.

Fang, H., H. Beaudoing, M. Rodell, W. Teng, and B. Vollmer. 2009. *Global Land Data Assimilation System (GLDAS) products, services and application from NASA Hydrology Data and Information Services Center (HDISC)*. Visited on 12/09/2018. http://ntrs.nasa.gov/search.jsp?R=20090005038.

Fang, Y.-h., X. Zhang, C. Corbari, M. Mancini, G.-y. Niu, and W. Zeng. 2017. "Improving the Xin'anjiang hydrological model based on mass – energy balance". *Hydrol. Earth Syst. Sci.* 21:3359–3375. doi:10.5194/hess-21-3359-2017.

FAO. 2018. *FAOSTAT Database*. Visited on 10/09/2018. http://www.fao.org/faostat/en/#data.

FAO-AQUASTAT. 2018. *FAO-AQUASTAT Database*. Visited on 12/09/2018. http://www.fao.org/nr/water/aquastat/data/query/index.html?lang=en.

FAO/AGL. 2003. *WRB Map of World Soil Resources*. Visited on 11/09/2018. http://www.fao.org/soils-portal/soil-survey/soil-classification/world-reference-base/en/.

FAO/IIASA/ISRIC/ISS-CAS/JRC. 2012. *Harmonized World Soil Database (version 1.2)*. 1–50. Laxenburg, Austria: FAO, Rome, Italy / IIASA.

Faramarzi, M., K. C. Abbaspour, S. Ashraf Vaghefi, M. R. Farzaneh, A. J. B. Zehnder, R. Srinivasan, and H. Yang. 2013. "Modeling impacts of climate change on freshwater availability in Africa". *Journal of Hydrology* 480:85–101. doi:10.1016/j.jhydrol.2012.12.016.

Faramarzi, M., K. C. Abbaspour, R. Schulin, and H. Yang. 2010. "Modelling blue and green water resources availibility in Iran". *Hydrological Processes* 23 (November 2008): 486–501. doi:10.1002/hyp.7160.

Fink, A., M. Christoph, K. Born, T. Bruecher, K. Piecha, S. Pohle, O. Schulz, and V. Ermert. 2010. "Climate". In *Impacts of Global Change on the Hydrological Cycle in West and Northwest Africa*, ed. by P. Speth, M. Christoph, and B. Diekkrüger, 54–58. Heidelberg: Springer.

Forootan, E., J. Kusche, I. Loth, W. D. Schuh, A. Eicker, J. Awange, L. Longuevergne, B. Diekkrüger, M. Schmidt, and C. K. Shum. 2014. "Multivariate Prediction of Total Water Storage Changes Over West Africa from Multi-Satellite Data". *Surveys in Geophysics* 35 (4): 913–940. doi:10.1007/s10712-014-9292-0.

Fujihara, Y., Y. Yamamoto, Y. Tsujimoto, and J.-i. Sakagami. 2014. "Discharge Simulation in a Data-Scarce Basin Using Reanalysis and Global Precipitation Data : A Case Study of the White Volta Basin". *Journal of Water Resource and Protection*, no. 6: 1316–1325. doi:10.4236/jwarp.2014.614121.

Fuka, D. R., M. T. Walter, C. Macalister, A. T. Degaetano, T. S. Steenhuis, and Z. M. Easton. 2013. "Using the Climate Forecast System Reanalysis as weather input data for watershed models". *Hydrological Processes*. doi:10.1002/hyp.10073.

Funk, C., P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell, and J. Michaelsen. 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. http://www.nature.com/articles/sdata201566.

Gaiser, T., H. Goldbach, S. Giertz, C. Hiepe, and A. Klose. 2010. "Soils". In *Impacts of Global Change on the Hydrological Cycle in West and Northwest Africa*, ed. by P. Speth, M. Christoph, and B. Diekkrüger. Heidelberg, Germany: Springer.

Gao, H., Q. Tang, X. Shi, C. Zhu, and T. Bohn. 2009. "Water budget record from Variable Infiltration Capacity (VIC) model". In *Algorithm Theoretical Basis Document, Version 1.2*, 120–173. http://eprints.lancs.ac.uk/89407/1/Gao_et_al_VIC_2014.pdf.

Gassman, P. W., M. R. Reyes, C. H. Green, and J. G. Arnold. 2007. "The Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Directions". *American Society of Agricultural and Biological Engineers* 50 (4): 1211–1250.

Gautier, D., D. Denis, and B. Locatelli. 2016. "Impacts of drought and responses of rural populations in West Africa: a systematic review". *WIREs Clim Change*. doi:10.1002/wcc.411.

Gavilán, P., I. J. Lorite, S. Tornero, and J. Berengena. 2006. "Regional calibration of Hargreaves equation for estimating reference et in a semiarid environment". *Agricultural Water Management* 81 (3): 257–281. doi:10.1016/j.agwat.2005.05.001.

Gessner, U., M. Niklaus, C. Kuenzer, and S. Dech. 2013. "Intercomparison of Leaf Area Index Products for a Gradient of Sub-Humid to Arid Environments in West Africa": 1235–1257. doi:10.3390/rs5031235.

Giertz, S., B. Diekkrüger, and G. Steup. 2006. "Physically-based modelling of hydrological processes in a tropical headwater catchment in Benin (West Africa) – process representation and multi-criteria validation". *Hydrology and Earth System Sciences* 3 (2): 595–651. doi:10.5194/hess-10-829-2006.

Gosset, M., J. Viarre, G. Quantin, and M. Alcoba. 2013. "Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks". *Quarterly Journal of the Royal Meteorological Society* 139 (673): 923–940. doi:10.1002/qj.2130.

Gourley, J. J., Y. Hong, Z. L. Flamig, J. Wang, H. Vergara, and E. N. Anagnostou. 2011. "Hydrologic Evaluation of Rainfall Estimates from Radar, Satellite, Gauge, and Combinations on Ft. Cobb Basin, Oklahoma". *Journal of Hydrometeorology* 12 (5): 973–988. doi:10.1175/2011JHM1287.1.

Grimes, D. I. F., E. Pardo-Igúzquiza, and R. Bonifacio. 1999. "Optimal areal rainfall estimation using raingauges and satellite data". *Journal of Hydrology* 222 (1-4): 93–108. doi:10.1016/S0022-1694(99)00092-X.

