Year : 2020

SPATIALLY EXPLICIT HYDROLOGICAL MODELLING FOR WATER ACCOUNTING UNDER CLIMATE CHANGE IN THE VOLTA RIVER BASIN IN WEST AFRICA

Dembélé Moctar

Dembélé Moctar, 2020, SPATIALLY EXPLICIT HYDROLOGICAL MODELLING FOR WATER ACCOUNTING UNDER CLIMATE CHANGE IN THE VOLTA RIVER BASIN IN WEST AFRICA

Originally published at : Thesis, University of Lausanne

Posted at the University of Lausanne Open Archive http://serval.unil.ch
Document URN : urn:nbn:ch:serval-BIB_08A0D80C87B74

UNIL | Université de Lausanne

Faculté des Géosciences et de l'Environnement
Institut des Dynamiques de la Surface Terrestre

SPATIALLY EXPLICIT HYDROLOGICAL MODELLING FOR WATER ACCOUNTING UNDER CLIMATE CHANGE IN THE VOLTA RIVER BASIN IN WEST AFRICA

THÈSE DE DOCTORAT

Présentée à la
Faculté des Géosciences et de l'Environnement
de l'Université de Lausanne
Par

Moctar Dembélé

Titulaire d'un
Master en Sciences et Ingénierie de l'Eau et de l'Environnement
De l'Institut International d'Ingénierie de l'Eau et de l'Environnement (2iE), Burkina Faso

JURY

Prof. Grégoire Mariéthoz	Directeur de thèse (Université de Lausanne)
Prof. Bettina Schaefli	co-Directrice de thèse (Université de Berne)
Prof. Emmanuel Reynard	Expert (Université de Lausanne)
Prof. Justin Sheffield	Expert (University of Southampton)
Prof. Nick van de Giesen	Expert (Delft University of Technology)
Prof. Dominique Arlettaz	Président du Jury (Université de Lausanne)

Lausanne, 2020

UNIL | Université de Lausanne
Faculté des géosciences et de l'environnement
bâtiment Géopolis bureau 4631

IMPRIMATUR

Vu le rapport présenté par le jury d'examen, composé de

Président de la séance publique :	M. le Professeur Dominique Arlettaz
Président du colloque :	M. le Professeur Dominique Arlettaz
Directeur de thèse :	M. le Professeur Grégoire Mariéthoz
Co-directrice de thèse :	Mme la Professeure Bettina Schaefli
Expert interne :	M. le Professeur Emmanuel Reynard
Expert externe :	M. le Professeur Justin Sheffield
Expert externe :	M. le Professeur Nick van de Giesen

Le Doyen de la Faculté des géosciences et de l'environnement autorise l'impression de la thèse de

Monsieur Moctar DEMBELE

Titulaire d'un
Master en ingénierie de l'eau
Du 2iE (Ouagadougou)

intitulée

Spacially Explicit Hydrological Modelling for Water Accounting under Climate Change in the Volta River Basin in West Africa

Lausanne, le 11 septembre 2020

Pour le Doyen de la Faculté des géosciences et de l'environnement

Professeur Dominique Arlettaz

This PhD research was supported by the Swiss Government Excellence Scholarship (2016.0533 / Burkina Faso / OP), and the Doc.Mobility fellowship (SNF, P1LAP2_178071) of the Swiss National Science Foundation.

A toi papa, pour m'avoir donné l'envie de faire de la recherche, car depuis l'enfance, j'ai voulu « faire comme toi ». De là-haut, tu as su veiller sur moi tout au long de cette aventure.

A toi maman, pour l'éducation et l'amour que tu as toujours eu à mon égard.

Acknowledgements - Remerciements

"If I have seen further than others, it is by standing upon the shoulders of giants". Isaac Newton

I am reaching the other riverbank. Yes, now I am there, but it seems to be the beginning of the story, I hope. Undertaking a doctoral research was a very enriching adventure for me. I had a dream that finally came true. I grew up under the mentorship and leadership of people whose support invigorated my courage to go up the hill (there is still some way to go…). I hope they remain satisfied with this piece of work as a reward of their contributions to my education.

I was honoured to do my PhD under the supervision of Prof. Bettina Schaefli and Prof. Grégoire Mariéthoz. Bettina and Grégoire, you have been of great help for the achievement of my PhD. I am grateful for the opportunity you gave me to join your research groups for four years, starting back in September 2016. If I were given another chance to choose my PhD supervisors, you would be my favourites. I have learnt a lot by your side. I appreciate your patience, availability and time for discussing research ideas and turning them into publications. The freedom I had for conducting my research was unique. You encouraged me to seize opportunities to fully explore and expand my research potential. You facilitated collaborations with other researchers and you opened doors for conferences, summer courses and mobility to research organizations all over the world. There was no limit to your enthusiasm to push me towards a successful PhD journey. I am wholeheartedly grateful for your relentless support and commitment. *Merci Bettina! Merci Grégoire!*

I thank Dr Sander Zwart for guiding me through my first steps in scientific research when I was a postgraduate fellow in his research group at AfricaRice Center for two years before starting my PhD. Sander, you encouraged me to pursue a PhD in water accounting and served as an external advisor during my PhD. I am sincerely thankful for your support.

Natalie, I greatly acknowledge your precious insights during the conception of my research proposal and your contribution to my PhD papers.

I thank Prof. Hubert Savenije for hosting me in his research group at TU Delft where I worked under the supervision of Prof. Markus Hrachowitz for one year. Markus, I am grateful for your time and your very constructive discussions. I thank Prof. Simon Dadson for the research discussions, the networking opportunities and the hospitality during my one-month stay at the UK CEH under the HydroJules visiting scientist fellowship.

I thank Dr. Jacob Tumbulto for welcoming me at the VBA in Ouagadougou, for sharing data and documentations on the Volta basin at the beginning of my PhD and for his time for discussing my research. I thank Dr. Fabio Oriani for spending time to guide me through his MATLAB code on Direct Sampling, which helped me a lot to take over the method. I appreciate the availability of Dr. Elga Salvadore who shared some evaporation data with me and spent time to discuss research questions on evaporation and WA+. I am grateful for the assistance that Dr. Mathieu Vrac provided to me for handling his multivariate bias correction method. I thank all my co-authors for the papers related to my PhD thesis. You provided invaluable contributions.

I thank the members of the Jury of my PhD (names on the first page) for taking the time to evaluate my work. I thank Prof. Wim Bastiaanssen for being an evaluator during my first year PhD committee.

I thank all my colleagues and friends for the lovely time we spent together at Unil: Anthony, Elisa, Fabio, Gustavo, Harsh, Mathieu, Inigo, Raphaël, Sassi, Thibault, Tom, Qiyu, and the list goes on.

My thanks go to Hilaire Sawadogo for sharing with me the announcement of the Swiss Confederation scholarship, and Prof. Mahamadou Koïta and Prof. Amadou Keïta for recommending my PhD application. This PhD research would not have been possible without the financial support of the Swiss Confederation (3 years) and the Swiss National Science Foundation (1 year). I am grateful for their assistance. I also thank the SASME (Marc Simond and his colleagues) for the administrative management of my PhD scholarship at the University of Lausanne. My thanks go to Sabrina Damiani, Carole Schrocker and Joëlle Richard for the administrative support they provided at the secretary's office of IDYST and FGSE.

I thank all the people who have been by my side to provide various forms of support during my PhD journey. Even not listed here, you remain loved and special to me. Words cannot overpower feelings.

Mes sincères remerciements vont à ma famille et à mes amis qui m'ont soutenu et encouragé tout au long de cette aventure.

Grand Merci! Awnitché!

Moctar

Summary

Competition for scarce water resources in the Volta River Basin (VRB) of West Africa will increase in the near future due to the combined effects of rapid population growth and climate change. Residents are dependent on subsistence, mainly rainfed agriculture that is sensitive to climate variabilities. Recurrent floods and droughts damage properties and take lives. Information on water resources and their future trends is fundamental for water actors, as the basis for proper management and implementation of adequate measures to bolster resilience to water scarcity and foster water security.

This PhD thesis proposes a novel and clear demonstration of combining the Water Accounting Plus (WA+) framework with hydrological modelling and climate change scenarios to report on the current and future states of water resources in the VRB. WA+ is a standardized framework that provides a comprehensive view of the water resources in terms of water availability and consumptive uses with respect to different land uses.

The adopted methodological framework addresses key challenges posed by large-scale hydrological modelling in data scarce environments such as the VRB. These challenges include the issue of missing data in streamflow records, the reliability of satellite and reanalysis data for forcing or calibrating hydrological models as an alternative to in-situ measurements, and the accuracy of the spatial and temporal representation of hydrological processes with spatially explicit models. A novel multivariate model calibration strategy is proposed to improve the representation of hydrological flux and state variables simulated with the fully distributed mesoscale Hydrologic Model (mHM). The proposed calibration strategy relies on the use of multiple satellite and reanalysis datasets from various sources. Then, a large ensemble of climate models are used to assess the impacts of climate change on water resources under various scenarios. The outputs of the mHM model are used to feed the WA+ framework to comprehensively report on the current and future conditions of water resources in the VRB.

The results show a clear increase in the projected exploitable water fraction while a decrease is expected in the available water fraction in the near future (2021-2050). Consequently, there is a clear need for adaptation measures to increase the water storage capacity in the VRB to facilitate a good exploitation of the projected increase in the net inflow, which would be beneficial for agriculture production and hydropower generation.

Résumé

La compétition pour l'usage de l'eau dans le bassin de la Volta (VRB) en Afrique de l'Ouest va s'intensifier dans un futur proche en raison des effets combinés de la croissance démographique galopante et du changement climatique. Les populations du bassin dépendent fortement d'une agriculture pluviale de subsistance qui demeure très sensible aux variabilités climatiques. Les inondations et les sécheresses récurrentes endommagent les infrastructures et créent des pertes en vie humaine. Les informations sur les ressources en eau et leurs tendances futures sont essentielles pour les acteurs de l'eau, car elles constituent la base d'une bonne gestion de l'eau et de la mise en œuvre de mesures adéquates pour renforcer la résilience à la pénurie d'eau et favoriser la sécurité de l'eau.

Cette thèse de doctorat propose une démonstration élaborée et innovante de la combinaison du Water Accounting Plus (WA+) avec la modélisation hydrologique et les scénarios de changement climatique pour faire un rapport de l'état actuel et futur des ressources en eau dans le VRB. WA+ ou "comptabilité de l'eau" est un outil standard de gestion de l'eau qui fournit une analyse complète des ressources en eau en termes de disponibilité et de consommation d'eau en fonction de différents usages du sol. La méthodologie adoptée aborde les principaux défis posés par la modélisation hydrologique à grande échelle dans des régions où les données sont rares, comme le VRB. Ces défis comprennent la question des données manquantes dans les séries temporelles de débit, la fiabilité des données satellitaires et de réanalyse pour faire tourner ou calibrer des modèles hydrologiques comme alternative aux mesures sur site, et la justesse de la représentation spatiale et temporelle des processus hydrologiques avec des modèles spatialement distribués. Une nouvelle stratégie de calibration multivariée est proposée pour améliorer la représentation des processus hydrologiques avec le modèle mesoscale Hydrologic Model (mHM). La nouvelle stratégie de calibration de modèle repose sur l'utilisation de plusieurs données satellitaires et de réanalyse provenant de diverses sources. Un large ensemble de modèles climatiques est utilisé pour évaluer les impacts du changement climatique sur les ressources en eau en considérant divers scénarios. Les sorties du modèle mHM sont utilisées pour alimenter le WA+ afin de faire une description complète des conditions actuelles et futures des ressources en eau dans le VRB.

Les résultats montrent une nette augmentation de la fraction d'eau exploitable dans un futur proche (2021-2050), tandis qu'une diminution de la fraction d'eau disponible est attendue. Par conséquent, il est opportun d'adopter des mesures d'adaptation pour augmenter la capacité de stockage de l'eau dans le bassin de la Volta afin de faciliter une bonne exploitation de l'augmentation prévue de l'apport net en eau, ce qui serait bénéfique pour la production agricole et la production hydroélectrique.

Contents

List of acronyms

ALEXI	Atmosphere-Land Exchange Inverse
ANN	Artificial Neural Networks
ARC	Africa Rainfall Estimate Climatology
CDF	Cumulative Distribution Function
CHIRPS	Climate Hazards group InfraRed Precipitation with Stations
CMIP	Coupled Model Intercomparison Project
CMORPH-CRT	Climate Prediction Centre MORPHing technique bias corrected
CMRSET	CSIRO MODIS Reflectance Scaling EvapoTranspiration
CORDEX	COordinated Regional-climate Downscaling Experiment
CPWF	CGIAR Challenge Program on Water and Food
CRU	Climatic Research Unit
DDS	Dynamically Dimensioned Search
DGRE	Direction Générale des Ressource en Eau
DHM	Distributed Hydrological Model
DS	Direct Sampling
ERA5	European Centre for Medium-range Weather Forecasts ReAnalysis 5
ESA CCI	European Space Agency Climate Change Initiative
ESGF	Earth System Grid Federation
EWEMBI	EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP
GADM	Database of Global Administrative Areas
GCM	Global Climate Models
GEF	Global Environment Facility
GIMMS	Global Inventory Modelling and Mapping Studies
GLEAM	Global Land Evaporation Amsterdam Model
GLiM	Global Lithological Map
GMTED	Global Multi-resolution Terrain Elevation Data
GPCC	Global Precipitation Climatology Centre
GRACE	Gravity Recovery and Climate Experiment
GSMaP	Global Satellite Mapping of Precipitation
HSD	Hydrological Service Department
IPCC	Intergovernmental Panel on Climate Change
ITCZ	Inter-Tropical Convergence Zone
JRA-55	Japanese 55 year ReAnalysis
LULC	Land Use and Land Cover
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications 2
mHM	mesoscale Hydrologic Model
MODIS	Moderate Resolution Imaging Spectroradiometer
MPS	Multiple-Point Statistics
MSWEP	Multi-Source Weighted-Ensemble Precipitation
PERSIANN-CDR	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Climate Data Record

PGF	Princeton university Global meteorological Forcing
RCM	Regional Climate Models
RCP	Representative Concentration Pathways
RFE	African Rainfall Estimate
SEBS	Surface Energy Balance System
SRPs	Satellite-based Rainfall Products
SRS	Satellite Remote Sensing
SSEBop	Operational Simplified Surface Energy Balance
TAMSAT	Tropical Applications of Meteorology using SATellite
TRMM	Tropical Rainfall Measuring Mission
UNEP	United Nations Environment Programme
VBA	Volta Basin Authority
VRB	Volta River Basin
WFDEI	WATCH Forcing Data methodology applied to ERA-Interim data
WMO	World Meteorological Organization

List of symbols

Symbols	Description
α	Spatial location measure
β	Bias measure
γ	Variability measure
λ	Latent heat of vaporization of water
μ	Mean value
Φ_{BA}	Objective function for case BA
Φ_{EA}	Objective function for E_a
Φ_{PW}	Objective function for calibration case pixel wise
Φ_{Q}	Objective function for calibration case streamflow-only
Φ_{SB}	Objective function for calibration case spatial bias-accounting
Φ_{SP}	Objective function for calibration case spatial bias-insensitive
σ	Standard deviation
AI	Aridity Index
D_E	Euclidian distance
E_a	Actual evaporation
E_{KG}	Kling-Gupta efficiency
E_{NS}	Nash-Sutcliffe efficiency
E_{NSlog}	Nash-Sutcliffe efficiency on logarithmic values
E_p	Potential evaporation
E_{ref}	Reference evaporation
E_{RMS}	Root mean squared error
E_{SP}	Spatial pattern efficiency
F_{DS}	Dynamical scaling function
I_{LA}	Leaf area index
P	Precipitation
Q	Streamflow
Q_{run}	Surface runoff
r	Pearson correlation coefficient
R_a	Extraterrestrial radiation
R_r	Groundwater recharge
r_s	Spearman correlation coefficient
S_t	Terrestrial water storage anomaly
S_u	Soil moisture
T_{avg}	Average air temperature
T_{max}	Maximum air temperature
T_{min}	Minimum air temperature
T_s	Land surface temperature

V_2	Second-order coefficient of variation
X_{mod}	Modelled variable X
X_{obs}	Observed variable X

Chapter 1

General Introduction

Keep exploring. Keep dreaming. Keep asking why.
Don't settle for what you already know.
Never stop believing in the power of your ideas,
your imagination, your hard work to change the world.
Barack Obama

Nothing in life is to be feared, it is only to be understood.
Now is the time to understand more, so that we may fear less.
Marie Curie

1.1 Research context

1.1.1 Water scarcity under climate change

The unavailability of reliable water resources for consumptive uses has been the limiting factor for development in many parts of the world (Hoekstra *et al.*, 2012; Milly *et al.*, 2005). Chronic water scarcity, exacerbated by climate change, is a big threat to prosperity in many regions, exposing countries to fragility and instability (Damania, 2020; World Bank Group, 2016). It is estimated that 4 billion people, or about one-half of the world's population, live in areas of severe physical water scarcity for at least one month per year (Mekonnen and Hoekstra, 2016). About 1.6 billion people, or almost one quarter of the world's population, face economic water shortage, which occurs when there is a lack of adequate infrastructure to ensure access to water (UN-Water, 2020). Moreover, it is foreseen that water scarcity will keep increasing in the future, with about 52% of the world's population living in water-stressed regions by 2050 (Kölbel *et al.*, 2018). Climate change, urbanization, population growth, deforestation and land degradation are likely to exacerbate water scarcity, which might lead to food insecurity and to conflicts between those who share the resources (Gaupp *et al.*, 2020; Zeitoun *et al.*, 2016). In addition, projected water scarcity by 2030 is expected to displace between 24 million and 700 million people living in arid and semi-arid regions (UN-Water, 2009). Reliable information on the current and future states of water resources is therefore essential for coping with water scarcity and strengthening water security (Vörösmarty *et al.*, 2010).

1.1.2 Water accounting for evidence-informed water management

While water problems around the world are increasing, information useful for water management seems to be decreasing (Bastiaanssen *et al.*, 2015). Decision in water resources management have long time been taken without reliable spatially explicit information on water resources, or without considering interaction among natural resources. Quantified and reliable information on the amount of water available, utilizable and utilized by means of consumptive use can help decision and policy makers to develop and implement mitigation measures (Karimi *et al.*, 2013b). Therefore, a systematic approach is needed to communicate how water is being used and how water resource developments will affect present and future water use patterns. This should lead to the efficient, effective and sustainable use of scarce water resources for food security and economic development, while maintaining the ecosystem services. Information on current and future water resources and their uses is thus fundamental for water actors. Water accounting frameworks are useful tools that can support water management as they serve to

provide a comprehensive view of the state of water resources. Among the existing water accounting frameworks (cf. Section 8.1), Water Accounting Plus (WA+) is a standardized framework that provides estimates of manageable and unmanageable water flows, stocks, consumption among users, and interactions with land use (Karimi *et al.*, 2013a). The results of WA+ are presented with standardized sheets accompanied with a set of indicators that summarize complex hydrological processes in a more accessible format to a broad panel of water professionals. Input data for WA+ are generally derived from satellite remote sensing observations, thereby making WA+ a suitable tool for data scarce regions. However, there are limitations in the use of earth observation data because of their usually coarse spatial resolutions, which limit the applicability of WA+ to large river basins. Moreover, using only satellite data poses challenges for closing the water balance (FAO and IHE Delft, 2020). Alternatively, hydrological models can be used to simulate hydrological flux and state variables (e.g., evaporation, runoff) necessary for analyses with the WA+ framework (Delavar *et al.*, 2020). In this case, climate change scenarios can be integrated in the hydrological modelling to predict future water resource conditions with WA+, which would be desirable for developing mitigation and adaptation policies (Hunink *et al.*, 2019).

1.2 Research objectives and challenges

1.2.1 Main research objective and hypothesis

The overarching goal of this PhD thesis is to demonstrate the applicability of the WA+ framework in combination with hydrological modelling and climate change scenarios in the Volta River basin (VRB). The VRB is a transboundary basin located in the semi-arid to sub-humid climatic zones of West Africa. It covers about 415,600 km^2, which are occupied at 82.5% by Burkina Faso and Ghana. The population in the VRB was estimated at 23.8 million people in 2010 and is expected to reach 38.4 million in 2030 (Williams *et al.*, 2016). Challenges for transboundary water resources management are increasing in the VRB as water demand is expected to increase by 1000% between 2000 and 2025, while the upstream and downstream countries have different priorities in terms of water uses (Biney, 2010). Burkina Faso increasingly builds reservoirs to develop small-scale irrigation upstream (de Fraiture *et al.*, 2014), while Ghana prioritises hydropower production downstream (Boadi and Owusu, 2019). The diverging national priorities for water uses remain a source of tension between the countries. In this context, an independent and unbiased assessment of the spatiotemporal

availability of water resources and their future evolutions could improve water governance and potentially alleviate the existing tensions.

The main research hypothesis is that WA+ can be combined with hydrological modelling and climate change scenarios to comprehensively report on the current and future states of water resources in the VRB. Reaching the main research objective requires answering several research questions and overcoming many challenges posed by large-scale hydrological modelling in data scarce environments such as the VRB. These challenges include: (i) the missing data issues in time series of streamflow records needed for hydrological modelling, (ii) the reliability of satellite and reanalysis data for forcing or calibrating hydrological models as an alternative to in-situ measurements, and (iii) the accuracy of the spatial and temporal representation of hydrological processes with spatially explicit models.

1.2.2 Missing streamflow data challenge

Missing data in streamflow records is a universal problem in hydrology, but exacerbated in developing countries where limited resources exist for data collection and management (Elshorbagy *et al.*, 2000; Serrat-Capdevila *et al.*, 2016). Missing values in streamflow records originate from various factors including the malfunctioning or failure of monitoring equipment, limited accessibility to measurement sites and human-induced errors. Water resources studies including hydrological modelling are limited by problems of missing values in observed streamflow data (Giustarini *et al.*, 2016; Tencaliec *et al.*, 2015). It is therefore essential to fill the gaps in streamflow time series before using them in water resources assessment studies (Bárdossy and Pegram, 2014; Enders, 2010). Gap-filling methods range from simple nearest neighbour data transfer methods (Bárdossy and Pegram, 2014; Giustarini *et al.*, 2016) to very sophisticated approaches such as dynamic state-space models (Amisigo and van de Giesen, 2005; Berendrecht and van Geer, 2016) and various forms of artificial neural networks (Dastorani *et al.*, 2010; Tfwala *et al.*, 2013). Each method has its advantages and limitations as described in different reviews (Harvey *et al.*, 2012; Marwala, 2009). An alternative to classical gap-filling methods is the Direct Sampling (DS) method (Mariethoz *et al.*, 2010), which is a multiple-point statistics (MPS) algorithm based on pattern reproduction. In contrast to deterministic methods, DS has the advantage of being a stochastic method that provides probabilistic estimates of the missing values, thereby allowing uncertainty quantification, which is very important for hydrograph estimation (Beven, 2016). This PhD thesis answers some questions related to the application of the DS method. *Can the DS method be used to*

simulate missing values under various lengths of data gaps? What is the performance of the DS method for different climatic zones and hydrological regimes in the VRB? How to choose the predictor station to obtain the most accurate simulations?

1.2.3 Reliable meteorological data challenge

Precipitation is the key input variable for hydrological modelling because it determines the spatiotemporal variability of other hydrological fluxes and state variables (Bárdossy and Das, 2008; Thiemig *et al.*, 2013). The in-situ measurements of precipitation have limitations, which include limited and uneven areal coverage, deficiencies in instruments and costly maintenance (Awange *et al.*, 2019; Harrison *et al.*, 2019; Kidd *et al.*, 2017). These drawbacks have led to the advent of precipitation estimation from space (Barrett and Martin, 1981), which produces spatially homogeneous estimates and covers inaccessible regions with uninterrupted records over time (Beck *et al.*, 2019b; Funk *et al.*, 2015). The advent of satellite-based rainfall products (SRPs) has opened up new avenues for large-scale hydrological modelling, especially in data sparse regions (Hrachowitz *et al.*, 2013; Serrat-Capdevila *et al.*, 2014; Sheffield *et al.*, 2018). However, SRPs also have limitations, which include measurement bias, inadequate spatiotemporal resolutions (e.g., for extreme event simulation) and shortness of the records for some applications (e.g., climate change impact assessment) (Marra *et al.*, 2019). Many SRPs have been developed over the past decades with different objectives and characteristics that are reviewed in the literature (e.g. Le Coz and van de Giesen, 2019; Maggioni *et al.*, 2016; Maidment *et al.*, 2014; Sun *et al.*, 2018). In addition to SRPs, precipitation estimates can be obtained from atmospheric retrospective analysis (or reanalysis) datasets (Lorenz and Kunstmann, 2012; Schröder *et al.*, 2018). Reanalysis datasets are sometimes preferred over SRPs because of their usually long-term records suitable for climate change studies and because of their higher performance in predicting complex climate systems (Potter *et al.*, 2018; Seyyedi *et al.*, 2015). While SRPs are subject to systematic biases, reanalysis products have uncertainties resulting from their model parametrization and low spatial resolution with poor representation of sub-grid processes (Bosilovich *et al.*, 2008; Laiti *et al.*, 2018). With the increasing number of SRPs and reanalysis datasets, it is fundamental to evaluate their adequacy for hydrological modelling. Choosing the best precipitation datasets that generate the most plausible hydrological processes in a region is a prerequisite to any water resources assessment studies. Research questions related to the reliability of meteorological datasets for large-scale hydrological modelling are addressed in this thesis. *What is the impact of different gridded*

rainfall and temperature datasets on the simulation of hydrological fluxes and state variables? How important is the choice of meteorological datasets for the representation of spatial patterns versus temporal dynamics of hydrological processes?

1.2.4 Accurate process representation challenge

With the development of distributed hydrological models that facilitate large-scale predictions (Clark *et al.*, 2017; Fatichi *et al.*, 2016; Ocio *et al.*, 2019), there is a growing interest in the accurate spatiotemporal representation of processes (Baroni *et al.*, 2019; Hrachowitz and Clark, 2017; Paniconi and Putti, 2015). Hydrological models are conventionally calibrated using streamflow data. However, a good prediction of streamflow does not guarantee a reliable spatial and temporal representation of other hydrological fluxes and state variables when using a spatially distributed hydrological model (Clark *et al.*, 2016; Minville *et al.*, 2014). This is explained by the fact that streamflow is the result of several inter-linked processes, thereby it masks spatial heterogeneity (Tobin and Bennett, 2017; Wambura *et al.*, 2018). The increasing availability of satellite-based earth observation data has promoted the development of spatial hydrology and large domain water management applications including the monitoring of floods and droughts (Revilla-Romero *et al.*, 2015; Senay *et al.*, 2015; Teng *et al.*, 2017; Wu *et al.*, 2014). Multivariate parameter estimation based on the simultaneous use of multiple satellite-derived remote sensing data sources can reduce the feasible parameter search space and lead to better model performances by improving the internal model dynamics and the representation of spatial differences in hydrological processes (Clark *et al.*, 2017; Shafii and Tolson, 2015). However, satellite data are not free of uncertainties and using different data from different satellite-derived remote sensing products simultaneously is not straightforward so that the value of such an approach remains rather under-explored (Rajib *et al.*, 2018b; Silvestro *et al.*, 2015). Combining multiple data sources in model calibration requires a meaningful integration of the datasets, which should harness their most reliable contents to avoid the accumulation of their uncertainties and mislead the parameter estimation procedure (Dembélé *et al.*, 2020a). Although satellite products are characterized by uncertainties, their most reliable key feature is the representation of spatial patterns, which is a unique and relevant source of information for distributed hydrological models (Dembélé *et al.*, 2020b; Stisen *et al.*, 2018). Key research questions are formulated to address the challenge of the spatial representation of dominant hydrological processes in a spatially distributed model. *Is it possible to simultaneously calibrate a distributed hydrological model with multiple non-commensurable variables without*

significantly deteriorating the streamflow signal? How can a distributed hydrological model be calibrated using only the spatial patterns of remote sensing data to improve the spatial patterns of hydrological fluxes and state variables? What is the importance of model calibration strategies for improving process representation in multivariate calibration with satellite and reanalysis data?

1.3 Research questions

The main fundamental research questions of this PhD thesis concern the applicability of the WA+ framework under climate change. *How can hydrological modelling and climate change scenarios be used to feed the WA+ framework for predicting water accounts? How will climate change impact on future water accounts in the VRB? Should we be prepared for a wet or a dry future?* To address these questions, the abovementioned challenges (cf. Section 1.2) on the accurate spatiotemporal representation of hydrological processes with the use of reliable satellite and reanalysis datasets need to be addressed first. A robust methodological framework is therefore deployed to address the research questions and achieve the main research goal.

1.4 Methodological framework

The methodological framework of this PhD thesis is articulated around different steps, which consist in addressing key identified challenges in hydrology and answering fundamental research questions. First, the Direct Sampling method (Mariethoz *et al.*, 2010) is used for the gap-filling of streamflow time series. Direct Sampling is a non-parametric and stochastic method that can generate new simulated values based on a conditional resampling of a provided training dataset. Secondly, a clear demonstration of the benefits of satellite and reanalysis datasets in improving the process representation of the fully distributed mesoscale Hydrologic Model (mHM) (Kumar *et al.*, 2013; Samaniego *et al.*, 2010) is proposed. Then, a large ensemble of global climate models (GCMs) and regional climate models (RCMs) under various representative concentration pathways (RCPs) is used to assess the impacts of climate change on water resources in the VRB. Finally, the WA+ framework is combined with hydrological modelling and climate change scenarios to provide an evidence-informed report on water conditions under current and future situations.

Overall, this PhD thesis contributes to the state-of-the-art by striving to provide answers to fundamental research questions that are identified as some of the key unsolved problems in

hydrology in the twenty-first century (i.e. uncertainty in large-scale measurements and data, spatial heterogeneity and modelling methods; Blöschl *et al.*, 2019; Wilby, 2019). The efforts deployed to address the identified challenges also respond to the recent call of the scientific community to embrace a fourth paradigm in hydrology (i.e. data-intensive science, Peters-Lidard *et al.*, 2017). Moreover, the main goal of this thesis is a scientific contribution to the UN sustainable development goals (SDGs), as it aligns with the SDG 6, and particularly the target 6.4, which is: "by 2030, substantially increase water-use efficiency across all sectors and ensure sustainable withdrawals and supply of freshwater to address water scarcity and reduce the number of people suffering from water scarcity" (UN-Water, 2015).

1.5 Outline of the thesis

The outline of the thesis is provided here with a brief summary of the chapters (Figure 1.1). Most of the results presented in this thesis have been presented to different conferences and published in international peer-reviewed scientific journals. A list of the presentations and papers resulting from this PhD thesis is provided at the Page 251. Therefore, the resulting thesis is a collection of papers published (Chapters 3, 5 and 6), submitted for publication (Chapter 4) and in preparation for publication (Chapters 7 and 8) in scientific journals. However, the content of the papers are modified to interconnect the chapters and avoid repetitions in the thesis. For instance, an entire chapter is dedicated to the description of the study area (Chapter 2), which is no longer presented in the subsequent chapters. Moreover, the description of the hydrological model and the modelling datasets are only presented in the Chapter 4, although the same model configuration is used from the Chapter 4 to the Chapter 8. Finally, a unique list of appendices and a unique list of references are provided after the general conclusion (Chapter 9). The thesis is organized as follows:

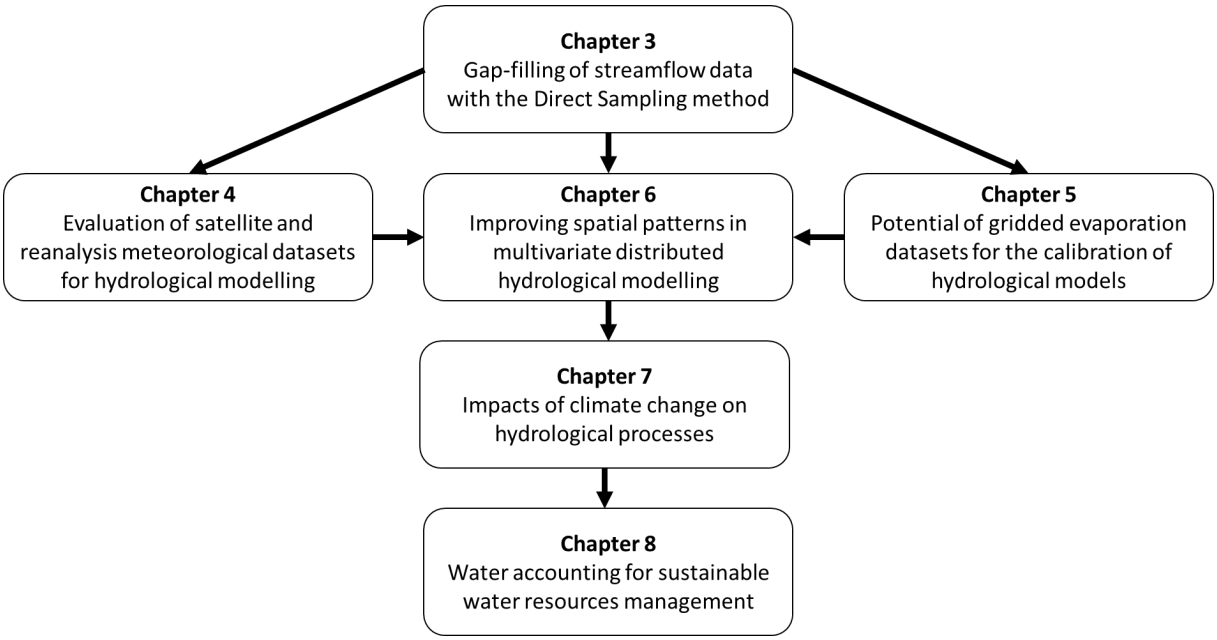


Figure 1.1. Outline of the main chapters of the thesis

Chapter 1 introduces the research undertaken in this thesis. It gives an overview of the research motivation, the research objectives, the fundamental questions addressed, the methodological development and the outline of the thesis.

Chapter 2 provides a comprehensive description of the Volta River basin, which is the study area for this thesis.

Chapter 3 proposes a robust framework for filling gaps in time series of streamflow data using the Direct Sampling (DS) approach. The methodological development includes the selection of predictor stations, the optimization of the DS parameters and the gap-filling of streamflow data collected in the Volta River basin. First, the performance of the method is assessed by applying it to various synthetic missing data scenarios. Then, a real-case application is done for the existing gaps in the streamflow records. Moreover, this study includes the assessment of the method for different climatic zones and hydrological regimes and for different upstream-downstream relations among the gauging stations used for gap filling (Dembélé *et al.*, 2019).

Chapter 4 assesses the skill of ten satellite-based precipitation products (TAMSAT, CHIRPS, ARC, RFE, MSWEP, GSMAP, PERSIANN-CDR, CMORPH-CRT, TRMM3B42, TRMM3B42-RT), and of seven reanalysis precipitation products (JRA55, EWEMBI, WFDEI-GPCC, WFDEI-CRU, MERRA2, PGF, ERA5) if used as forcing data in the mHM hydrological model (Dembélé *et al.*, 2020c). The subsequent ability of the meteorological datasets to

reproduce the temporal and spatial dynamics of several hydrological flux and state variables (i.e., streamflow, actual evaporation, soil moisture and terrestrial water storage) is evaluated. In total, 102 rainfall-temperature input data combinations are tested using six different temperature datasets (JRA55, EWEMBI, WFDEI, MERRA2, PGF, ERA5) for the calculation of potential evaporation. The mHM hydrological model is recalibrated for each of the 102 input data combinations and the outputs are evaluated with in-situ streamflow data, soil moisture data from ESA CCI, evaporation data from GLEAM and terrestrial water storage data from GRACE.

Chapter 5 evaluates the performance of twelve different global evaporation products (MOD16A2, SSEBop, ALEXI, CMRSET, SEBS, GLEAM v3.2a, GLEAM v3.3a, GLEAM v3.2b, GLEAM v3.3b, ERA5, MERRA-2 and JRA55) when used as calibration variables in hydrological modelling. The ability of the gridded evaporation datasets in improving the spatiotemporal variability of multiple hydrological state and flux variables (i.e. streamflow, evaporation, soil moisture and terrestrial water storage) under multivariate calibration strategies with streamflow data is assessed using the mHM model (Dembélé *et al.*, 2020a). Four different calibration strategies that differently harness the information content of the evaporation products are tested: (i) temporal basin average, (ii) temporal pixel-wise, (iii) spatial bias-accounting, and (iv) spatial bias-insensitive. A set of 48 combinations of evaporation products and calibration strategies are tested and compared to the benchmark model calibrated only on streamflow. Independent datasets of soil moisture (ESA CCI) and terrestrial water storage (GRACE) are used for a process-diagnostic evaluation.

Chapter 6 tests a novel multivariate parameter estimation approach for modelling daily hydrological flux and state variables based on the simultaneous incorporation of spatial patterns derived from three remote sensing products in the mHM model. Each of the remote sensing products describes a different component of the hydrological system (i.e., evaporation from GLEAM, soil moisture from ESA-CCI and terrestrial water storage change from GRACE). A new bias insensitive multicomponent spatial pattern metric used as objective function is developed to incorporate the different data sources in the calibration procedure. The goal is to investigate the potential improvement of the predictions of the spatial and temporal patterns together with the resulting effect on streamflow predictions, as compared to traditional model calibration on streamflow data alone (Dembélé *et al.*, 2020b).

Chapter 7 provides a comprehensive evaluation of the impacts of climate change on water resources in the VRB. A large ensemble of twelve GCMs that are dynamically downscaled by five RCM from CORDEX-Africa under three representative concentration pathways (RCP2.6,

RCP4.5 and RCP8.5) is used. A multivariate bias correction is applied to the climate projection datasets before using them to force the mHM hydrological model. The past and future trends in climatic variables (i.e., precipitation, temperature and potential evaporation) and hydrological variables (i.e., streamflow, surface runoff, actual evaporation, groundwater recharge, soil moisture and terrestrial water storage) are analysed over the period from 1991 to 2100.

Chapter 8 demonstrates the possibility of using the Water Accounting Plus (WA+) framework combined with hydrological modelling to comprehensively report on the state of water resources under climate change. Water fluxes, stocks and flows are predicted with the fully distributed mHM hydrological model for the historical period 1999-2020 and the near term future period 2021-2050. A large ensemble of nine GCMs and four RCMs under the RCP8.5 scenario is used. The WA+ framework is fed with hydrological processes derived from the mHM model. An evidence-informed reporting on the state and trends of water resources is proposed for the VRB.

Chapter 9 summarizes the key findings of this thesis, draws conclusions and offers suggestions for future studies.