Grippa, M., L. Kergoat, F. Frappart, Q. Araud, A. Boone, P. De Rosnay, J. M. Lemoine, S. Gascoin, G. Balsamo, C. Ottlé, B. Decharme, S. Saux-Picart, and G. Ramillien. 2011. "Land water storage variability over West Africa estimated by Gravity Recovery and Climate Experiment (GRACE) and land surface models". *Water Resources Research* 47 (5): 1–18. doi:10.1029/2009WR008856.

Gruber, A., W. A. Dorigo, W. Crow, and W. Wagner. 2017. "Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals". *IEEE Transactions on Geoscience and Remote Sensing* 55 (12): 6780–6792. doi:10.1109/TGRS.2017.2734070.

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

Gyau-Boakye, P. 2001. "Environmental impacts of the Akosombo dam and effects of climate change on the lake levels". *Environment, Development and Sustainability* 3 (1): 17–29. doi:10.1023/A:1011402116047.

Hargreaves, G. H., and Z. A. Samani. 1985. "Reference Crop Evapotranspiration from Temperature". *Applied Engineering in Agriculture*: 96–99.

Hassan, A., and S. Jin. 2016. "Water storage changes and balances in Africa observed by GRACE and hydrologic models". *Geodesy and Geodynamics* 7 (1): 39–49. doi:10.1016/j.geog.2016.03.002.

Hattermann, F., V. Krysanova, S. N. Gosling, R. Dankers, P. Daggupati, C. Donnelly, M. Flörke, S. Huang, Y. Motovilov, S. Buda, T. Yang, C. Müller, G. Leng, Q. Tang, F. T. Portmann, S. Hagemann, D. Gerten, Y. Wada, Y. Masaki, T. Alemayehu, Y. Satoh, and L. Samaniego. 2017. "Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins". *Climatic Change* 141 (3): 561–576. doi:10.1007/s10584-016-1829-4.

Hollinger, F., and J. M. Staatz. 2015. *Agricultural Growth in West Africa Market and policy drivers*. Ed. by F. Hollinger and J. M. Staatz. African Development Bank/Food and Agriculture Organization of the United Nations.

Hsu, K.-l., X. Gao, S. Sorooshian, and H. V. Gupta. 1997. "Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks". *Journal of Applied Meteorology* 36 (9): 1176–1190. doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Huang, S., R. Kumar, M. Flörke, T. Yang, Y. Hundecha, P. Kraft, C. Gao, A. Gelfan, S. Liersch, A. Lobanova, M. Strauch, F. van Ogtrop, J. Reinhardt, U. Haberlandt, and V. Krysanova. 2017. "Erratum to: Evaluation of an ensemble of regional hydrological models in 12 large-scale river basins worldwide (Climatic Change, 10.1007/s10584-016-1841-8)". *Climatic Change* 141 (3): 399–400. doi:10.1007/s10584-016-1895-7.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker. 2007. "The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales". *Journal of Hydrometeorology* 8 (1): 38–55. doi:10.1175/JHM560.1.

Hughes, D. A. 2006. "Comparison of satellite rainfall data with observations from gauging station networks". *Journal of Hydrology* 327 (3-4): 399–410. doi:10.1016/j.jhydrol.2005.11.041.

Ibrahim, M., D. Wisser, A. Ali, B. Diekkrüger, O. Seidou, A. Mariko, and A. Afouda. 2017. "Water Balance Analysis over the Niger Inland Delta-Mali: Spatio-Temporal Dynamics of the Flooded Area and Water Losses". *Hydrology* 4 (3): 40. doi:10.3390/hydrology4030040.

Igué, A. 2000. *The Use of Soil and Terrain Database for Land Evaluation Procedures: Case Study of Central Benin*. Hohenheime. Stuttgart, Germany: University of Hohenheim.

Iloeje, N. 1986. *A New Geography of West Africa*. 6th ed. Harlow, United Kingdom: Longman.

Itiveh, K., and G. Bigg. 2008. "The variation of discharge entering the Niger Delta system, 1951–2000, and estimates of change under global warming". *International Journal of Climatology* 28:659–666. doi:10.1002/joc.

IUSS Working Group WRB. 2006. *World reference base for soil resources 2006*, 116. Rome: Food and Agriculture Organization of the United Nations.

Jackisch, C., E. Zehe, L. Samaniego, and A. K. Singh. 2014. "An experiment to gauge an ungauged catchment: rapid data assessment and eco-hydrological modelling in a data-scarce rural catchment". *Hydrological Sciences Journal* 59 (12): 2103–2125. doi:10.1080/02626667.2013.870662.

Jalloh, A., M. D. Faye, H. Roy-Macauley, P. Sérémé, R. Zougmoré, T. S. Thomas, and G. C. Nelson. 2013. "Overview". In *West African agriculture and climate change*, ed. by A. Jalloh, G. C. Nelson, T. S. Thomas, R. Zougmoré, and H. Roy-Macauley. Washington, DC: International Food Policy Research Institute.

Janssen, P., Z. Deng, V. Mulindabigwi, and J. Röhrig. 2010. "Agriculture and food". In *Impacts of Global Change on the Hydrological Cycle in West and Northwest Africa*, ed. by P. Speth, M. Christoph, and B. Diekkrüger. Heidelberg, Germany: Springer.

Jobard, I., F. Chopin, J. C. Berges, and R. Roca. 2011. "An intercomparison of 10-day satellite precipitation products during West African monsoon". *International Journal of Remote Sensing* 32 (9): 2353–2376. doi:10.1080/01431161003698286.

Jovanovic, N., Q. Mu, R. D. H. Bugan, and M. Zhao. 2015. "Dynamics of MODIS evapotranspiration in South Africa". *Water SA* 41 (1): 79–91. doi:10.4314/wsa.v41i1.11.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie. 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.

Kasei, R., B. Diekkrger, and C. Leemhuis. 2010. "Drought frequency in the Volta Basin of West Africa". *Sustain Sci*, no. 5: 89–97. doi:10.1007/s11625-009-0101-5.

Katsanos, D., A. Retalis, and S. Michaelides. 2016. "Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period". *Atmospheric Research* 169:459–464. doi:10.1016/j.atmosres.2015.05.015.

Kauffeldt, A., F. Wetterhall, F. Pappenberger, P. Salamon, and J. Thielen. 2016. "Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level". *Environmental Modelling and Software* 75:68–76. doi:10.1016/j.envsoft.2015.09.009.

Kebede, A., B. Diekkrüger, and S. a. Moges. 2014. "Comparative study of a physically based distributed hydrological model versus a conceptual hydrological model for assessment of climate change response in the Upper Nile, Baro-Akobo basin: a case study of the Sore watershed, Ethiopia". *International Journal of River Basin Management*: 1–20. doi:10.1080/15715124.2014.917315.

Key, R. M. 1992. "An introduction to the crystalline basement of Africa". *Geological Society, London, Special Publications* 66 (1): 29–57. doi:10.1144/GSL.SP.1992.066.01.02.

Knyazikhin, Y., J. Glassy, J. Privette, Y. Tian, A. Lotsch, Y. Zhang, Y. Wang, J. Morisette, P. Votava, R. Myneni, R. Nemani, and S. Running. 1999. *MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document*. Visited on 11/22/2018. https://modis.gsfc.nasa.gov/data/atbd/atbd_mod15.pdf.

Koster, R. D., and M. J. Suarez. 1996. *Energy and Water Balance Calculations in the Mosaic LSM*, 9:69. Greenbelt, Maryland: Goddard Space Flight Center. https://gmao.gsfc.nasa.gov/pubs/docs/Koster130.pdf.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel. 2006. "World map of the Köppen-Geiger climate classification updated". *Meteorologische Zeitschrift* 15 (3): 259–263. doi:10.1127/0941-2948/2006/0130.

Koutsouris, A. J., D. Chen, and S. W. Lyon. 2016. "Comparing global precipitation data sets in eastern Africa: A case study of Kilombero Valley, Tanzania". *International Journal of Climatology* 36:2000–2014. doi:10.1002/joc.4476.

Krysanova, V., T. Vetter, S. Eisner, S. Huang, I. Pechlivanidis, M. Strauch, A. Gelfan, R. Kumar, V. Aich, B. Arheimer, A. Chamorro, A. van Griensven, D. Kundu, A. Lobanova, V. Mishra, S. Plötner, J. Reinhardt, O. Seidou, X. Wang, M. Wortmann, X. Zeng, and F. F. Hattermann. 2017. "Intercomparison of regional-scale hydrological models and climate change impacts projected for 12 large river basins worldwide—a synthesis". *Environmental Research Letters* 12 (10): 105002. doi:10.1088/1748-9326/aa8359.

Kumar, R., B. Livneh, and L. Samaniego. 2013a. "Toward computationally efficient large-scale hydrologic predictions with a multiscale regionalization scheme". *Water Resources Research* 49 (9): 5700–5714. doi:10.1002/wrcr.20431.

Kumar, R., L. Samaniego, and S. Attinger. 2013b. "Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations". *Water Resources Research* 49 (1): 360–379. doi:10.1029/2012WR012195.

Kumar, R., L. Samaniego, and S. Attinger. 2010. "The effects of spatial discretization and model parameterization on the prediction of extreme runoff characteristics". *Journal of Hydrology* 392 (1-2): 54–69. doi:10.1016/j.jhydrol.2010.07.047.

Kusche, J. 2007. "Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models". *Journal of Geodesy* 81 (11): 733–749. doi:10.1007/s00190-007-0143-3.

Lawrence, D. M., K. W. Oleson, M. G. Flanner, P. E. Thornton, S. C. Swenson, P. J. Lawrence, X. Zeng, Z.-L. Yang, S. Levis, K. Sakaguchi, G. B. Bonan, and A. G. Slater. 2011. "Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model". *Journal of Advances in Modeling Earth Systems* 3 (1). doi:10.1029/2011MS00045.

Lehner, B., K. Verdin, and A. Jarvis. 2013. *HydroSHEDS Technical Documentation Version 1.2*. Washington, DC: World Wildlife Fund US. http://www.hydrosheds.org.

Lehner, B., C. R. Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, J. Magome, C. Nilsson, J. C. Robertson, R. Rödel, N. Sindorf, and D. Wisser. 2011. "High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management". *Frontiers in Ecology and the Environment* 9 (9): 494–502. doi:10.1890/100125.

Lehner, B., K. Verdin, and A. Jarvis. 2008. "New global hydrography derived from spaceborne elevation data". *Eos* 89 (10): 93–94. doi:10.1029/2008EO100001.

Li, D., X. Ding, and J. Wu. 2015. "Simulating the regional water balance through hydrological model based on TRMM satellite rainfall data". *Hydrology and Earth System Sciences Discussions* 12 (2): 2497–2525. doi:10.5194/hessd-12-2497-2015.

Li, K. Y., M. T. Coe, and N. Ramankutty. 2005. "Investigation of hydrological variability in West Africa using land surface models". *Journal of Climate* 18 (16): 3173–3188. doi:10.1175/JCLI3452.1.

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

Liu, Y. Y., W. A. Dorigo, R. M. Parinussa, R. A. M. De Jeu, W. Wagner, M. F. McCabe, J. P. Evans, and A. I. J. M. Van Dijk. 2012. "Trend-preserving blending of passive and active microwave soil moisture retrievals". *Remote Sensing of Environment* 123 (October 2006): 280–297. doi:10.1016/j.rse.2012.03.014.

Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. De Jeu, W. Wagner, A. I. J. M. Van Dijk, M. F. McCabe, and J. P. Evans. 2011. "Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals". *Hydrology and Earth System Sciences* 15 (2): 425–436. doi:10.5194/hess-15-425-2011.

Livneh, B., R. Kumar, and L. Samaniego. 2015. "Influence of soil textural properties on hydrologic fluxes in the Mississippi river basin". *Hydrological Processes* 29 (21): 4638–4655. doi:10.1002/hyp.10601.

Long, D., L. Longuevergne, and B. R. Scanlon. 2015. "Global analysis of approaches for deriving total water storage changes from GRACE satellites". *Water Resources Research* 51 (4): 2574–2594. doi:10.1002/2014WR016853.

Longuevergne, L., B. R. Scanlon, and C. R. Wilson. 2010. "GRACE hydrological estimates for small basins: Evaluating processing approaches on the High Plains aquifer, USA". *Water Resources Research* 46 (11): 1–15. doi:10.1029/2009WR008564.

López, P. L., E. H. Sutanudjaja, J. Schellekens, G. Sterk, and M. F. Bierkens. 2017. "Calibration of a large-scale hydrological model using satellite-based soil moisture and evapotranspiration products". *Hydrology and Earth System Sciences* 21 (6): 3125–3144. doi:10.5194/hess-21-3125-2017.