Chapter 2

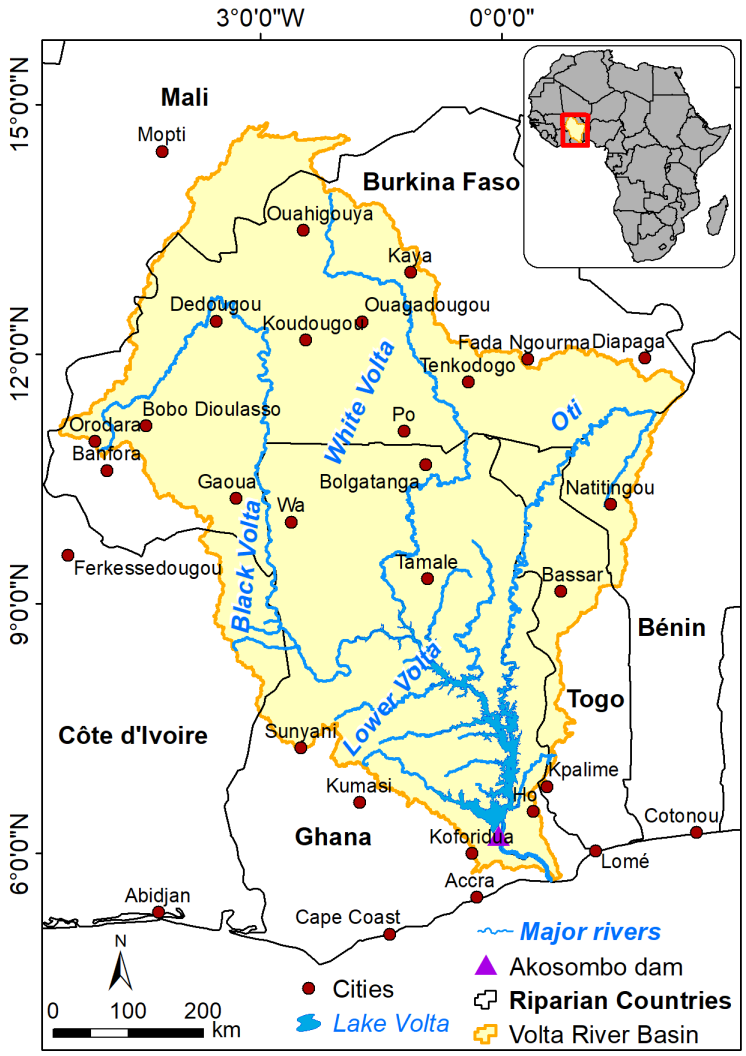
Case Study: Volta River Basin

How could drops of water know themselves
to be a river? Yet the river flows on.
Antoine de Saint-Exupery

When you put your hand in a flowing stream,
you touch the last that has gone before
and the first of what is still to come.
Leonardo da Vinci

2.1 Location and overview

The Volta River Basin (VRB) is a major transboundary basin in West Africa (Figure 2.1). It is the ninth largest drainage basin in sub-Saharan Africa (UNEP-GEF, 2013), and is located between latitudes 5°40' N and 14°55' N and longitudes 2°20' E and 5°25' W. The VRB covers approximately 415,600 km^2 and represents almost 28% of the African West Coast (Frenken, 1997). The basin area lies across parts of six countries, namely, Benin, Burkina Faso, Côte d'Ivoire, Ghana, Mali and Togo. Burkina Faso and Ghana have the largest portions with respectively 42.3% and 40.2% of the basin area (Table 2.1). Togo has 6.4% of the basin area and the remaining 11.1% is shared among Benin, Côte d'Ivoire and Mali. The VRB was home to 23.8 million people in 2010 and projected to reach 38.4 million in 2030, and 56.1 million in 2050 (Williams *et al.*, 2016).



Data source: GADM & HydroSHEDS, 2016

Figure 2.1. Volta River Basin in West Africa

Table 2.1. Proportions of the Volta River Basin in the riparian countries

Countries	Area of basin (km^2)	% of basin	% of country area in basin	% of basin population
Benin	15,375	3.7	12.1	2.6
Burkina Faso	175,740	42.3	62.4	47.7
Cote d'Ivoire	13,422	3.2	3.1	2.2
Ghana	167,165	40.2	70.1	35.5
Mali	17,440	4.2	1	3.4
Togo	26,463	6.4	45	8.6
Total	415,605			

Source: adapted from Williams *et al.* (2016)

2.2 Physical features

2.2.1 Relief

The relief of the VRB is predominately flat with altitudes ranging from zero to 940 m (Figure 2.2b). Most of the basin topography lies below 400 m above sea level with a mean altitude of about 255 m (Dembélé *et al.*, 2020c). The global slope index is between 25 and 50 cm km^{-1} (Moniod *et al.*, 1977).

The eastern border of the VRB is delimited by a series of hills and mountain ranges, namely from south to north, the Akwapim Range, the Togo Mountains, the Fazao Mountains and the Atacora Range in Benin, which reaches the Niger basin, with the Oti River flowing along its western side. The Kwahu plateau starts at the Akosombo Gorge and leads northwestwardly. The Banfora plateau is the main relief on the north-western part of the basin (De Condappa and Lemoalle, 2009).

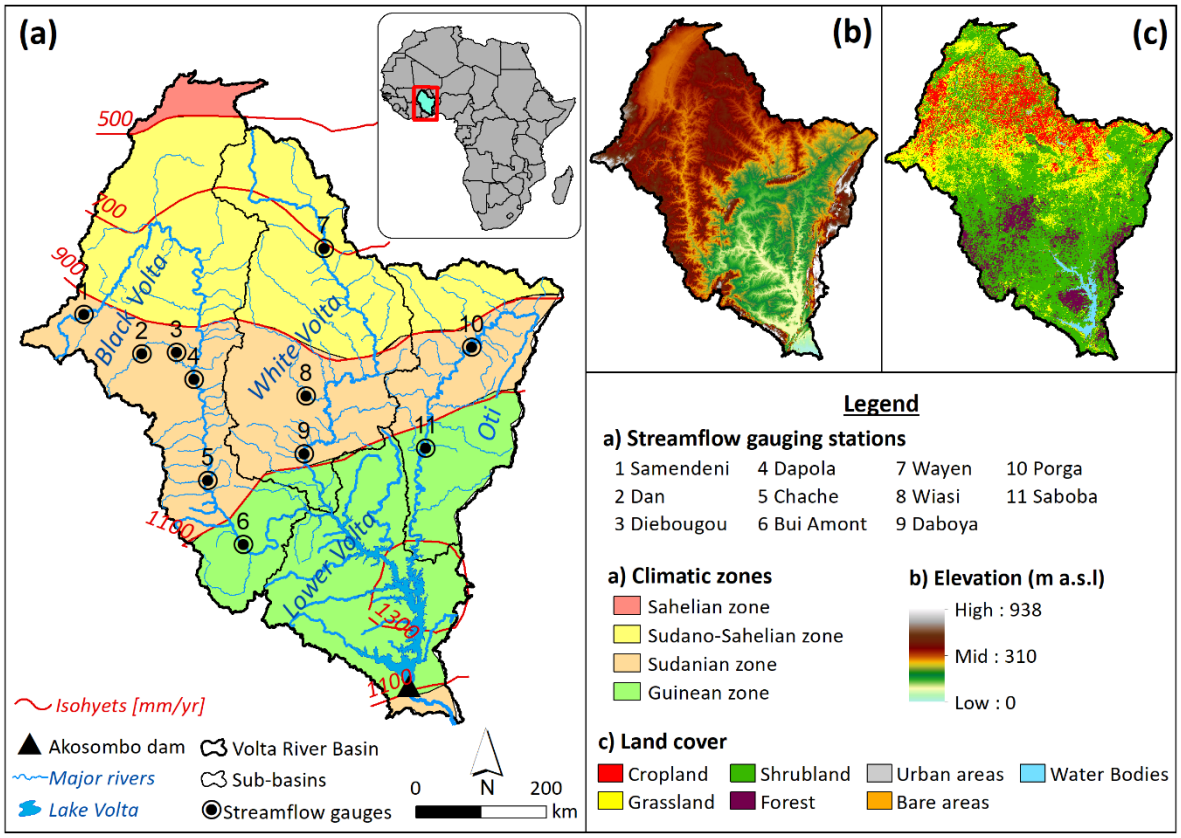


Figure 2.2 Physical and hydroclimatic characteristics of the Volta River basin. Adapted from Dembélé *et al.* (2020a).

2.2.2 Geology

Three major geological provinces underlie the VRB. Two of them underlie over 90 % of the basin (Mul *et al.*, 2015), namely the basement crystalline province (crystalline and metamorphic rock formations) of the Precambrian age associated with the West African Craton, and the consolidated sedimentary province (Voltaian formations) of the Proterozoic to Palaeozoic ages. The third province is the unconsolidated sedimentary province of the Proterozoic to Palaeozoic ages and tertiary sedimentary rocks (Continental Terminal), consisting mainly of sandstones, schists, conglomerate and dolomite (Barry *et al.*, 2005; Martin, 2006; Obuobie and Barry, 2012; Sandwidi, 2007). The most dominant geological province in the VRB is the basement crystalline province, which underlies 80% of Burkina Faso and 54% of Ghana (Mul *et al.*, 2015). It is mainly composed of metamorphic, igneous, granite-gneiss-greenstone rocks and anorogenic intrusions. The consolidated sedimentary province mainly underlies the central-west and southwestern parts of the VRB. It consists of sandstone intercalated with politic schist, arkoses, shale, dolomitic limestone and conglomerate, sandy and pebbly beds, and mudstones

(Williams *et al.*, 2016). Figure 2.3 shows the lithological map of the VRB based on the Global Lithological Map (GLiM v1.0; Hartmann and Moosdorf, 2012).

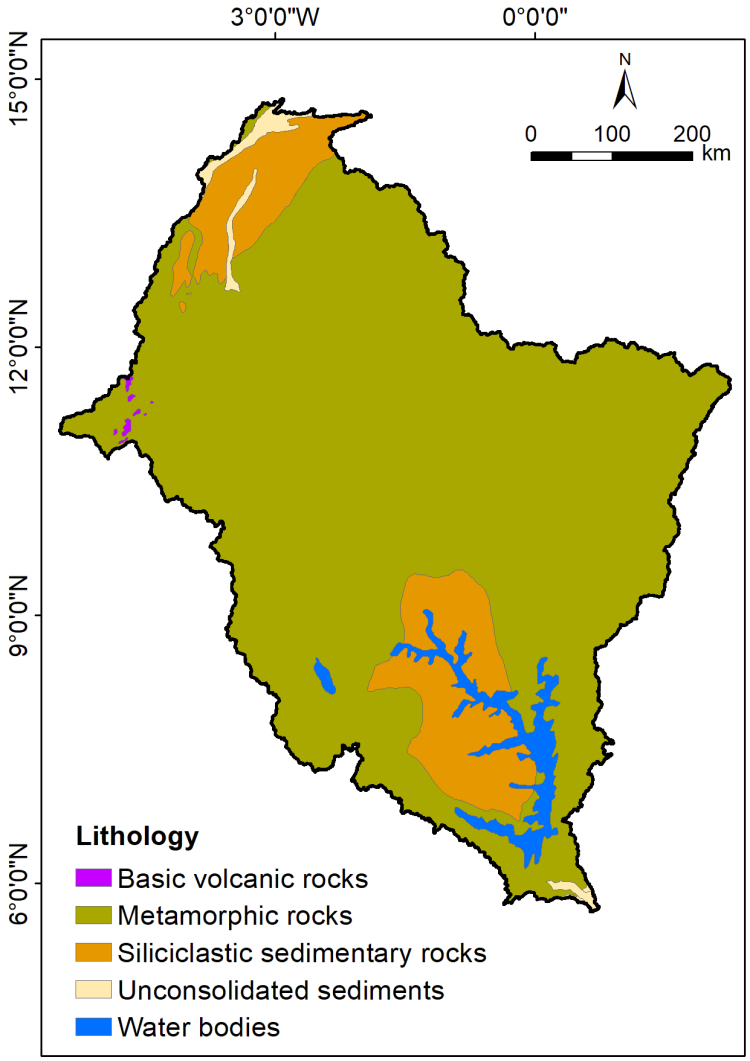


Figure 2.3. Lithological classes in the Volta River Basin based on GLiM v1.0 data

2.2.3 Soil

The formation of soils in the VRB resulted from weathered parent materials of the mid-Palaeozoic age or older, with a long period of leaching (Andah *et al.*, 2003; Benneh *et al.*, 1990). The dominant soil types in the VRB are Luvisols except for the northern areas where Regosols and Arenosols are dominant. These Luvisols have unstable structures and with low nutrient content, which facilitate their erosion and make them susceptible to slaking (Mul *et al.*, 2015). The coarse-textured Regosols resulting from unconsolidated materials are prone to erosion due to low coherence of the soil matrix material (Mando, 1997). Their high permeability and low water-holding capacity make them vulnerable to drought. The Arenosols with high

clay content are present in the upland areas. Their low water-holding capacity and low fertility limit their suitability for agriculture. Lithosols are the most important soils after Luvisols and they are found in the southern parts of the VRB. They are largely well-drained, gravelly with a light-textured matrix. The organic content and fertility of the Lithosols increase southwards (Mul *et al.*, 2015).

2.2.4 Land Cover and Land Use

The dominant land cover in the VRB is savannah, which is composed of grassland interspersed with shrubs and trees over 75% of the basin area, followed by cropland (13%), forest (9%), water bodies (2%) and bare land and settlements (1%) (Figure 2.2c). Extensive farming, wood extraction and overgrazing are limiting factors for large coverage of trees (Abubakari *et al.*, 2012). The density and vigour of the vegetation follow a south-north decreasing gradient, which is similar to the rainfall pattern. Grassy savannah is mostly found in the northern parts of the basin, while woody savannah is found in the south. Major cultivated staple food crops in the VRB are millet, sorghum, maize, rice, cassava, yam, cowpea and groundnut (Williams *et al.*, 2016). Protected areas represent about 10.7% of the basin area and include national parks and wildlife reserves.

2.3 Climate

Climate in West Africa is unique and complex (Berthou *et al.*, 2019; Bichet and Diedhiou, 2018; Nicholson *et al.*, 2018a). Rainfall is characterized by interannual and multidecadal variabilities (Biasutti *et al.*, 2018; Nicholson *et al.*, 2018b; Thorncroft *et al.*, 2011). The seasonal and latitudinal oscillation of the Inter-Tropical Convergence Zone (ITCZ) is the predominant rainfall generation mechanism in West Africa (Biasutti, 2019), thereby depicting a south-north gradient of increasing aridity in the VRB. The ITCZ is a narrow belt of clouds associated with low pressure and intense convective activities resulting from the near-surface convergence of warm and moist trade winds (Dezfuli, 2017; Schneider *et al.*, 2014). The warm northeasterly Harmattan winds emanate from the Sahara and the moist southwest monsoon winds originate in the Atlantic ocean (Nicholson, 2013; Vizy and Cook, 2018). Actual evaporation exceeds 80% of annual rainfall in the basin (Andreini *et al.*, 2000; De Condappa and Lemoalle, 2009; Dembélé *et al.*, 2020a).

The VRB extends over four eco-climatic zones identified as follows:

- The Sahelian zone located in the northern part of the basin and above the 14° N parallel;
- The Sudano-Sahelian zone located between the 11° 30' N and 14 °N parallels. It covers a greater part of Burkina Faso;
- The Sudanian zone located below 11° 30' N parallel. It comprises the northern part of Ghana and some parts of Côte d'Ivoire, Benin and Togo;
- The Guinean zone extending from approximately 8° N to 11° N. It covers the southern part of Ghana.

The four eco-climatic zones are characterized by increasing vegetation density and receive increasing precipitation from north to south (Figure 2.2; Table 2.2). The information provided in Table 2.2 is obtained from the global aridity index database (Trabucco and Zomer, 2018), and the WFDEI meteorological data (see Table 4.1; Weedon *et al.*, 2014) for the period 1979-2016. The maps of spatial patterns of rainfall and temperature in the VRB for different datasets are shown in Figure 4.2 and Figure 4.3. The climatology of rainfall and temperature per climatic zones are provided in Appendix 8 to Appendix 11.

Table 2.2. Characteristics of the four eco-climatic zones in the Volta River basin with the Aridity Index (*AI*). The annual mean value with the range ([min-max]) is provided for Precipitation (*P*) and average air Temperature (T_{avg}). Adapted from Dembélé *et al.* (2020b) and Dembélé *et al.* (2020c).

Eco-climatic zones	**Climate class**	***AI* (-)**	***P* (mm/year)**	**T_{avg} (°C)**
Sahelian zone	Arid	0.16 [0.12-0.20]	570 [470-610]	29.3 [29.0-29.7]
Sudano-Sahelian zone	Semi-arid	0.29 [0.16-0.43]	790 [570-980]	28.6 [28.0-29.3]
Sudanian zone	Semi-arid/ Dry sub-humid	0.47 [0.33-0.98]	1010 [890-1290]	28.1 [26.4-28.8]
Guinean zone	Dry sub-humid/ Humid	0.70 [0.49-1.22]	1190 [1030-1420]	27.6 [26.0-28.6]

2.4 Hydrology and surface water

The Volta River was named by early Portuguese traders because of its twisting and meandering course. It flows from north to south from Burkina Faso to Ghana over a distance of 1,850 km

before it drains into the Atlantic Ocean at the Gulf of Guinea through the Volta Estuary (Williams *et al.*, 2016). Before reaching the Atlantic Ocean, the Volta River transits in the Lake Volta, which is the World's largest manmade lake in terms of area. The Lake Volta has a dendritic shape covering an area of 8,502 km^2 with a storage capacity of 148 km^3, an average depth of 19 m, and a residence time of 4.3 years. The Lake Volta has been formed because of the construction of the Akosombo dam in Ghana, which begun in 1961 and was completed in 1965. It is a rockfill dam with an embankment of 119 m high and 640 m long resulting in a total storage volume of 7.94 10^6 m^3 (ILEC, 2017).

The drainage system of the VRB is composed of four major sub-basins and rivers (Figure 2.4; Table 2.3), namely:

1- The Black Volta (Mouhoun in Burkina Faso) is the westerly tributary shared by Mali, Burkina Faso, Ghana, and Côte d'Ivoire. It originates from the Kong Mountains in the Dinderesso Forest Reserve in the southwest of Burkina Faso at the altitude 500 m above sea level (Moniod *et al.*, 1977), flows in a north-easterly direction and receives on its left bank side the Sourou, a tributary originating in Mali, before changing its path into a south-easterly direction. The main tributaries of the Black Volta are the Sourou, Benchi, Chuko, Laboni, Gbalon, Pale, Kamba and Tain rivers. After reaching Bamboi, the river takes a west-easterly direction towards the Lake Volta.

2- White Volta (Nakambe in Burkina Faso) originates in northern Burkina Faso at 335 m above sea level (Moniod *et al.*, 1977), flows southeastward, and then westwards before flowing in a southerly direction. It receives water from one of its most important tributaries, the Sissili, and flows towards the Lake Volta by crossing the border between Burkina Faso and Ghana. The Red Volta (Nazinon in Burkina Faso) originates in the central part of Burkina Faso, near Ouagadougou, and flows southeastward to the border with Ghana. After crossing the border, it joins the White Volta near Gambaga in Ghana. The main tributaries of the White Volta are the Mole, Kulpawn, Sisili, Red Volta, Asibilika, Agrumatue, Nasia, Morago, Tamne and Nabogo rivers.

3- The Oti River (Pendjari in Benin), the easterly tributary, originates at 600 m above sea level (Moniod *et al.*, 1977) in the north of Benin. It flows near the Benin's rainy Atakora Mountains, which are an extension of the Togo's Akwapim Ranges, where it is known as the Pendjari River. It crosses Togo, where it is known as the Oti River, and flows down to Ghana border where it joins the Volta River near Kete Krachi and continues into the Lake Volta. The main tributaries of the Oti River are the Koumongou, Kéran,

Kara, Mô, Kpanlé, Wawa, Ménou, Danyi, Afram, Obosom, Sene, Pru, Kulurakun, Daka, and Asukawkaw rivers.

4- The Lower Volta is shared between Ghana and Togo, and encompasses the Lake Volta, which is fed by the Black Volta, the White Volta and the Oti River. It also includes a series of small streams flowing directly into Lake Volta, and the portion of the river downstream from the Kpong dam. The Lower Volta discharges into the Gulf of Guinea in Ghana, near the town of Ada Foah (Mul *et al.*, 2015).

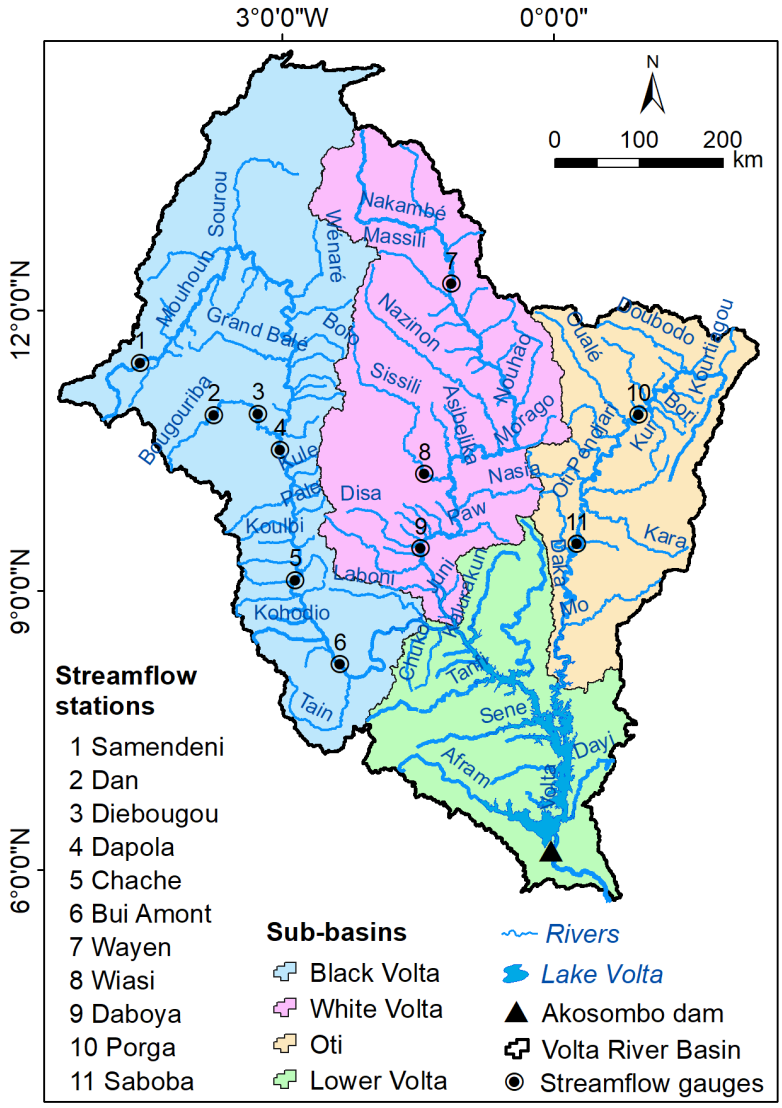


Figure 2.4. Hydrography of the Volta River Basin

Figure 2.5 shows the hydrographs of daily streamflow at gauging stations located at the outlets of the Black Volta (Bui amont), the White Volta (Daboya) and the Oti River (Saboba).

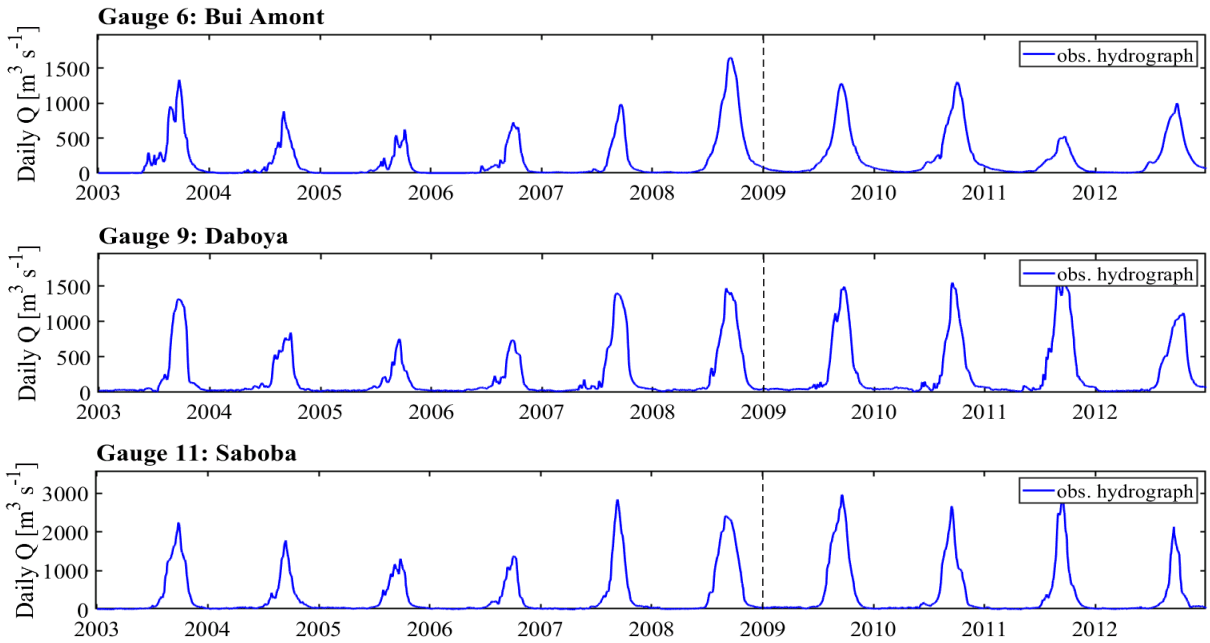


Figure 2.5. Hydrographs of daily streamflow at the outlets of the major sub-basins in the Volta basin

The VRB is one of the rare basins in the world where upstream is drier than downstream. Most of the streams at higher latitudes are not perennial in the White Volta, unlike those at more southern latitudes (UNEP-GEF, 2013). The river network is poorly gauged with an estimated average density of about one station per 3,000 km^2 while the World Meteorological Organization (WMO) recommends a maximum of about 1,900 km^2 per station (WMO, 2008).

The annual groundwater recharge rate of the Volta River system varies between 5% and 10% of the mean annual rainfall (Williams *et al.*, 2016).

Table 2.3. Characteristics of the sub-basins in the Volta River basin

Sub-basins	Area (km^2)	% of the Volta River basin	Mean annual flow (km^3/year)	Mean annual runoff coefficient (%)
Black Volta	152,829	36.8	8.6	4.8
White Volta	113,423	27.3	7.6	6.9
Oti	74,501	17.9	9.8	10.5
Lower Volta	74,852	18.0	9.6	10.3

Source: adapted from MWH (1998); UNEP-GEF (2013)

Chapter 3

Gap-filling of Streamflow Data with Direct Sampling*

Errors using inadequate data are much less

than those using no data at all.

Charles Babbage

You can have data without information,

but you cannot have information without data.

Daniel Keys Moran

* This chapter is based on the following publication:
Dembélé, M., F. Oriani, J. Tumbulto, G. Mariethoz, and B. Schaefli (2019), Gap-filling of daily streamflow time series using Direct Sampling in various hydroclimatic settings, Journal of Hydrology, 569, 573-586, https://doi.org/10.1016/j.jhydrol.2018.11.076

Abstract

Complete hydrological time series are necessary for water resources management and modelling. This can be challenging in data scarce environments where data gaps are ubiquitous. In many applications, repetitive gaps can have unfortunate consequences including ineffective model calibration, unreliable timing of peak flows, and biased statistics. Here, Direct Sampling (DS) is used as a non-parametric stochastic method for infilling gaps in daily streamflow records. A thorough gap-filling framework including the selection of predictor stations and the optimization of the DS parameters is developed and applied to data collected in the Volta River basin, West Africa. Various synthetic missing data scenarios are developed to assess the performance of the method, followed by a real-case application to the existing gaps in the flow records. The contribution of this study includes the assessment of the method for different climatic zones and hydrological regimes and for different upstream-downstream relations among the gauging stations used for gap filling. Tested in various missing data conditions, the method allows a precise and reliable simulation of the missing data by using the data patterns available in other stations as predictor variables. The developed gap-filling framework is transferable to other hydrological applications, and it is promising for environmental modelling.

3.1 Introduction

In many regions around the world, observed streamflow records contain missing values that hinder their use for water resources management, engineering applications or hydrological modelling (Giustarini *et al.*, 2016; Tencaliec *et al.*, 2015). When riddled with gaps, the usability of streamflow records dwindles substantially and therefore infilling gaps in time series is an essential step in planning, design and operation of water resources systems (Bárdossy and Pegram, 2014; Enders, 2010).

Missing data can occur because of the malfunctioning or failure of monitoring equipment, extreme weather conditions, limited accessibility to measurement sites, fortuitous absence of observers, human-induced errors, budget restraints, and public turmoil or political conflicts among other factors (Elshorbagy *et al.*, 2000; Serrat-Capdevila *et al.*, 2016). Depending on the usually unpredictable factors for which missing data occurs, the gaps in the time series can vary from one time step (i.e. sub-daily to one day) to several months, or even a complete lack of data for several years. Problems related to missing data are universal in hydrology, but exacerbated in developing countries where limited resources exist for data collection, quality assessment, repository provision, and maintenance.

For many applications, the recommended approach is to infill the gaps and flag the corresponding values (Harvey *et al.*, 2010). Gap-filling methods range from simple linear models to complex deterministic or stochastic techniques. The most common approaches include the simple nearest neighbour method by data transfer (Bárdossy and Pegram, 2014; Giustarini *et al.*, 2016), interpolation techniques (Pappas *et al.*, 2014; Piazza *et al.*, 2015; Teegavarapu, 2014), autoregressive models (Bennis *et al.*, 1997; Tencaliec *et al.*, 2015), simple and multiple regressions (Hirsch, 1979; 1982; Woodhouse *et al.*, 2006), classification and regression trees (Sidibe *et al.*, 2018), recession methods (Gyau-Boakye and Schultz, 1994), recursive models (Lambert, 1969), nonlinear and storage models (Coulibaly and Baldwin, 2005; Dawdy and O'Donnell, 1965), satellite data applications (Papadakis *et al.*, 1993), dynamic state-space models (Amisigo and van de Giesen, 2005; Berendrecht and van Geer, 2016), and various forms of artificial neural networks (Dastorani *et al.*, 2010; Tfwala *et al.*, 2013) among others (Dumedah and Coulibaly, 2011). Different studies provided a review of these methods (Harvey *et al.*, 2012; Marwala, 2009), which all have their limitations depending on the application. For instance, the nearest neighbour method brings discontinuity in the time series (Lepot *et al.*, 2017; Peterson and Western, 2018). Interpolation techniques offer a limited representation of the space-time structure of the time series, being therefore unsuitable for

periods with floods, major rainfall events, or long sequences of gaps (Di Piazza *et al.*, 2011). Linear regression methods assume linearity between variables, which may not always be valid (Mwale *et al.*, 2012). Multiple regression approaches ignore existing information in the target variable and need many explanatory variables, which can lead to multicolinearity issues (Miaou, 1990). Autoregression and recession models require considerable amount of data for training and validation (Amisigo and van de Giesen, 2005). Dynamic state-space models require prior knowledge of the model parameters and the modelling system while the conditioning is done only on past observations (Berendrecht and van Geer, 2016). Regression trees like Random Forests (Breiman, 2001) suffer of a lack of understanding of the construction of the trees (Sidibe *et al.*, 2018). Artificial neural networks (ANNs) have complex formulations leading to intense calculations with high computational cost (Dawson *et al.*, 2002), and the resulting model parameters generally have no physical interpretation. Findings from comparison studies revealed that methods can outperform each other depending on the dataset used (Campozano *et al.*, 2015). The choice of an appropriate infilling technique depends on factors including the length of the gaps, the seasons of gap occurrence, the climatic region of the measurement sites, the length and characteristics of the existing records, the availability of ancillary or proxy data, the accuracy of the estimates versus the complexity of the approach, and the purpose of the use of the gap-filled records .

Here, the Direct Sampling (DS) method (Mariethoz *et al.*, 2010) is proposed as an alternative to the above approaches. It is a multiple-point statistics (MPS) algorithm suited to pattern reproduction (Guardiano and Srivastava, 1993). The main advantage of DS, as a stochastic method, is its ability to provide probabilistic estimates of the missing values, which allows uncertainty quantification, a very important feature in hydrograph estimation (Beven, 2016), in contrast to deterministic methods. Moreover, as a data driven approach, it can fit various data structures and simulate the outcome of a complex natural process without assuming a specific statistical model (Oriani *et al.*, 2016). Other advantages of DS are the simplicity of its application, the multisite capability, the possibility to use auxiliary variables and predictors that contains gaps, and the ability to condition the simulation on both past and future observations for a given gap. In the present study, DS is used to infill gaps in daily streamflow data in the transboundary Volta River basin (VRB), located in West Africa. It is a data scarce region with a poor streamflow gauging density, and the available data often present long gaps in the time series. On average, many stations show gaps up to 80% of daily records per year, while about 10% of them have complete time series in some years between 1950 and 2016 (Figure 3.1).

Some studies have proposed valuable approaches to infill gaps in streamflow data in the VRB. However, they often present limitations (Dembélé *et al.*, 2019). The aim of this study is to formulate a simple infilling method that is relevant for the entire VRB and accessible to a large audience so that different water resources practitioners can adopt it. The method should be user-friendly and only require basic analytic skills, limited computational power, and solely streamflow records to avoid dependence on supplemental data. Therefore, a good candidate to that demand is the DS technique. The current study extends the gap-filling approach with DS to a large number of stations and assesses for the first time the performance of the method at a large scale. The benefit of this work is the development of a thorough gap-filling framework with a stepwise implementation of the method that includes the choice of the predictor station and the setup of the algorithm. This involves the consideration of different climatic zones, the upstream-downstream and spatiotemporal relations among a large number of gauging stations with different flow regimes, the strong streamflow seasonality, and the application to an entire river basin, which altogether constitute the novelty of this study. The resulting complete daily streamflow dataset can be used for a better understanding of the water balance in the entire VRB, and more specifically for water accounting purposes.

3.2 Streamflow data

Data acquisition in West Africa is a rather complicated task (Paturel *et al.*, 2003; Taylor *et al.*, 2006). For this study, daily streamflow records were obtained from different national hydrological services, regional project databases, and global online platforms (see the acknowledgement section). The time series are available between 1950 and 2016 and, for 81 streamflow gauging stations spread all over the VRB (Figure 3.1), with different completeness per year and per station. In the pre-processing phase, the data collected from different sources have been quality-checked and merged into a unique dataset. As the datasets were gathered from different sources, it often happens that a gauging station has data from several sources. Therefore, a first comparison was done among the sources to complete the missing portions of the time series when they had similar data in overlapping periods. The data from the official or national data centre was kept when differences were found in the data from several sources. Moreover, errors in floating point notation were corrected. No-data values were uniformed in all sources and streamflow units were converted from cfs to $m^3 s^{-1}$ for some stations.

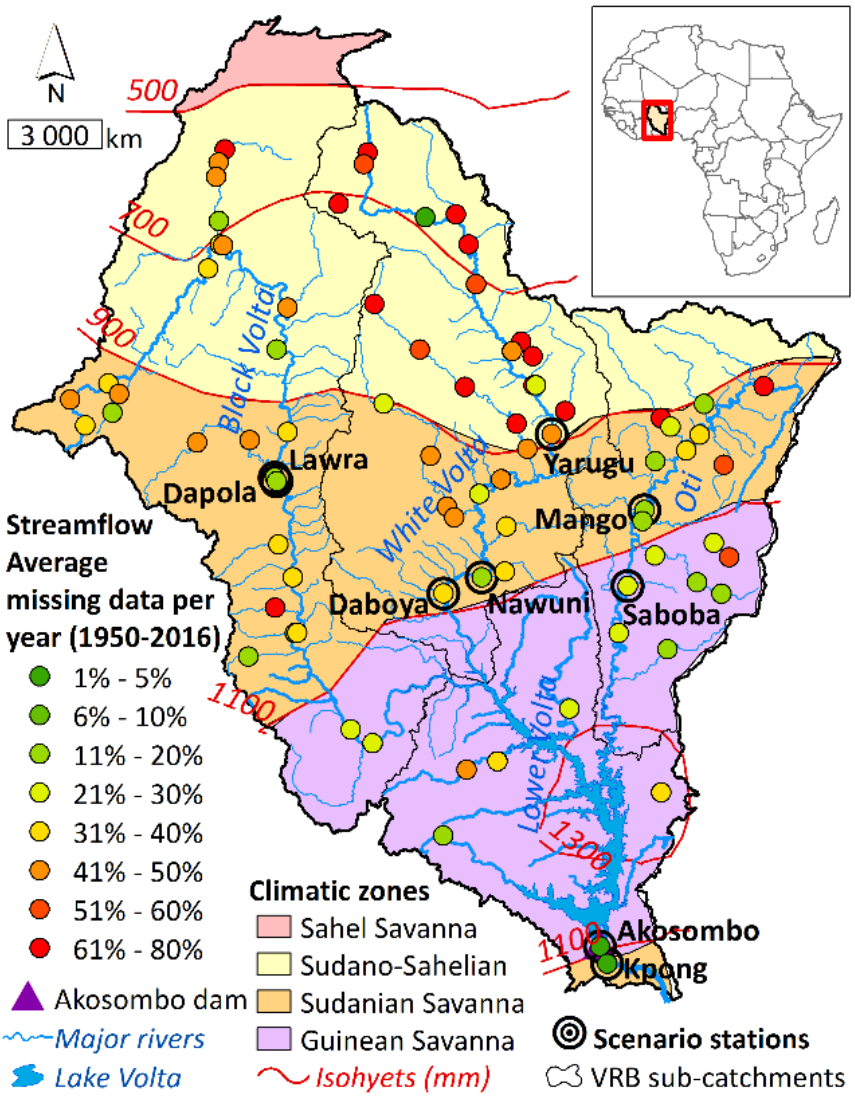


Figure 3.1. Proportion of missing data between 1950 and 2016 per streamflow gauging stations located in different climatic zones in the VRB

3.3 Gap-filling with Direct Sampling

3.3.1 Direct Sampling method

The principle of the Direct Sampling (DS) algorithm (Mariethoz *et al.*, 2010) is that it generates new simulated values based on a conditional resampling of a provided training dataset (TI). The newly simulated values are called simulation grid (SG). The methodology is generic and can be adapted to various application cases of stochastic conditional or unconditional simulations, requiring the definition of a specific simulation framework, i.e. the choice of the TI, the SG, the nature of the auxiliary variables, and the parametrization of the algorithm. In the proposed framework, the simulated data (stored in the SG) are sampled from the historical record of the same station (used as TI), where a similar neighbourhood data pattern occurs. Only uninformed

time steps (i.e. missing values in the time series) are simulated in the SG, and the existing data is used for conditioning. Other auxiliary variables can be used as conditioning dataset to better inform the simulation. The auxiliary variables can be streamflow time series from other stations. Those stations can be chosen based on their proximity and their similarity with the target station (Rees, 2008; Wagener *et al.*, 2007). The DS implementation used in this study is DeeSse (Straubhaar, 2017). The code is available upon request to the Randlab team at the University of Neuchâtel. Here, the target variable is simulated with one or more auxiliary variables that are used for conditioning, and can be partially or fully informed (i.e. with or without gaps). The workflow for continuous variables is described hereafter.