Maidment, R. I., D. I. F. Grimes, R. P. Allan, H. Greatrex, O. Rojas, and O. Leo. 2013. "Evaluation of satellite-based and model re-analysis rainfall estimates for Uganda". *Meteorological Applications* 20 (3): 308–317. doi:10.1002/met.1283.

Maidment, R. I., D. Grimes, R. P. Allan, E. Tarnavsky, M. Stringer, T. Hewison, R. Roebeling, and E. Black. 2014. "The 30 year TAMSAT African Rainfall Climatology And Time series (TARCAT) data set". *Journal of Geophysical Research: Atmospheres* 119:10619–10644. doi:10.1002/2014JD021927.

Malagò, A., D. Efstathiou, F. Bouraoui, N. P. Nikolaidis, M. Franchini, G. Bidoglio, and M. Kritsotakis. 2016. "Regional scale hydrologic modeling of a karst-dominant geomorphology: The case study of the Island of Crete". *Journal of Hydrology* 540:64–81. doi:10.1016/j.jhydrol.2016.05.061.

Martens, B., D. G. Miralles, H. Lievens, R. Van Der Schalie, R. A. De Jeu, D. Fernández-Prieto, H. E. Beck, W. A. Dorigo, and N. E. Verhoest. 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.

Masih, I., S. Maskey, and P. Trambauer. 2014. "A review of droughts on the African continent : a geospatial and long-term perspective". *Hydrology and Earth System Science*, no. 18: 3635–3649. doi:10.5194/hess-18-3635-2014.

Mayer-Gürr, T., S. Behzadpour, M. Ellmer, A. Kvas, B. Klinger, and N. Zehentner. 2016. *ITSG-Grace2016 - Monthly and Daily Gravity Field Solutions from GRACE*. Tech. rep. Graz, Austria: GFZ Data Services. https://www.tugraz.at/institute/ifg/downloads/gravity-field-models/itsg-grace2016/.

Milzow, C., P. E. Krogh, and P. Bauer-Gottwein. 2011. "Combining satellite radar altimetry, SAR surface soil moisture and GRACE total storage changes for hydrological model calibration in a large poorly gauged catchment". *Hydrology and Earth System Sciences* 15 (6): 1729–1743. doi:10.5194/hess-15-1729-2011.

Miralles, D. G., T. R. Holmes, R. A. De Jeu, J. H. Gash, A. G. Meesters, and A. J. Dolman. 2011. "Global land-surface evaporation estimated from satellite-based observations". *Hydrology and Earth System Sciences* 15 (2): 453–469. doi:10.5194/hess-15-453-2011.

Mishra, A. K., and V. P. Singh. 2010. "Review paper A review of drought concepts". *Journal of Hydrology* 391:202–216. doi:10.1016/j.jhydrol.2010.07.012.

Moazami, S., S. Golian, M. R. Kavianpour, and Y. Hong. 2013. "Comparison of PERSIANN and V7 TRMM Multi-satellite Precipitation Analysis (TMPA) products with rain gauge data over Iran". *International Journal of Remote Sensing* 34 (April 2015): 8156–8171. doi:10.1080/01431161.2013.833360.

Morgan, W., and J. Pugh. 1969. *West Africa*. Frome and London, Great Britain: Butler & Tanner.

Moriasi, D. N., J. G. Arnold, M. W. V. Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith. 2007. "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations". *Transactions of the ASABE 2007* 50 (3): 885–900.

Mu, Q., F. Heinsch, M. Zhao, and S. Running. 2007. "Development of a global evapotranspiration algorithm based on MODIS and global meteorology data". *Remote Sensing of Environment* 106 (3): 285–304. doi:10.1016/j.rse.2006.07.007.

Mu, Q., M. Zhao, and S. W. Running. 2011. "Improvements to a MODIS global terrestrial evapotranspiration algorithm". *Remote Sensing of Environment* 115 (8): 1781–1800. doi:10.1016/j.rse.2011.02.019.

Mul, M., E. Obubie, R. Appoh, K. Kankam-Yeboah, E. Bekoe-Obeng, B. Amisigo, F. Logah, B. Ghansah, and M. McCartney. 2015. *Water Ressources Assessment of the Volta River Basin*. 78. Colombo, Sri Lanka. doi:10.5337/2015.220.

Myneni, R., Y. Knyazikhin, and T. Park. 2015a. *MCD15A2 MODIS/Combined Terra+Aqua Leaf Area Index/FPAR Daily L4 Global 1km SIN Grid*. Visited on 11/22/2018. https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/lai-and-fpar/MCD15A2/.

— . 2015b. *MOD15A2H MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006 [LAI 500m]*. NASA EOSDIS Land Processes DAAC. doi: 10.5067/MODIS/MOD15A2H.006. doi:10.5067/MODIS/MOD15A2H.006.

Nash, J. E., and J. V. Sutcliffe. 1970. "River Flow Forecasting Through Conceptual Models Part I-a Discussion of Principles". *Journal of Hydrology* 10:282–290. doi:10.1016/0022-1694(70)90255-6.

Ndehedehe, C., J. Awange, N. Agutu, M. Kuhn, and B. Heck. 2016. "Understanding Changes in Terrestrial Water Storage over West Africa between 2002 and 2014". *Advances in Water Resources* 88 (January): 211–230. doi:10.1016/j.advwatres.2015.12.009.

Neitsch, S., J. Arnold, J. Kiniry, R. Srinivasan, and J. Williams. 2002. *Soil and Water Assessment Tool User's Manual*. College Station, TX, USA: Texas Water Resources Institute. http://swat.tamu.edu/media/1294/swatuserman.pdf.

Neitsch, S., J. Arnold, J. Kiniry, and J. Williams. 2011. *Soil & Water Assessment Tool Theoretical Documentation Version 2009*. Tech. rep. College Station, TX, USA.

Nicholson, S. E., B. Some, J. McCollum, E. Nelkin, D. Klotter, Y. Berte, B. M. Diallo, I. Gaye, G. Kpabeba, O. Ndiaye, J. N. Noukpozounkou, M. M. Tanu, A. Thiam, A. A. Toure, and A. K. Traore. 2003. "Validation of TRMM and Other Rainfall Estimates with a High-Density Gauge Dataset for West Africa. Part II: Validation of TRMM Rainfall Products". *Journal of Applied Meteorology* 42 (10): 1355–1368. doi:10.1175/1520-0450(2003)042<1355:VOTAOR>2.0.CO;2.

Nikulin, G., C. Jones, F. Giorgi, G. Asrar, M. Büchner, R. Cerezo-Mota, O. B. Christensen, M. Déqué, J. Fernandez, A. Hänsler, E. van Meijgaard, P. Samuelsson, M. B. Sylla, and L. Sushama. 2012. "Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations". *Journal of Climate* 25 (18): 6057–6078. doi:10.1175/JCLI-D-11-00375.1.

Novella, N. S., and W. M. Thiaw. 2013. "African rainfall climatology version 2 for famine early warning systems". *Journal of Applied Meteorology and Climatology* 52 (3): 588–606. doi:10.1175/JAMC-D-11-0238.1.

Ojo, O. 1977. *The Climates of West Africa*. London, England: Heinemann.

Oldenborgh, G. van, M. Collins, J. Arblaster, J. Christensen, J. Marotzke, S. Power, M. Rummukainen, and T. Zhou. 2013. "IPCC, 2013: Annex I: Atlas of Global and Regional Climate Projections". In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, ed. by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, 1311–1394. Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9781107415324.029.

OMVS, AFD, and WMO. 2012. *Senegal HYCOS*. Tech. rep. Lyon, France: Compagnie Nationale du Rhône.

Orth, R., M. Staudinger, S. I. Seneviratne, J. Seibert, and M. Zappa. 2015. "Does model performance improve with complexity? A case study with three hydrological models". *Journal of Hydrology* 523:147–159. doi:10.1016/j.jhydrol.2015.01.044.

Perrin, C., C. Michel, and V. Andréassian. 2003. "Improvement of a parsimonious model for streamflow simulation". *Journal of Hydrology* 279:275–289. doi:10.1016/S0022-1694(03)00225-7.

Pfeifroth, U., J. Trentmann, A. H. Fink, and B. Ahrens. 2016. "Evaluating satellite-based diurnal cycles of precipitation in the African tropics". *Journal of Applied Meteorology and Climatology*, no. January: 23–39. doi:10.1175/JAMC-D-15-0065.1.

Pignotti, G., H. Rathjens, R. Cibin, I. Chaubey, and M. Crawford. 2017. "Comparative analysis of HRU and grid-based SWAT models". *Water* 9 (4): 272. doi:10.3390/w9040272.

Pokhrel, P., H. V. Gupta, and T. Wagener. 2008. "A spatial regularization approach to parameter estimation for a distributed watershed model". *Water Resources Research* 44 (12): 1–16. doi:10.1029/2007WR006615.

Pombo, S., R. P. de Oliveira, and A. Mendes. 2015. "Validation of remote-sensing precipitation products for Angola". *Meteorological Applications* 22 (3): 395–409. doi:10.1002/met.1467.

Poméon, T., B. Diekkrüger, A. Springer, J. Kusche, and A. Eicker. 2018. "Multi-Objective Validation of SWAT for Sparsely-Gauged West African River Basins—A Remote Sensing Approach". *Water* 10 (451): 22. doi:10.3390/w10040451.

Poméon, T., D. Jackisch, and B. Diekkrüger. 2017. "Evaluating the performance of remotely sensed and reanalysed precipitation data over West Africa using HBV light". *Journal of Hydrology* 547:222–235. doi:10.1016/j.jhydrol.2017.01.055.

Qi, W., C. Zhang, G. Fu, C. Sweetapple, and H. Zhou. 2016. "Evaluation of global fine-resolution precipitation products and their uncertainty quantification in ensemble discharge simulations". *Hydrology and Earth System Sciences* 20 (2): 903–920. doi:10.5194/hess-20-903-2016.

Qiao, L., R. B. Herrmann, and Z. Pan. 2013. "Parameter Uncertainty Reduction for SWAT Using Grace, Streamflow, and Groundwater Table Data for Lower Missouri River Basin". *Journal of the American Water Resources Association* 49 (2): 343–358. doi:10.1111/jawr.12021.

Radchenko, I., L. Breuer, I. Forkutsa, and H.-G. Frede. 2014. "Simulating Water Resource Availability under Data Scarcity—A Case Study for the Ferghana Valley (Central Asia)". *Water* 6 (11): 3270–3299. doi:10.3390/w6113270.

Rahaman, M. M., and O. Varis. 2005. "Integrated water resources management: evolution, prospects and future challenges". *Sustainability: Science, Practice and Policy* 1 (1): 15–21. doi:10.1080/15487733.2005.11907961.

Rajib, M. A., V. Merwade, and Z. Yu. 2016. "Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed / in-situ soil moisture". *Journal of Hydrology* 536:192–207. doi:10.1016/j.jhydrol.2016.02.037.

Rakovec, O., R. Kumar, S. Attinger, and L. Samaniego. 2016a. "Improving the realism of hydrologic model functioning through multivariate parameter estimation". *Water Resources Research* 52 (10): 7779–7792. doi:10.1002/2016WR019430.

Rakovec, O., R. Kumar, J. Mai, M. Cuntz, S. Thober, M. Zink, S. Attinger, D. Schäfer, M. Schrön, and L. Samaniego. 2016b. "Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins". *Journal of Hydrometeorology* 17 (1): 287–307. doi:10.1175/JHM-D-15-0054.1.

Ramarohetra, J., B. Sultan, C. Baron, T. Gaiser, and M. Gosset. 2013. "How satellite rainfall estimate errors may impact rainfed cereal yield simulation in West Africa". *Agricultural and Forest Meteorology* 180:118–131. doi:10.1016/j.agrformet.2013.05.010.

Rateb, A., C.-Y. Kuo, M. Imani, K.-H. Tseng, W.-H. Lan, K.-E. Ching, and T.-P. Tseng. 2017. "Terrestrial Water Storage in African Hydrological Regimes Derived from GRACE Mission Data: Intercomparison of Spherical Harmonics, Mass Concentration, and Scalar Slepian Methods". *Sensors* 17 (3): 566. doi:10.3390/s17030566.

Rathjens, H., and N. Oppelt. 2012. "SWATgrid: An interface for setting up SWAT in a grid-based discretization scheme". *Computers and Geosciences* 45:161–167. doi:10.1016/j.cageo.2011.11.004.

Rientjes, T. H. M., L. P. Muthuwatta, M. G. Bos, M. J. Booij, and H. A. Bhatti. 2013. "Multi-variable calibration of a semi-distributed hydrological model using streamflow data and satellite-based evapotranspiration". *Journal of Hydrology* 505:276–290. doi:10.1016/j.jhydrol.2013.10.006.