The inputs required for the simulation are: 1) a TI that contains the target variable *Z* at informed time steps, as well as auxiliary variables, and 2) a simulation grid (SG) that is a time vector hosting the simulated target variable. The time steps are uniformly spaced and identical in both TI and SG. The simulation of the target variable follows a random path and completes the SG at non-consecutive time steps. The SG is filled progressively and becomes the actual output of the simulation.

With the time vectors $x = [x_1, \dots, x_{\mathrm{n}}]$ allocated to the SG and $y = [y_1, \dots, y_{\mathrm{m}}]$ allocated to the TI, the algorithm runs through the following steps:

1. Each continuous variable is linearly normalized to a range of [0,1] using the transformation $Z \to Z \cdot (max(Z) - min(Z))^{-1}$ for patterns comparison at step 6.

2. A random simulation path $t \in \{1, \cdots, M\}$ of same length as the SG is generated by doing a random permutation of the index vector in the SG.

3. An uninformed time step t of the SG is chosen for simulation by following the random simulation path.

4. A data event $d(x_{\mathrm{t}}) = \{Z(x_{\mathrm{t}+l_1}), \cdots, Z(x_{\mathrm{t}+l_{\mathrm{n}}})\}$, representing a pattern of neighbouring data of t, is retrieved from the SG according to a radius R centred on x_{t}. It consists of at most N informed time steps, closest to x_{t}. This defines a set of lags $L = \{l_1, \dots, l_{\mathrm{n}}\}$ with $|l_{\mathrm{i}}| \leq R$ and $n \leq N$. For example, if $R = 50$ and $N = 10$, the pattern is composed by the 10 informed time steps closest to t inside the time span $[t \pm 50]$. The size of $d(x_{\mathrm{t}})$ is therefore limited by N and the available informed time steps inside the search neighbourhood window of $2R$.

5. A random time step y_i is scanned and the corresponding data event $d(y_i)$ is retrieved to be compared with $d(x_t)$ based on the same time lags.

6. A distance $D\big(d(x_t), d(y_i)\big)$ is calculated as a measure of dissimilarity between both data events, defined as follows:

$$D\big(d(x_t), d(y_i)\big) = \frac{1}{n}\sum_{j=1}^{n}\big|Z(x_j) - Z(y_j)\big|, \tag{3.1}$$

   where n is the number of elements of the data event. Independently from their position, the elements of $d(x_t)$ play an equivalent role in conditioning the simulation of $Z(x_t)$. The normalization at step 1 ensures the distances to be in the range $[0,1]$.

7. If $D\big(d(x_t), d(y_i)\big)$ is below a defined threshold $T \in [0,1]$, meaning that the two data events are sufficiently similar, the iteration stops and the datum $Z(y_i)$ is assigned to $Z(x_t)$ for all uninformed variables. Otherwise, the procedure is repeated from step 5 to 7 until a suitable $d(y_i)$ is found or a prescribed TI fraction F is scanned.

8. In case no time step corresponding to $D\big(d(x_t), d(y_i)\big) < T$ is found, the datum $Z(y_i^*)$ minimizing this distance is assigned to $Z(x_t)$.

9. The procedure from step 3 to 8 is iterated for the simulation at each x_t until the entire SG is completely informed.

10. The variables are linearly back transformed to their original range.

The auxiliary variables simultaneously undergo the same simulation steps as $Z(y_t)$. Figure 3.2 gives a synthesized description of the DS algorithm. More details on the method can be found in the work of Mariethoz *et al.* (2010) and Oriani *et al.* (2016).

The main DS user-defined parameters are the search window radius R, the maximum number of conditioning neighbour data N, and the threshold T for the dissimilarity measure. Each of them can have a different value for each variable. Another parameter is the maximum TI fraction scanned, called F.

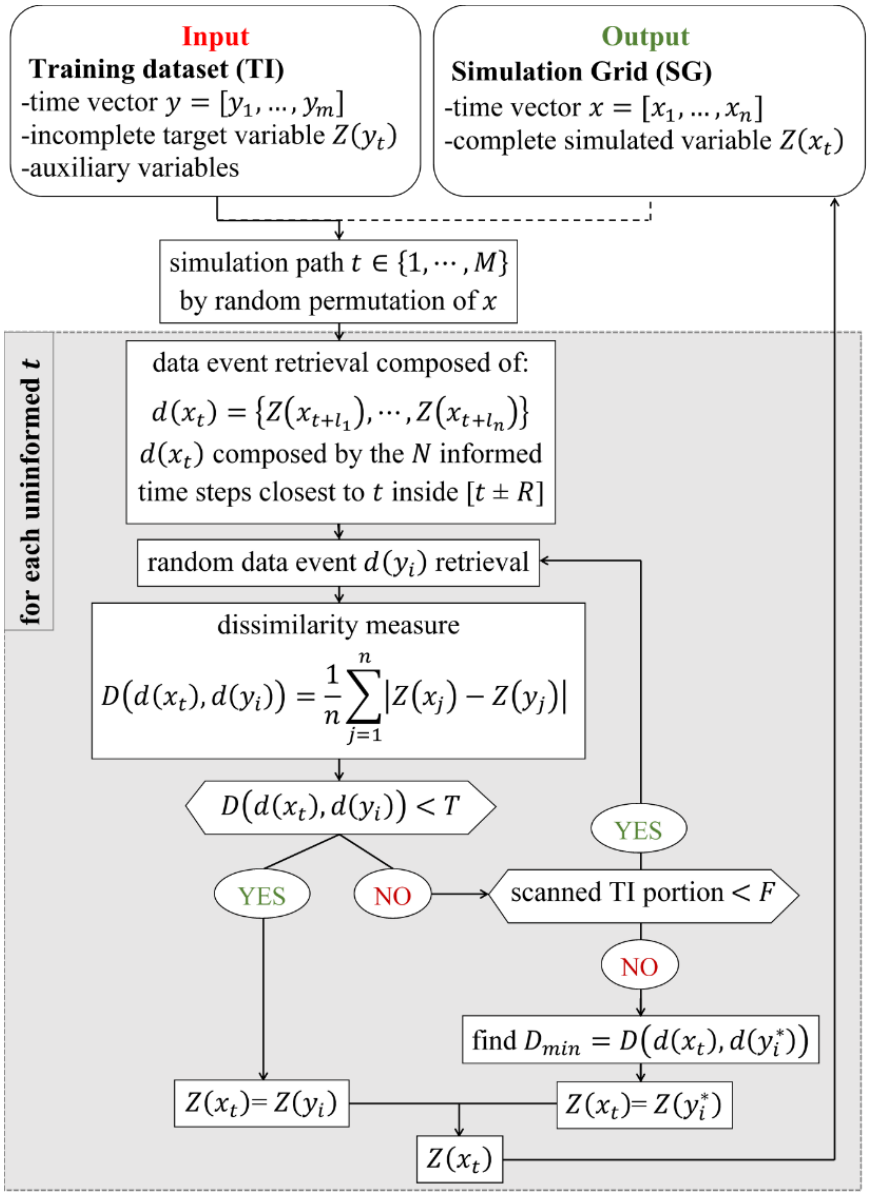


Figure 3.2. Synthesized workflow of the DS algorithm for continuous variables. R is the radius of the search neighbourhood window ($[t \pm R]$) composed by a number of neighbours N closest to x_t, with lags $L = \{l_1, \dots, l_n\}$, such that $l_i \leq R$ and $n \leq N$. D_{min} is the minimum dissimilarity found in the TI. F is the maximum TI fraction scanned.

The resampling technique used in DS has two main features that make it different from the k-nearest neighbour bootstrap (k-NN) technique (Efron, 1992). First, the algorithm uses a random path during the simulation and completes the SG at non-consecutive time steps. Consequently, the simulation at each x_t allows conditioning on both past and future time steps, as opposed to the classical Markov chain techniques that uses a linear simulation path conditioned on past data only. Secondly, the simulation follows a variable conditioning scheme that uses the n informed neighbours closest to x_t. Large-scale patterns are used at the beginning of the simulation and denser small-scale patterns toward the final iterations as the SG becomes more

populated. This variable time-dependence allows the preservation of the statistical structure of the data without a prior formulation of a high-dimensional statistical model (Oriani *et al.*, 2014). Figure 3.3 shows the simulation procedure with the DS technique.

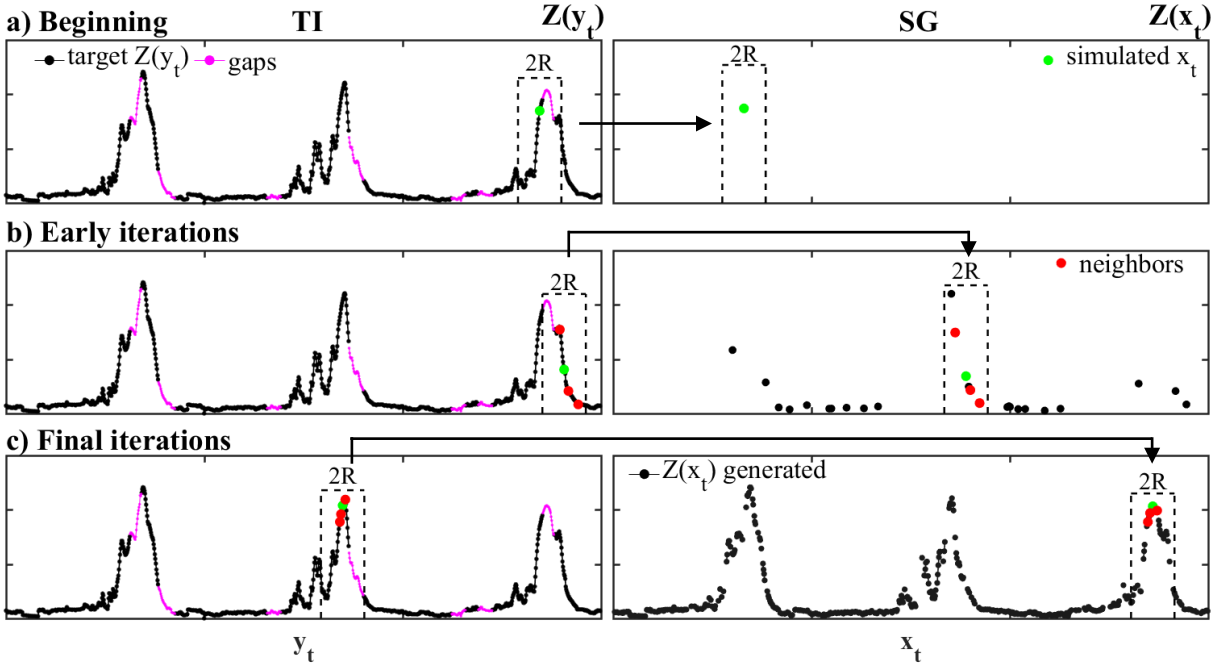


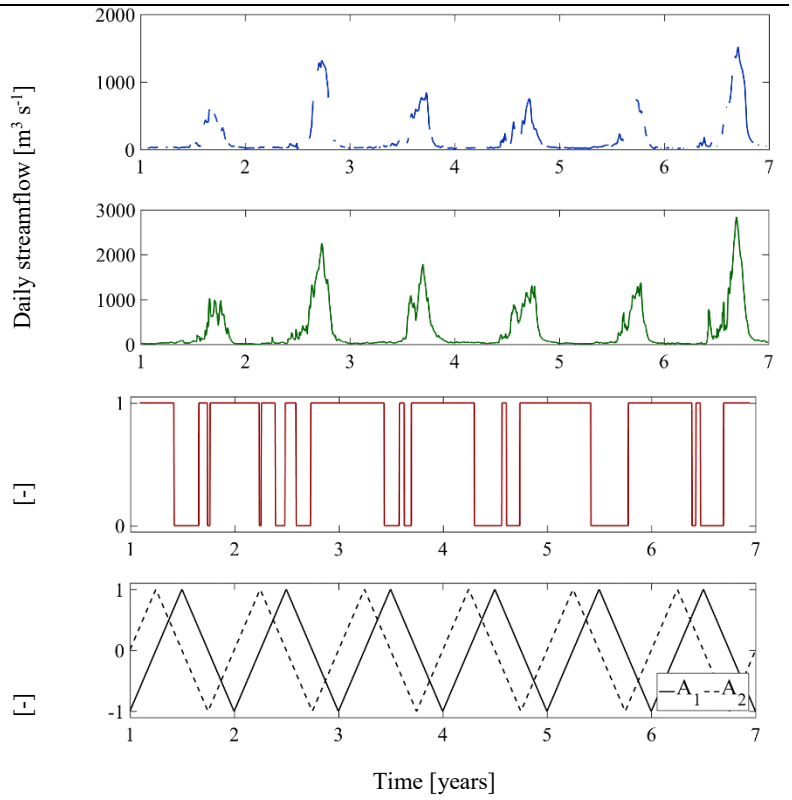
Figure 3.3. Sequential simulation of streamflow time series with DS. The dashed rectangle represents the search window defined as twice the radius R, and contains the data event that is formed by the simulated datum in green and the N neighbouring data in red.

3.3.2 DS setup for gap-filling of streamflow data

The DS setup adopted for this study is composed of the target variable $Z(y)$ and several auxiliary variables. A typical example of the setup is described in Table 3.1. $Q(y)$ is the streamflow at another correlated station, preferably located nearby, called the predictor station. The target variable contains the missing portions that are generated during the simulation and its informed time steps are considered as conditioning data. The correlated streamflow variable $Q(y)$ is given as auxiliary variable. In the example of Table 1, $Q(y)$ is complete, but in practice it can also present missing data. It helps restricting the uncertainty and provides guidance around the missing values. Considering the strong annual seasonality of streamflow in the VRB, two out-of-phase periodic triangular functions $\left(A_1(y) \text{ and } A_2(y)\right)$, each with a period of 365.25 days, are used as auxiliary variables to constrain the position of the simulated values inside the annual cycle.

Table 3.1. Example of a training dataset (TI) with DS parameters (R: search window radius; N: maximum number of conditioning neighbour data; T: distance threshold; and F: maximum TI fraction scanned, $F = 0.5$).

Variables	R	N	T	*TI* example
$Z(y)$	200	3	0.005	
$Q(y)$	200	7	0.01	
$H(y)$	200	10	0.1	
$A_1(y)$	1	1	0.05	
$A_2(y)$	1	1	0.05	



Another auxiliary variable is the recession indicator $H(y)$, computed based on $Q(y)$ and used to constrain the occurrence of high and low flows. It takes a value of $H = 0$ for a rising hydrograph limb and $H = 1$ during a recession. Its computation requires two user-defined parameters (w, v) as follows: the minimum and maximum (local extremes) of $Q(y)$ are identified inside a moving temporal window $[y \pm w]$. Each new extreme is considered only when its difference with the previous extreme is greater than v, and the next extreme is of a different type (minimum or maximum). A logical test is finally used to obtain H. A rising limb is activated with a local minimum until the occurrence of a local maximum activates a recessing limb, with a continuous alternation of both states (Oriani *et al.*, 2016). The values $w \in [10 - 30]$ and $v \in [80 - 120]$ are used in the current setup. The DS parameters for each variable for an example of TI are given in Table 3.1. Following Meerschman *et al.* (2013), the DS parameters values were taken in the ranges: $R \in [100 - 300]$, $N \in [3 - 20]$ and $T \in [0.002 - 0.1]$. R and N are in the unit of the time series in the TI, and T is unitless. For variables $A_1(y)$ and $A_2(y)$ which have a predictable behavior, high-order conditioning is not necessary

and therefore R and N are set to 1. The maximum TI fraction scanned is set to $F = 0.5$ for all the simulations. This value allows sampling a sufficient portion of the TI and avoids the phenomenon of verbatim copy, which results in mimicking portions of the TI.

3.3.3 Calibration and evaluation

In a first step, a calibration procedure is carried out to obtain an adequate set of DS parameters. Contrary to the DS setup in (Oriani *et al.*, 2016), where the target variable $Z(y)$ and the predictor variable $Q(y)$ have the same DS parameters, in this case study all variables can have different parameters values. Consequently, several sets of parameters are used to simulate the same time series and the resulting realizations are compared to the reference data using six pairwise statistical indicators. The statistical indicators, as described in the work of Dembélé and Zwart (2016), are: the Pearson correlation coefficient (r), the Spearman rank-order correlation coefficient (r_s), the Mean Error (*ME*), the *Bias*, the root mean square error (*RMSE*), and the Nash–Sutcliffe Efficiency coefficient (*NSE*). The best set of parameters is chosen as the one yielding the smallest prediction error. The rank of each set of parameters is calculated for each of the statistical indicators. The best set of parameters is taken as the one with the best cumulative rank. All results are evaluated based on visual comparison and statistical indicators. Quantile-Quantile (QQ) plots (Chambers, 2017) are used to compare the probability distribution of the simulated gap portions and reference data portions, for daily QQ-plots. The monthly and yearly QQ-plots represent the 1-month and 1-year moving average of daily values for the entire reconstructed time series.

3.3.4 Experimental set-up for artificial gap scenarios

To assess the DS performance, artificial random gaps are created inside the streamflow records of some stations having long and complete records. Representative stations are selected based on their locations in different sub-catchments with specific climatic and physical characteristics to account for different hydroclimatic settings in the VRB (Figure 2.2a). For each station, a predictor station $Q(y)$ is selected, which can be fully or partially informed. The Daboya station is chosen as the target station for all the gap-filling scenarios proposed hereafter. It is located at the outlet of the White Volta sub-catchment and receives the water drained from the semi-arid to the sub-humid zone of the VRB. A strong annual seasonality of the streamflow is depicted in Figure 3.4, which also shows the artificial gaps. The randomly created gaps vary in size from 5 to 30 consecutive days and represent 50% of the whole time series over a period of 6 years.

Ten scenarios are developed to assess the performance of the DS method (Table 3.3). Among them, five scenarios are developed by simulating $Z(y)$ with the time series of a predictor station $Q(y)$ that is fully informed (i.e. without gaps). Those scenarios are numbered with odd numbers between 1 and 9. The predictor stations in scenarios 1 and 3, the Yarugu and Nuwuni stations respectively, belong to the same sub-catchment as the target station Daboya. However, in scenarios 5, 7 and 9, the predictor stations, Akosombo, Lawra and Saboba respectively, belong to different sub-catchments from Daboya. They are located downstream of the Lower Volta, Black Volta and Oti sub-catchments (Figure 3.1), respectively. Therefore, they all receive streamflow from different sub-catchments and with different characteristics from that of Daboya.

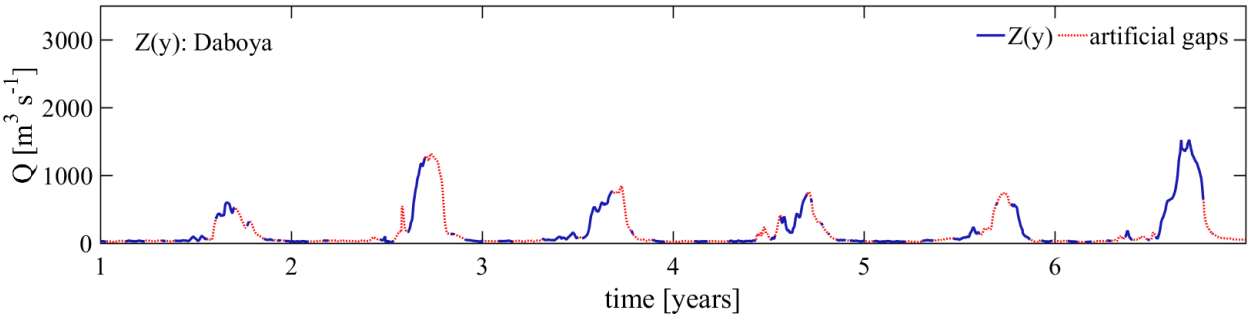


Figure 3.4. Time series of the target station with artificial gaps

The five remaining scenarios, numbered with even numbers between 2 and 10, are duplicates of the scenarios in the first group with the exception that the time series of the predictor stations are not fully informed (i.e. contain gaps). For each scenario, $Z(y)$ always contains 50% of gaps, and when $Q(y)$ contains gaps, they represent 30% of the time series or about 2 out of 6 years of missing values. To determine whether the dependence between target and predictor variables is maintained in the simulated values, the corresponding correlations before and after simulation are given in Table 3.3 for each scenario. The predictor variables are chosen purposely to have some with weak correlation and others with strong correlation with the target variable. For each scenario, ten realizations of the target variable are produced. A comparison of DS with the linear regression method is carried out and the results can be found in the Appendix 2. After demonstrating the performance of the DS for infilling artificial gaps in streamflow time series with various scenarios, the method is used to infill real gaps.

3.3.5 Predictor station selection for real gaps reconstruction

The selection of the most appropriate predictor station is not a trivial task . Here, the most similar station to the target station is chosen as predictor. Only one predictor is used to condition the target station, which avoids conflicting information among potential predictors, prevents unrealistic simulated time-series structures, and reduces computational time. Therefore, the choice of the predictor station is crucial and depends on its similarity with the target station (Wagener *et al.*, 2007). Both r and r_s are used to estimate the statistical similarity of the variables. When several stations show the same correlation with the target station, the one with the longest overlapping time steps, the smallest proportion of gaps and the highest spatial proximity with the target is chosen as the predictor station. A priority index based on these criteria allows ranking all the stations of the river network from the one that is easiest to be gap-filled, to the one that is the most challenging. In a first run, only the overlapping periods are simulated, i.e. common years of data between the target and the predictor stations. Thereafter, the next best predictor candidate is considered to simulate the remaining portions. This procedure is executed with different predictors until a fully informed SG is obtained. In case there is no other predictor that overlaps the remaining non-covered portions of the data after the first run, the first predictor is used to fully inform the SG.

3.3.6 Simulation procedure for real gaps

The simulation is done in two steps: (1) additional gaps are created in the target variable before the simulation. A calibration is done based on the simulation of these newly created gaps only. It allows obtaining a suitable set of parameters that yield the best scores for the statistical indicators in section 3.3.3. This step is similar to the scenarios (section 3.3.4), but additional gaps are created in the target variable that already contains gaps; (2) the suitable set of parameters obtained in the first step is used to simulate the real missing values in the target variable, and therefore to reconstruct the entire time series. Six applications for real gaps reconstruction are presented and discussed using representative stations of the hydrological regimes, the agro-climatic zones, and the sub-catchments of the VRB (Table 3.2).

Table 3.2. Applications for real gaps reconstruction (App: application, WV: White Volta, BV: Black Volta, LV: Lower Volta, Q position: predictor station position relative to the target)

App	Target (Z)	Predictor (Q)	Q position	Z length (years)	missing in Z (%)	catchment	$r(Z,Q)$	$r_s(Z,Q)$
1	Daboya	Nawuni	upstream	26	18	WV	0.85	0.86
2	Saboba	Mango	upstream	39	21	Oti	0.95	0.94
3	Dapola	Lawra	upstream	60	15	BV	0.93	0.93
4	Lawra	Dapola	downstream	60	22			
5	Kpong	Akossombo	upstream	31	4	LV	0.89	0.89
6	Akosombo	Kpong	downstream	31	0.05			

The fourth and the sixth applications are duplicates of the third and the fifth, respectively, with the exception that the predictor station became the target and vice versa. This inversion is done to assess the influence of the upstream-downstream position of the predictor station on the performance of the simulation. Five hundred realizations are produced per application.

3.4 Results and discussions

3.4.1 Gap-filling scenarios evaluation

In Table 3.3, the minimum and maximum values of the Pearson ($r(Z,Q)$) and the Spearman ($r_s(Z,Q)$) correlation coefficients among the ten realizations highlight the possibility of obtaining a similar target variable with the same strength and degree of association with the predictor variable as before the simulation and before creating the artificial gaps.

The correlation values $r(Z,Q)$ and $r_s(Z,Q)$ before the simulation usually lie in the same range as after the simulation, with an average error of 2%. Therefore, DS is able to preserve the correlation between the target and the predictor stations. However, this statistical relation is preserved more or less faithfully depending on the performance in reproducing the missing portions. The capability of the DS to reproduce the missing portions of the target variable is highlighted in Figure 3.5. DS usually performs better when the predictor station is fully informed. For that reason and due to the recurrence of gaps in the collected data in the VRB, attention is paid to the even-numbered scenarios from 2 to 10 (Table 3.3), which represent the scenarios with gaps in the predictor variable.

Table 3.3. Pearson ($r(Z,Q)$) and the Spearman ($r_s(Z,Q)$) correlation coefficients between the target (Z) and the predictor (Q) variables (SC: same catchment; indicates where the target and the predictor stations are located. WV: White Volta, BV: Black Volta, LV: Lower Volta). Before gaps, Z and Q are fully informed. After gaps, Z (Daboya station) contains 50% of gaps for all scenarios, while Q contains 30% of gaps for only even-numbered scenarios.

SC	Scenario	Q station	Catchment	Created gaps in Q (%)	Before simulation				After simulation			
					Before gaps		After gaps					
					r	r_s	r	r_s	r (min-max)		r_s (min-max)	
yes	1	Yarugu	WV	0	0.593	0.669	0.656	0.680	0.580	0.604	0.611	0.619
	3	Nawuni	WV		0.984	0.913	0.988	0.908	0.988	0.989	0.910	0.916
no	5	Akosombo	LV		-0.309	-0.302	-0.475	-0.356	-0.318	-0.289	-0.256	-0.238
	7	Lawra	BV		0.871	0.809	0.920	0.814	0.840	0.876	0.752	0.781
	9	Saboba	Oti		0.947	0.822	0.951	0.828	0.944	0.950	0.789	0.819
yes	2	Yarugu	WV	30	0.593	0.669	0.689	0.660	0.550	0.606	0.608	0.629
	4	Nawuni	WV		0.984	0.913	0.988	0.904	0.983	0.986	0.881	0.899
no	6	Akosombo	LV		-0.309	-0.302	-0.430	-0.265	-0.301	-0.277	-0.270	-0.260
	8	Lawra	BV		0.871	0.809	0.911	0.824	0.852	0.869	0.720	0.739
	10	Saboba	Oti		0.947	0.822	0.952	0.857	0.939	0.950	0.795	0.816

Weakest Strongest

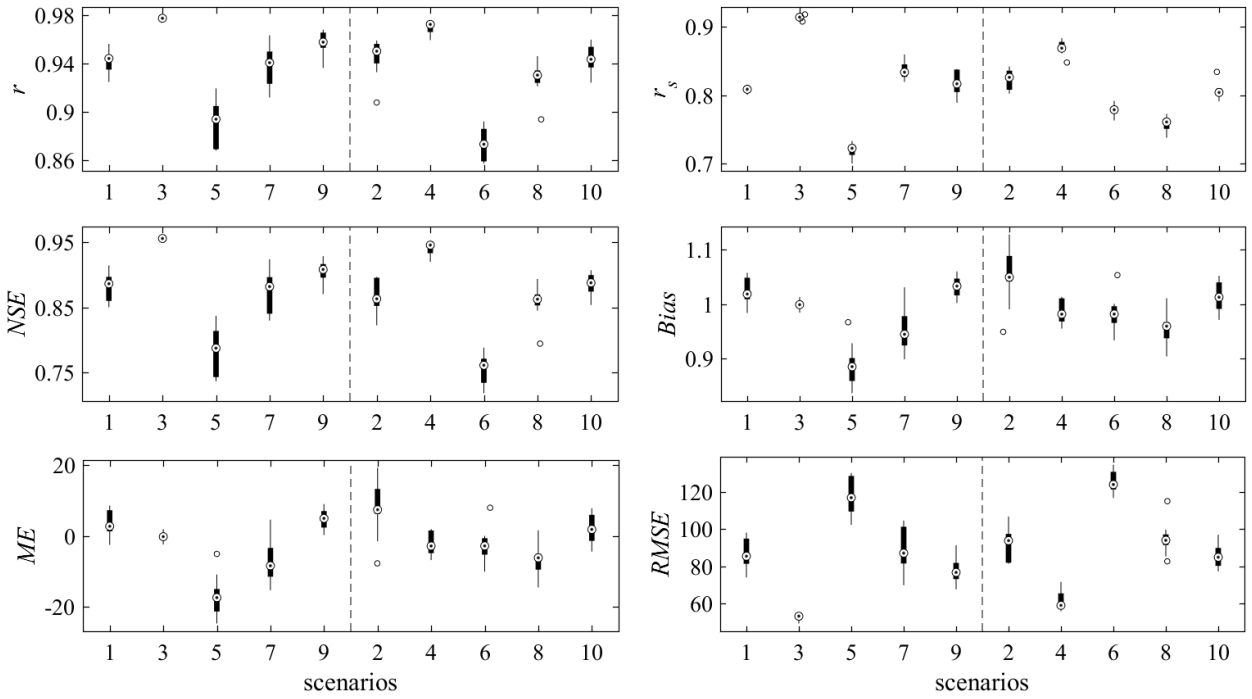


Figure 3.5. Boxplots of six statistical indicators (Y-axis) of the realizations for the ten gap-filling scenarios (X-axis). The statistical indicators (r, r_s, NSE, $Bias$, ME, and $RMSE$) are computed only between the simulated missing portions of the newly simulated target variable (Z_{sim}) and their corresponding values in the reference target variable (Z).

In this regard, scenario 4 outperforms the other scenarios for all statistical indicators. It is followed by scenarios 10, 2, and 8 in increasing order. Scenario 6 gives the weakest scores. This performance in scenario 6 can be linked to the negative and weak correlation between the target and predictor station (Table 3.3). Notwithstanding that scenario 6 gives the lowest performance, it still has good scores in terms of prediction. Such results might have been possible due to the parameter optimization step preceding the simulation. Therefore, a relevant set of parameters is obtained for a good prediction. The overall performance of the DS is good and for all considered scenarios according to the statistical indicators calculated only between the simulated portions and the corresponding reference values (Figure 3.5). The parameters used for the simulation are given in Table 3.4 for each scenario.

Table 3.4. DS parameters for the gap-filling scenarios. (see section 3.3.2 for parameters and variables description).

Scenarios	DS parameters for all variables in TI											
	R				***N***				***T***			
	A_1, A_2	H	Q	Z	A_1, A_2	H	Q	Z	A_1, A_2	H	Q	Z
1	1	200	200	200	1	10	7	3	0.05	0.1	0.02	0.02
3	1	200	200	200	1	10	7	3	0.05	0.1	0.002	0.005
5	1	200	200	200	1	10	3	7	0.05	0.1	0.02	0.005
7	1	200	200	200	1	10	7	3	0.05	0.1	0.02	0.002
9	1	200	200	200	1	10	7	3	0.05	0.1	0.01	0.005
2	1	200	200	200	1	10	3	3	0.05	0.1	0.07	0.005
4	1	200	200	200	1	10	3	3	0.05	0.1	0.002	0.002
6	1	200	200	200	1	10	7	3	0.05	0.1	0.07	0.02
8	1	200	200	200	1	10	3	7	0.05	0.1	0.05	0.02
10	1	200	200	200	1	10	7	3	0.05	0.1	0.05	0.02

3.4.2 Time series reconstruction for different scenarios

Figure 3.6 shows the time series of the reconstructed target variable and the corresponding predictor station. The average scores of the statistical indicators are given alongside the plots. Flow duration curves for the scenarios are provided in Appendix 3. Except for scenario 1, only scenarios where the predictor variable contains artificial gaps are presented because they represent the most challenging cases and are the most likely situation in data scarce regions. A comparison between scenarios 1 and 2 highlights an increase in the width of the simulation ensemble, meaning a higher uncertainty in the prediction, and a deviation of the mean of the realizations from the reference time series ($Z(y)$), when the predictor variable ($Q(y)$) is not

fully informed (scenario 2). A visual check of the scenarios shows that the realizations ensemble usually contains the target variable. Mismatches often depend on the predictor station used.

In scenario 2, the target and predictor stations are located in the same sub-catchment but they are weakly correlated ($r(Z,Q)$=0.69). The realizations mostly cover the reference time series but miss some local flow peaks, predict some false high flows, and show lags in the prediction. This is due to the inability to sample a similar data pattern in the predictor variable, therefore suboptimal patterns are chosen. In scenario 4, the reconstructions are of high quality because the target and predictor stations are in the same sub-catchment and strongly associated ($r(Z,Q)$=0.99). The overall prediction is excellent in scenario 4. The average magnitude of the simulation error ($RMSE$=61.44 $m^3\ s^{-1}$) is due to the unmet peak flows but this is balanced by the alternation of under- and overestimated high flows, which finally results in a small ME=-2.3 $m^3\ s^{-1}$.

In scenarios 6, 8 and 10, the target and the predictor stations are located in different catchments. In scenario 6, a negative and low correlation is observed between the variables ($r(Z,Q)$=-0.43). The predictor station measures the discharge of the Akosombo dam located downstream of the Lower Volta sub-catchment (Figure 3.1). This scenario is the most challenging and is not recommended even if the results show acceptable prediction performance (NSE= 0.76). The $RMSE$ is higher as observed during the second year of the simulation period. The under- and overestimated flow portions balance and result in a small average error with ME=-2.36 $m^3\ s^{-1}$. However, this might not be the case for a shorter or longer period of simulation. Consequently, simulating the target variable with a poorly informative predictor station is not advisable. In scenario 8 ($r(Z,Q)$=0.91), the predictor station is located in the Black Volta sub-catchment, where the main river is perennial and characterized by low flows. Conversely, the river gauged by the target station is mainly not perennial in its upstream part and is characterized by a long low flow period followed by a short high flow period with a long recession time. The mean of the realizations matches well the reference time series, except for some extreme flows as in the third, fourth and fifth years. Despite a narrow ensemble range, the evaluation scores are lower than those of scenario 4.

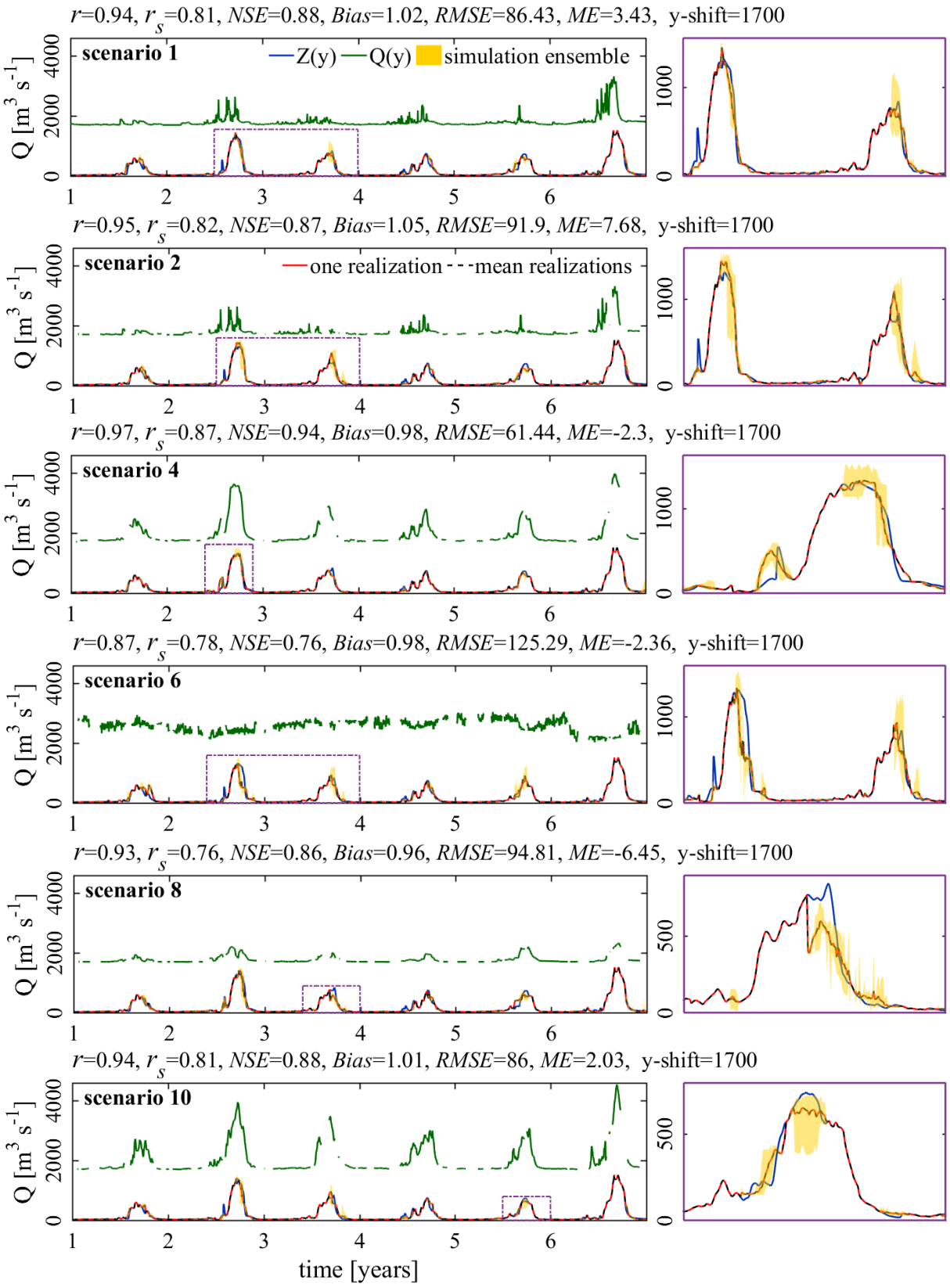


Figure 3.6. Selected gap-filling scenarios. The six statistical indicators compare the reference to the simulated values only, and represent the average for ten realizations produced per scenario. $Q(y)$ is shifted on the y-axis (y-shift) for display purpose.

In scenario 10 ($r(Z,Q)$=0.95), the predictor station is located downstream of the Oti sub-catchment where the river drains the steepest region of the VRB, and is characterized by a long low flow period followed by a short high flow period with a short recession time. The prediction shows good average evaluation scores, with some mismatches of local flow peaks and time lags in the second, third, and fifth years. The scores in scenario 10 are slightly better than those in scenario 8. This can be due to the higher proximity of the predictor station to the target station in scenario 10, which presents similar hydrological regimes and climate conditions.

The alternating under- and overestimation of flow extremes, when they occur, are balanced over time and result in a low average error. This brings to attention a possible improvement or conservation of the statistical content of the predicted time series at a lower temporal resolution. This assumption is investigated in the next section. In addition, DS performs generally well compared to the linear regression (Appendix 2).

3.4.3 Probability distributions for gap-filling scenarios

For some applications (e.g. water balance), it is important that the simulated time series reproduce the statistical content at larger temporal scales. Consequently, the probability distribution of the reference and simulated time series at different temporal resolutions are analysed. In Figure 3.7, the QQ-plots depict the probability distribution of the simulated missing values against the reference values for each scenario at daily scale, while the entire reconstructed time series is considered for monthly and yearly scales.