Rietbroek, R., S. E. Brunnabend, J. Kusche, and J. Schröter. 2012a. "Resolving sea level contributions by identifying fingerprints in time-variable gravity and altimetry". *Journal of Geodynamics* 59-60:72–81. doi:10.1016/j.jog.2011.06.007.

Rietbroek, R., M. Fritsche, S. E. Brunnabend, I. Daras, J. Kusche, J. Schröter, F. Flechtner, and R. Dietrich. 2012b. "Global surface mass from a new combination of GRACE, modelled OBP and reprocessed GPS data". *Journal of Geodynamics* 59-60:64–71. doi:10.1016/j.jog.2011.02.003.

Rubel, F., K. Brugger, K. Haslinger, and I. Auer. 2017. "The climate of the European Alps: Shift of very high resolution Köppen-Geiger climate zones 1800-2100". *Meteorologische Zeitschrift* 26 (2): 115–125. doi:10.1127/metz/2016/0816.

Rusli, S. R., D. Yudianto, and J.-t. Liu. 2015. "Effects of temporal variability on HBV model calibration". *Water Science and Engineering* 8 (4): 291–300. doi:10.1016/j.wse.2015.12.002.

Saha, S., S. Moorthi, H. L. Pan, X. Wu, J. Wang, S. Nadiga, P. Tripp, R. Kistler, J. Woollen, D. Behringer, H. Liu, D. Stokes, R. Grumbine, G. Gayno, J. Wang, Y. T. Hou, H. Y. Chuang, H. M. H. Juang, J. Sela, M. Iredell, R. Treadon, D. Kleist, P. Van Delst, D. Keyser, J. Derber, M. Ek, J. Meng, H. Wei, R. Yang, S. Lord, H. Van Den Dool, A. Kumar, W. Wang, C. Long, M. Chelliah, Y. Xue, B. Huang, J. K. Schemm, W. Ebisuzaki, R. Lin, P. Xie, M. Chen, S. Zhou, W. Higgins, C. Z. Zou, Q. Liu, Y. Chen, Y. Han, L. Cucurull, R. W. Reynolds, G. Rutledge, and M. Goldberg. 2010. "The NCEP climate forecast system reanalysis". *Bulletin of the American Meteorological Society* 91 (8): 1015–1057. doi:10.1175/2010BAMS3001.1.

Samaniego, L., R. Kumar, L. Breuer, A. Chamorro, M. Flörke, I. G. Pechlivanidis, D. Schäfer, H. Shah, T. Vetter, M. Wortmann, and X. Zeng. 2017a. "Propagation of forcing and model uncertainties on to hydrological drought characteristics in a multi-model century-long experiment in large river basins". *Climatic Change* 141 (3): 435–449. doi:10.1007/s10584-016-1778-y.

Samaniego, L., R. Kumar, and C. Jackisch. 2011. "Predictions in a data-sparse region using a regionalized grid-based hydrologic model driven by remotely sensed data". *Hydrology Research* 42 (5): 338–355. doi:10.2166/nh.2011.156.

Samaniego, L., J. Brenner, M. Cuntz, C. M. Demirel, M. Kaluza, R. Kumar, B. Langenberg, J. Mai, O. Rokovec, D. Schäfer, M. Schrön, S. Stisen, S. Thober, and M. Zink. 2017b. *The mesoscale Hydrologic Model - Documentation for verion 5.8*. 1099. Leipzig: Helmholtz Centre for Environmental Research - UFZ.

Samaniego, L., R. Kumar, and S. Attinger. 2010. "Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale". *Water Resources Research* 46 (5): 1–25. doi:10.1029/2008WR007327.

Samaniego, L., R. Kumar, and M. Zink. 2013. "Implications of Parameter Uncertainty on Soil Moisture Drought Analysis in Germany". *Journal of Hydrometeorology* 14 (1): 47–68. doi:10.1175/JHM-D-12-075.1.

Schamm, K., M. Ziese, K. Raykova, A. Becker, P. Finger, A. Meyer-Christoffer, and U. Schneider. 2015. *GPCC Full Data Daily Version 1.0 at 1.0°: Daily Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data*. Visited on 06/07/2016. ftp://ftp.dwd.de/pub/data/gpcc/html/fulldata-daily_v1_doi_download.html.

Schlüter, T. 2006. *Geological Atlas of Africa*. Berlin, Heiderberg, New York: Springer.

Schneider, U., M. Ziese, A. Becker, A. Meyer-Christoffer, and P. Finger. 2015. *Global Precipitation Analysis Products of the GPCC*. Visited on 01/01/2016. ftp://ftp.dwd.de/pub/data/gpcc/PDF/GPCC_intro_products_2008.pdf.

Schulla, J., and K. Jasper. 2007. *Model description WaSiM-ETH (Water balance Simulation Model ETH)*. 1–181. Zurich, Switzerland: ETH Zurich.

Schuol, J., and K. C. Abbaspour. 2006a. "Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa". *Advances in Geosciences* 9:137–143. doi:10.5194/adgeo-9-137-2006.

— . 2006b. "Using monthly weather statistics to generate daily data in a SWAT model application to West Africa". *Ecological Modelling* 201 (3-4): 301–311. doi:10.1016/j.ecolmodel.2006.09.028.

Schuol, J., K. C. Abbaspour, R. Srinivasan, and H. Yang. 2008a. "Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model". *Journal of Hydrology* 352 (1-2): 30–49. doi:10.1016/j.jhydrol.2007.12.025.

Schuol, J., K. C. Abbaspour, H. Yang, R. Srinivasan, and A. J. B. Zehnder. 2008b. "Modeling blue and green water availability in Africa". *Water Resources Research* 44 (7): 1–18. doi:10.1029/2007WR006609.

Sebastian and Kate. 2009. *Agro-ecological Zones of Africa*. Washington, DC: International Food Policy Research Institute (datasets). http://hdl.handle.net/1902.1/22616.

Seibert, J. 2000. "Multi-criteria calibration of a conceptual runoff model using a genetic algorithm". *Hydrology and Earth System Science* 4 (2): 215–224. doi:10.5194/hess-4-215-2000.

Seibert, J., and K. J. Beven. 2009. "Gauging the ungauged basin: how many discharge measurements are needed?" *Hydrol. Earth Syst. Sci. Discuss.* doi:10.5194/hessd-6-2275-2009.

Seibert, J., and M. J. P. Vis. 2012. "Teaching hydrological modeling with a user-friendly catchment-runoff-model software package". *Hydrology and Earth System Sciences* 16 (9): 3315–3325. doi:10.5194/hess-16-3315-2012.

Seibert, J. 1997. "Estimation of Parameter Uncertainty in the HBV Model". *Nordic Hydrology* 28 (1982): 247–262.

— . 2005. *HBV light version 2 User 's Manual*. November. Stockholm, Sweden: Stockholm University.

— . 1999. "Regionalisation of parameters for a conceptual rainfall-runoff model". *Agricultural and Forest Meteorology* 98-99:279–293. doi:10.1016/S0168-1923(99)00105-7.

Shanahan, T. M., J. T. Overpeck, K. J. Anchukaitis, J. W. Beck, J. E. Cole, D. L. Dettman, J. A. Peck, C. A. Scholz, and J. W. King. 2009. "Atlantic Forcing of Persistent Drought in West Africa". *Science* 324 (April): 377–380.

Singh, V. 1995. *Computer models of watershed hydrology*. Baton Rouge: Water Resources Publications.

Srinivasan, R., T. S. Ramanarayanan, J. G. Arnold, and S. T. Bednarz. 1998. "Large area hydrologic modeling and assessment part II: model application". *Journal of the American Water Resources Association* 34 (1): 91–101.

Strauch, M., and M. Volk. 2013. "SWAT plant growth modification for improved modeling of perennial vegetation in the tropics". *Ecological Modelling* 269:98–112. doi:10.1016/j.ecolmodel.2013.08.013.

Sun, Z., M. Gebremichael, J. Ardö, A. Nickless, B. Caquet, L. Merboldh, and W. Kutschi. 2012. "Estimation of daily evapotranspiration over Africa using MODIS/Terra and SEVIRI/MSG data". *Atmospheric Research* 112:35–44. doi:10.1016/j.atmosres.2012.04.005.

Tanaka, M., T. A. Adjadeh, S. Tanaka, and T. Sugimura. 2002. "Water surface area measurement of Lake Volta using SSM/I 37-GHz polarization difference in rainy season". *Advances in Space Research* 30 (11): 2501–2504. doi:10.1016/S0273-1177(02)80320-9.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber. 2004. "The gravity recovery and climate experiment: Mission overview and early results". *Geophysical Research Letters* 31 (9): 1–4. doi:10.1029/2004GL019920.

Tarnavsky, E., D. Grimes, R. Maidment, E. Black, R. P. Allan, M. Stringer, R. Chadwick, and F. Kayitakire. 2014. "Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present". *Journal of Applied Meteorology and Climatology* 53 (12): 2805–2822. doi:10.1175/JAMC-D-14-0016.1.

The NOAA Climate Prediction Center. 2002. "African Rainfall Estimation Algorithm Version 2.0". http://www.cpc.ncep.noaa.gov/products/fews/RFE2.0_tech.pdf.

Thiemig, V., R. Rojas, M. Zambrano-Bigiarini, and A. De Roo. 2013. "Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin". *Journal of Hydrology* 499:324–338. doi:10.1016/j.jhydrol.2013.07.012.

Thiemig, V., R. Rojas, M. Zambrano-Bigiarini, V. Levizzani, and A. De Roo. 2012. "Validation of Satellite-Based Precipitation Products Over Sparsely-Gauged African River Basins". *Journal of Hydrometeorology*: 120713072538009. doi:10.1175/JHM-D-12-032.1.

Thober, S., R. Kumar, J. Sheffield, J. Mai, D. Schäfer, and L. Samaniego. 2015. "Seasonal Soil Moisture Drought Prediction over Europe Using the North American Multi-Model Ensemble (NMME)". *Journal of Hydrometeorology* 16 (6): 2329–2344. doi:10.1175/JHM-D-15-0053.1.

Tobin, K. J., and M. E. Bennett. 2014. "Satellite precipitation products and hydrologic applications". *Water International*, no. June 2015: 1–21. doi:10.1080/02508060.2013.870423.

— . 2009. "Using SWAT to model streamflow in two river basins with ground and satellite precipitation data". *Journal of the American Water Resources Association* 45 (1): 253–271. doi:10.1111/j.1752-1688.2008.00276.x.

Tolson, B. A., and C. A. Shoemaker. 2007. "Dynamically dimensioned search algorithm for computationally efficient watershed model calibration". *Water Resources Research* 43 (1): 1–16. doi:10.1029/2005WR004723.

Tramblay, Y., V. Thiemig, A. Dezetter, and L. Hanich. 2016. "Evaluation of satellite-based rainfall products for hydrological modelling in Morocco". *Hydrological Sciences Journal* 61 (14): 2509–2519. doi:10.1080/02626667.2016.1154149.

Uebbing, B., J. Kusche, and E. Forootan. 2015. "Waveform Retracking for Improving Level Estimations From TOPEX/Poseidon, Jason-1, and Jason-2 Altimetry Observations Over African Lakes". *IEEE Transactions on Geoscience and Remote Sensing* 53 (4): 2211–2224. doi:10.1109/TGRS.2014.2357893.

Uhlenbrook, S., J. Seibert, C. Leibundgut, and A. Rodhe. 1999. "Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure". *Hydrological Sciences Journal* 44 (5): 779–797. doi:10.1080/02626669909492273.

USDA SCS. 1972. *National Engineering Handbook, Hydrology Section 4*. Ed. by V. Mockus. https://directives.sc.egov.usda.gov/OpenNonWebContent.aspx?content=18393.wba.

Wagner, S., H. Kunstmann, A. Bárdossy, C. Conrad, and R. R. Colditz. 2009. "Water balance estimation of a poorly gauged catchment in West Africa using dynamically downscaled meteorological fields and remote sensing information". *Physics and Chemistry of the Earth* 34 (4-5): 225–235. doi:10.1016/j.pce.2008.04.002.

Wagner, W., W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, and M. Ertl. 2012. "Fusion of Active and Passive Microwave Observations To Create an Essential Climate Variable Data Record on Soil Moisture". *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* I-7:315–321.

Wahr, J., M. Molenaar, and F. Bryan. 1998. "Time variability of the Earth's gravity field' Hydrological and oceanic effects and their possible detection using GRACE". *Journal of Geophysical Research* 103:30205–30229.

Weiß, M., and L. Menzel. 2008. "A global comparison of four potential evapotranspiration equations and their relevance to stream flow modelling in semi-arid environments". *Adv. Geosci* 18:15–23. doi:10.5194/adgeo-18-15-2008.