Daily, monthly and yearly scales are considered to assess the temporal scale effect on the DS simulations. The comparison between the daily probability distribution of the scenarios in which the predictor variable is fully informed (odd-numbered scenarios, Figure 3.7a) and those with the predictor variables partially informed (even-numbered scenarios, Figure 3.7b), shows higher deviations from the first bisector when the predictor variable is partially informed. Moreover, the uncertainty of the prediction, indicated by the range between the 5th and 95th percentiles, increases in those scenarios. At daily scale (Figure 3.7a-b), the simulations considerably overestimate (scenarios 2, 3, 7, 9 and 10) or underestimate (scenarios 1, 4, 5, 6 and 8) streamflow in the range 700-1200 m^3 s^{-1}, which is seemingly the flow peak during the second year (Figure 3.4). With a short time series, finding a similar pattern to infill this gap at the peak of the hydrograph has been challenging. This was even more acute for scenarios 1, 2, 5 and 6, in which the predictor is not strongly correlated with the target variable (Table 3.3). Moreover, the comparison between scenario 3 versus 7, 9 and scenario 4 versus 8, 10 reveals

that the simulation is better when the predictor and the target are located in the same catchment. This might be due to common factors influencing flow generation at both the predictor and target stations locations.

In general, the probability distribution is better preserved at the monthly scale (Figure 3.7c-d) and yearly scale (Figure 3.7e-f). These results suggest that the daily simulation is not significantly biased and that it preserves the large-scale variability by generating realistic daily patterns. However, the yearly QQ-plots unveil greater uncertainties of simulation compared to the monthly scale. This can be explained by the interannual variability of the streamflow.

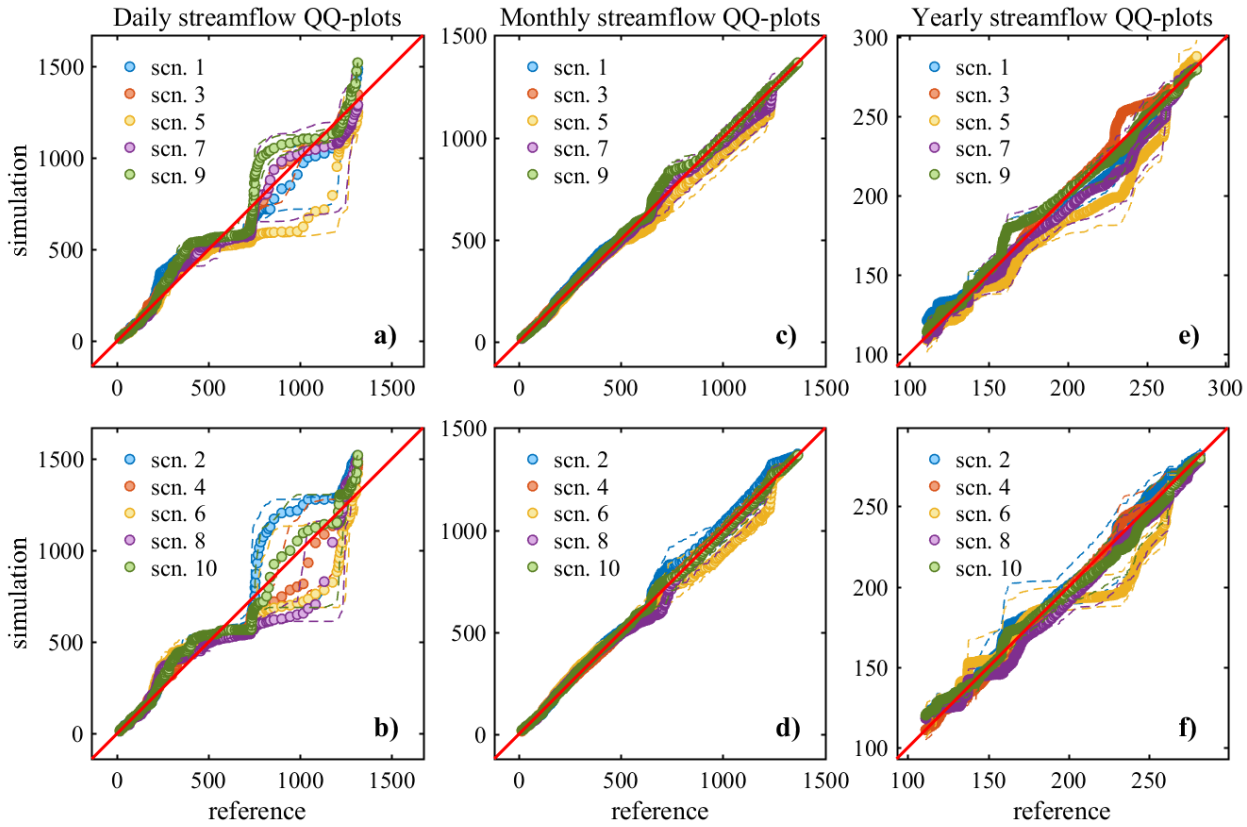


Figure 3.7. Comparison of the empirical probability distribution of the reconstructed artificial gaps (Y-axis) against the corresponding reference data (X-axis) for each scenario, using daily (a-b), monthly (c-d), and yearly (e-f) QQ-plots. The dots represent the median of the realizations and the dashed lines indicate the 5th and 95th percentiles. The predictor variable is partially informed in even-numbered scenarios (b-d-f) and fully informed in odd-numbered scenarios (a-c-e). All units are in m^3 s^{-1}.

3.4.4 Plausibility of the parameters set for simulating real gaps

Table 3.5 presents the DS parameters obtained from the calibration step.

Table 3.5. DS parameters for the reconstruction of real gaps. (see section 3.3.2 for parameters and variables description).

App	Target (Z)	Predictor (Q)	DS parameters for all variables in TI											
			R				***N***				***T***			
			A_1, A_2	H	Q	Z	A_1, A_2	H	Q	Z	A_1, A_2	H	Q	Z
1	Daboya	Nawuni	1	200	200	200	1	10	5	5	0.05	0.1	0.01	0.002
2	Saboba	Mango	1	200	200	200	1	10	5	5	0.05	0.1	0.01	0.002
3	Dapola	Lawra	1	200	200	200	1	10	5	5	0.05	0.1	0.01	0.01
4	Lawra	Dapola	1	200	200	200	1	10	5	5	0.05	0.1	0.01	0.002
5	Kpong	Akosombo	1	200	200	200	1	10	5	10	0.05	0.1	0.002	0.002
6	Akosombo	Kpong	1	200	200	200	1	10	10	5	0.05	0.1	0.002	0.002

The results of the calibration carried out before the reconstruction of real gaps are evaluated with QQ-plots shown in Figure 3.8.

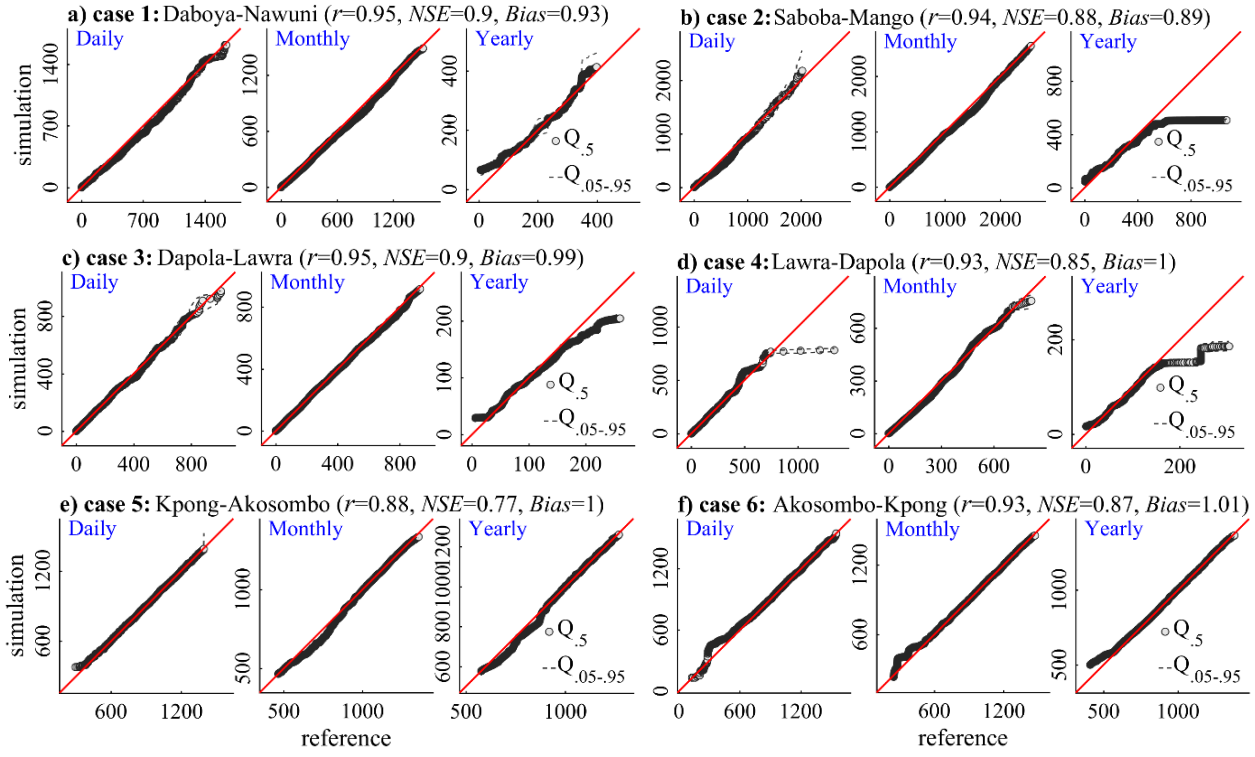


Figure 3.8. Probability distribution obtained for the calibration step, for the predicted values in six gap-filling applications with daily, monthly and yearly QQ-plots. The dots represent the median of the realizations and the dashed lines indicate the 5th and 95th percentiles. All units are in m^3 s^{-1}.

For all the considered application cases, the proposed setup gives fairly sharp predictions with average NSE scores between 0.77 and 0.90, and average $Bias$ scores between 0.89 and 1.01. The probability distribution of the predicted values matches very well that of the reference at daily scale and is further improved at monthly and yearly scales. However, some noteworthy deviations occur at daily scale for cases 4 and 6. In those cases, as the predictor station is located downstream of the target station, its streamflow records exceed those of the target station, with some flow extremes that might not be identified in the target variable. These differences can result in false high flow signals, and ultimately result in large deviations in the prediction. At yearly scale, the underestimated simulations in cases 2, 3 and 4 can be attributed to the interannual variability of the streamflow. This underestimation was predictable mainly for case 4 where the simulation is not able to predict values above 700 $m^3 s^{-1}$ at daily scale. For cases 1, 2 and 3, the bias in the yearly QQ-plots can be explained by the bias for extreme values in the daily QQ-plots. The good conservation of the probability distribution at yearly scale for case 5 and 6 is due to the high fluctuation of the flow at both the target and predictor stations that measure the discharge of the Akosombo and Kpong dams. Therefore, similar patterns to the gap portions can be easily identified in the remaining time series to match the peak flows. While for other cases, the peak flows usually occur once a year due to the annual seasonality (Figure 3.9).

3.4.5 Reconstruction of real gaps in time series

Figure 3.9 shows the reconstructed streamflow time series. While the full time series span decades, only some portions containing long gaps are shown for visualization. The characteristics of the time series are given in Table 3.2. The shape of the predicted hydrograph is well reproduced with a good timing and representation of the annual seasonality. The mean of 500 realizations has a shape that is consistent with the predictor variable. Such consistency reveals a conservation of the high strength of association between the target and the predictor variable, which is critical for the choice of the predictor variable and remains the base of the simulation. The prediction is more challenging when several consecutive years of data are missing (Figure 3.9c-d). In this case, the nearest neighbours to the uninformed time step to be simulated can be located in very different years during which other conditions might have contributed to streamflow generation. Consequently, higher uncertainty is expected and mostly during high flows (Figure 3.9a, c, d). However, this effect is minimized when flow peaks periods are exhaustively represented in the data, for example during the fifth year in Figure

3.9b. Infilling gaps in the fifth reconstruction (Figure 3.9e), should be easier due to the small proportion of missing data. The roughly one year of missing data seems to be well predicted but the strong increase in discharge at the end of the second year is suspicious in the absence of trend in the predictor variable. However, DS provides prediction with very low uncertainty for short gaps (Figure 3.9f). Flow duration curves for each of the cases are provided in the Appendix 4.

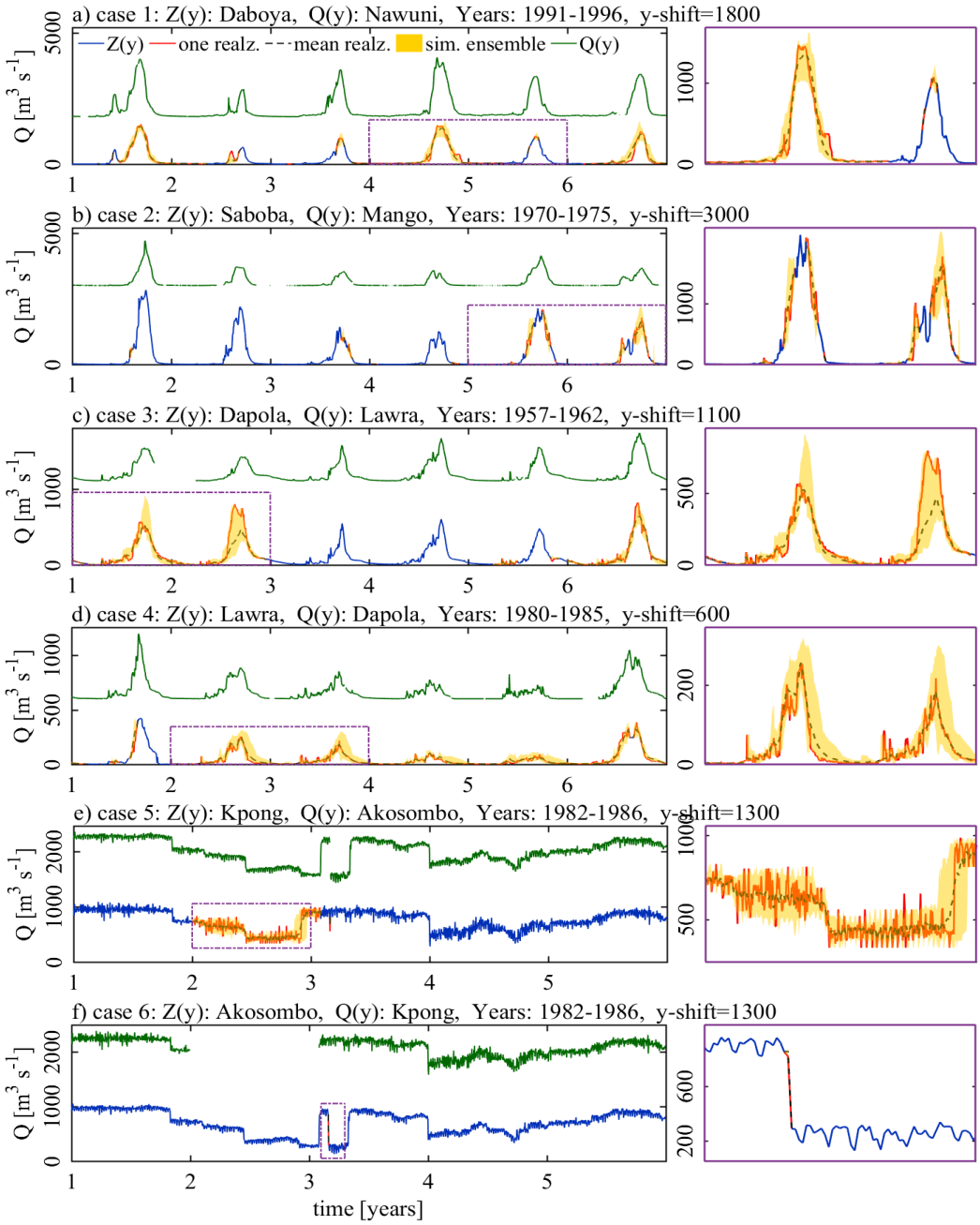


Figure 3.9. Reconstruction of streamflow records of selected stations in different sub-catchments of the VRB. Note that the time series have been shifted on the y-axis (y-shift) for display purpose and the number of years is different for each simulation.

3.5 Conclusion

This study proposes a robust framework for gap filling of streamflow data with the Direct Sampling (DS) technique, which is a data-driven approach relying on a small set of parameters. It presents the first application of DS for comprehensive gap filling of streamflow data along a complex river network in a large basin. The key elements of the developed gap-filling framework are the performance assessment based on artificial data gaps, the selection of predictor stations and the calibration of the small set of parameters.

The overall performance of the framework is satisfactory after a test for different hydroclimatic settings of the Volta River basin (VRB), West Africa. As for all data-driven approaches, the main limitation of DS is the need for historical records that are sufficiently informative for the simulation of missing data. However, as shown in this study, the proposed gap-filling framework is able to yield good predictions for large data gaps in streamflow time series. It shows a satisfying predictive performance in terms of sharpness and reliability: the statistical content of the target variable is preserved, the probability distribution of the simulation matches accurately the reference, and the shape of the hydrograph shows a good timing with a strong preservation of the annual seasonality. Even if some local over- or underestimations of flow extremes do occur, they are usually balanced over time and result in a small bias. The statistical behaviour is preserved from daily to monthly scale, giving the possibility to use the output at a temporal scale higher than the one of the simulation. The outcomes highlight that better results are obtained in the following conditions:

- the target and the predictor stations are located in the same sub-catchment;
- the predictor station is well correlated with the target variable;
- the predictor station is located upstream of the target station;
- the predictor station is fully informed;
- the target station contains rather short gaps.

For the application to time series gap filling, the technique relies on a small set of parameters that can be set up by a simple calibration procedure. Moreover, an optimization of the parameters yields better results even in case of moderate correlation between the target and predictor stations. The results obtained for the VRB with the proposed framework clearly show that DS is a promising approach for time series simulation in environmental sciences.

The DS performs better than the simple linear regression method. Although DS has several advantages compared to competing gap-filling methods, future studies should compare DS to other methods and investigate the benefit of using complex methods as compared to DS.

The uncertainty of the DS results increases with the complexity of the gap-filling scenarios (e.g. large gaps, high and low flows). Therefore, the gap-filled data should be used with caution. When using the gap-filled data for further applications such as hydrological modelling, the most reliable gap-filled datasets (i.e. short gaps to be filled) should be preferred to limit the propagation of uncertainties into subsequent applications.

The possibility of using the DS without an exogenous auxiliary variable (e.g. rainfall, evaporation) and its ability to use a predictor variable that also contains gaps make it a powerful tool that can be easily used in data scarce regions and elsewhere. Further development of the current gap-filling framework might focus on the simultaneous use of multiple predictor variables and exogenous auxiliary variable to inform the simulation.

Chapter 4

Evaluation of Satellite and Reanalysis Meteorological Datasets for Hydrological Modelling*

It is presumed that there exists a great unity in nature,
in respect of the adequacy of a single cause to account for
many different kinds of consequences.
Immanuel Kant

Doubt is not a pleasant condition, but certainty is absurd.
Voltaire

* This chapter is based on the following publication:
Dembélé, M., B. Schaefli, N. van de Giesen, and G. Mariéthoz (under revirew in HESS), Suitability of 17 rainfall and temperature gridded datasets for large-scale hydrological modelling in West Africa, Hydrol Earth Syst Sci. https://doi.org/10.5194/hess-2020-68

Abstract

This study evaluates the ability of different gridded rainfall datasets to plausibly represent the spatiotemporal patterns of multiple hydrological processes (i.e. streamflow, actual evaporation, soil moisture and terrestrial water storage) for large-scale hydrological modelling in the predominantly semi-arid Volta River Basin (VRB) in West Africa. Seventeen precipitation products based on satellite data (TAMSAT, CHIRPS, ARC, RFE, MSWEP, GSMaP, PERSIANN-CDR, CMORPH-CRT, TRMM 3B42, TRMM 3B42RT) and on reanalysis (JRA-55, EWEMBI, WFDEI-GPCC, WFDEI-CRU, MERRA-2, PGF and ERA5) are compared as input for the fully distributed mesoscale Hydrologic Model (mHM). To assess the model sensitivity to meteorological forcing during rainfall partitioning into evaporation and runoff, six different temperature reanalysis datasets are used in combination with the precipitation datasets, which results in evaluating 102 combinations of rainfall-temperature input data. The model is recalibrated for each of the 102 input combinations, and the model responses are evaluated by using in-situ streamflow data and satellite remote sensing datasets from GLEAM evaporation, ESA CCI soil moisture, and GRACE terrestrial water storage. A bias-insensitive metric is used to assess the impact of meteorological forcing on the simulation of the spatial patterns of hydrological processes. The results of the process-based evaluation show that the rainfall datasets have contrasting performances across the four climatic zones present in the VRB, which suggests cautiousness in performance generalizability to different spatial domains. No single rainfall or temperature dataset consistently ranks first in reproducing the spatiotemporal variability of all hydrological processes. A dataset that is best in reproducing the temporal dynamics is not necessarily the best for the spatial patterns. In addition, the results suggest that there is more uncertainty in representing the spatial patterns of hydrological processes than their temporal dynamics. Finally, some region-tailored datasets outperform the global datasets, thereby stressing the necessity and importance of regional evaluation studies for satellite and reanalysis meteorological datasets.

4.1 Introduction

Our understanding of environmental systems is underpinned by observational data whose unavailability and uncertainties hinder research and operational applications. Among other factors, atmospheric data quality is of prime importance for the reliability of hydro-meteorological and climatological studies (Ledesma and Futter, 2017; Zandler *et al.*, 2019). Precipitation is one of the major components of the water cycle, which has led to numerous initiatives on understanding its generation, and estimating its amount and variability on Earth (Cui *et al.*, 2019; Maidment *et al.*, 2015). In hydrological modelling (Beven, 2019b; Singh, 2018), precipitation is the most important driver variable that determines the spatiotemporal variability of other hydrological fluxes and state variables (Bárdossy and Das, 2008; Thiemig *et al.*, 2013).

With the development of distributed hydrological models that facilitate large-scale predictions (Clark *et al.*, 2017; Fatichi *et al.*, 2016; Ocio *et al.*, 2019), there is a growing need to inform and evaluate those models with distributed observational datasets to improve spatiotemporal process representation (Baroni *et al.*, 2019; Hrachowitz and Clark, 2017; Paniconi and Putti, 2015). A key challenge is the spatiotemporal intermittency of precipitation, which is a major challenge for its measurement and its spatial interpolation (Acharya *et al.*, 2019; Bárdossy and Pegram, 2013; Tauro *et al.*, 2018; Wagner *et al.*, 2012a) , especially in regions with particular features such as complex topography, convection-driven precipitation or snowfall occurrence. A comprehensive description of precipitation measurement techniques can be found in previous studies (e.g. Kidd and Huffman, 2011; Stephens and Kummerow, 2007; Tapiador *et al.*, 2012). The drawbacks of in-situ measurements of precipitation include limited and uneven areal coverage, deficiencies in instruments and costly maintenance (Awange *et al.*, 2019; Harrison *et al.*, 2019; Kidd *et al.*, 2017), and have led to the advent of precipitation estimation from space (Barrett and Martin, 1981). Precipitation estimates from space are spatially homogeneous and cover inaccessible regions with uninterrupted records over time (Beck *et al.*, 2019b; Funk *et al.*, 2015).

The advent of satellite-based rainfall products (SRPs) has opened up new avenues for water resources monitoring and prediction, especially in data sparse regions (Hrachowitz *et al.*, 2013; Serrat-Capdevila *et al.*, 2014; Sheffield *et al.*, 2018). Although, the use of SRPs in hydrology is increasing (Chen and Wang, 2018; Xu *et al.*, 2014), they have not been fully adopted for operational purposes yet (Ciabatta *et al.*, 2016; Kidd and Levizzani, 2011). The limited uptake of SRPs in hydrology is due to measurement bias, inadequate spatiotemporal resolutions (e.g.

for extreme event simulation) and shortness of the records for some applications (e.g., climate change impact assessments), and the skepticism of some potential users with regard to the data quality (Marra *et al.*, 2019). In the past decades, a large number of SRPs have been developed with different objectives, spatial and temporal resolutions, input sources, algorithms and acquisition methods (Ashouri *et al.*, 2015; Brocca *et al.*, 2019; Ciabatta *et al.*, 2018). Several studies provide a review of SRPs (e.g. Le Coz and van de Giesen, 2019; Maggioni *et al.*, 2016; Maidment *et al.*, 2014; Sun *et al.*, 2018).

In addition to SRPs, there are also atmospheric retrospective analysis (or reanalysis) datasets of precipitation. A reanalysis system is composed of a forecast model and a data assimilation scheme that integrates spatiotemporal observations of meteorological variables (i.e. temperature, humidity, wind and pressure) to generate gridded atmospheric data (Lorenz and Kunstmann, 2012; Schröder *et al.*, 2018). Precipitation is one of the reanalysis model-generated fields that generally has more uncertainties than the meteorological state fields (Roca *et al.*, 2019). Reanalysis datasets are often used in hydrological modelling (Duan *et al.*, 2019; Gründemann *et al.*, 2018; Tang *et al.*, 2019), and sometimes they are preferred over SRPs because of their usually long-term records suitable for climate change studies, and because of their higher performance in predictable large-scale stratiform systems (Potter *et al.*, 2018; Seyyedi *et al.*, 2015).

Despite the progress in satellite instruments, which has led to substantial advances in improving precipitation estimates (Sorooshian *et al.*, 2011; Tang *et al.*, 2019), there are known inconsistencies among the available SRPs (Sun *et al.*, 2018; Tapiador *et al.*, 2017). SRPs are subject to inherent errors originating mainly from precipitation retrieval instruments and algorithms, sampling frequency, and inadequate representation of cloud physics in some regions (Alazzy *et al.*, 2017; Laiti *et al.*, 2018; Romilly and Gebremichael, 2011). While on the one hand SRPs are subject to systematic biases, reanalysis products on the other hand have uncertainties resulting from their model forcing parameters, low spatial resolution with poor representation of sub-grid processes, and the model physics (Bosilovich *et al.*, 2008; Laiti *et al.*, 2018). Uncertainty quantification both in SRPs and reanalysis data is subject to intense research (e.g. Awange *et al.*, 2016; Gebremichael, 2010; Maggioni *et al.*, 2016; Westerberg and Birkel, 2015). The errors quantification of SRPs and reanalysis products is usually done by comparing them with in-situ measurements (e.g. Beck *et al.*, 2019a; Caroletti *et al.*, 2019; Dembélé and Zwart, 2016; Satgé *et al.*, 2020; Thiemig *et al.*, 2012), or by assessing their reliability as forcing for hydrological models (e.g.Duethmann *et al.*, 2013; Nkiaka *et al.*, 2017;

Pan *et al.*, 2010). Other evaluation approaches include triple collocation, which is a technique that estimates the variance of unknown errors of three independent variables without a reference or observed variable (e.g. Alemohammad *et al.*, 2015; Massari *et al.*, 2017; McColl *et al.*, 2014; Roebeling *et al.*, 2012). Compared to the ground-truthing approach, the hydrological evaluation approach has received limited attention (Camici *et al.*, 2018; Poméon *et al.*, 2017).

The non-linearity of hydrological processes in rainfall-runoff modelling can reduce or amplify the errors in the used input rainfall data and result in a satisfactory or poor representation of the hydrological responses (Beven, 2011; Blöschl and Zehe, 2005; Clark *et al.*, 2009; Maggioni and Massari, 2018; Nijssen, 2004; Peel and McMahon, 2020; Wagener *et al.*, 2004). Consequently, the hydrological model can give a good representation of a hydrological state or flux variable for the wrong reasons (cf. Kirchner, 2006), thereby potentially leading to unfortunate consequences for water resources management (Zambrano-Bigiarini *et al.*, 2017). When testing models as hypotheses (Beven, 2018; Pfister and Kirchner, 2017), type I errors (i.e. false positive model acceptability; Beven, 2010) should be avoided to ensure a high predictive skill of the model and its correctness for good decision-making. This sheds light on the importance of assessing the reliability of hydrological predictions generated with the use of SRPs and reanalysis products (Behrangi *et al.*, 2011; Kuczera *et al.*, 2010). In this context, knowing the adequacy and coherence of meteorological data in reproducing hydrological processes is a prerequisite to data selection for water resources management (Casse *et al.*, 2015; Laiti *et al.*, 2018).

In the context of hydrological evaluation of precipitation datasets, some limitations can be identified in previous studies. Some studies only evaluate a small number of precipitation datasets or do not consider reanalysis products (e.g. Bhattacharya *et al.*, 2019; Bitew and Gebremichael, 2011; Liu *et al.*, 2017; Ma *et al.*, 2018). Usually, the influence of temperature datasets in combination with rainfall datasets is not tested (e.g. Camici *et al.*, 2018; Casse *et al.*, 2015; Qi *et al.*, 2016; Satgé *et al.*, 2019; Zhang *et al.*, 2019), with the exception of a few studies (e.g. Laiti *et al.*, 2018; Lauri *et al.*, 2014), despite the importance of this interaction for evaporation simulation. Most studies evaluate a single hydrological state or flux variable, generally streamflow (e.g. Li *et al.*, 2012c; Poméon *et al.*, 2017; Seyyedi *et al.*, 2015; Shayeghi *et al.*, 2020), or soil moisture (e.g. Brocca *et al.*, 2013). Some studies use lumped or semi-distributed models, therefore averaging the rainfall amount on large areas (e.g. Duan *et al.*, 2019; Gosset *et al.*, 2013; Shawul and Chakma, 2020; Tang *et al.*, 2019; Tobin and Bennett, 2014), which reduces the bias effect that could occur at the pixel level with a fully distributed

model. Often, the model is not recalibrated for each precipitation dataset (e.g. Li *et al.*, 2012b; Su *et al.*, 2008; Tramblay *et al.*, 2016; Voisin *et al.*, 2008), which is, however, a prerequisite for reliable input field assessment (Stisen *et al.*, 2012). Moreover, some studies perform a global-scale analysis and ignore regionally tailored products (e.g. Beck *et al.*, 2017b; Fekete *et al.*, 2004; Mazzoleni *et al.*, 2019), which can outperform global products (e.g. Thiemig *et al.*, 2013). Finally, no study evaluated the simultaneous impact of various precipitation and temperature datasets on the spatial patterns of several hydrological processes (i.e. soil moisture and evaporation).

In light of the above, this study assesses the adequacy of different combinations of 17 precipitation datasets (10 SRPs and 7 reanalysis products) and 6 temperature datasets from reanalysis, when used as forcing data for a fully distributed hydrological model, in reproducing the spatiotemporal variability of multiple hydrological processes (i.e. streamflow, actual evaporation, soil moisture, and terrestrial water storage). In total, 102 rainfall-temperature input data combinations are tested with the mesoscale Hydrologic Model (mHM) by recalibrating the model for each of the input data combinations. The experiment is carried out in the poorly gauged and predominantly semi-arid Volta River Basin (VRB) located in West Africa, over the period 2003-2012. It is noteworthy that the goal of this study is not to estimate the intrinsic quality of the meteorological forcing (i.e. precipitation and temperature) but rather to understand the impact of the propagation of associated uncertainties on the simulation of hydrological processes (Bhuiyan *et al.*, 2019; Falck *et al.*, 2015; Marthews *et al.*, 2020).

The VRB case study is particularly interesting from both scientific and societal perspectives. On the one hand, precipitation modelling in tropical monsoon climates is a challenging task due to strong seasonality and diurnal variations of rainfall (Cook and Vizy, 2019; Pfeifroth *et al.*, 2016; Turner *et al.*, 2011), and due to isolated convection systems in semi-arid regions (Mathon *et al.*, 2002; Parker and Diop-Kane, 2017; Taylor *et al.*, 2017). On the other hand, open access and good quality datasets are needed for water resources management in West Africa (Di Baldassarre *et al.*, 2010; Dinku, 2019; Roudier *et al.*, 2014a; Serdeczny *et al.*, 2017). The following research questions are addressed:

- What is the impact of different gridded rainfall and temperature datasets on the simulation of hydrological fluxes and state variables?
- How important is the choice of meteorological datasets for the representation of spatial patterns versus temporal dynamics?

A growing interest in using satellite remote sensing data in hydrological modelling is expected (McCabe *et al.*, 2017; Peters-Lidard *et al.*, 2017; Wilkinson *et al.*, 2016). Therefore, knowing the suitability of the input data for hydrological modelling is a prerequisite for reliable spatiotemporal predictions, as the goal is to increase model performance with minimum uncertainty (Beven, 2016; McMillan *et al.*, 2018; Savenije, 2009).

4.2 Overview of the modelling experiment

The adequacy of the rainfall and temperature datasets to plausibly reproduce various hydrological processes is tested with all the 102 possible combinations of 17 rainfall and 6 temperature datasets used as meteorological forcing (see section 4.3). Different temperature datasets are used to allow flexibility in rainfall partitioning into evaporation and runoff because temperature is a key variable for the calculation of potential evaporation (Kirchner and Allen, 2020; Van Stan *et al.*, 2020; Zheng *et al.*, 2019). The hydrological model is recalibrated for each of the 102 combinations of rainfall-temperature datasets (Figure 4.1).

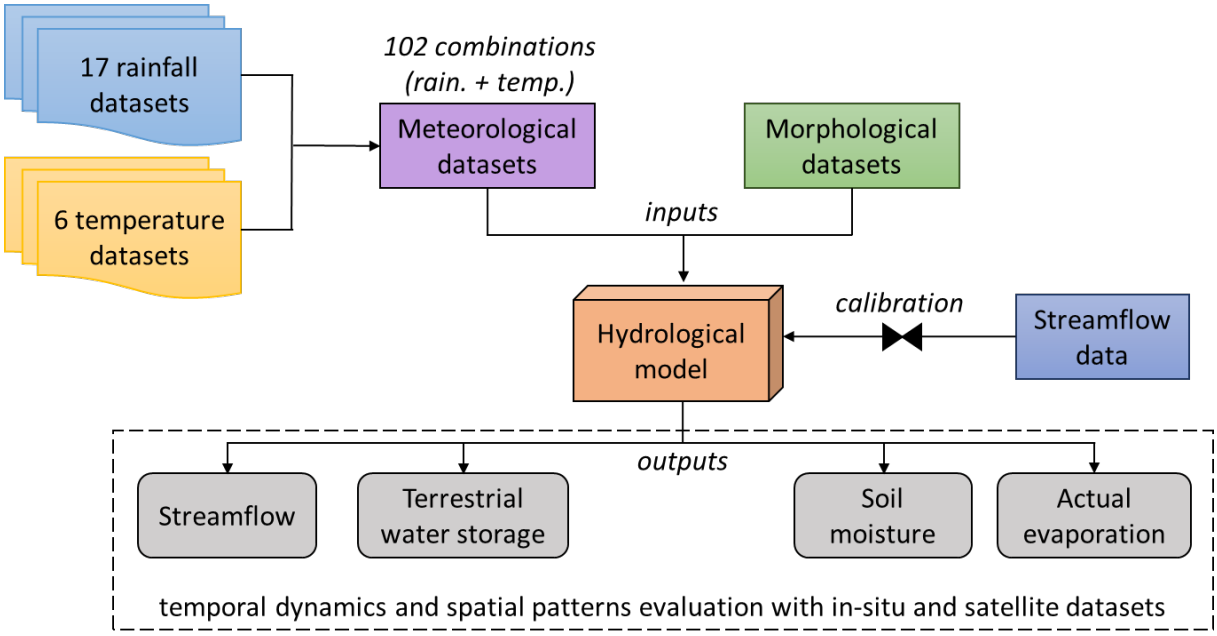


Figure 4.1. Flowchart of the methodology used to evaluate the suitability of meteorological datasets in reproducing plausible hydrological processes.

The differences in the performance of model outputs are assumed to result from the propagation of the input data uncertainty through the model simulations (Fallah *et al.*, 2020; Nikolopoulos *et al.*, 2010). In case of uncertainties resulting from the hydrological model structure, these uncertainties can be assumed to remain consistent for all the input datasets and therefore it

should not hinder the interpretation of the results, because only the parameters change during model calibration and not the model structure (Raimonet *et al.*, 2017).

4.3 Meteorological datasets

This study evaluates 17 rainfall products composed of 10 satellite-based products: TAMSAT, CHIRPS, ARC, RFE, MSWEP, GSMaP, PERSIANN-CDR, CMORPH-CRT, TRMM 3B42 and TRMM 3B42RT; and 7 reanalysis products: JRA-55, EWEMBI, WFDEI-GPCC, WFDEI-CRU, MERRA-2, PGF and ERA5 (Table 4.1). The data access portals and their full names are provided in Appendix 5. Widely used global and Africa-tailored datasets were selected based on their availability in the period for which streamflow data is available for the hydrological modelling (2000-2012). For SRPs having multiple versions, the gauge-corrected version was selected to avoid the known systematic biases found in the SRPs as compared to ground measurements (Jiang and Wang, 2019; Pellarin *et al.*, 2020). The selected rainfall datasets include single and multi-sensor, with various merged and gauge-corrected products obtained from rain gauges, microwave sensors on low Earth orbits and infrared sensors on geostationary satellites (Golian *et al.*, 2019; Maggioni and Massari, 2018; Thiemig *et al.*, 2013). Moreover, six different datasets of air temperature (at 2 m above ground) are used for the calculation of potential evaporation and they are obtained from the reanalysis products: JRA-55, EWEMBI, WFDEI, MERRA-2, PGF and ERA5.

The maps of spatial patterns of rainfall and temperature in the VRB for different datasets are shown in Figure 4.2 and Figure 4.3. The climatology of rainfall and temperature per climatic zones are provided in Appendix 8 to Appendix 11.

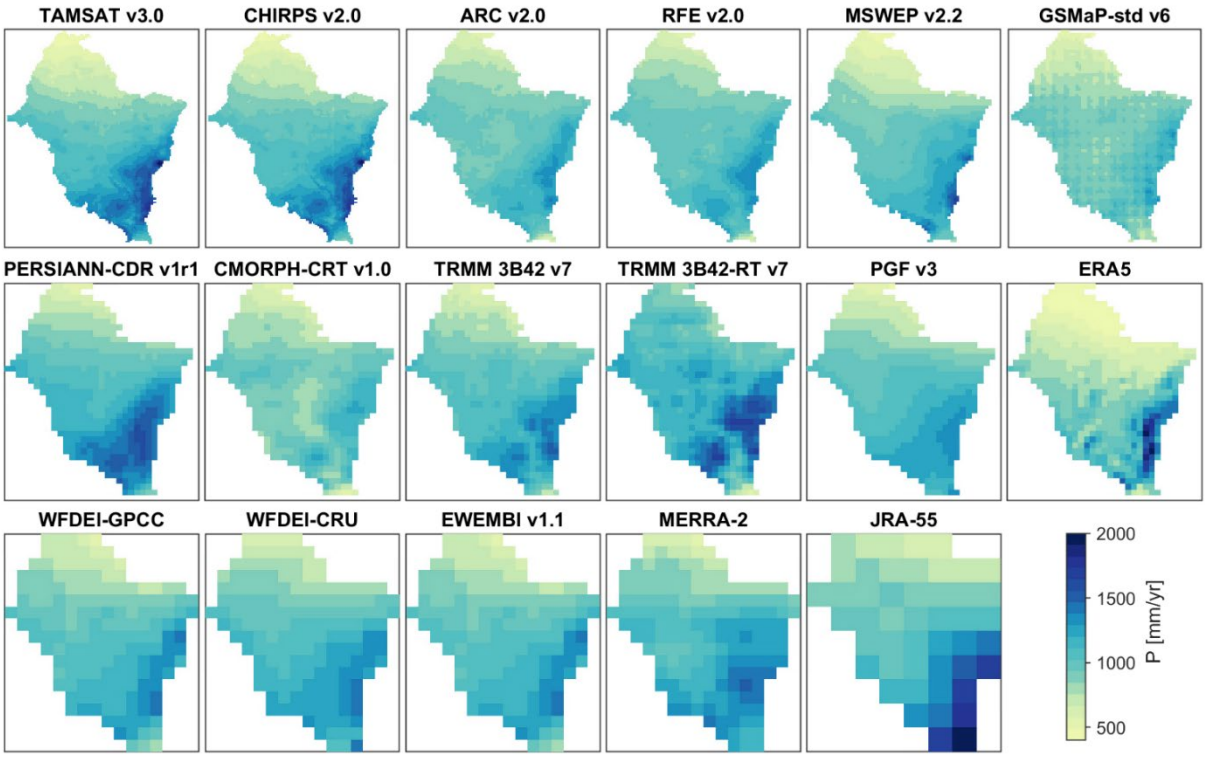


Figure 4.2. Mean annual rainfall totals over the period 2003-2012 for 17 rainfall datasets the Volta River basin

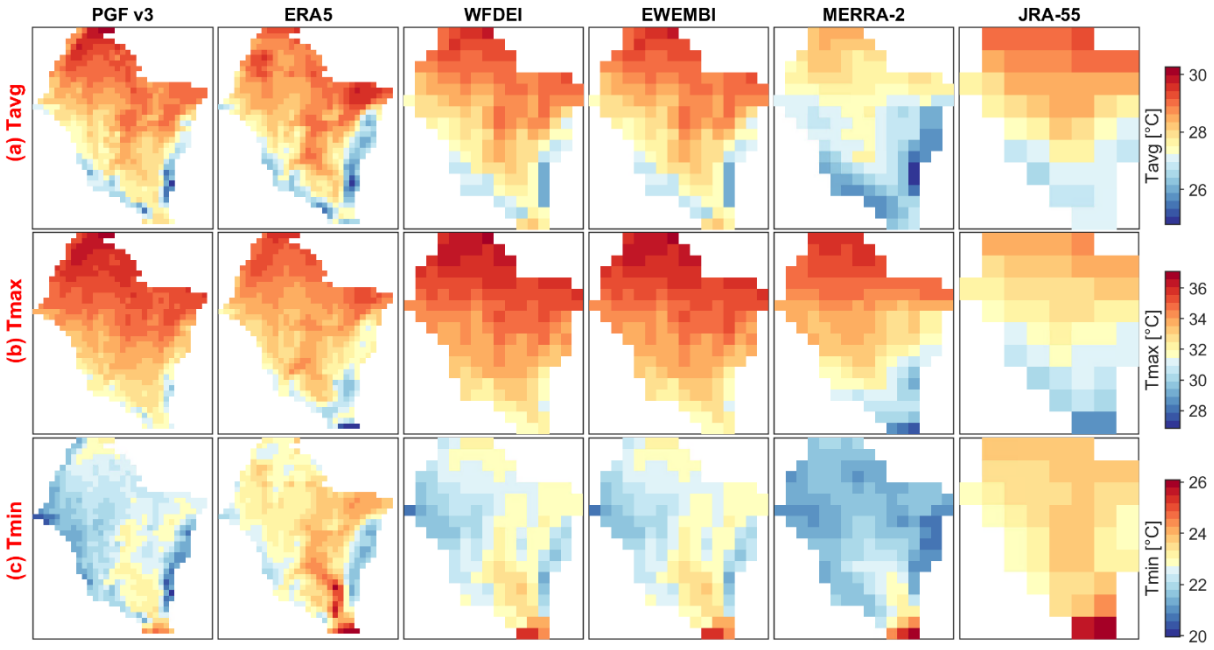


Figure 4.3. Mean annual air temperature (average (a), maximum (b) and minimum (c)) over the period 2003-2012 for 6 temperature datasets in the Volta River basin

Table 4.1. Meteorological datasets with used spatial resolution; the table presents the characteristics of the datasets used in this study, although different spatial and temporal resolutions can be available from the data providers. G: gauge, S: satellite, R: reanalysis, NP: near-present.

Datasets	Data sources	Spatial coverage	Spatial resolution	Temporal coverage	Temporal resolution	References
TAMSAT v3.0	S, G	Africa 38°N – 36°S, 19°W – 52°E	0.0375°	1983-NP	daily	Maidment *et al.* (2017), Tarnavsky *et al.* (2014), Maidment *et al.* (2014)
CHIRPS v2.0	S, G, R	Land 50° N/S, 180° E/W	0.05°	1981-NP	daily	Funk *et al.* (2015)
ARC v2.0	S, G	Africa 40°N – 40°S, 20°W – 55°E	0.1°	1983-NP	daily	Novella and Thiaw (2013)
RFE v2.0	S, G	Africa 40°N – 40°S, 20°W – 55°E	0.1°	2001-NP	daily	Xie and Arkin (1996), Herman *et al.* (1997)
MSWEP v2.2	S, G, R	Global	0.1°	1979-NP	3-hourly	Beck *et al.* (2017a)
GSMaP-std v6	R, G	60◦ N/S, 180° E/W	0.1°	2001-2013	daily	Ushio *et al.* (2009), Ushio *et al.* (2019)
PERSIANN-CDR v1r1	S, G	60◦ N/S, 180° E/W	0.25°	1983-2016	6-hourly (daily)	Ashouri *et al.* (2015)
CMORPH-CRT v1.0	S, G	60◦ N/S, 180° E/W	0.25°	1998-2015	daily	Joyce *et al.* (2004), Xie *et al.* (2017)
TRMM 3B42 v7	S, G	50◦ N/S, 180° E/W	0.25°	2000-2017	3-hourly	Huffman *et al.* (2007)
TRMM 3B42 RT v7	S	50◦ N/S, 180° E/W	0.25°	2000-NP	3-hourly	Huffman *et al.* (2007)
WFDEI-CRU	R, G	Global	0.5°	1979-2018	3-hourly	Weedon *et al.* (2014)
WFDEI-GPCC	R, G	Global	0.5°	1979-2016	3-hourly	Weedon *et al.* (2014)
PGF v3	R, G	Global	0.25°	1948-2012	3-hourly	Sheffield *et al.* (2006)
ERA5	R	Global	0.25°	1979-NP	hourly	Hersbach *et al.* (2018)
MERRA-2	S, G, R	Global	0.625° x 0.5°	1980-NP	hourly	Gelaro *et al.* (2017), Reichle *et al.* (2017)
EWEMBI v1.1	R, G	Global	0.5°	1976-2013	daily	Lange (2016)
JRA-55	R	Global	1.25°	1959-NP	3-hourly	Kobayashi *et al.* (2015)

4.4 Modelling datasets

In addition to the meteorological datasets (Table 4.1), an ensemble of datasets is required for the set-up and the calibration and evaluation of the hydrological model (Table 4.2; Appendix 6).

Table 4.2. Modelling datasets. ESA CCI SM: European Space Agency Climate Change Initiative Soil Moisture; GIMMS: Global Inventory Modelling and Mapping Studies; GLEAM: Global Land Evaporation Amsterdam Model; GLiM: Global Lithological Map; GMTED: Global Multi-resolution Terrain Elevation Data; GRACE: Gravity Recovery and Climate Experiment; WFDEI: WATCH Forcing Data methodology applied to ERA-Interim data.

Variables	Products	Spatial resolution	Temporal resolution	References
Morphological data				
Terrain characteristics (elevation, slope, aspect, flow direction and flow accumulation)	GMTED 2010	225 m (0.0021°)	static	Danielson and Gesch (2011)
Soil properties (horizon depth, bulk density, sand and clay content,)	SoilGrids	250 m (0.0023°)	static	Hengl *et al.* (2017)
Geology	GLiM v1.0	0.5°	static	Hartmann and Moosdorf (2012)
Land use land cover	Globcover 2009	300 m (0.0028°)	static	Bontemps *et al.* (2011)
Phenology (leaf area index)	GIMMS	8 km (0.0833°)	bimonthly	Tucker *et al.* (2005), Zhu *et al.* (2013)
Model calibration/evaluation				
Streamflow	-	point	daily	VBA and DGRE in Burkina Faso, HSD in Ghana
Terrestrial water storage anomaly (S_t)	GRACE TellUS v5.0	1°	monthly	Tapley *et al.* (2004), Landerer and Swenson (2012)
Surface soil moisture (S_u)	ESA CCI SM v4.2	0.25°	daily	Dorigo *et al.* (2017)
Actual evaporation (E_a)	GLEAM v3.2a	0.25°	daily	Martens *et al.* (2017), Miralles *et al.* (2011)
Land surface temperature (T_s) *(only for model evaluation)*	MYD11A2 v6	1 km (0.0083°)	8-day	Wan *et al.* (2015)

The streamflow datasets obtained from national and international organizations were pre-processed (i.e. gap-filling and quality control) in the work of Dembélé *et al.* (2019). Although

a large number of streamflow datasets was considered in the gap-filling procedure (cf. Chapter 3), only 11 streamflow stations that are found to have the highest data quality are used for hydrological modelling. This limits the propagation of the gap-filled streamflow data uncertainty in the hydrological modelling.

4.5 Distributed hydrological model setup

The fully distributed mesoscale Hydrologic Model (mHM) is a conceptual model that simulates dominant hydrological processes (e.g., evaporation, soil moisture and discharge) per grid cell in the modelling domain (Kumar *et al.*, 2013; Samaniego *et al.*, 2010). Samaniego *et al.* (2011) provide a schematic representation of the processes accounted for in mHM. A multiscale routing model based on the Muskingum-Cunge method (Cunge, 1969) is used for the routing of the total grid-generated runoff through the river network (Thober *et al.*, 2019). The sub-grid variability of the basin physical characteristics (e.g., soil texture, land cover) is accounted for with a multiscale parameter regionalization technique (Samaniego *et al.*, 2017). The model parameters (e.g., hydraulic conductivity, soil porosity) are linked to the basin physical characteristics via pedo-transfer functions and global parameters. Thirty-six global parameters (cf. Pokhrel *et al.*, 2008) are tuned during model calibration for this study (Appendix 17). The version 5.9 of mHM is used in this study.

Reference evaporation (E_{ref}) is calculated following the method of Hargreaves and Samani (1985), which was found to be reliable for semi-arid regions like the VRB (Bai *et al.*, 2016; Er-Raki *et al.*, 2010; Gao *et al.*, 2017). E_{ref} is formulated as follows:

$$E_{\mathrm{ref}} = \kappa \frac{R_{\mathrm{a}}}{\lambda} \left(T_{\mathrm{avg}} + 17.8\right)\left(T_{\mathrm{max}} - T_{\mathrm{min}}\right)^{0.5} \tag{4.1}$$

where R_{a} (MJ/m^2/day) is the extra-terrestrial radiation computed based on the latitude of the location and the day of the year (Allen et al., 1998), λ = 2.45 MJ/kg is the latent heat of vaporization of water. The unit of radiation is converted into equivalent water evaporation in mm/day with the ratio R_{a}/λ. The differences in advection or vapour transfer effect are compensated by the constant κ = 0.0023, and T_{avg}, T_{max} and T_{min} represent the daily average, maximum and minimum air temperature in degrees Celcius (°C) at a given location.

Potential evaporation (E_{p}) is calculated by adjusting E_{ref} with a dynamical scaling function (F_{DS}) based on leaf area index (Allen *et al.*, 1998; Demirel *et al.*, 2018), therefore accounting for vegetation-climate interactions (Bai *et al.*, 2018a; Birhanu *et al.*, 2019; Jiao *et al.*, 2017). E_{p} is formulated as follows:

$$E_{\mathrm{p}} = F_{\mathrm{DS}} \cdot E_{\mathrm{ref}}, \text{ with} \tag{4.2}$$

$$F_{\mathrm{DS}} = a + b\left(1 - e^{(c \cdot I_{\mathrm{LA}})}\right) \tag{4.3}$$

where a is the intercept term, b is the vegetation dependent component, and c represents the degree of nonlinearity of the leaf area index (I_{LA}). The coefficients a, b, and c are determined through model calibration.

In this study, actual evaporation (E_{a}) is defined as the sum of transpiration and evaporation from interception, land and water bodies (Coenders-Gerrits *et al.*, 2020; Shuttleworth, 1993). E_{a} is calculated as a fraction of E_{p} from soil layers depending on soil moisture availability and the rooting depth (Feddes *et al.*, 1976). Soil moisture is estimated by a multi-layer infiltration capacity approach adopting a three-layer soil scheme (0-5, 5-30 and 30-100 cm depths). Terrestrial water storage at each grid cell is the sum of the surface and subsurface water storage (i.e., lakes, wetlands, soil moisture reservoirs, interflow and baseflow).

4.6 Multisite model calibration on streamflow data

A multisite calibration strategy is adopted by simultaneously constraining the model with the 11 streamflow (Q) gauging stations (Figure 2.2) to infer a unique parameter set for the whole basin. The objective function Φ_Q combines the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) of streamflow (E_{NS}) and the Nash-Sutcliffe efficiency of the logarithm of streamflow (E_{NSlog}), and it is formulated such that it has to be minimized:

$$\Phi_Q = \frac{1}{g}\sum_1^g \sqrt{(1 - E_{NS})^2 + \left(1 - E_{NSlog}\right)^2}, \text{ with} \tag{4.4}$$

$$E_{NS} = 1 - \frac{\sum_1^t \left(Q_{mod}(t) - Q_{obs}(t)\right)^2}{\sum_1^t (Q_{obs}(t) - \overline{Q_{obs}})^2} \text{ and} \tag{4.5}$$

$$E_{NSlog} = 1 - \frac{\sum_1^t \left[\log\left(Q_{mod}(t)\right) - \log\left(Q_{obs}(t)\right)\right]^2}{\sum_1^t \left[\log\left(Q_{obs}(t)\right) - \overline{\log(Q_{obs})}\right]^2} \tag{4.6}$$

where Q_{mod} and Q_{obs} are the modelled and observed streamflow data, t is the number of time steps of the calibration period, and g is the number of streamflow gauging stations present within the modelling domain. Φ_Q is obtained by equally weighing the streamflow gauging stations, and it ranges from its ideal value that is 0 to positive infinity.

The model is calibrated solely with Q data because it is the only available in-situ measurement, and to avoid potential trade-offs of a multivariate calibration that would result in difficulties in identifying the source of variation in the model performance (i.e. input data vs. model

parametrization). The parameter estimation is done with the dynamically dimensioned search algorithm (Tolson and Shoemaker, 2007) using 4,000 iterations for each of the 102 rainfall-temperature dataset combinations.

4.7 Multivariable model evaluation with streamflow and satellite data

4.7.1 The Kling-Gupta efficiency

The Kling-Gupta efficiency (E_{KG}) (Kling *et al.*, 2012) is used to evaluate the model performance for streamflow.

$$E_{KG} = 1 - \sqrt{(r_{KG} - 1)^2 + (\beta_{KG} - 1)^2 + (\gamma_{KG} - 1)^2} \quad (4.7)$$

Where r_{KG} is the Pearson correlation coefficient, β_{KG} is the bias term (i.e. the ratio of the means), and γ_{KG} is the variability term (i.e. the ratio of the coefficients of variation) between the observed (Q_{obs}) and modelled (Q_{mod}) streamflow. The E_{KG} ranges from negative infinity to its optimal value that is unity. As a reference, $E_{KG} > -0.41$ indicates that the model is better than the mean observed flow (Knoben *et al.*, 2019). Φ_Q ranges from its optimal value that is 0 to positive infinity.

In addition to Q, several non-commensurable and satellite-based variables are used for model evaluation (Table 4.2). The model performance for Q is evaluated with E_{KG}. The bias-insensitive Pearson's correlation coefficient (r) is used to assess the temporal dynamics of S_t, S_u and E_a because the model is not calibrated on these variables, and their evaluation datasets are satellite-derived products that encompass uncertainties and can be biased.

4.7.2 Spatial pattern efficiency

The spatial pattern representation of hydrological processes is assessed by using a bias-insensitive and multi-component metric developed by Dembélé *et al.* (2020b). The proposed spatial pattern efficiency (E_{SP}) metric only considers the spatial pattern of the underlying variables and ignores their absolute values. E_{SP} is an integrated measure of the dynamics, the spatial variability, and the locational matching of grid cells between the modelled (X_{mod}) and observed (X_{obs}) variables. With X_{obs} and X_{mod} composed of n cells, E_{SP} is defined as follows:

$$E_{SP} = 1 - \sqrt{(r_s - 1)^2 + (\gamma - 1)^2 + (\alpha - 1)^2}, \text{ with} \quad (4.8)$$

$$r_s = 1 - \frac{6\sum_1^n d_i^2}{n(n^2-1)}, \quad (4.9)$$

$$\gamma = \frac{\sigma_{\mathrm{mod}}/\mu_{\mathrm{mod}}}{\sigma_{\mathrm{obs}}/\mu_{\mathrm{obs}}} \text{ and} \tag{4.10}$$

$$\alpha = 1 - E_{\mathrm{RMS}}\left(Z_{\mathrm{X_{mod}}}, Z_{\mathrm{X_{obs}}}\right), \text{ with} \tag{4.11}$$

where r_s is the Spearman rank-order correlation coefficient with d_i the difference between the ranks of the i^{th} cell of X_{mod} and X_{obs}. γ is the variability ratio (i.e., the ratio of the coefficients of variation) that assesses the similarity in the dispersion of the probability distributions of X_{mod} and X_{obs}, with μ and σ representing the mean and the standard deviation of the variable. α is the spatial location matching term calculated as the root mean squared error (E_{RMS}) of the standardized values (z-scores, Z_X) of X_{mod} and X_{obs} (Dembélé *et al.*, 2020b).

The standardized values (z-scores, Z_X) of X_{mod} and X_{obs} and the root mean squared error (E_{RMS}) of a modelled variable (X_{mod}) and an observed variable (X_{obs}) of n elements, are defined as follows:

$$Z_X = \frac{X-\mu}{\sigma} \text{ and} \tag{4.12}$$

$$E_{\mathrm{RMS}}(X_{\mathrm{mod}}, X_{\mathrm{obs}}) = \sqrt{\frac{1}{n}\sum_1^n (X_{\mathrm{mod}} - X_{\mathrm{obs}})^2} \tag{4.13}$$

where μ and σ are the mean and the standard deviation of a given variable X.

The z-score is a standardization of the scale of a distribution that facilitates its comparison with another distribution. The z-scores identify and describe the exact location of each observation in a distribution (Gravetter and Wallnau, 2013). For a given variable with values represented spatially as a 2-D matrix, the z-scores represent the number of standard deviation the value in each grid cell is from the population mean (Oyana and Margai, 2015). Consequently, forcing the z-scores of X_{mod} and X_{obs} to be equal (i.e. minimizing their E_{RMS}) corresponds to matching their grid cell locations (i.e. spatial patterns). E_{SP} ranges from negative infinity to its optimal value that is unity. E_{SP} = 0 when there is a moderate relationship between the ranks of the observed and modelled variables (i.e., r_s = 0.55), and E_{SP} = -0.67 when the ranks are not related (i.e., r_s = 0). A comparison of E_{SP} to *SPAEF*, another spatial pattern metric, is provided in the Appendix 7.

The spatial pattern evaluation is done for S_u and E_a, while only the temporal dynamics of S_t are assessed due to the coarse spatial resolution of the GRACE data.

4.7.3 Second-order coefficient of variation

The relative variation in model performance is assessed with the second-order coefficient of variation (V_2) (Kvålseth, 2017). V_2 is an alternative to the classic Pearson's coefficient of variation (CV), which has significant limitations that are comprehensively discussed by Kvålseth (2017). For all sample data $x = (x_1,\ldots, x_n) \in R^n$, with $R = (-\infty, \infty)$, V_2 is defined as follows:

$$V_2 = \left(\frac{s^2}{s^2 + \bar{x}^2}\right)^{1/2} \tag{4.14}$$

where s is the standard deviation and $\bar{x}$ is the mean of x. V_2 varies from 0 to 1 or 0% to 100%, and represents the distance between x and $\bar{x}$ relative to the distance between x and the origin zero.

4.8 Results

The results are presented and discussed for the entire simulation period (2003-2012, i.e. combined calibration and evaluation periods) because reliable meteorological datasets are expected to produce a plausible representation of hydrological processes independently of the modelling period (Bisselink *et al.*, 2016). Additional results are provided in the supporting information of Dembélé *et al.* (2020c).

4.8.1 Model performance for streamflow

Similar model performance patterns are obtained with E_{KG}, E_{NS} and E_{NSlog} of daily streamflow (Q) (Figure 4.4). Therefore, only E_{KG} is retained for the description of the results. All input dataset combinations show a median $E_{KG} > 0.5$, except those having JRA-55 as rainfall input (Figure 4.4); this can be justified by the coarse spatial resolution of that product. The ranking of the rainfall and temperature datasets based on the model performance for Q is provided in Figure 4.6. The analysis of model performance for Q is done for the entire VRB and not per climatic zone due to the limited number of stations. As expected, the discrepancies in median E_{KG} are more pronounced across rainfall datasets than across temperature datasets, as visible in the colour-coded ranking of the products in Figure 4.4. For a given rainfall product, the ranking among all rainfall products hardly varies with different temperature products. The ranking of all the datasets for the model performance for Q is also summarized in Figure 4.6. For rainfall datasets, the second-order coefficient of variations (V_2) across temperature datasets varies

between 0.5% for GSMaP-std and 4% for JRA-55, with an average V_2 of 2%. For temperature datasets, the V_2 of median E_{KG} of Q across rainfall datasets varies between 10% for MERRA-2 and 12% for ERA5, with an average V_2 of 11%. This result suggests that the choice of a rainfall dataset has a stronger impact on the E_{KG} of Q than the choice of a temperature dataset.

The analysis of the components of E_{KG} (i.e. the Pearson correlation r_{KG}, the bias β_{KG} and the variation γ_{KG}) reveals that, when choosing a rainfall dataset, there is more uncertainty in the bias of Q (V_2 = 14%) than in its variability (V_2 = 6%) and in its dynamics (V_2 = 3%), which is in agreement with the work of Thiemig *et al.* (2013).

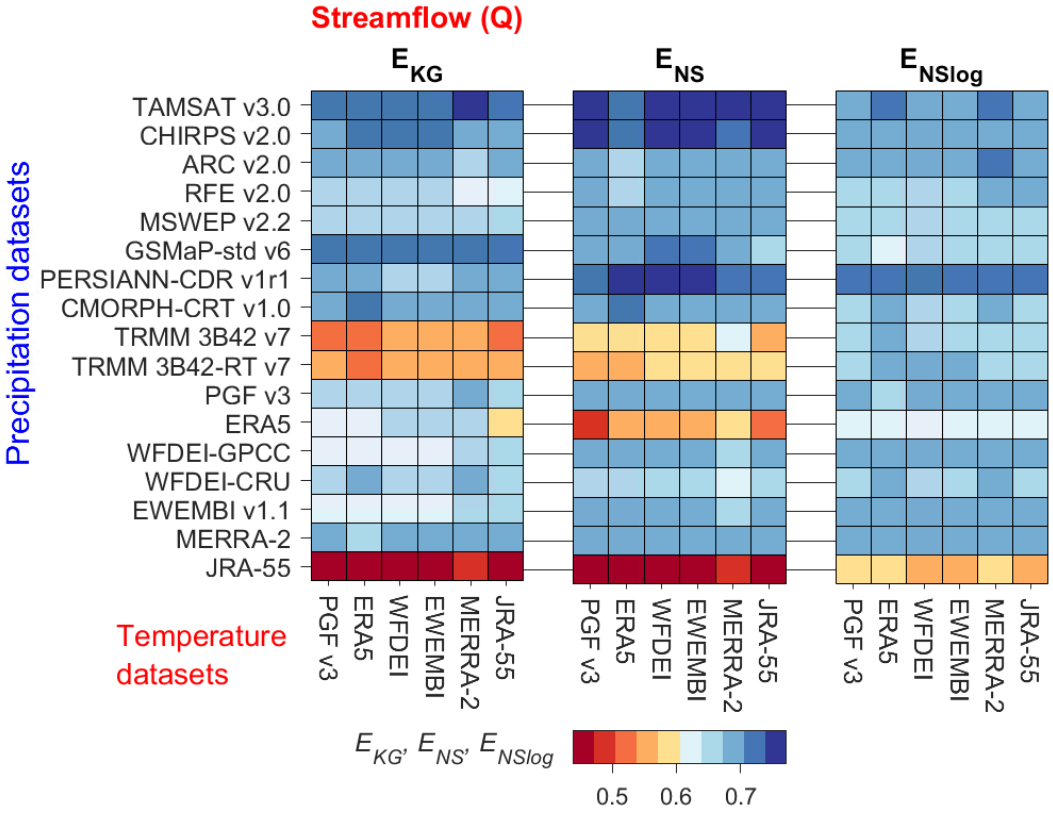


Figure 4.4. Kling-Gupta efficiency (E_{KG}), Nash-Sutcliffe efficiency (E_{NS}) and Nash-Sutcliffe efficiency of the logarithm (E_{NSlog}) of daily streamflow (Q) over the simulation period (2003-2012) for 102 combinations of 17 rainfall datasets (y-axis) and 6 temperature datasets (x-axis) used as forcing for the hydrological model.

4.8.2 Model performance for terrestrial water storage

The model performance for the temporal dynamics of monthly terrestrial water storage (S_t) compared to the GRACE product is shown in Figure 4.5. The average Pearson correlation coefficient (r) of S_t for all datasets in the entire VRB is 0.80, with discrepancies across climatic zones. The driest and wettest climatic zones show the lowest performances, i.e. Sahelian (r =

0.67) and Guinean (r = 0.60) zones, compared to the intermediate climatic zones, i.e. Sudano-Sahelian (r = 0.72) and Sudanian (r = 0.79) zones. Figure 4.6 provides the ranking of all the meteorological datasets for the model performance for S_t.

The rainfall datasets show different performances across climatic zones, with ARC showing the highest score for all the climatic zones except the Guinean zone, where CMORPH-CRT ranks first. The choice of the rainfall dataset leads to an average V_2 of 15% for the r of S_t, while the average V_2 is 5% for the choice of the temperature dataset.

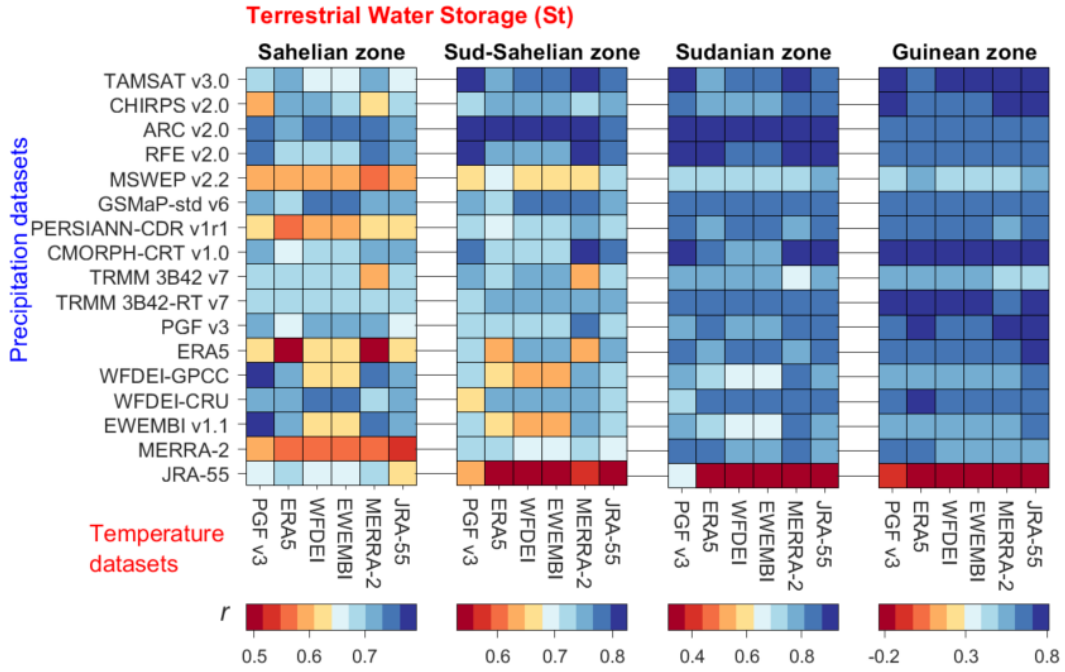


Figure 4.5. Pearson correlation coefficient (r) of modelled terrestrial water storage compared to GRACE data in four climatic zones in the Volta River basin over the simulation period (2003-2012) considering 102 combinations of rainfall (y-axis) and temperature datasets (subplots on x-axis) used as forcing for the hydrological model.

		VRB								Sahelian zone					Sudano-Sahelian zone					Sudanian zone					Guinean zone				
		Temporal dynamics						Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns	
	Variables	*Q*	*Q*	*Q*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*
	Performance metrics	E_{NS}	E_{NSlog}	E_{KG}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}
Rainfall datasets	TAMSAT v3.0	0.745	0.704	0.731	0.856	0.937	0.945	-0.044	0.210	0.686	0.931	0.929	-0.223	-1.215	0.777	0.936	0.954	-0.313	-0.294	0.857	0.927	0.916	-0.371	-0.216	0.740	0.901	0.808	-0.503	-0.612
	CHIRPS v2.0	0.738	0.698	0.703	0.863	0.933	0.919	-0.114	0.178	0.671	0.923	0.958	-0.337	-1.601	0.740	0.915	0.920	-0.359	-0.244	0.824	0.919	0.880	-0.411	-0.357	0.715	0.891	0.819	-0.318	-0.820
	ARC v2.0	0.682	0.693	0.680	0.910	0.932	0.965	-0.055	0.026	0.732	0.968	0.952	-0.477	-0.580	0.799	0.948	0.969	-0.330	-0.241	0.918	0.926	0.940	-0.337	-0.437	0.667	0.840	0.887	-0.416	-0.586
	RFE v2.0	0.689	0.664	0.646	0.882	0.920	0.964	-0.066	0.098	0.709	0.969	0.948	-0.546	-0.490	0.780	0.935	0.976	-0.343	-0.217	0.890	0.913	0.938	-0.354	-0.350	0.648	0.833	0.886	-0.380	-0.671
	MSWEP v2.2	0.690	0.650	0.655	0.719	0.918	0.937	-0.003	0.258	0.579	0.945	0.965	-0.365	-1.898	0.681	0.937	0.950	-0.299	-0.325	0.728	0.911	0.904	-0.359	-0.186	0.490	0.822	0.863	-0.270	-0.528
	GSMaP-std v6	0.695	0.646	0.713	0.842	0.945	0.949	-0.100	0.076	0.718	0.973	0.971	-0.582	-0.404	0.756	0.956	0.970	-0.372	-0.192	0.845	0.938	0.924	-0.360	-0.359	0.656	0.878	0.856	-0.371	-0.805
	PERSIANN-CDR v1r1	0.735	0.711	0.682	0.804	0.957	0.938	-0.082	0.167	0.594	0.952	0.974	-0.405	-1.125	0.718	0.956	0.954	-0.361	-0.224	0.807	0.950	0.905	-0.347	-0.291	0.637	0.921	0.820	-0.453	-0.651
	CMORPH-CRT v1.0	0.696	0.667	0.690	0.875	0.939	0.937	-0.110	-0.122	0.691	0.953	0.966	-0.549	-0.716	0.764	0.940	0.956	-0.383	-0.275	0.856	0.926	0.905	-0.352	-0.525	0.775	0.915	0.809	-0.293	-1.037
	TRMM 3B42 v7	0.590	0.658	0.539	0.769	0.907	0.880	-0.157	-0.062	0.671	0.934	0.965	-0.599	-0.787	0.717	0.918	0.897	-0.373	-0.338	0.780	0.890	0.833	-0.448	-0.551	0.516	0.842	0.776	-0.276	-0.894
	TRMM 3B42-RT v7	0.574	0.669	0.540	0.860	0.907	0.887	-0.295	-0.163	0.677	0.929	0.943	-0.613	-0.921	0.744	0.905	0.890	-0.399	-0.590	0.834	0.890	0.842	-0.486	-0.846	0.724	0.849	0.752	-0.376	-1.613
	PGF v3	0.695	0.679	0.661	0.823	0.960	0.928	-0.080	0.159	0.688	0.954	0.904	-0.540	-0.611	0.729	0.973	0.968	-0.421	-0.327	0.824	0.951	0.888	-0.334	-0.151	0.715	0.900	0.762	-0.431	-0.577
	ERA5	0.540	0.631	0.627	0.823	0.907	0.867	-0.195	-0.228	0.574	0.868	0.903	-0.396	-1.416	0.704	0.893	0.888	-0.398	-0.503	0.814	0.874	0.813	-0.500	-0.749	0.691	0.870	0.697	-0.327	-1.431
	WFDEI-GPCC	0.677	0.687	0.638	0.748	0.965	0.948	-0.065	0.173	0.707	0.974	0.909	-0.517	-0.795	0.678	0.975	0.980	-0.384	-0.393	0.756	0.955	0.912	-0.356	-0.296	0.546	0.884	0.806	-0.496	-0.615
	WFDEI-CRU	0.646	0.666	0.666	0.829	0.958	0.945	-0.091	0.142	0.720	0.967	0.927	-0.545	-0.693	0.730	0.972	0.977	-0.415	-0.355	0.829	0.946	0.909	-0.367	-0.196	0.697	0.902	0.791	-0.470	-0.595
	EWEMBI v1.1	0.677	0.687	0.638	0.748	0.965	0.948	-0.065	0.173	0.707	0.974	0.909	-0.517	-0.795	0.678	0.975	0.980	-0.384	-0.393	0.756	0.955	0.912	-0.356	-0.296	0.546	0.884	0.806	-0.496	-0.615
	MERRA-2	0.687	0.691	0.684	0.800	0.974	0.932	-0.112	0.198	0.558	0.934	0.959	-0.274	-1.997	0.712	0.973	0.952	-0.318	-0.480	0.807	0.970	0.904	-0.429	-0.446	0.615	0.932	0.809	-0.530	-0.598
	JRA-55	0.460	0.581	0.453	0.377	0.830	0.838	-0.185	-0.178	0.658	0.904	0.909	-0.501	-0.731	0.558	0.894	0.865	-0.413	-0.547	0.385	0.839	0.804	-0.494	-0.826	-0.185	0.505	0.748	-0.395	-0.845
Temperature datasets	JRA-55	0.654	0.670	0.640	0.811	0.943	0.928	-0.115	0.070	0.682	0.950	0.934	-0.494	-0.841	0.733	0.950	0.948	-0.373	-0.332	0.817	0.933	0.892	-0.411	-0.390	0.609	0.869	0.801	-0.371	-0.768
	MERRA-2	0.663	0.672	0.656	0.785	0.924	0.921	-0.101	0.067	0.659	0.943	0.943	-0.452	-1.139	0.706	0.934	0.938	-0.361	-0.375	0.785	0.913	0.886	-0.388	-0.439	0.594	0.848	0.810	-0.393	-0.761
	EWEMBI	0.663	0.666	0.642	0.781	0.931	0.925	-0.102	0.061	0.663	0.944	0.942	-0.454	-0.959	0.713	0.940	0.944	-0.368	-0.335	0.776	0.920	0.891	-0.398	-0.416	0.593	0.855	0.804	-0.403	-0.811
	WFDEI	0.663	0.666	0.642	0.781	0.931	0.925	-0.102	0.061	0.663	0.944	0.942	-0.454	-0.958	0.713	0.940	0.944	-0.368	-0.335	0.776	0.920	0.891	-0.400	-0.410	0.593	0.855	0.804	-0.403	-0.811
	ERA5	0.658	0.671	0.644	0.808	0.940	0.927	-0.105	0.076	0.665	0.944	0.939	-0.478	-1.015	0.736	0.947	0.946	-0.367	-0.342	0.808	0.932	0.892	-0.370	-0.453	0.587	0.861	0.807	-0.396	-0.831
	PGF v3	0.657	0.671	0.640	0.808	0.927	0.925	-0.116	0.057	0.670	0.940	0.944	-0.484	-1.008	0.728	0.936	0.943	-0.374	-0.378	0.807	0.916	0.888	-0.385	-0.390	0.620	0.854	0.807	-0.434	-0.781

Figure 4.6. Model performance for streamflow (Q), terrestrial water storage (S_t), soil moisture (S_u) and actual evaporation (E_a) using various rainfall-temperature dataset combinations as model inputs. Each score for a given rainfall product represents the average over individual combinations with 6 temperature datasets, while the score is the average over combinations with 17 rainfall datasets for each temperature dataset. The skill scores of the temporal dynamics are obtained with the Kling-Gupta efficiency (E_{KG}), the Nash-Sutcliffe efficiency (E_{NS}) and the Nash-Sutcliffe efficiency of the logarithm (E_{NSlog}) for Q, and the Pearson's correlation coefficient (r) for S_t, S_u and E_a. The spatial pattern efficiency (E_{SP}) is used to assess the spatial representation of S_u and E_a. The skill scores are ranked from the best (blue) to the worst (red). The results are shown for the four climatic zones in the Volta River basin (VRB) over the simulation period (2003-2012).

4.8.3 Model performance for soil moisture

Figure 4.7 shows the model performance for the temporal dynamics of monthly soil moisture (S_u) compared to the ESA CCI product. The average r of S_u for the entire VRB over all datasets is 0.93. The r of S_u decreases from the drier to the wetter climatic zones: Sahelian (r = 0.94), Sudano-Sahelian (r = 0.94), Sudanian (r = 0.92) and Guinean (r = 0.86). The ranking of the meteorological datasets based on the model performance for S_u is provided in Figure 4.6. EWEMBI and WFDEI-GPCC show the highest performance in the Sahelian and Suadano-Sahelian zones respectively, while MERRA-2 shows the highest performance in the Sudanian and Guinean zones. The choice of the rainfall dataset leads to an average V_2 of 4% for the temporal dynamics of S_u, while the average V_2 is 2% for the choice of the temperature dataset.

The spatial patterns of S_u show considerable differences when using different combinations of rainfall and temperature input datasets, as illustrated in Figure 4.8 (see Appendix 13 for all datasets). The south-north gradient of increasing aridity is not similarly spread among the rainfall-temperature dataset combinations. More interestingly, west-east differences in the spatial patterns of S_u can be observed. These differences in spatial pattern reproduction can also be seen in the spatial pattern efficiency metric (E_{SP}) of S_u for the 102 rainfall-temperature dataset combinations (Figure 4.9). The average E_{SP} of S_u in the VRB over all datasets is -0.11.