Werth, S., D. White, and D. W. Bliss. 2017. "GRACE Detected Rise of Groundwater in the Sahelian Niger River Basin". *Journal of Geophysical Research: Solid Earth* 122:10459–10477. doi:10.1002/2017JB014845.

WHO and UNICEF. 2017. *Progress on Drinking Water, Sanitation and Hygiene 2017*. Geneva, Switzerland: WHO / UNICEF. https://www.unicef.org/publications/index_96611.html.

Winchell, M., R. Srinivasan, M. Di Luzio, and J. Arnold. 2010. *ArcSWAT Interface for SWAT 2009*. Temple, Texas: USDA Agricultural Research Service.

Windmeijer, P., and W. Andriesse. 1993. *Inland Valleys in West Africa: An Agro-Ecological Characterization of Rice-Growing Environments*. Wageningen, The Netherlands: International Institute for Land Reclamation and Improvement.

WMO. 2018. *World Meteorological Organization World Weather - Station 325, Port-Harcourt*. Visited on 12/09/2018. http://worldweather.wmo.int/en/city.html?cityId=325.

Worqlul, A. W., B. Maathuis, A. A. Adem, S. S. Demissie, S. Langan, and T. S. Steenhuis. 2014. "Comparison of TRMM, MPEG and CFSR rainfall estimation with the ground observed data for the Lake Tana Basin, Ethiopia". *Hydrology and Earth System Sciences* 11 (7): 8013–8038. doi:10.5194/hess-18-4871-2014.

Worqlul, A. W., H. Yen, A. S. Collick, S. A. Tilahun, S. Langan, and T. S. Steenhuis. 2017. "Evaluation of CFSR, TMPA 3B42 and ground-based rainfall data as input for hydrological models, in data-scarce regions: The upper Blue Nile Basin, Ethiopia". *Catena* 152:242–251. doi:10.1016/j.catena.2017.01.019.

Wösten, J. H. M., Y. a. Pachepsky, and W. J. Rawls. 2001. "Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics". *Journal of Hydrology* 251 (3-4): 123–150. doi:10.1016/S0022-1694(01)00464-4.

Xie, H., L. Longuevergne, C. Ringler, and B. R. Scanlon. 2012. "Calibration and evaluation of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data". *Hydrology and Earth System Sciences* 16 (9): 3083–3099. doi:10.5194/hess-16-3083-2012.

Xie, P., S.-h. Yoo, R. Joyce, and Y. Yarosh. 2011. *Bias-Corrected CMORPH: A 13-Year Analysis of High-Resolution Global Precipitation*. Visited on 03/13/2018. http://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/REF/EGU_1104_Xie_bias-CMORPH.pdf.

Yéo, W. E., B. Tié, A. Goula, B. Diekkrüger, and A. Afouda. 2016. "Vulnerability and adaptation to climate change in the Comoe River Basin (West Africa)". *SpringerPlus*. doi:10.1186/s40064-016-2491-z.

Yira, Y., B. Diekkrüger, G. Steup, and A. Yaovi Bossa. 2017. "Impact of climate change on hydrological conditions in a tropical West African catchment using an ensemble of climate simulations". *Hydrology and Earth System Sciences* 21 (4): 2143–2161. doi:10.5194/hess-21-2143-2017.

Zelelew, M. B., and K. Alfredsen. 2013. "Hydrological model parameter transferability studies to estimate runoff at ungauged catchments". *Hydrological Sciences Journal*, no. August: 130827075642009. doi:10.1080/02626667.2013.838003.

Zink, M., R. Kumar, M. Cuntz, and L. Samaniego. 2017. “A high-resolution dataset of water fluxes and states for Germany accounting for parametric uncertainty”. *Hydrology and Earth System Sciences* 21 (3): 1769–1790. doi:10.5194/hess-21-1769-2017.

Zink, M., J. Mai, M. Cuntz, and L. Samaniego. 2018. “Conditioning a Hydrologic Model Using Patterns of Remotely Sensed Land Surface Temperature”. *Water Resources Research* 54 (4): 2976–2998. doi:10.1002/2017WR021346.

Zink, M., L. Samaniego, R. Kumar, S. Thober, J. Mai, D. Schäfer, and A. Marx. 2016. “The German drought monitor”. *Environmental Research Letters* 11 (7). doi:10.1088/1748-9326/11/7/074002.

A Appendix

Table A.1: mHM parameters.

Parameter	Lower Threshold	Upper Threshold
	Interception	
canopyInterceptionFactor	0.15	0.4
	Soil Moisture	
orgMatterContent_forest	0	20
orgMatterContent_impervious	0	1
orgMatterContent_pervious	0	4
PTF_lower66_5_constant	0.6462	0.9506
PTF_lower66_5_clay	0.0001	0.0029
PTF_lower66_5_Db	-0.3727	-0.1871
PTF_higher66_5_constant	0.5358	1.1232
PTF_higher66_5_clay	-0.0055	0.0049
PTF_higher66_5_Db	-0.5513	-0.0913
PTF_Ks_constant	-1.2	-0.285
PTF_Ks_sand	0.006	0.026
PTF_Ks_clay	0.003	0.013
rootFractionCoefficient_forest	0.9	0.999
rootFractionCoefficient_impervious	0.9	0.95
rootFractionCoefficient_pervious	0.001	0.09
infiltrationShapeFactor	1	4
	Direct Sealed Area Runoff	
imperviousStorageCapacity	0	5
	Potential Evapotranspiration	
minCorrectionFactorPET	0.7 (0.9)	1.3 (0.96)
maxCorrectionFactorPET	0 (0.17)	0.2
aspectTresholdPET	160	200
HargreavesSamaniCoeff	0.0016 (0.0021)	0.003 (0.0027)
	Interflow	
interflowStorageCapacityFactor	75	200
interflowRecession_slope	0	10
fastInterflowRecession_forest	1	3
slowInterflowRecession_Ks	1	30
exponentSlowInterflow	0.05	0.3
	Percolation	
rechargeCoefficient	0	50 (200)
	Geological Parameter	
GeoParam(1)	1	1000 (1500)

Given are the parameters calibrated in mHM and their standard ranges. Values in brackets signify modified ranges to improve actual evapotranspiration simulations of the Q calibration scheme. During the Q/ET calibration, default ranges were used. For further details, see Samaniego et al. (2017b) and www.ufz.de/mhm.