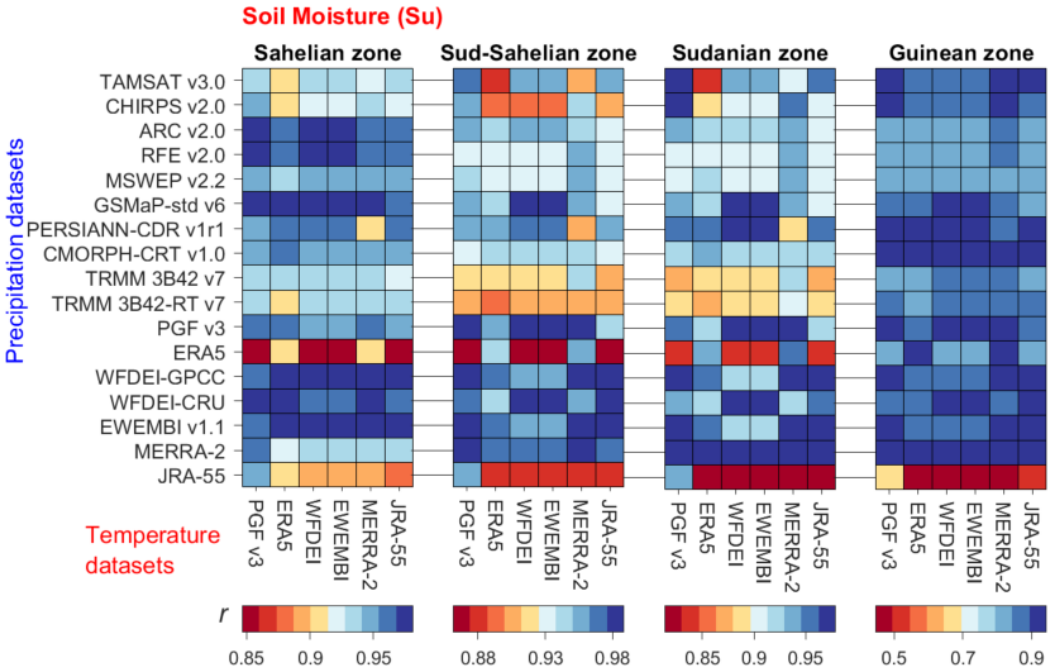


Figure 4.7. Pearson correlation coefficient (r) of modelled soil moisture (S_u) compared to ESA CCI data over the simulation period (2003-2012) considering 102 combinations of rainfall (y-axis) and temperature datasets (subplots on x-axis) used as forcing for the hydrological model.

For the entire VRB, the choice of the rainfall dataset leads to an average variation of 61% for the E_{SP} of S_u, while the choice of the temperature dataset involves a variation of 45%. Lower impacts of data choices are observed in the climatic zones where the climate is homogeneous as compared to the entire VRB. The choice of a rainfall dataset is more critical for the E_{SP} of S_u in the driest and wettest climatic zones, i.e. Sahelian (E_{SP} = -0.47, V_2 = 25%) and Guinean (E_{SP} = -0.40, V_2 = 26%) zones, than the intermediate zones, i.e. Sudano-Sahelian (E_{SP} = -0.37, V_2 = 11%) and Sudanian (E_{SP} = -0.39, V_2 = 17%) zones. A smaller impact on the E_{SP} of S_u is observed for the choice of the temperature dataset: Sahelian (V_2 = 8%), Guinean (V_2 = 19%), Sudano-Sahelian (V_2 = 5%) and Sudanian (V_2 = 9%) zones.

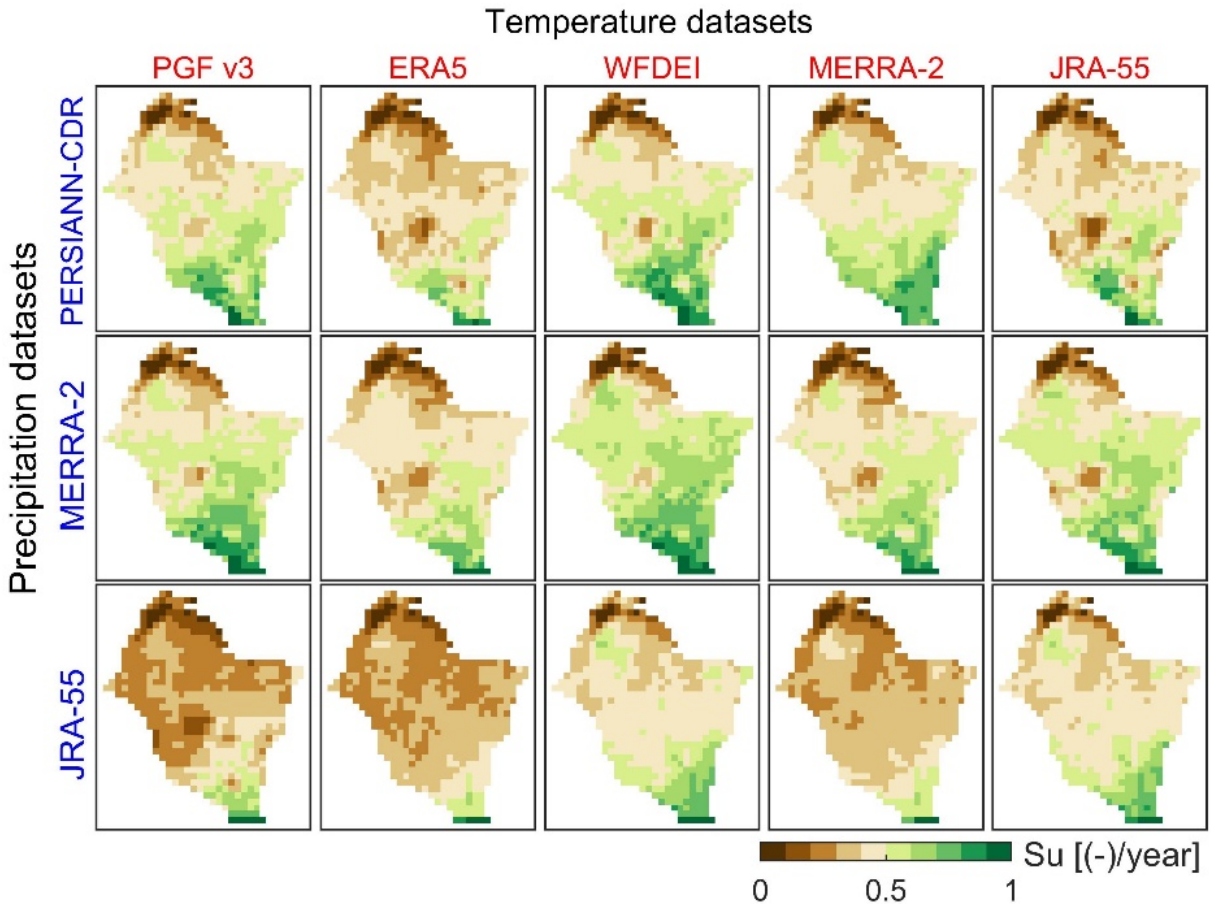


Figure 4.8. Maps of long-term (2003-2012) average of annual soil moisture (S_u) obtained with different forcing of rainfall (y-axis, blue font) and temperature (x-axis, red font) datasets. The values are normalized between 0 and 1 to emphasize spatial patterns and to use a unique colour scale.

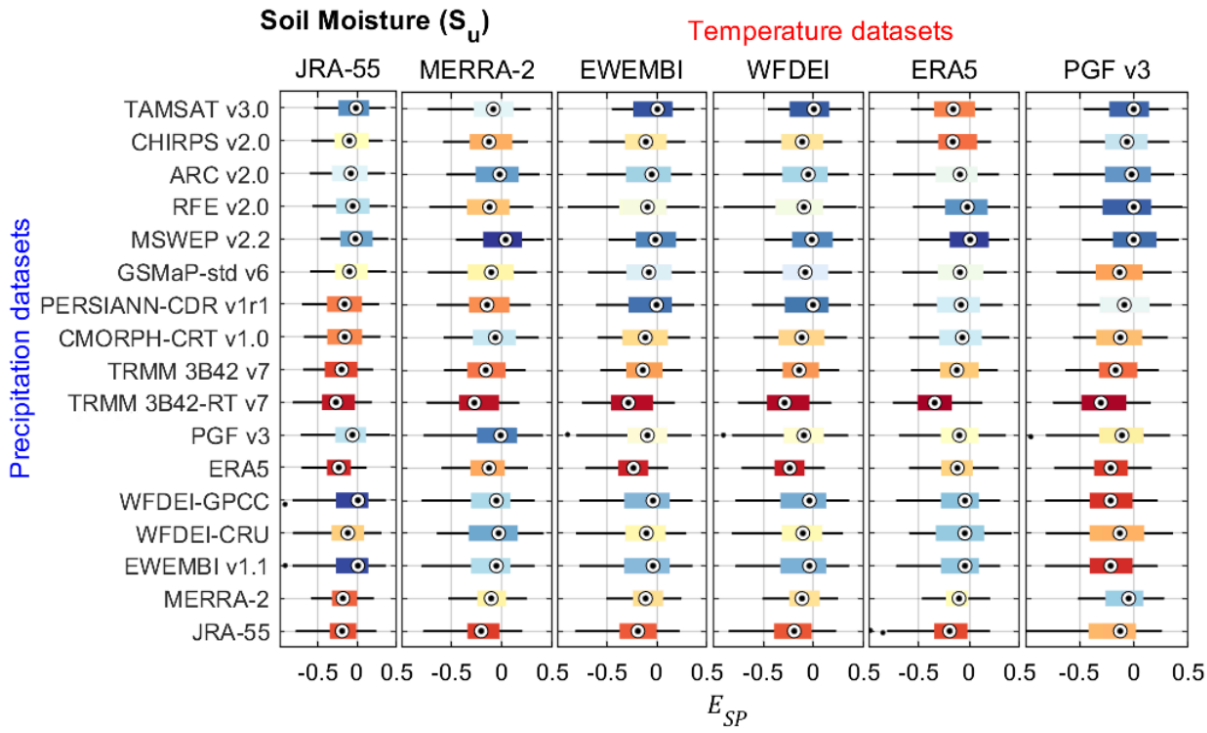


Figure 4.9. Spatial pattern efficiency (E_{SP}) of soil moisture (S_u) over the entire simulation period (2003-2012) for the Volta River basin (VRB) using different combinations of precipitation and temperature datasets used as input for hydrological modelling. Each boxplot has 120 values corresponding to the number of months. The boxplots are coloured from the best (blue) to the worst performance (red) based on the median value.

4.8.4 Model performance for actual evaporation

The model performance for the temporal dynamics of monthly actual evaporation (E_a) compared to the GLEAM product is shown in Figure 4.10. The average r of E_a for the entire VRB over all datasets is 0.93. Similarly to S_u, the r of E_a is higher is the driest climatic zones: Sahelian (r = 0.94), Sudano-Sahelian (r = 0.94), Sudanian (r = 0.89) and Guinean (r = 0.81). However, the predictive skill of the model for the temporal dynamics of E_a is higher than its predictive skill for E_a in the wetter climatic zones. Figure 4.6 shows the ranking of all the meteorological datasets for the model performance for E_a. The rainfall datasets show different performances across climatic zones, with the following best datasets: PERSIANN-CDR in the Sahelian zone, EWEMBI and WFDEI-GPCC in the Soudano-Sahelian zone, ARC in the Sudanian and Guinean zones. The choice of the rainfall dataset leads to an average V_2 of 4% for the temporal dynamics of E_a, while the average V_2 is 1.5% for the choice of the temperature dataset, which aligns with the findings of Jung *et al.* (2019).

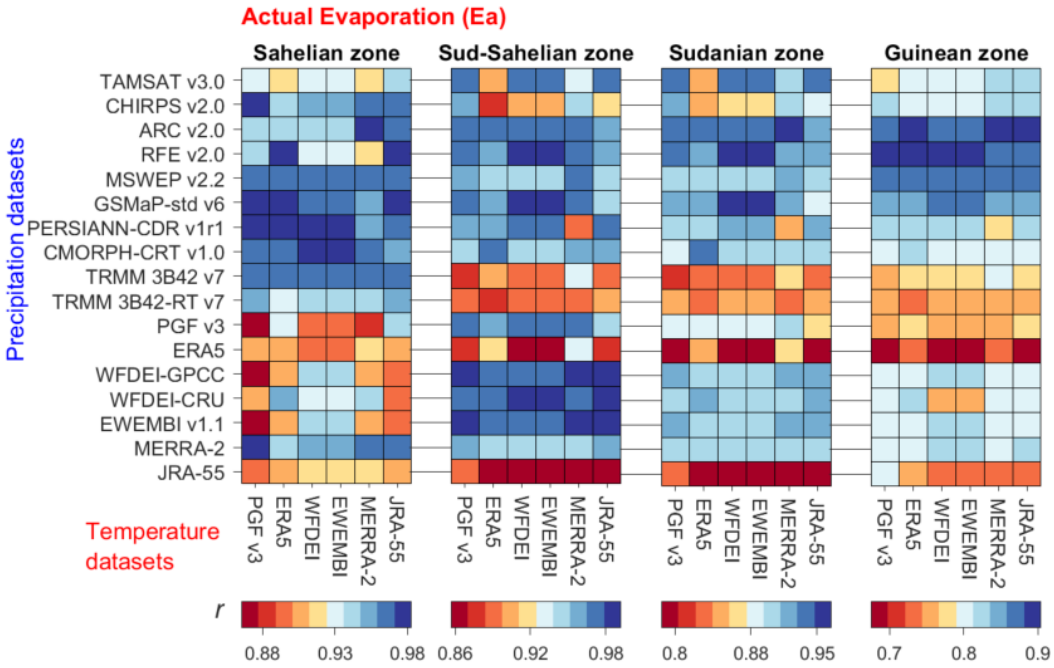


Figure 4.10. Pearson correlation coefficient (r) of modelled actual evaporation (E_a) compared to GLEAM data over the simulation period (2003-2012) considering 102 combinations of rainfall (y-axis) and temperature datasets (subplots on x-axis) used as forcing for the hydrological model.

As for S_u, the choice of input datasets has a considerable impact on the reproduction of the spatial patterns of E_a (Figure 4.11; see Appendix 14 for all datasets). It can be observed that different rainfall-temperature combinations used to force the model result in large discrepancies in the spatial pattern of E_a, especially in the southern region. The south-north gradient of increasing aridity with west-east differences is represented differently among the rainfall-temperature dataset combinations (see e.g., the difference between the first two columns of the first row in Figure 4.11).

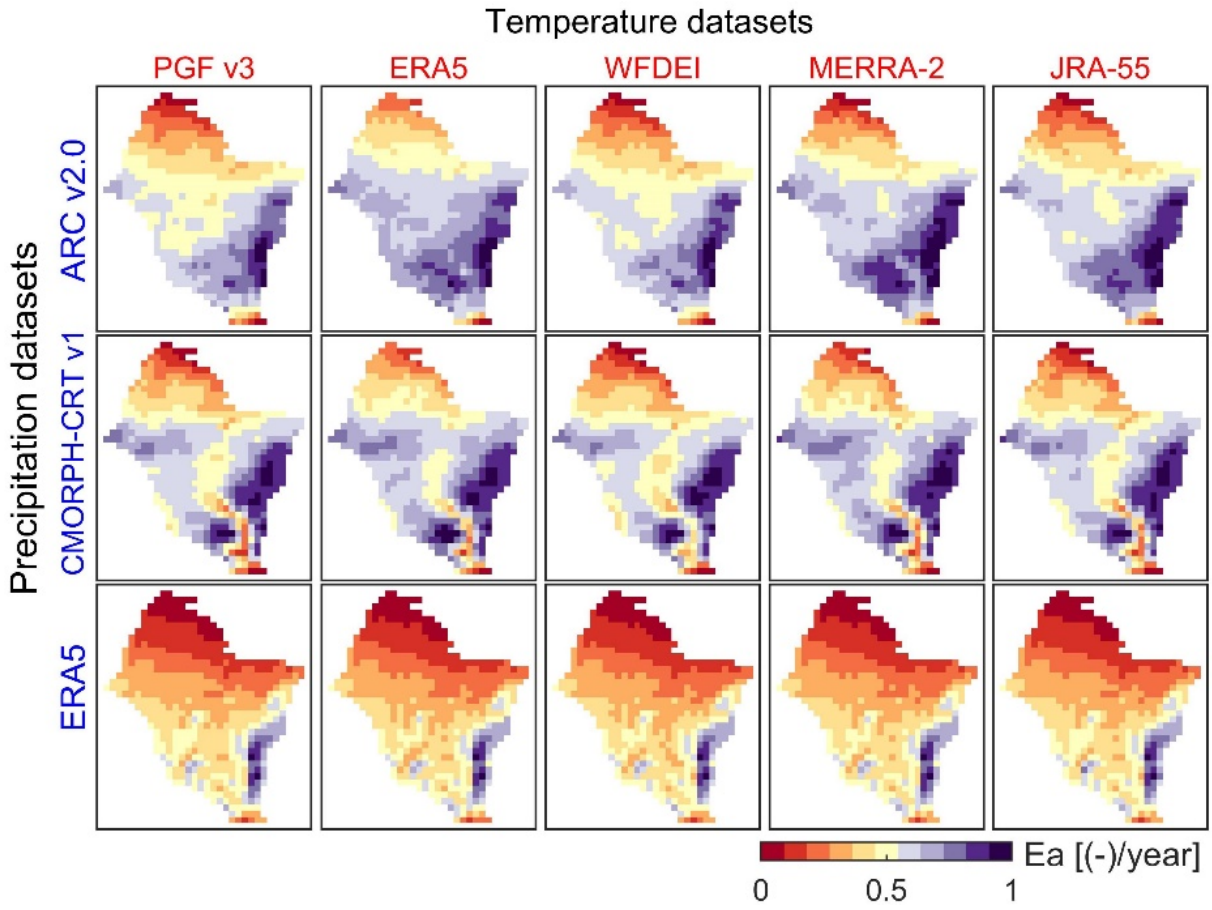


Figure 4.11. Maps of long-term (2003-2012) average of annual actual evaporation (E_a) obtained with different forcing of rainfall (y-axis, blue font) and temperature (x-axis, red font) datasets. The values are normalized between 0 and 1 to emphasize spatial patterns and to use a unique colour scale.

The E_{SP} of E_a for the 102 rainfall-temperature dataset combinations in the VRB and the climatic zones is given in Figure 4.12. The average E_{SP} of E_a in the VRB over all datasets is 0.07, which is higher than for S_u (E_{SP} = -0.11). The choice of the rainfall dataset for the VRB affects the E_{SP} of E_a on average by 93%, while the choice of the temperature dataset involves a variation 33%. However, lower impacts of data choices are observed in the climatic zones. The choice of rainfall dataset is more critical for the E_{SP} of E_a in the driest and wettest climatic zones, i.e. Sahelian (E_{SP} = -0.99, V_2 = 49%) and Guinean (E_{SP} = -0.79, V_2 = 37%) zones, than the intermediate zones, i.e. Sudano-Sahelian (E_{SP} = -0.35, V_2 = 36%) and Sudanian (E_{SP} = -0.42, V_2 = 49%) zones. A smaller impact on the E_{SP} of E_a is observed for the choice of the temperature dataset: Sahelian (V_2 = 21%), Guinean (V_2 = 10%), Sudano-Sahelian (V_2 = 17%) and Sudanian (V_2 = 21%) zones.

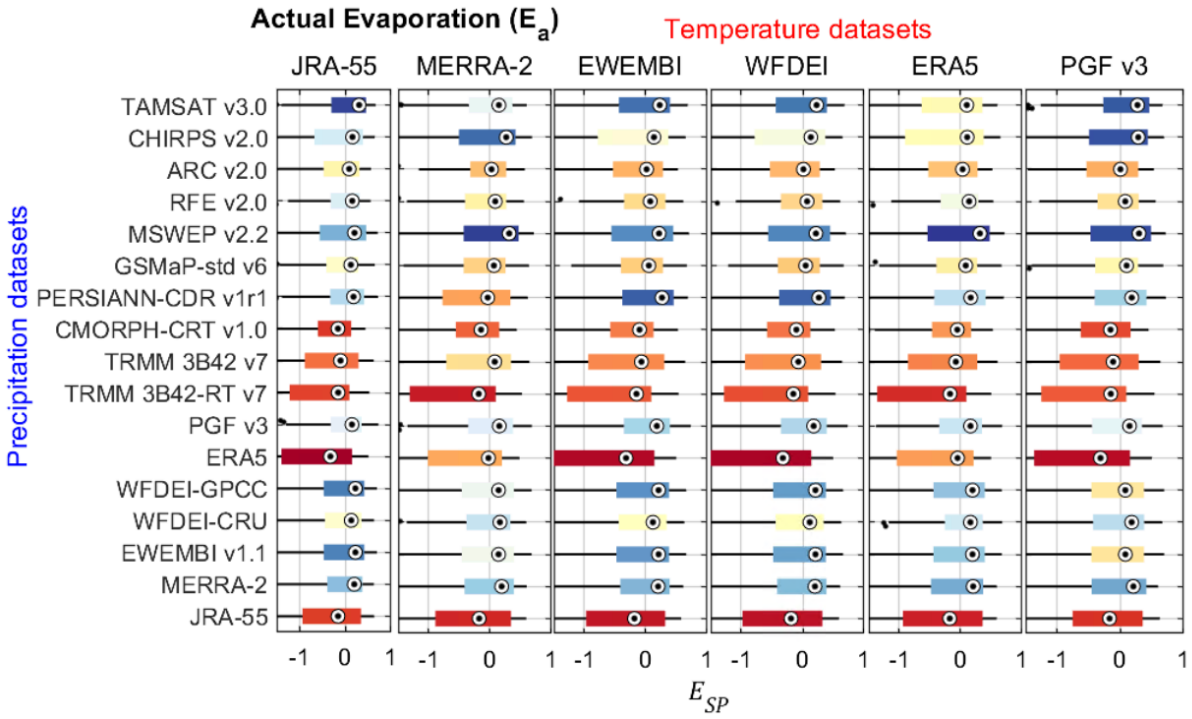


Figure 4.12. Spatial pattern efficiency (E_{SP}) of actual evaporation (E_a) over the entire simulation period (2003-2012) for the Volta River basin (VRB) and its climatic zones, using different combinations of precipitation and temperature datasets used as input for hydrological modelling. Each boxplot has 120 values corresponding to the number of months. The boxplots are coloured from the best (blue) to the worst performance (red) based on the median value.

4.9 Discussions

This study builds upon and expands existing research studies on the evaluation of meteorological datasets in several ways:

(i) The evaluation of the spatial patterns of multiple hydrological processes (i.e. streamflow, actual evaporation, soil moisture, and terrestrial water storage) in addition to the more classically evaluated temporal dynamic.

(ii) The evaluation of a high number of both satellite-based and reanalysis rainfall datasets considered in combination with different temperature datasets.

(iii) The assessment of the model performance across four considerably different climatic zones from semi-arid to sub-humid.

The overall outcome of this analysis is the ranking of all the meteorological datasets based on their ability to simulate various hydrological processes across different climatic zones in the VRB (Figure 4.6). It is worth noting that the overall ranking shows which product is best or worst at simulating a given hydrological flux or state variable. However, the ranking does not

systematically tell whether a dataset is good or bad. Only the skill scores can be used to draw a judgement on the adequacy of a given dataset to produce plausible model outputs.

The results show that there is no single rainfall dataset outperforming the others in reproducing all hydrological processes across different climatic zones. These findings align with previous studies in the sense that there is no rainfall dataset that is the best everywhere (Beck *et al.*, 2017b; Sylla *et al.*, 2013). For datasets providing both rainfall and temperature data, the combination of the two variables as model input is not necessarily the best option for obtaining the highest performance in modelling a given hydrological state or flux variable. The best rainfall-temperature combinations for the spatiotemporal representation of each hydrological flux and state variable are provided in the Appendix 12.

The results can be considered valid for West Africa and regions with similar hydroclimatic and physical features. A wider generalization of the findings should be done with caution and after repeating similar evaluation studies in other places. Nevertheless, the key message is that: *there is no rainfall dataset of all hydrological processes* and *the best rainfall dataset for temporal dynamics might not be the best for spatial patterns*.

Despite the efforts to produce a comprehensive evaluation of the meteorological datasets, the results obtained might be subject to uncertainties related to the potential model structural deficiencies as well as errors in the observational datasets used for the model evaluation (Gupta and Govindaraju, 2019; McMillan *et al.*, 2010; Renard *et al.*, 2010). The distribution of the final model parameters (Appendix 18; Appendix 19) highlights the possibility of obtaining equally good model performances for different parameter sets (i.e. equifinality), which can be a justification for model recalibration. Moreover, it can be noticed that most of the model parameters are sensitive to the change in meteorological input datasets (Appendix 18). A detailed analysis of parameter variability as a function of input data is beyond the scope of the current study, but could build the basis of future research, namely to identify data errors by analyzing parameter patterns (e.g. rooting depth), and resolve potential structural deficiencies of the mHM model. However, the mHM is chosen because of its adequacy for the experiment of this study (for model selection, see Addor and Melsen, 2019). The structure of mHM allows the representation of seamless spatial patterns of hydrological processes through the MPR scheme (Samaniego *et al.*, 2017). In addition, mHM facilitates parameter regionalization and is therefore convenient for large-scale modelling, and it harnesses the full potential of the forcing datasets as it is a fully distributed model that has performed well in previous studies including those in the VRB (e.g. Dembélé *et al.*, 2020b; Poméon *et al.*, 2018). Regarding the model

evaluation, the comparison between the observed and modelled hydrological processes is done only on their temporal dynamics and spatial patterns using bias-insensitive metrics, except for streamflow, which limits the potential impact of satellite data uncertainty.

The model is calibrated only on *Q* data despite the known limitations of the *Q*-only calibration (Demirel *et al.*, 2018). However, regarding the goal of this study, that was the best option to obtain the impact of various meteorological forcing datasets on the plausibility of hydrological processes. As no rainfall dataset ranks first in simulating all the hydrological processes, this study confirms that model calibration on multiple variables is a way forward in improving the overall representation of the hydrological system and increasing the predictive skill of hydrological models (Dembélé *et al.*, 2020a; Dembélé *et al.*, 2020b). The domain-wide calibration strategy adopted in this study generates a unique parameter set for the simulation of multiple hydrological processes across several catchments with different hydroclimatic features, which has the consequence of having local differences in model performance. However, domain-wide calibration has proved to perform similarly to domain-split calibration in previous studies (Mizukami *et al.*, 2017), and it was ideal for this study because of the interest in simulating seamless spatial patterns, which might have not been possible with separately simulated portions of the basin. Moreover, the main goal of this study is to assess the adequacy of the meteorological datasets for large-scale hydrological modelling, knowing that these datasets usually have a coarse spatial resolution with pixels often averaged over regions with strong sub-grid variability.

Finally, the importance of regional evaluation is emphasized by this study because some region-tailored datasets (e.g. TAMSAT and ARC) which are not included in global scale studies (e.g. Beck *et al.*, 2017b; Essou *et al.*, 2016; Mazzoleni *et al.*, 2019) outperform global datasets. The decision to use a given dataset is not only motivated by the availability or the accuracy of the data, but also by data accessibility (i.e. storage platforms, openness, format, pre-processing requirement, etc.). The findings of this study provide further awareness for the data users and improvement avenues for data producers in their quest of the most accurate products (Beck *et al.*, 2017a; Berg *et al.*, 2018; Brocca *et al.*, 2014; Contractor *et al.*, 2020; Cucchi *et al.*, 2020; Massari *et al.*, 2020).

4.10 Conclusion

This modelling study evaluates the ability of multiple combinations of rainfall-temperature datasets to reproduce plausible hydrological processes and patterns. The experiment is done in the Volta River basin with the fully distributed mesoscale Hydrologic Model (mHM) over a 10-year period (2003-2012), using 17 rainfall and 6 temperature datasets from satellite and reanalysis sources. The spatial and temporal representation of streamflow, terrestrial water storage, soil moisture and actual evaporation are evaluated using in-situ and satellite remote sensing observational datasets. The key findings are:

- No rainfall dataset consistently outperforms all the others in reproducing the highest model performance for all hydrological processes, and the best dataset for the temporal dynamics is not necessarily the best for the spatial patterns.
- Rainfall datasets have a higher impact on the spatiotemporal representation of hydrological processes than temperature datasets, but the later have a higher influence on the spatial patterns of soil moisture.
- The large-scale performance for the meteorological datasets is not always valid for sub-regions in the same basin.

The findings of this study give a critical insight of the performance for several meteorological datasets in the challenging hydroclimatic environment of West Africa. They are expected to foster further research initiatives in improving the gridded meteorological datasets and further draw users' attention on the contrasting performances of these datasets in modelling hydrological fluxes and state variables. Efforts should be devoted in reporting on the impact of data uncertainties on process representation in hydrological modelling, especially when model outputs are used for decision-making.

Future studies can test the transferability of the model's global parameters across different input datasets, i.e. how reliable a parameter set obtained with a given input dataset is for running the same model with a different input dataset. The answer to this research question will shed light on the necessity of model recalibration when using different meteorological forcing. Furthermore, the predictive skill of the model can be improved with a parameter sensitivity analysis to determine parameters that affect the spatiotemporal representation of each hydrological flux and state variable.

Chapter 5

Potential of Gridded Evaporation Datasets for the Calibration of Hydrological Models*

Data is the new science. Big Data holds the answers.

Are you asking the right questions?

Patrick P. Gelsinger

Exactness and neatness in moderation is a virtue,

but carried to extremes narrows the mind.

Francois Fenelon

* This chapter is based on the following publication:
Dembélé, M., Ceperley, N., Zwart, S. J., Salvadore, E., Mariethoz, G., & Schaefli, B. (2020). Potential of Satellite and Reanalysis Evaporation Datasets for Hydrological Modelling under Various Model Calibration Strategies. Advances in Water Resources, https://doi.org/10.1016/j.advwatres.2020.103667

Abstract

Twelve actual evaporation datasets are evaluated for their ability to improve the performance of the fully distributed mesoscale Hydrologic Model (mHM). The datasets consist of satellite-based diagnostic models (MOD16A2, SSEBop, ALEXI, CMRSET, SEBS), satellite-based prognostic models (GLEAM v3.2a, GLEAM v3.3a, GLEAM v3.2b, GLEAM v3.3b), and reanalysis (ERA5, MERRA-2, JRA-55). Four distinct multivariate calibration strategies (basin-average, pixel-wise, spatial bias-accounting and spatial bias-insensitive) using actual evaporation and streamflow are implemented, resulting in 48 scenarios whose results are compared with a benchmark model calibrated solely with streamflow data. A process-diagnostic approach is adopted to evaluate the model responses with in-situ data of streamflow and independent remotely sensed data of soil moisture from ESA-CCI and terrestrial water storage from GRACE. The method is implemented in the Volta River basin, which is a data scarce region in West Africa, for the period from 2003 to 2012.

The results show that the evaporation datasets have a good potential for improving model calibration, but this is dependent on the calibration strategy. All the multivariate calibration strategies outperform the streamflow-only calibration. The highest improvement in the overall model performance is obtained with the spatial bias-accounting strategy (+29%), followed by the spatial bias-insensitive strategy (+26%) and the pixel-wise strategy (+24%), while the basin-average strategy (+20%) gives the lowest improvement. On average, using evaporation data in addition to streamflow for model calibration decreases the model performance for streamflow (-7%), which is counterbalance by the increase in the performance of the terrestrial water storage (+11%), temporal dynamics of soil moisture (+6%) and spatial patterns of soil moisture (+89%). In general, the top three best performing evaporation datasets are MERRA-2, GLEAM v3.3a and SSEBop, while the bottom three datasets are MOD16A2, SEBS and ERA5. However, performances of the evaporation products diverge according to model responses and across climatic zones. These findings open up avenues for improving process representation of hydrological models and advancing the spatiotemporal prediction of floods and droughts under climate and land use changes.

5.1 Introduction

Assessing the spatiotemporal variability of hydrological processes is the crux of effective water resource management. Global warming is expected to intensify (i.e., accelerate) the hydrological cycle, thus increasing or decreasing evaporation depending on places (Donat *et al.*, 2016; Famiglietti and Rodell, 2013; Huntington, 2006).

Evaporation is a dominant flux of the water cycle (Martens *et al.*, 2018; Oki and Kanae, 2006). It represents the exchange of water and energy between terrestrial ecosystems and the atmosphere. Therefore, evaporation can be used as a proxy for moisture availability and its consumption rate (He *et al.*, 2019; Joiner *et al.*, 2018; Van der Ent *et al.*, 2010). The basic theories and estimation methods of evaporation are widely documented (e.g., Chen and Liu, 2020; Liou and Kar, 2014; McMahon *et al.*, 2013; Pan *et al.*, 2019; Zhang *et al.*, 2016). As evaporation is the central flux that defines land-atmosphere interactions, improving its simulation in hydrological models is a prerequisite for reliable studies that link climate and land use change (Fisher *et al.*, 2017; Mueller *et al.*, 2011). This is particularly the case for catchments with strong anthropogenic influences (e.g., irrigation schemes, dams, etc.), and in data scarce regions (Becker *et al.*, 2019; Jiang and Wang, 2019), where there is insufficient data on human water use for robust model calibration. Accordingly, calibrating hydrological models solely based on streamflow is not sufficient to guarantee a plausible representation of the hydrological system because streamflow is the result of several inter-linked processes, thereby it masks spatial heterogeneity (Tobin and Bennett, 2017; Wambura *et al.*, 2018). This limitation in model implementation has been overcome by the advent of multivariate calibration techniques using satellite-based datasets, which offers models a chance for better spatial heterogeneity (Efstratiadis and Koutsoyiannis, 2010; Rakovec *et al.*, 2016b).

In the quest to improve process representation in hydrological models, large scale distributed hydrological modelling faces the challenging requirement of spatially explicit observational datasets for model setup and performance evaluation (Clark *et al.*, 2017; Fatichi *et al.*, 2016; Hrachowitz and Clark, 2017). Correspondingly, Satellite Remote Sensing (SRS) and reanalysis datasets of various hydrological processes have been used as input data (Beck *et al.*, 2017b; Maggioni and Massari, 2018; Sheffield *et al.*, 2018) or as calibration and evaluation data for hydrological models (Koppa and Gebremichael, 2020; McCabe *et al.*, 2017). Evaporation estimates from SRS are increasingly used in multivariate calibration of hydrological models. Because it is a key indicator of surface water availability, evaporation is an essential source of information for better constraining the spatiotemporal representation of processes in

hydrological models (Bai and Liu, 2018; Cui *et al.*, 2019; Talsma *et al.*, 2018). The increasing availability and diversity of gridded evaporation datasets has triggered many evaluation and comparison studies (Long *et al.*, 2014; Vinukollu *et al.*, 2011b), which highlight significant differences between the datasets and thereby indicate underlying uncertainty in the evaporation estimates (Baik *et al.*, 2018; López *et al.*, 2017). The uncertainty stems from the strong variability of bio-geophysical variables that drive evaporation (e.g., albedo, net radiation, surface roughness and temperature) and the diversity of the model structures, model parametrizations and input datasets used to estimate evaporation (Badgley *et al.*, 2015; Wang and Dickinson, 2012; Zhang *et al.*, 2020). Therefore, the choice and use of SRS evaporation data in hydrological modelling should be done cautiously, particularly in catchments with strong anthropogenic influences (Senkondo *et al.*, 2019; Yang *et al.*, 2016). Generally, the following four approaches are adopted to evaluate gridded evaporation products:

- Analysis of the variance between several products (e.g., Jimenez *et al.*, 2011; Khan *et al.*, 2018; Mueller *et al.*, 2011; Senkondo *et al.*, 2019; Trambauer *et al.*, 2014);
- Point-to-pixel comparison with ground-based measurements (e.g., Chen *et al.*, 2014; McCabe *et al.*, 2015; Michel *et al.*, 2016; Ramoelo *et al.*, 2014; Velpuri *et al.*, 2013);
- Hydrological consistency by water balance calculation (e.g., Liu *et al.*, 2016; McCabe *et al.*, 2008; Miralles *et al.*, 2016; Wang *et al.*, 2018; Weerasinghe *et al.*, 2019); and
- Assessing the ability of evaporation datasets in improving the parameter estimation of hydrological models (e.g., Demirel *et al.*, 2018; Immerzeel and Droogers, 2008; Jiang *et al.*, 2020; Poméon *et al.*, 2018; Winsemius *et al.*, 2008).

Assessing the uncertainty of evaporation estimates at large-scale is challenging due to the limited availability of ground-based measurements (Bhattarai *et al.*, 2019; Ceperley *et al.*, 2017). The uncertainty in evaporation varies in space and according to climate regions (Jung *et al.*, 2019; Long *et al.*, 2014; Vinukollu *et al.*, 2011a). In evaluating the contribution of gridded evaporation datasets to hydrological model calibration, some limitations can be denoted in previous studies. Most previous studies only use or compare few evaporation datasets (e.g., Kunnath-Poovakka *et al.*, 2016; Vervoort *et al.*, 2014), and rarely (if any) investigate the use of reanalysis datasets (i.e., retrospective analysis; cf. Bosilovich *et al.*, 2008), which are an important source of spatial evaporation estimates (Feng *et al.*, 2019). Usually, a lumped or semi-distributed model is used (e.g., Odusanya *et al.*, 2019; Rientjes *et al.*, 2013), which does not harness the full potential of the gridded evaporation datasets that is their spatial patterns

(Armstrong *et al.*, 2019; Stisen *et al.*, 2018). Most studies do not test different model calibration strategies, with some exceptions that use a semi-distributed model (e.g., Herman *et al.*, 2018; Rajib *et al.*, 2018a). Finally, a few studies use a bias-insensitive metric to focus only on the spatial patterns of gridded evaporation products (Dembélé *et al.*, 2020b; Koch *et al.*, 2018).

This study aims to fill current knowledge gaps by evaluating the utility of nine satellite-based and three reanalysis evaporation datasets in improving the performance of a distributed hydrological model using four distinct calibration strategies. This study does not intend to quantify the intrinsic accuracy of the evaporation products nor determine whether a product is better than the others in terms of absolute values. Rather it strives to evaluate their ability to improve the simulations of a distributed hydrological model when used as a calibration variable. Besides the high number and diversity of gridded evaporation datasets evaluated, the novelty of this study is the implementation of four distinct model calibration strategies with a fully distributed hydrological model, the evaluation of the model responses with multiple variables (i.e., streamflow, soil moisture and terrestrial water storage) to test evaporation error propagation on other hydrological processes, and the application of the experiment in a large basin spread across four eco-climatic zones with considerable anthropogenic influence. This study strives to answer two inter-related research questions. Firstly, what is the ability of satellite and reanalysis evaporation datasets to improve the overall predictive skill of a fully distributed hydrological model? Secondly, how important is the model calibration strategy in improving the representation of hydrological processes? The proposed research is carried out in the Volta River basin located in West Africa, using the mesoscale Hydrologic Model (mHM) (cf. Section 4.5, Chapter 4) over a period of ten consecutive years (2003-2012).

5.2 Evaporation datasets

Twelve gridded actual evaporation datasets including nine SRS-based products (MOD16A2, SSEBop, ALEXI, CMRSET, SEBS, GLEAM v3.2a, GLEAM v3.3a, GLEAM v3.2b, GLEAM v3.3b) and three reanalysis products (ERA5, MERRA-2, JRA-55) are evaluated in this study. Based on evaporation modelling approaches (Yilmaz *et al.*, 2014), the SRS-based datasets can be further classified as diagnostic products (MOD16A2, SSEBop, ALEXI, CMRSET, SEBS) or prognostic products (GLEAM v3.2a, GLEAM v3.3a, GLEAM v3.2b, GLEAM v3.3b), while the reanalysis datasets are all prognostic products. A summary of the evaporation datasets is provided in Table 5.1. The data access portals and their full names are provided in Appendix 20. Four versions of the GLEAM product are evaluated. They differ in terms of input data used

for their production and in terms of their spatiotemporal coverage (cf. Table 1 in Martens *et al.*, 2017). The version v3.3 differs from the v3.2 in the following forcing datasets: surface radiation, near-surface air temperature and land cover maps. The versions v3.3a and v3.2a are produced with reanalysis, satellite and gauge-based datasets, while the versions v3.3b and v3.2b are mainly produced with satellite datasets.

Considerable differences can be observed both in the temporal dynamics and the spatial patterns of the 12 gridded evaporation datasets across the climatic zones in the VRB (Figure 5.1 and Figure 5.2).

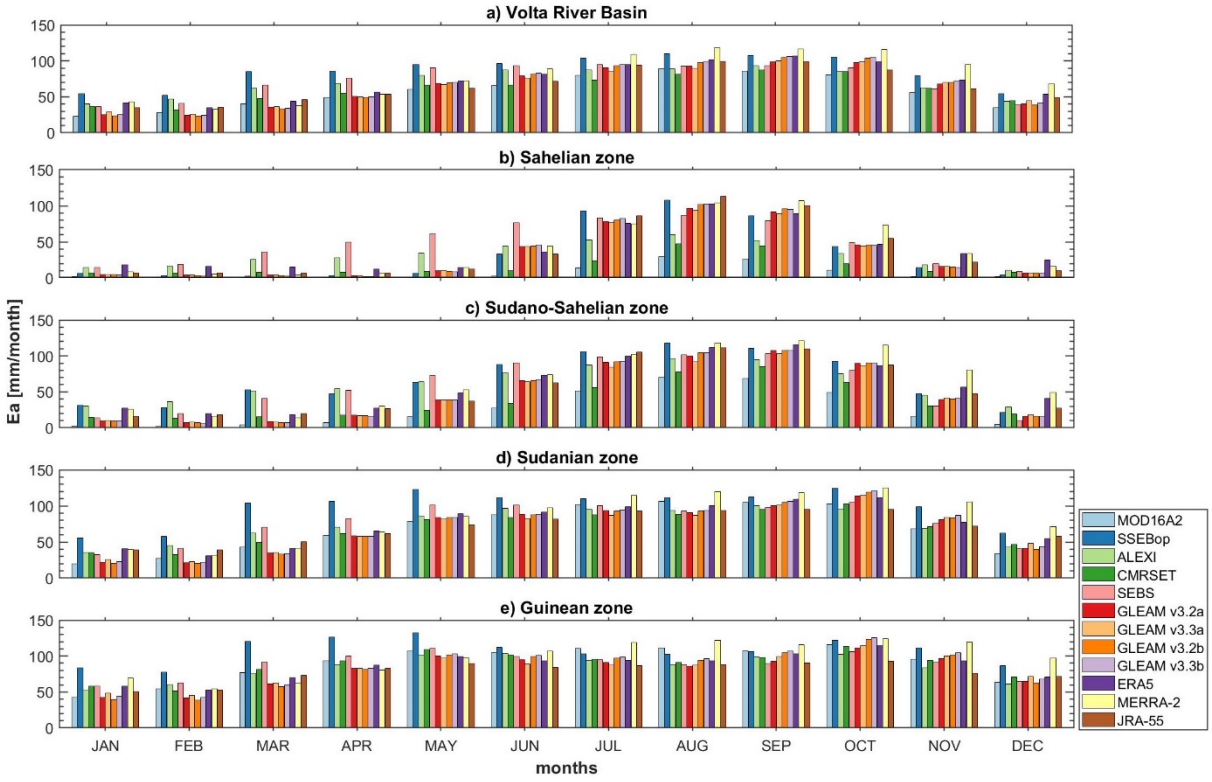


Figure 5.1. Mean monthly total actual evaporation (E_a) of 12 gridded evaporation datasets averaged for the four climatic zones (sub-figures b-e) in the Volta River basin over the period 2003-2012.

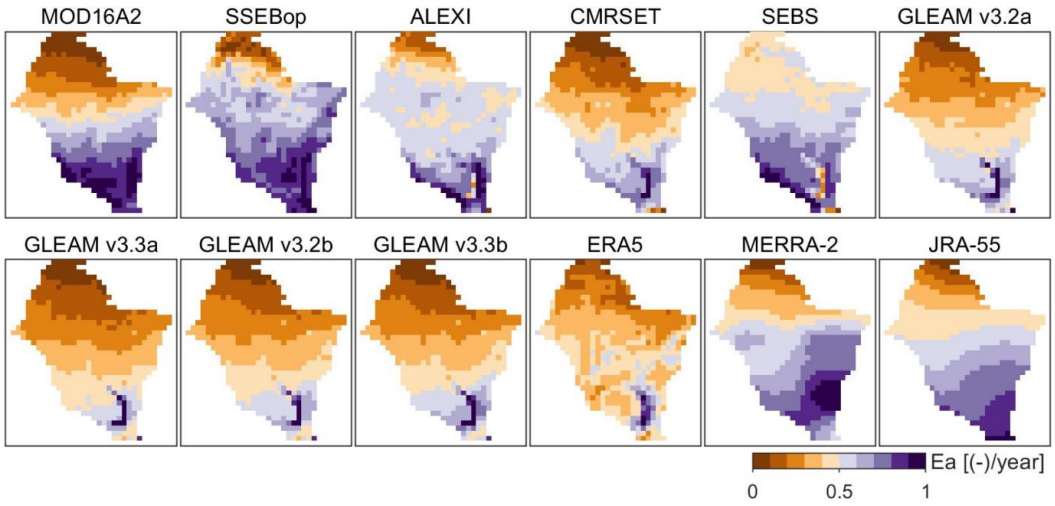


Figure 5.2. Spatial distribution of total annual actual evaporation (E_a) for 12 gridded evaporation datasets in the Volta River basin averaged for the period 2003-2012. A min-max normalization of the values allows rescaling them between 0 and 1 to facilitate the intercomparison of the spatial patterns among the datasets.

Table 5.1. Gridded actual evaporation datasets. The characteristics of the datasets are those used in this study although the same datasets can be available from the data providers with different versions and spatiotemporal resolutions.

Datasets	Spatial coverage	Spatial resolution	Temporal coverage	Temporal resolution	References
MOD16A2	Global	0,0085° (~1 km)	2001-2014	Monthly	Mu *et al.* (2011)
SSEBop	Global	0.0083° (~1 km)	2003-2014	Monthly	Senay *et al.* (2007), Senay *et al.* (2013)
ALEXI	70° N–60° S	0.05° (~5.6 km)	2003-2015	Monthly	Anderson *et al.* (1997), Anderson *et al.* (2007)
CMRSET	Global	0.05° (~5.6 km)	2001-2013	Monthly	Guerschman *et al.* (2009)
SEBS	40° N–40° S	0.05° (~5.6 km)	2001-2012	Monthly	Su (2002)
GLEAM v3.2a	Global	0.25° (~28 km)	1980-present	Daily	Martens *et al.* (2017), Miralles *et al.* (2011)
GLEAM v3.3a					
GLEAM v3.2b	50° N–50° S		2003-present		
GLEAM v3.3b					
ERA5	Global	0.25° (~28 km)	1979-present	Hourly	Hersbach *et al.* (2018)
MERRA-2	Global	0.5° x 0.625° (~56 km)	1980-present	Hourly	Gelaro *et al.* (2017), Reichle *et al.* (2017)
JRA-55	Global	1.25° (~140 km)	1959-present	3-hourly	Kobayashi *et al.* (2015), Harada *et al.* (2016)

5.3 Model calibration and evaluation strategies

The description of the hydrological model setup is provided in the Section 4.5 of Chapter 4. The modelling period extends from 2000 to 2012 with 3 years of model warm-up (2000-2002), 6 years for calibration (2003-2008), and 4 years for evaluation (2009-2012). Available daily in-situ streamflow datasets from 11 gauging locations are used for model calibration and evaluation, while monthly datasets of E_a (Table 5.1) are used for model calibration, and monthly datasets of S_u (ESA CCI) and S_t (GRACE) are used for model evaluation. All the E_a datasets are rescaled to 0.25° using bilinear interpolation to match the modelling spatial resolution, and sub-monthly data are aggregated to monthly resolution. Only the first soil layer of mHM is compared to the ESA CCI data, which represents the surface soil moisture.

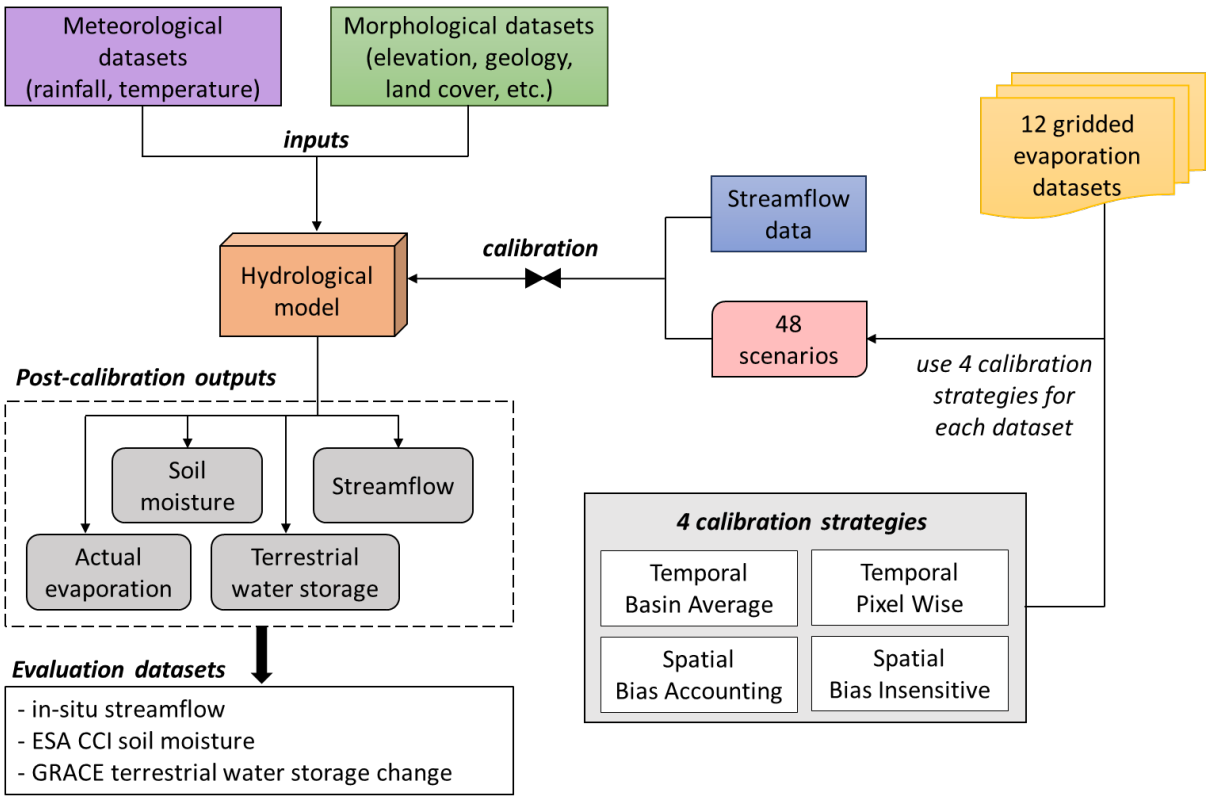


Figure 5.3. Overview of the modelling approach to evaluate the reanalysis and satellite-based evaporation datasets.

First, a streamflow-only calibration is adopted as benchmark. Then, the contribution of evaporation datasets in improving hydrological model calibration is tested by simultaneously constraining the model with streamflow and each of the twelve gridded evaporation datasets using four calibration strategies for each (Figure 5.3). Therefore, 48 scenarios (i.e., 12 datasets times 4 calibration strategies) are developed and compared to the benchmark calibration to

evaluate the impact of different calibration strategies on model performance. The dynamically dimensioned search algorithm (Tolson and Shoemaker, 2007) is used for parameter estimation, using 5,000 iterations for each of the 48 scenarios and for the benchmark model calibration. The computational runtime is about 6 days for each of the 49 model simulations on a computer Intel Xeon Processor E5-2697 v3 with 64 GB of RAM.

5.3.1 Calibration on streamflow data – benchmark

The benchmark calibration (case Q) is elaborated by calibrating the hydrological model solely with streamflow (Q) data. The objective function for case Q (Φ_{Q}) to be minimized is obtained by calculating the average Kling-Gupta efficiency (E_{KG}) over the 11 gauging points for Q in the basin, and subtracting it from 1. Φ_{Q} ranges from its optimal value that is 0 to positive infinity, and is formulated as follows:

$$\Phi_{\mathrm{Q}} = 1 - \left[\frac{1}{g}\sum_{i=1}^{g} E_{\mathrm{KG,i}}\left(Q_{\mathrm{mod,i}}, Q_{\mathrm{obs,i}}\right)\right], \text{ with} \tag{5.1}$$

$$E_{\mathrm{KG}} = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{5.2}$$

where g is the total number of streamflow gauging stations in the basin, E_{KG} is the modified Kling-Gupta efficiency (Kling *et al.*, 2012) calculated for the observed ($Q_{\mathrm{obs,i}}$) and modelled ($Q_{\mathrm{mod,i}}$) streamflow of the i$^{\mathrm{th}}$ gauging point. E_{KG} is composed of the Pearson correlation coefficient (r), the bias term (β, i.e., the ratio of the means) and the variability term (γ, i.e., the ratio of the coefficients of variation). E_{KG} ranges from negative infinity to its optimal value that is 1. A model is better than the mean observed flow if $E_{\mathrm{KG}} > -0.41$ (Knoben *et al.*, 2019).

5.3.2 Multivariate calibration with evaporation and streamflow

A bias-insensitive and multi-component metric developed by Dembélé *et al.* (2020b) is used to quantify the degree of reproduction of the spatial patterns of hydrological processes. The proposed spatial pattern efficiency (E_{SP}) metric only considers the spatial pattern of the underlying variables and ignores their absolute values (cf. Eq. 4.8 in Section 4.7, Chapter 4).

Four multivariate calibration strategies with distinct objective functions are proposed to simultaneously consider Q and E_{a} data as calibration variables (Equations 5.4, 5.8, 5.10 and 5.12). Each objective function is formulated based on the Euclidean distance approach, in which all elements are equally weighted (Khu and Madsen, 2005). The Euclidian distance (D_{E})

between two points X and Y of coordinates ($x_1, x_2, \ldots, x_n$) and ($y_1, y_2, \ldots, y_n$) in an n-dimensional space (Upton and Cook, 2014) is given by:

$$D_{\mathrm{E}} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \tag{5.3}$$

The four multivariate calibration strategies differ from each other based on the formulation of the sub-objective function for E_a (i.e., $\Phi_{\mathrm{E_a}}$), while Φ_{Q} remains unchanged. The observed variable ($E_{\mathrm{a,obs}}$) and the modelled variable ($E_{\mathrm{a,mod}}$) of actual evaporation are represented each by a 3D array of dimension [M x N x T], with M the number of rows, N the number of columns, and T the number of time steps (Figure 5.4). For this study, M represents 40 latitude rows, N represents 36 longitude columns and T represents 72 months (i.e. calibration period). The modelling domain has 1440 grid cells of which 619 are active (i.e. grid cells representing the basin area). The inactive grid cells are masked out during the calculation of the performance metrics.

In the following, $\mathbf{E}_{\mathrm{a}}^{t}$ is a 2D array of actual evaporation represented by all cells (i,j) in the spatial domain Ω, and E_{a}^{ij} is the time series of actual evaporation for a given cell at row i and column j. The four multivariate calibration strategies are defined as follows:

1- *Temporal basin average* (BA): the matching of the observed and modelled E_a is done on basin-averaged time series. The sub-objective function ($\Phi_{\mathrm{E_a\_BA}}$) to be minimized is obtained by calculating the E_{KG} for the observed and modelled time series of basin average actual evaporation ($\overline{E_{\mathrm{a,obs}}}$ and $\overline{E_{\mathrm{a,mod}}}$) and subtracting it from 1. The objective function (Φ_{BA}) is formulated as follows:

$$\Phi_{\mathrm{BA}} = \sqrt{{\Phi_{\mathrm{Q}}}^2 + {\Phi_{\mathrm{E_a\_BA}}}^2}, \text{ with} \tag{5.4}$$

$$\Phi_{\mathrm{E_a\_BA}} = 1 - E_{\mathrm{KG}}\left(\overline{E_{\mathrm{a,mod}}}, \overline{E_{\mathrm{a,obs}}}\right), \text{ where} \tag{5.5}$$

$$\overline{E_{\mathrm{a,mod}}}(t) = \frac{1}{M \times N} \sum_{\substack{t=1 \\ \forall\, i,j \in \Omega}}^{T} \mathbf{E}_{\mathrm{a,mod}}^{t}(i,j)\,, \text{and} \tag{5.6}$$

$$\overline{E_{\mathrm{a,obs}}}(t) = \frac{1}{M \times N} \sum_{\substack{t=1 \\ \forall\, i,j \in \Omega}}^{T} \mathbf{E}_{\mathrm{a,obs}}^{t}(i,j) \tag{5.7}$$

2- *Temporal pixel-wise* (PW): the matching of the modelled and observed E_a is done individually on the time series of each grid cell in the basin. The E_{KG} is calculated for the observed and modelled time series of E_a at each grid cell in the basin, and the sub-objective function ($\Phi_{E_{a\_PW}}$) to be minimized is the average of the E_{KG} calculated for all grids, subtracted from 1. The objective function (Φ_{PW}) is formulated as follows:

$$\Phi_{PW} = \sqrt{\Phi_Q{}^2 + \Phi_{E_{a\_PW}}{}^2}, \text{ with} \tag{5.8}$$

$$\Phi_{E_{a\_PW}} = 1 - \left[\frac{1}{M \times N} \sum_{\substack{1 \le i \le M \\ 1 \le j \le N \\ \forall\, t \in T}} E_{KG}\left(E^{ij}_{a,mod}, E^{ij}_{a,obs}\right) \right] \tag{5.9}$$

3- *Spatial bias-accounting* (SB): the matching of the modelled and observed E_a is done for all pixels at each time step. The E_{KG} is calculated at each time step between all the pixels of the observed and modelled E_a. The sub-objective function ($\Phi_{E_{a\_SB}}$) to be minimized is the average of the E_{KG} calculated for all time steps, subtracted from 1. The objective function (Φ_{SB}) is formulated as follows:

$$\Phi_{SB} = \sqrt{\Phi_Q{}^2 + \Phi_{E_{a\_SB}}{}^2}, \text{ with} \tag{5.10}$$

$$\Phi_{E_{a\_SB}} = 1 - \frac{1}{T} \sum_{\substack{t=1 \\ \forall\, i,j \in \Omega}}^{T} E_{KG}\left(\mathbf{E}^t_{a,mod}(i,j), \mathbf{E}^t_{a,obs}(i,j)\right) \tag{5.11}$$

4- *Spatial bias-insensitive* (SP): the sub-objective function ($\Phi_{E_{a\_SP}}$) to be minimized is similarly calculated as for the SB calibration except that a bias-insensitive metric (i.e., E_{SP}) is used as skill score instead of E_{KG}. E_{SP} is a bias insensitive metric of spatial patterns developed by Dembélé *et al.* (2020b), cf. Eq. 4.8 (Section 4.7, Chapter 4). The objective function (Φ_{SP}) is formulated as follows:

$$\Phi_{SP} = \sqrt{\Phi_Q{}^2 + \Phi_{E_{a\_SP}}{}^2}, \text{ with} \tag{5.12}$$

$$\Phi_{E_{a\_SP}} = 1 - \frac{1}{T} \sum_{\substack{t=1 \\ \forall\, i,j \in \Omega}}^{T} E_{SP}\left(\mathbf{E}^t_{a,mod}(i,j), \mathbf{E}^t_{a,obs}(i,j)\right) \tag{5.13}$$

All the objective functions (Φ_{BA}, Φ_{PW}, Φ_{SB} and Φ_{SP}) vary between their optimal value that is 0 and positive infinity. Except case SP that is a bias-insensitive approach, other calibration strategies consider the absolute values (i.e. raw data) of evaporation.

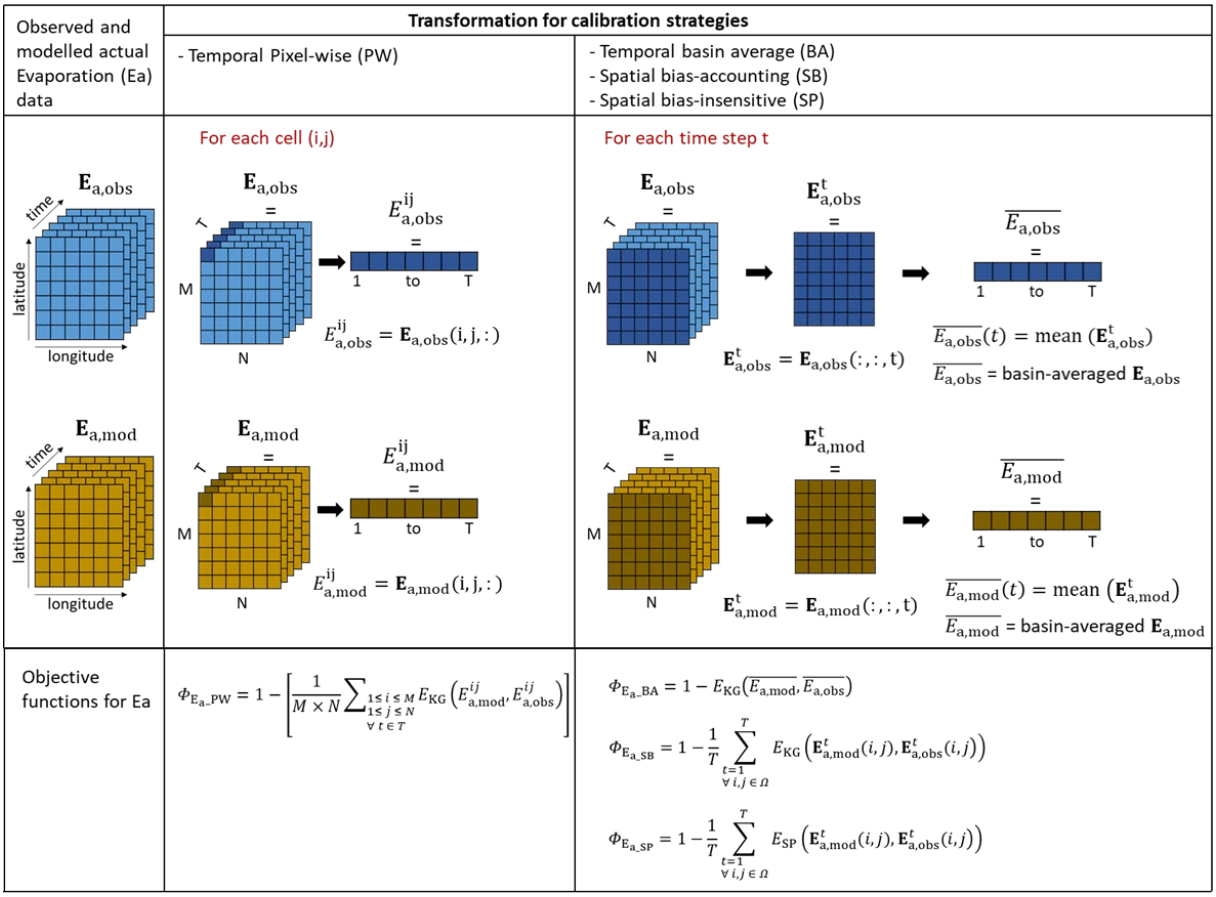


Figure 5.4. Graphical and programming syntax description of the use of evaporation data in the four calibration strategies. The blue matrix represents the observed data while the brown matrix represents the modelled data. For the syntax, (i, j, :) means ith element of the first dimension (latitude), jth element of the second dimension (longitude) and all elements of the third dimension (time). Similarly, (:, :, t) means tth element of the third dimension and all elements of the first and second dimensions. The objective functions are calculated based on the Kling-Gupta efficiency (E_{KG}) and the spatial pattern efficiency (E_{SP}) metrics.

5.3.3 Model output evaluation

In addition to daily streamflow data, independent monthly datasets of satellite-based soil moisture and terrestrial water storage (Table 4.2) are used for model evaluation. The temporal dynamics of streamflow is evaluated with E_{KG}. Both the temporal dynamics and the spatial patterns of modelled S_u are evaluated using the Pearson correlation coefficient (r) and the spatial pattern efficiency metric (E_{SP}), while only the temporal dynamics of modelled S_t is assessed using r, due to the coarse spatial resolution of the GRACE data. The skill scores for the temporal dynamics are calculated for each pixel (or gauging point) in the basin, while spatial skill scores are calculated per time step.

5.4 Results

In general, the trend of the model performance (high vs. low scores) among the scenarios (i.e., evaporation datasets vs. calibration strategies) is conserved between the calibration and the evaluation periods for all variables. Therefore, the following results are presented for the entire simulation period that comprises the calibration and evaluation periods. Additional and detailed results are provided in the supplementary material of Dembélé et al., (2020a).

5.4.1 Model performance for multiple hydrological processes in the VRB

The model performance for various hydrological processes in the VRB reveals the potential of SRS and reanalysis evaporation datasets to improve the model responses if the appropriate calibration strategy is used (Figure 5.5).

For Q, the benchmark model (i.e., Q-only) yields a median E_{KG} of 0.69. In the multivariate calibration scenarios (i.e., Q+E_a), the E_{KG} of Q varies between 0.42 for SEBS with case PW to 0.73 for CMRSET with case SP. The best multivariate calibration strategy is the case SB with an average E_{KG} of 0.68 and 75 % of the evaporation datasets producing a higher model performance than the benchmark, followed by case SP (E_{KG} = 0.67), case PW (E_{KG} = 0.63) and case BA (E_{KG} = 0.60). The top 3 best evaporation datasets for the average E_{KG} of Q over the calibration strategies are GLEAM v3.2b (E_{KG} = 0.71), GLEAM v3.3b (E_{KG} = 0.71) and GLEAM v3.3a (E_{KG} = 0.70), while the worst are ALEXI (E_{KG} = 0.61), SEBS (E_{KG} = 0.56) and ERA5 (E_{KG} = 0.51). The decrease in the model performance for Q in the multivariate calibration might be an artefact caused by equifinality (i.e., non-uniqueness of model parameters; Beven, 2006b; Savenije, 2001) that occurred with the Q-only calibration, which gives more degrees of freedom for constraining the model parameter space (Dembélé *et al.*, 2020b). In fact, Appendix 24 to Appendix 27 show that the high performance for Q achieved with the Q-only calibration is obtained at the expense of poor performance for other hydrological processes, while the multivariate calibrations with Q+E_a result in parameter sets that provide equivalent model performance for Q but higher model performance for S_t and S_u. Appendix 23 shows the distribution of the global parameters.

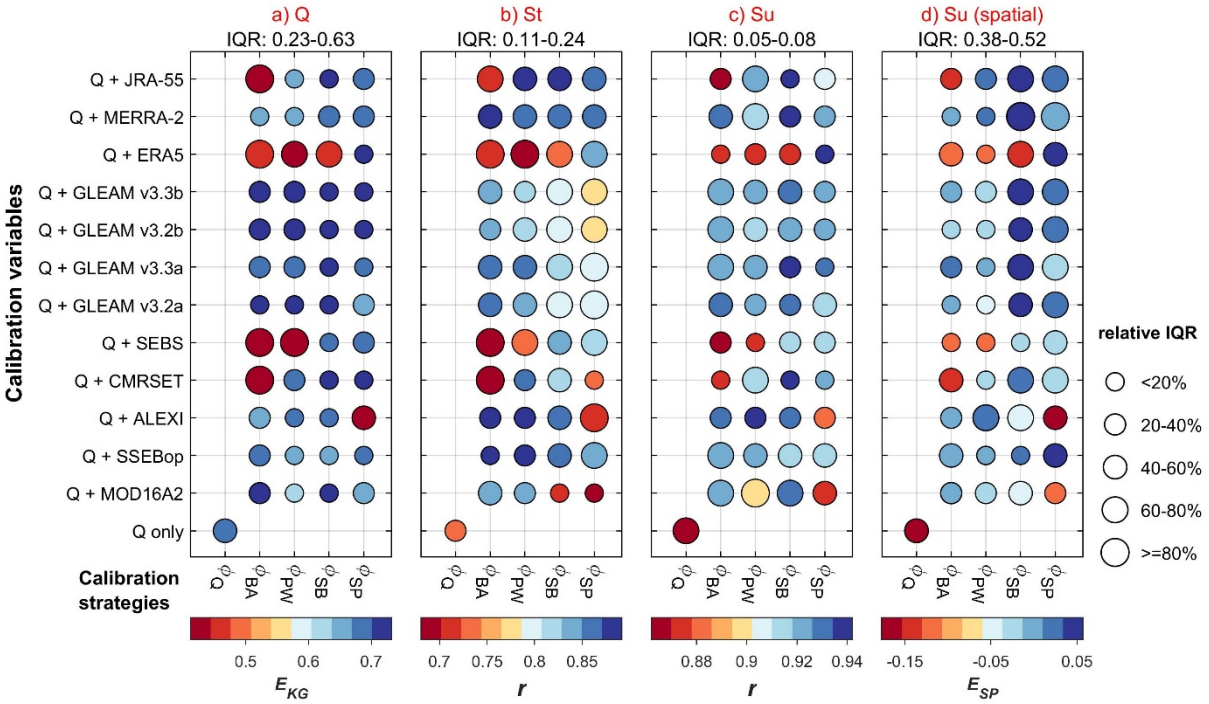


Figure 5.5. Model performance in the entire simulation period (2003-2012) for the temporal dynamics (a, b, c) of streamflow (Q), terrestrial water storage (S_t) and soil moisture (S_u), and the spatial patterns (d) of S_u in the Volta River basin. The x-axis gives the objective functions for different model calibration strategies. The y-axis indicates the variables used for the model calibration. Q-only is the benchmark calibration. Circle colour represents the median model performance obtained with 11 gauging points for Q, with 52 pixels for S_t, with 619 pixels for S_u (c) or with 120 months for S_u (d). The colour bars show the skill score (i.e., E_{KG}, r, E_{SP}). Circle size represents model performance variability in terms of relative Interquartile Range (IQR) computed as $(IQR - IQR_{min})/(IQR_{max} - IQR_{min})$. The IQR range for the 49 scenarios is given in the subplot titles. The best model is the bluest and smallest circle, while the worst model is the reddest and largest circle.

The evaporation datasets show a high potential to improve the temporal dynamics of modelled S_t as 79% of the multivariate calibration scenarios outperform the case Q (r = 0.73) with an average r of 0.81. The lowest average r of S_t among the calibration strategies is given by case SP (r = 0.79), but interestingly, it outperforms the case Q. The highest performances for median r of S_t are obtained with MERRA-2 (r = 0.87), SSEBop (r = 0.87) and ALEXI (r = 0.84) while lowest performances are given by SEBS (r = 0.77), MOD16A2 (r = 0.76) and ERA5 (r = 0.74).

The temporal dynamics of modelled S_u show a higher model performance than that of S_t, with an average r of 0.91 and 98% of the multivariate calibration scenarios that outperform the case Q (r = 0.86). The case BA performs similarly to case PW and case SP with an average r of 0.91 across evaporation datasets, and outperformed by case SB (r = 0.93). The highest r of S_u is 0.93,

and it is obtained with GLEAM v3.3a, MERRA-2 and GLEAM v3.2a, while the weakest scores are obtained with MOD16A2 (r = 0.91), SEBS (r = 0.89) and ERA5 (r = 0.89).

The representation of the spatial patterns of S_u improves for all the multivariate scenarios as compared to the case Q. However, case SB (E_{SP} = 0.0) has the highest average performance across the evaporation datasets, while case BA (E_{SP} = -0.04) has the lowest performance. The best evaporation datasets for the simulation of the spatial patterns of S_u considering all the calibration strategies are SSEBop (E_{SP} = 0.02), MERRA-2 (E_{SP} = 0.01) and GLEAM v3.3a (E_{SP} = 0.01), while the worst are MOD16A2 (E_{SP} = -0.04), SEBS (E_{SP} = -0.07) and ERA5 (E_{SP} = -0.09).

In general, it is observed that the model performances for S_t and S_u improve for any multivariate calibration scenario. Among the multivariate calibration strategies, case SB gives the best results considering the average model performance for all variables (Q, S_t and S_u). In Figure 5.6b, the highest average relative change in model performance by multivariate calibration strategies as compared to the Q-only calibration is obtained with case SB is (+29%), followed by case SP (+26%), case PW (+24%) and case BA (+20%). Consequently, all grid-based model calibration strategies outperform the basin-average calibration (case BA), which gives the lowest average model performance. These results highlight the value of calibrating hydrological models on the full extent of gridded evaporation datasets. It is noted that, in most of the scenarios, calibrating the model only on the spatial patterns (case SP) of the evaporation datasets, thereby ignoring their absolute values, improves the predictive skill of the model with a higher performance than that of case Q and case BA. With these findings, case SP (i.e., only spatial patterns of evaporation datasets) is preferred to case SB (i.e., absolute values of evaporation datasets) for the calibration of hydrological models because the propagation of the errors of the evaporation estimates into the modelling process can be case specific and depends on the model structure. This observation also draw attention on the adequacy of the model for a given experiment (Addor and Melsen, 2019). Moreover, for a given evaporation dataset, the spatial variation in biases in the estimates can lead to contrasting performance across regions (Jung *et al.*, 2019; Nicholson, 2000).

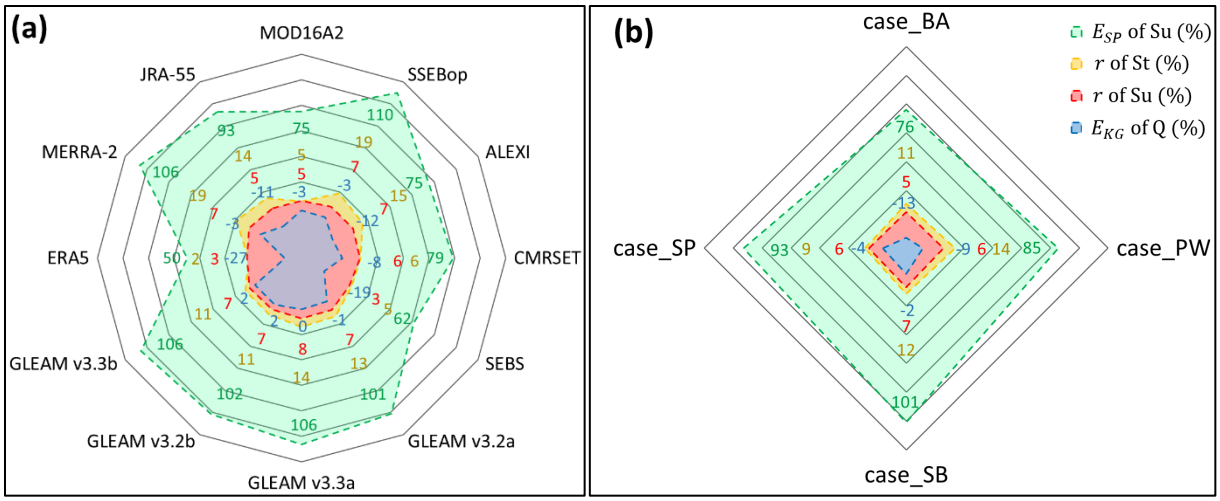


Figure 5.6. Relative change in the hydrological model performance as compared to the Q-only calibration strategy when (a) adopting a multivariate calibration strategy (Q+E_a) with twelve evaporation datasets, or (b) using four different calibration strategies. The values on the line from each vertex to the center of the polygons give the relative difference in model performance as compared to the Q-only calibration. The change in model performance is given in percentage of the performance metrics (i.e., E_{KG}, r, E_{SP}) for streamflow (Q), terrestrial water storage (S_t), and soil moisture (S_u). For instance in (a), using SSEbop in multivariate calibration increases E_{SP} of S_u by 110%, while in (b), the calibration strategy case BA decreases E_{KG} of Q by of -13%, as compared to the Q-only calibration.

The spatial patterns of ALEXI, followed by those of MOD16A2 and CMRSET, were the least informative for the model calibration in case SP. Considering all hydrological processes and model calibration strategies, the top three best performing evaporation datasets are MERRA-2, GLEAM v3.3a and SSEBop, while the bottom three datasets are MOD16A2, SEBS and ERA5. However, it is noteworthy that they outperform the Q-only calibration when used in case SP, meaning that only their spatial patterns improve the model performance. In general, the versions 3.3 of GLEAM show a slightly higher model performance than the versions 3.2. Moreover, it is noteworthy that the reanalysis E_a products (i.e. MERRA-2 and JRA-55, with the exception of ERA5) perform better than the satellite-based products for the temporal dynamics of S_t, which can be justified by the fact that access to deep groundwater is not considered in the satellite-based products.

5.4.2 Impact of calibration strategies on spatial patterns

Different model calibration strategies result in different spatial patterns of modelled E_a and S_u, as shown for selected scenarios in Figure 5.8, and for the Q-only calibration in Figure 5.7. In general, the south to north gradient of increasing aridity observed with modelled E_a and S_u is well depicted for all the calibration strategies. However, considerable mismatches in the variability of the patterns are observed among the calibration strategies. Such discrepancies in spatial patterns have implications for water resources assessment including water accounting, flood and drought monitoring and prediction (AghaKouchak *et al.*, 2015; Klemas, 2014; Teng *et al.*, 2017; West *et al.*, 2019). Knowing when flood or drought events occur is important, but knowing the spatial extent of the event is crucial for deploying efficient adaptation and mitigation strategies (Diaz *et al.*, 2019; He *et al.*, 2020). Consequently, improving the representation of the spatial patterns of hydrological processes should be a key consideration in modelling with spatially distributed models. The comparison of the maps of modelled E_a (Figure 5.8a) with the reference ALEXI dataset (Figure 5.2) reveals that the spatial patterns of E_a in the case BA show the highest mismatch with the reference, thereby unveiling the potential pitfalls of the basin-average calibration.

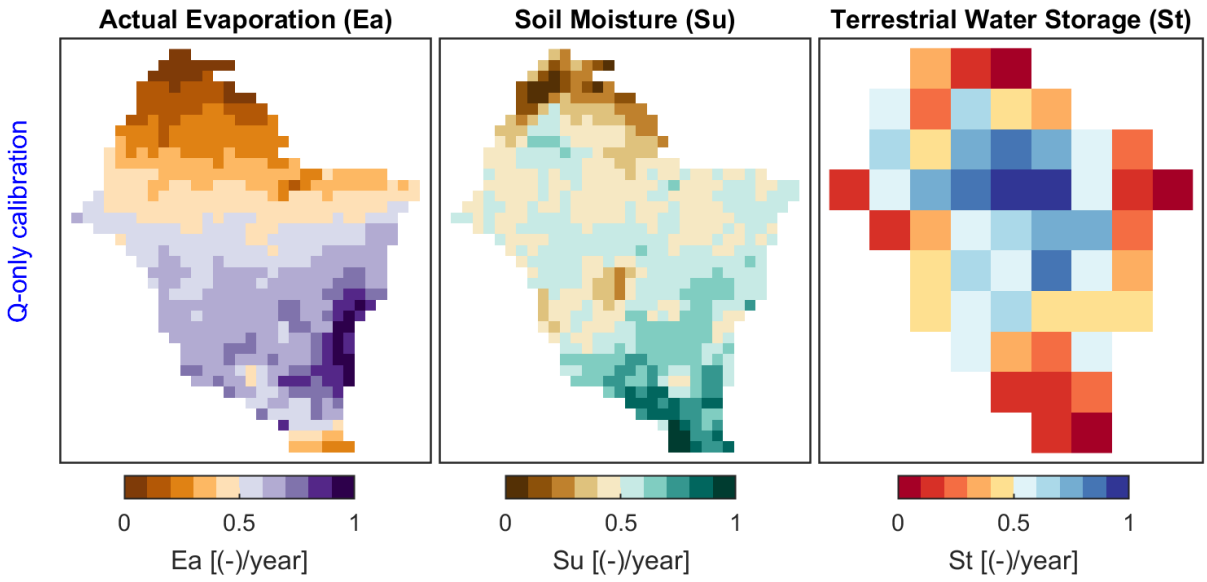


Figure 5.7. Maps of long-term (2003-2012) annual average of modelled actual evaporation (*E*a), soil moisture (*S*u) and terrestrial water storage (*S*t) using streamflow alone to calibrate the mHM model. The values are normalized by their range to better emphasize on patterns and use a unique color scale.

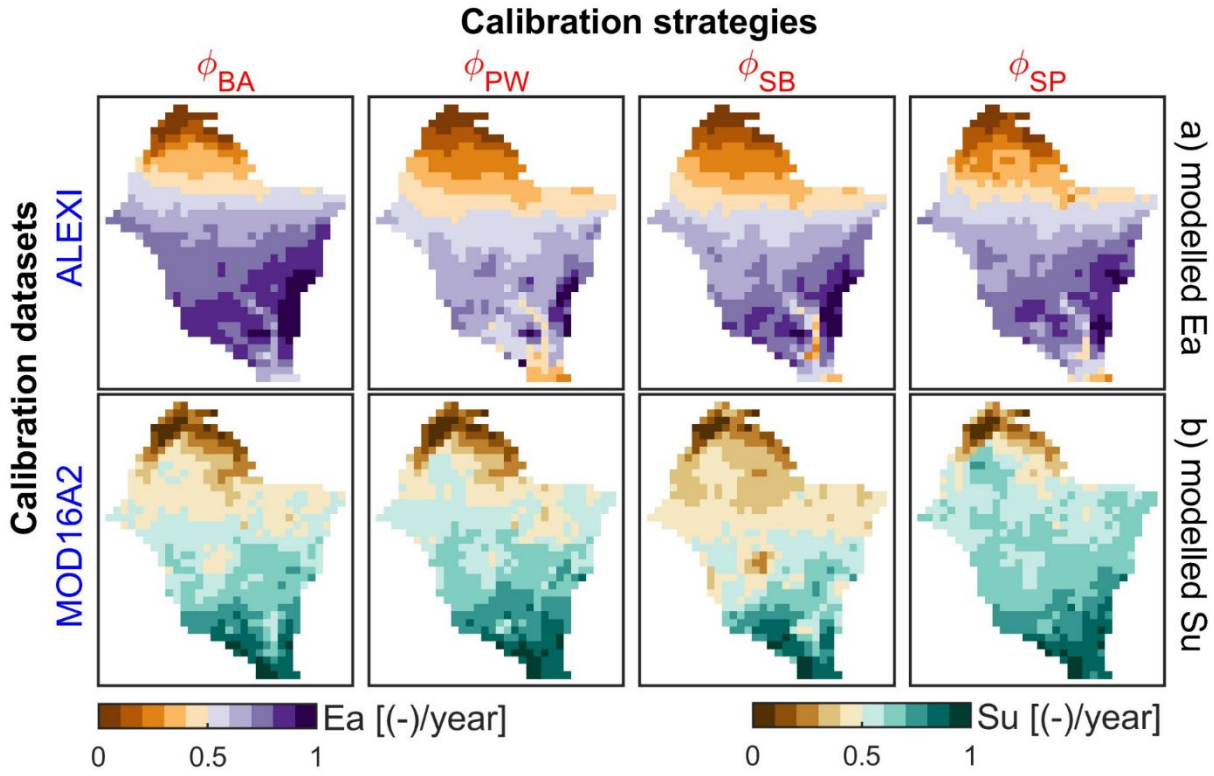


Figure 5.8. Maps of long-term (2003-2012) average of annual (a) actual evaporation (E_a) and (b) soil moisture (S_u) obtained by calibrating the mHM model with two evaporation datasets (y-axis, blue font) and different calibration strategies (x-axis, red font). The values are normalized between 0 and 1 for to emphasize the spatial patterns and use a unique colour scale.

5.4.3 Analysis per climatic zone

The analysis of the model performance according to climatic zones is done for all hydrological processes except Q because of the few gauging stations, which are unevenly distributed across the VRB (cf. Figure 2.2). The results reveal contrasting model performances across climatic zones (Figure 5.9).

In average, the model performance at predicting the temporal dynamics of S_t and the spatial patterns of S_u is higher in the intermediate climatic regions (i.e., Sudano-Sahelian and Sudanian zones) than in the driest and wettest regions (i.e., Sahelian and Guinean zones) of the VRB. In terms of the temporal dynamics of S_u, the model performance decreases slightly from the driest to the wettest regions.

In general, the multivariate calibration scenarios with evaporation datasets lead to a higher model performance for all hydrological processes in all climatic zones, as compared to the Q-only calibration. Although the GLEAM products globally perform well, they are the least effective at predicting the temporal dynamics of S_t in the Sahelian zone. The top three best

evaporation datasets for improving the average model performance across climatic zones are SSEBop, MERRA-2 and ALEXI for the temporal dynamics of S_t, MERRA-2, GLEAM v3.3a and GLEAM v3.2a for the temporal dynamics of S_u, and GLEAM v3.3b, GLEAM v3.3a and GLEAM v3.2a for the spatial patterns of S_u. Similar to results obtained at the entire VRB scale (Section 5.4.1), MOD16A2, SEBS and ERA5 still show the lowest contribution in improving the model performance across different climatic zones.

Contrary to the basin scale analysis, case SP is the least efficient calibration strategy per climatic zones. This result can be justified by the fact that the model calibration on the spatial patterns of evaporation datasets is done at the scale of the VRB. Consequently, the spatial variability of hydrological processes at large scale is not representative of the climatic zones where the patterns are more homogenous. For regions with strong spatial variability, sub-region model calibration on spatial patterns can be a way forward in overcoming the pitfalls of domain-wide calibration, thereby ultimately resulting in a higher model performance.

		Skill scores For temporal dynamics: EKG of Q, r of St and Su | For spatial patterns: ESP of Su															
		VRB				Sahelian zone			Sudano-Sahelian zone			Sudanian zone			Guinean zone		
		Temporal dynamics			Spatial patterns	Temporal dynamics		Spatial patterns	Temporal dynamics		Spatial patterns	Temporal dynamics		Spatial patterns	Temporal dynamics		Spatial patterns
		Q	St	Su	Su	St	Su	Su	St	Su	Su	St	Su	Su	St	Su	Su
Q only		0.69	0.73	0.86	-0.18	0.73	0.91	-0.36	0.72	0.88	-0.35	0.75	0.86	-0.40	0.60	0.78	-0.50
Q + Evaporation datasets	MOD16A2	0.68	0.76	0.91	-0.04	0.70	0.93	-0.37	0.75	0.92	-0.35	0.82	0.91	-0.36	0.69	0.83	-0.41
	SSEBop	0.67	0.87	0.92	0.02	0.71	0.92	-0.46	0.84	0.93	-0.33	0.92	0.92	-0.38	0.83	0.85	-0.38
	ALEXI	0.61	0.84	0.92	-0.04	0.73	0.94	-0.48	0.83	0.93	-0.34	0.89	0.92	-0.36	0.78	0.86	-0.47
	CMRSET	0.64	0.77	0.91	-0.04	0.68	0.94	-0.44	0.77	0.92	-0.35	0.83	0.92	-0.38	0.75	0.86	-0.36
	SEBS	0.56	0.77	0.89	-0.07	0.70	0.91	-0.48	0.79	0.90	-0.35	0.83	0.89	-0.42	0.72	0.84	-0.29
	GLEAM v3.2a	0.69	0.82	0.93	0.00	0.61	0.94	-0.46	0.78	0.94	-0.34	0.89	0.93	-0.35	0.80	0.86	-0.39
	GLEAM v3.3a	0.70	0.83	0.93	0.01	0.60	0.94	-0.46	0.80	0.94	-0.33	0.89	0.93	-0.37	0.80	0.86	-0.40
	GLEAM v3.2b	0.71	0.81	0.92	0.00	0.63	0.94	-0.49	0.78	0.93	-0.33	0.88	0.92	-0.35	0.78	0.85	-0.41
	GLEAM v3.3b	0.71	0.81	0.92	0.01	0.65	0.94	-0.48	0.79	0.93	-0.34	0.88	0.92	-0.34	0.77	0.85	-0.41
	ERA5	0.51	0.74	0.89	-0.09	0.72	0.92	-0.43	0.76	0.90	-0.36	0.80	0.88	-0.40	0.68	0.83	-0.30
	MERRA-2	0.68	0.87	0.93	0.01	0.66	0.95	-0.48	0.83	0.94	-0.33	0.93	0.93	-0.37	0.83	0.86	-0.45
	JRA-55	0.62	0.83	0.91	-0.01	0.67	0.93	-0.43	0.81	0.92	-0.33	0.89	0.91	-0.38	0.78	0.85	-0.38
Calibration strategies	ϕ_BA	0.60	0.81	0.91	-0.04	0.69	0.92	-0.41	0.81	0.92	-0.34	0.86	0.91	-0.38	0.74	0.84	-0.40
	ϕ_PW	0.63	0.83	0.91	-0.03	0.68	0.92	-0.39	0.82	0.93	-0.34	0.89	0.92	-0.38	0.78	0.85	-0.37
	ϕ_SB	0.68	0.81	0.93	0.00	0.66	0.94	-0.51	0.79	0.94	-0.34	0.88	0.93	-0.36	0.78	0.86	-0.38
	ϕ_SP	0.67	0.79	0.91	-0.01	0.66	0.95	-0.52	0.76	0.92	-0.34	0.86	0.91	-0.36	0.76	0.85	-0.40

Figure 5.9. Model performance for streamflow (Q), terrestrial water storage (S_t) and soil moisture (S_u) using various evaporation datasets and model calibration strategies with the mHM model. The results are shown for the four climatic zones in the Volta River Basin (VRB) over the simulation period (2003-2012). Each score for a given evaporation product represents the average over the scores obtained with 4 calibration strategies, while the score is the average over scenarios with 12 evaporation datasets for each calibration strategy. The skill scores of the temporal dynamics are obtained with the Kling-Gupta efficiency (E_{KG}) for Q and the Pearson's correlation coefficient (r) for S_t and S_u. The spatial pattern efficiency (E_{SP}) is used to assess the spatial representation of S_u. The benchmark model performance is given for the Q-only calibration as a reference. The skill scores are ranked from the best (bluest) to the worst (reddest).

5.5 Discussions

The results presented are primarily valid for the VRB because errors in evaporation estimates are known to vary according to region (Hartanto *et al.*, 2017; Sörensson and Ruscica, 2018). However, the innovative model evaluation approach proposed in this study is not location-specific and can be applied to other regions using a grid-based hydrological model. The robust evaluation approach for evaporation datasets is based on multiple hydrological processes, and responds to the recent call of the scientific community for process-oriented diagnostics of earth system models (Maloney *et al.*, 2019; Melsen *et al.*, 2016). Moreover, this study highlights the potential of satellite and reanalysis evaporation datasets to improve the representation of various hydrological processes, which might guide modellers in choosing the adequate product for their applications, and support the data developers in their quest of improving global estimates of evaporation (McCabe *et al.*, 2019). It must be noted that the overall ranking in Figure 5.9 does not systematically determine whether a dataset is good or bad, rather it shows which evaporation product provides the highest or lowest model performance for a given hydrological flux or state variable. Only the skill scores allow a judgement on the ability of a given dataset to yield a good model performance.

There might be uncertainties related to the rescaling of the evaporation datasets from their native spatial resolutions to that of the hydrological modelling resolution. However, for a fair comparison, all the datasets should be used at the same spatiotemporal resolution. The high performance of the GLEAM datasets is most likely due to the integration of soil moisture information in the calculation of actual evaporation. Therefore, GLEAM is expected to perform well for hydrological modelling. Surprisingly, MERRA-2 ranks among the best evaporation datasets notwithstanding its coarse spatial resolution. The high performance of MERRA-2 can be explained by its high temporal resolution, which might compensate its lower spatial resolution. In general, there is no clear tendency of SRS datasets to outperform the reanalysis datasets, and vice versa. Thus, besides their use as forcing data, reanalysis datasets represent a valuable source of information for the calibration of hydrological models. Satellite data of soil moisture and terrestrial water storage used for model evaluation in this study are not free of errors. However, at large scale and in poorly gauged regions, they are the only source of data that can be used for spatial model evaluation (Peters-Lidard *et al.*, 2019).

Overall, this study strives to provide solutions to some of the current challenges in hydrology (i.e., modelling methods, uncertainty and spatial variability; Blöschl et al., 2019). The proposed methodology represents an innovative way to use satellite and reanalysis datasets to improve

process representation in hydrological models (Clark et al., 2015; Peters-Lidard et al., 2017), develop hydrological water accounting (Hunink et al., 2019) and advance prediction in ungauged basins (Hrachowitz et al., 2013; Zhang et al., 2019).

Future studies should test the advantages of a multi-scale calibration framework that accounts for both domain-wide and sub-domain (i.e. climatic regions or sub-basins) spatial heterogeneity, which might lead to better prediction of spatial patterns across climatic zones at large scale. Calibration strategies with multiple non-commensurable variables as well as spatial patterns is a way forward in advancing process representation in hydrological models (Dembélé *et al.*, 2020b; Nijzink *et al.*, 2018; Zink *et al.*, 2018). More importantly, the choice of the calibration strategy or the objective function is determinant for high model performance, mainly under a changing environment (Fowler *et al.*, 2018b; Schaefli *et al.*, 2010). Uncertainties in the structure of the mHM model might influence the modelled hydrological processes. Further work should investigate the impact of different model structures on the methodology proposed in this study, and assess the parameter sensitivity of mHM depending on calibration strategies and variables, which is beyond the scope of the current study. As the accuracy of the satellite and reanalysis datasets depends on the quality of the input meteorological datasets used for their production, it is also important to assess the impact of precipitation datasets on evaporation modelling (Mao and Wang, 2017; Or and Lehmann, 2019).

Finally, an ensemble product that merges different evaporation datasets is a potential way forward in reducing regional uncertainties and thereby improving global estimates (da Motta Paca *et al.*, 2019; Jiménez *et al.*, 2018). Such advances in evaporation product development can facilitate prediction in ungauged basins using earth observations.

5.6 Conclusion

Four model calibration strategies are used to evaluate twelve satellite and reanalysis datasets in the large transboundary Volta River basin located in West Africa. The experiment is done with the mHM model over the period 2003-2012. The key findings can be summarized as follows:

- Satellite and reanalysis datasets can improve the predictive skill of the hydrological model if the appropriate calibration strategy is used.
- Overall, MERRA-2, GLEAM v3.3a and SSEBop individually provide the highest contribution in improving the model performance.

- Model calibration on the full extent of the gridded evaporation datasets result in a higher model performance than calibration on basin-average estimates.
- Using only the spatial patterns of gridded evaporation data for model calibration, and not their absolute values, yields higher model performance than classical approaches based on basin average evaporation signal or based only on streamflow.
- Contrasting spatial patterns of soil moisture are obtained depending on the modelling scenarios, with differences in the model performances according to climatic zones.

These findings contribute to solving current challenges related to large-scale hydrological modelling and provide avenues for improving process representation with the use of increasingly available satellite and reanalysis datasets. Moreover, the results provide insights to the developers of the evaporation datasets and might serve of guidance for future developments. However, a replication of the proposed methodology to evaluate evaporation datasets should be applied in other regions with different hydro-climatic conditions, and with different hydrological and land surface models. Future work should also investigate the possibility of prediction in ungauged basins solely from earth observation datasets, which are increasingly and readily becoming available.

Chapter 6

Improving Spatial Patterns in Multivariate Distributed Hydrological Modelling*

Nature uses only the longest threads to weave her patterns,
so that each small piece of her fabric reveals
the organization of the entire tapestry.
Richard P. Feynman

Continuous improvement is better
than delayed perfection.
Mark Twain

* This chapter is based on the following publication:
Dembélé, M., M. Hrachowitz, H. H. G. Savenije, G. Mariéthoz, and B. Schaefli (2020), Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, Water Resources Research, 56(1), https://doi.org/10.1029/2019wr026085

Abstract

Hydrological model calibration combining Earth observations and in situ measurements is a promising solution to overcome the limitations of the traditional streamflow-only calibration. However, combining multiple data sources in model calibration requires a meaningful integration of the data sets, which should harness their most reliable contents to avoid accumulation of their uncertainties and mislead the parameter estimation procedure. This study analyses the improvement of model parameter selection by using only the spatial patterns of satellite remote sensing data, thereby ignoring their absolute values. Although satellite products are characterized by uncertainties, their most reliable key feature is the representation of spatial patterns, which is a unique and relevant source of information for distributed hydrological models. This study proposes a novel multivariate calibration framework exploiting spatial patterns and simultaneously incorporating streamflow and three satellite products (i.e., Global Land Evaporation Amsterdam Model [GLEAM] evaporation, European Space Agency Climate Change Initiative [ESA CCI] soil moisture, and Gravity Recovery and Climate Experiment [GRACE] terrestrial water storage). The Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature data set is used for model evaluation. A bias-insensitive and multicomponent spatial pattern matching metric is developed to formulate a multiobjective function. The proposed multivariate calibration framework is tested with the mesoscale Hydrologic Model (mHM) and applied to the poorly gauged Volta River basin located in a predominantly semiarid climate in West Africa. Results of the multivariate calibration show that the decrease in performance for streamflow (-7%) and terrestrial water storage (-6%) is counterbalanced with an increase in performance for soil moisture (+105%) and evaporation (+26%). These results demonstrate that there are benefits in using satellite data sets, when suitably integrated in a robust model parametrization scheme.

6.1 Introduction

One of the key challenges in hydrological modelling (Beven, 2019b; Singh, 2018) is the reliable representation of the spatiotemporal variability of natural processes, to which the footprint of human activity is often superimposed. In most places, available in-situ observations are not sufficient to capture the spatiotemporal heterogeneity of dominant hydrological processes (AghaKouchak *et al.*, 2015; Hrachowitz and Clark, 2017). With the upswing in development of distributed hydrological models that offer spatially explicit predictions as an essential tool for decision-making (Fatichi *et al.*, 2016; Paniconi and Putti, 2015; Semenova and Beven, 2015), there is a growing interest in the plausibility of their spatial patterns (Ko *et al.*, 2019; Koch *et al.*, 2018; Wealands *et al.*, 2005).

Most commonly, hydrological models are calibrated using streamflow data alone (Becker *et al.*, 2019; Yassin *et al.*, 2017). The streamflow signal represents an integrated response of the hydrological system to a set of natural drivers (e.g. climate and landscape) and anthropogenic influences (e.g. deforestation and reservoirs) occurring upstream of the measurement's location. Although streamflow is key to understanding the temporal dynamics of a system, it does not disclose much information on the system-internal spatial heterogeneity of the hydrological processes (McDonnell *et al.*, 2007; Rajib *et al.*, 2018a). It therefore has little discriminatory power to constrain the feasible parameter space of a distributed model, i.e. the boundary flux or closure problem (Beven, 2006a). Consequently, a spatially distributed hydrological model (DHM) calibrated only on streamflow is very unlikely able to reproduce a reliable spatiotemporal representation of other hydrological fluxes and states (Clark *et al.*, 2016; Minville *et al.*, 2014), even if a multiscale parameter regionalization scheme is used (Rakovec *et al.*, 2016a). Mismatches between temporal and spatial patterns should therefore be expected when comparing hydrological models outputs to other distributed observational datasets (Vereecken *et al.*, 2008; Xu *et al.*, 2014).

For a few decades, satellite remote sensing (SRS) has opened up new avenues for the development of spatial hydrology (Cui *et al.*, 2018; Lettenmaier *et al.*, 2015; Pasetto *et al.*, 2018). The increasing and unprecedented availability of SRS data at increasingly finer spatial and temporal resolutions has triggered the development of large domain water management applications including flood and drought monitoring (Revilla-Romero *et al.*, 2015; Senay *et al.*, 2015; Teng *et al.*, 2017; Wu *et al.*, 2014). The use of SRS data in water resources monitoring is promising and it has led to an increasing number of studies on a variety of topics in hydrology, including precipitation, evaporation and soil moisture estimation (Cazenave *et al.*, 2016; Chen

and Wang, 2018; NASEM, 2019; Schultz and Engman, 2012). SRS data complement in-situ hydro-meteorological data (Balsamo *et al.*, 2018), which are typically scarce and whose unavailability hinders the understanding of environmental systems (Tang *et al.*, 2009). This aspect is particularly relevant for developing countries where research for development initiatives have been increasing in the recent years (Montanari *et al.*, 2015).

Besides direct use of SRS data for water resources monitoring and management (Sheffield *et al.*, 2018), an increasing body of literature addresses the question of how these datasets can be used to improve hydrological modelling (Baroni *et al.*, 2019; Liu *et al.*, 2012a; Parajka *et al.*, 2009). The scientific community has, in fact, long been advocating the use of spatial data for DHM evaluation (Beven and Feyen, 2002; Grayson and Bloschl, 2001). SRS datasets have the potential to improve models either via data assimilation (Leroux *et al.*, 2016; Tian *et al.*, 2017) or via calibration (Li *et al.*, 2018; Rientjes *et al.*, 2013). In this context, data assimilation is used to update the states of a given model, e.g. to compensate for model structural deficiencies (Spaaks and Bouten, 2013). For parameter estimation (i.e. model calibration) with SRS data, the existing approaches consist in using SRS data alone or in combination with in-situ data, usually streamflow data (Immerzeel and Droogers, 2008; Wambura *et al.*, 2018). Calibration of hydrological models without concomitant streamflow data remains challenging and attempts to do so have only shown limited success (Sutanudjaja *et al.*, 2014; Wanders *et al.*, 2014).

The simultaneous calibration of hydrological models with streamflow and different combinations of complementary data from SRS is increasingly discussed in recent literature (Stisen *et al.*, 2018). Multivariate (i.e. multiple variables) parameter estimation (Efstratiadis and Koutsoyiannis, 2010) can substantially reduce the feasible model and parameter space and lead to more realistic internal model dynamics and related hydrological signatures (Clark *et al.*, 2017; Shafii and Tolson, 2015), which can ultimately enhance the overall representation of catchment functioning. Furthermore and intimately linked to the above, multivariate calibration strategies can considerably reduce equifinality (i.e. non-identifiable model parameters in inverse modelling approaches; Beven, 2006b; Savenije, 2001) and reduce prediction uncertainty (Fenicia *et al.*, 2008; Fovet *et al.*, 2015; Gupta *et al.*, 2008; Schoups *et al.*, 2005). However, important open questions remain with respect to the combination of SRS data with streamflow data for model parameter estimation. While some studies observed a significant improvement in the representation of model outputs after SRS data incorporation (Chen *et al.*, 2017; Leroux *et al.*, 2016; Werth *et al.*, 2009), others found minor changes or even major deteriorations (Demirel *et al.*, 2018; Tangdamrongsub *et al.*, 2017). Such apparently

contradictory conclusions are case-study specific and need to be understood as resulting from model structures, model parametrizations, and trade-offs between improving water balance estimates and related streamflow dynamics and better representing other hydrological fluxes and states (Euser *et al.*, 2013). More generally, the key challenge results from the integration of several data sources (SRS or in-situ) in parameter estimation, which can be attributed to conflicting information from different types of SRS data. Nonetheless, multivariate parameter estimation with SRS data remains promising, especially when streamflow data availability is limited or the data quality is questionable.

Although SRS data are more accessible with higher spatiotemporal resolution compared to in-situ observations, they are generally not direct measurements of hydrological processes, which adds a level of uncertainty to any SRS based parameter estimation study (Ehlers *et al.*, 2018; Knoche *et al.*, 2014; Ma *et al.*, 2018). However, they provide spatial information on hydrological processes, which makes them a unique and relevant information source for spatially distributed representations of the system in models (Stisen *et al.*, 2018). For instance, many studies report different model performances when using different satellite-based products as input (e.g. precipiation; Dembélé *et al.*, 2020c; Poméon *et al.*, 2018; Thiemig *et al.*, 2013) or as calibration variables (e.g. evaporation, soil moisture, terrestrial water storage; Bai *et al.*, 2018b; Nijzink *et al.*, 2018). Nevertheless, for a given region, different products can give considerably different absolute values of a specific variable while they may exhibit plausible and similar spatial patterns (Beck *et al.*, 2017b; Dembélé and Zwart, 2016). Additionally, retaining only the spatial pattern information of SRS data can substantially mitigate the uncertainty resulting from the fact that they are not direct observations, as long as their relative values are used rather than their absolute values (Dembélé *et al.*, 2020a; Mendiguren *et al.*, 2017).

In the context of using SRS data for DHM calibration, the simultaneous use of more than one SRS product to constrain several hydrological state or flux variables is uncommon (Lopez *et al.*, 2017), as is the incorporation of spatial pattern in the calibration scheme using bias-insensitive metrics (Demirel *et al.*, 2018; Zink *et al.*, 2018). Using different variables from SRS products simultaneously in parameter estimation is in general not straightforward because they all have limitations (e.g. spatiotemporal resolutions and accuracy), which can lead to significant trade-offs in multivariate calibration (Rajib *et al.*, 2018b; Silvestro *et al.*, 2015).

In light of the above, this study proposes to test a novel multivariate calibration strategy in which a DHM will be trained to simultaneously reproduce spatial pattern, i.e. relative spatial

differences, of three variables from different SRS products describing different components of the hydrological system (i.e. evaporation, soil moisture and terrestrial water storage), as well as in-situ observations of streamflow. The proposed calibration framework combines simultaneously four non-commensurable variables and a new bias-insensitive metric for spatial pattern representation, which as a whole is different from previous studies (e.g. Demirel *et al.*, 2018; Koppa *et al.*, 2019; Nijzink *et al.*, 2018; Rakovec *et al.*, 2016a; Zink *et al.*, 2018) and therefore makes the novelty of this study. The following research hypotheses are tested:

- Simultaneously calibrating a DHM on four non-commensurable variables and spatial patterns of satellite data considerably improves the predictive skill of the model, even for a DHM integrating a multiscale parameter regionalization scheme;
- The newly proposed bias-insensitive metric based on pixel-by-pixel locational matching can be used to improve the calibration of a DHM on observed spatial patterns of hydrological processes.

The overall goal of this study is to improve the spatial representation of dominant hydrological processes of a DHM without significantly deteriorating the streamflow signal, and reproducing plausible dynamics of the hydrological system using spatial pattern information from SRS datasets. Such improvement will be an asset for spatial hydrology and large domain water management applications (e.g. water accounting, drought monitoring and flood prediction), and might subsequently lead to advances in prediction in ungauged basins (Blöschl *et al.*, 2013; Hrachowitz *et al.*, 2013; Sivapalan, 2003) with the use of readily accessible SRS data (Butler, 2014; Wulder and Coops, 2014). This work embraces the fourth paradigm for hydrology (i.e. data-intensive science, Peters-Lidard *et al.*, 2017), and contributes to solving some of the issues (e.g. spatial variability and modelling methods) recently identified as the twenty-three unsolved problems in hydrology (UPH) in the twenty-first century (Blöschl *et al.*, 2019). The proposed multivariate calibration framework is tested with the mesoscale Hydrologic Model (mHM) (cf. Section 4.5, Chapter 4), with a case study in the poorly gauged Volta River basin in West Africa.

6.2 Distributed hydrological modelling

The setup of the mHM hydrological model is described in Section 4.5 of Chapter 4. In-situ streamflow data and complementary data from SRS are used to calibrate and to evaluate the model performance. A description of the modelling datasets with their characteristics and their

sources is given in Table 4.2 (Section 4.4, Chapter 4). The streamflow data was quality checked and missing data portions were gap-filled (Dembélé *et al.*, 2019).

Concerning the SRS products, the terrestrial water storage (S_t) anomaly data derived from changes in surface mass, which is related to the Earth's gravity field, is obtained from GRACE (Landerer and Swenson, 2012; Tapley *et al.*, 2004). Over land, S_t is the sum of snow, ice, surface water, soil moisture and groundwater. The data release RL05 (Swenson, 2012) is used in this study. It is a simple arithmetic mean of different solutions from three processing centers: Jet Propulsion Laboratory (JPL), Center for Space Research at University of Texas (CSR), and Geoforschungs Zentrum Potsdam (GFZ). Sakumura *et al.* (2014) found this ensemble mean product more effective in reducing noise in the Earth's gravity signal compared to the individual products. As the original baseline for GRACE-derived S_t anomaly data is the period 2004-2009, the S_t data is converted to a new baseline corresponding to the modelling period (2003-2012) used in this study, by averaging each grid point over the new baseline and subtracting that value from all time steps (NASA, 2019).

The surface soil moisture (S_u) data for a soil layer depth of 2-5 cm is obtained from ESA CCI (Dorigo *et al.*, 2017). The combined product used in this study is a blended product of both active and passive microwave products derived from scatterometer (ERS AMI and ASCAT) and radiometer (SMMR, SSM/I, TMI, AMSR-E, WindSat, AMSR2, and SMOS) retrievals (Liu *et al.*, 2012b; Wagner *et al.*, 2012b). The merging algorithm of the combined product version 4.2 is described by Gruber *et al.* (2017).

Actual evaporation (E_a) data is obtained from the GLEAM land surface model that uses satellite data as input (Martens *et al.*, 2017; Miralles *et al.*, 2011). It separately estimates the components of terrestrial evaporation (i.e. transpiration, bare soil evaporation, open-water evaporation, interception loss, and sublimation) based on the fraction of land cover types (i.e. bare soil, low vegetation, tall vegetation and open water) before aggregating them for each grid cell. In GLEAM, potential evaporation (E_p) is calculated based on the Priestley-Taylor (1972) equation and thereafter converted into transpiration or bare soil evaporation using a stress factor, which is a parameter that accounts for environmental conditions limiting evaporation. The stress factor is estimated from microwave vegetation optical depth (i.e. water content in vegetation) and root-zone soil moisture that is calculated with a multi-layer water balance algorithm. The fraction of open-water evaporation is assumed to equal E_p. The Gash (1979) analytical model further refined by Valente *et al.* (1997) is used to calculate rainfall interception by forests.

Land surface temperature (T_s) data from the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument of the NASA satellites is used as an independent data for model evaluation, and it is not used during model calibration. The daytime product from the Aqua platform is used because that satellite passes over our study region around 13:45 in local time, corresponding to the highest daily temperature period with a clear-sky coverage (Wan *et al.*, 2015).

6.3 Model calibration and evaluation strategies

The modelling period spans from 2000 to 2012 and consists of 3 years (2000-2002) model warm-up period, 6 years (2003-2008) calibration period, and 4 years (2009-2012) evaluation period. Based on data availability and quality in the VRB (Dembélé *et al.*, 2019), 11 streamflow gauging stations are chosen to have a good coverage of the river network (Figure 2.2), and the calibration is done on them simultaneously to obtain a single parameter set for the whole VRB. The domain-wide calibration, which was proven to give similar performance as the domain-split calibration (Mizukami *et al.*, 2017), is preferred here because of the limited number of streamflow stations and for seamless spatial pattern representation across sub-basins (Figure 2.4).

Two main calibration approaches are adopted to evaluate the benefit of including spatial patterns in multivariate parameter estimation with SRS data. The first approach is the streamflow-only calibration, and the second approach uses multiple SRS datasets in addition to streamflow. In both cases, the formulation of the objective functions follows the Euclidian distance approach in which all elements are equally weighted (Khu and Madsen, 2005), cf. Eq. 5.3 (Section 5.3.2, Chapter 5).

6.3.1 Calibration on streamflow – benchmark

The first calibration approach is the benchmark calibration case (case *Q*) where the hydrological model is constrained with in-situ streamflow (*Q*) data only. The objective function Φ_Q combines the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) of streamflow (E_{NS}) and the Nash-Sutcliffe efficiency of the logarithm of streamflow (E_{NSlog}), as in Eq. 4.4 (Section 4.6, Chapter 4).

6.3.2 Calibration on multiple variables with spatial patterns

6.3.2.1 Multivariate calibration strategies

In contrast to the first calibration strategy, which only considers Q as target variable (case Q), the second calibration strategy involves multiple variables (case MV). The potential improvement of the modelled fluxes and states is estimated by constraining the parameter estimation with a simultaneous combination of three variables from SRS products (S_t, S_u and E_a), in addition to Q. Here, the spatial patterns of E_a and S_u are used in the multivariate calibration, while the temporal dynamics of S_t are averaged over the whole basin due to the relatively coarse spatial resolution of the GRACE data. The temporal dynamics of S_t per grid cell is also assessed during model evaluation.

The degree of reproduction of the spatial patterns of E_a and S_u is quantified with a new spatial pattern matching metric, denoted E_{SP} (cf. Eq. 4.8 in Section 4.7, Chapter 4), developed by Dembélé *et al.* (2020b). The multivariate objective function Φ_{MV} (Eq. 6.1) is defined as follows:

$$\Phi_{\mathrm{MV}} = \sqrt{{\Phi_{\mathrm{Q}}}^2 + {\Phi_{\mathrm{S_t}}}^2 + {\Phi_{\mathrm{S_u}}}^2 + {\Phi_{\mathrm{E_a}}}^2}, \text{ with} \tag{6.1}$$

$$\Phi_{\mathrm{E_a}} = 1 - \frac{1}{t}\sum_1^t E_{\mathrm{SP}}\left(E_{\mathrm{a,mod}}(t), E_{\mathrm{a,obs}}(t)\right), \tag{6.2}$$

$$\Phi_{\mathrm{S_u}} = 1 - \frac{1}{t}\sum_1^t E_{\mathrm{SP}}\left(S_{\mathrm{u,mod}}(t), S_{\mathrm{u,obs}}(t)\right), \text{ and} \tag{6.3}$$

$$\Phi_{\mathrm{S_t}} = E_{\mathrm{RMS}}\left(Z_{\mathrm{S_{t,mod}}}(t), Z_{\mathrm{S_{t,obs}}}(t)\right) \tag{6.4}$$

where t is the number of time steps of the calibration period. The sub-objective functions $\Phi_{\mathrm{E_a}}$ and $\Phi_{\mathrm{S_u}}$ are based on the E_{SP} (Eq. 4.8) of modelled and observed E_a and S_u, while $\Phi_{\mathrm{S_t}}$ denotes the root mean squared error (E_{RMS}) of the z-scores of the modelled and observed basin-averaged S_t anomaly.

Consequently, Φ_{Q} ensures a reliable prediction of streamflow signatures (i.e. high and low flows), $\Phi_{\mathrm{E_a}}$ and $\Phi_{\mathrm{S_u}}$ serve to improve the spatial patterns of the modelled E_a and S_u, while $\Phi_{\mathrm{S_t}}$ constrains the temporal dynamics of the modelled S_t, which should contribute to a better prediction of the water balance at monthly and annual scales. In fact, $\Phi_{\mathrm{E_a}}$ and $\Phi_{\mathrm{S_u}}$ are calculated such that the spatial pattern efficiencies of E_a and S_u are determined over the grid cells at each monthly time step, before averaging them over the calibration period, while $\Phi_{\mathrm{S_t}}$ is calculated for the basin-averaged S_t over the calibration period. Note that in the E_{RMS} metric (Eq. 4.13), n denotes the number of grid cells in the spatial domain when calculating $\Phi_{\mathrm{E_a}}$ and $\Phi_{\mathrm{S_u}}$ (i.e. α in E_{SP}), while it corresponds to the length of the calibration period (i.e. $n = t$) in the calculation of

Φ_{S_t}. All constituents of Φ_{MV} (i.e. $\Phi_Q, \Phi_{E_a}, \Phi_{S_u}, \Phi_{S_t}$) vary in the same range from 0 to positive infinity. Therefore, Φ_{MV} has the same range of values with an optimal value of zero. The Dynamically Dimensioned Search (DDS) algorithm (Tolson and Shoemaker, 2007) is used for parameter estimation using 3,000 iterations. Daily streamflow data are used for model calibration while monthly SRS data are preferred to avoid errors due to potential time lags among satellite sensor measurements and model simulations.

6.3.2.2 Contribution of individual variables to multivariate calibration

Assessing the individual contribution of variables used in multivariate calibration is rarely done, whereas it can help quantify trade-offs in modelling flux and states variables (Koppa *et al.*, 2019). Here, the contribution of each SRS dataset to the multivariate calibration case is investigated with a leave-one-out approach. The procedure consists in removing one SRS data type from the calibration case MV and evaluating the predictive skill of the model. In addition, multivariate calibration without streamflow data, and thus exclusively based on SRS data, is tested to determine the potential of SRS data for hydrological model calibration in regions where little or no streamflow data are available. Consequently, four additional objective functions (Table 6.1) are used for multivariate calibration cases without E_a (case MV-E_a), without S_u (case MV-S_u), without S_t (case MV-S_t) and without streamflow (case MV-Q).

Table 6.1. Variants of multivariate calibration cases in the leave-one-out approach

Calibration cases	Calibration variables	Objective functions	Specificities	
case MV-E_a	Q, S_t, S_u	$\Phi_{MV-E_a} = \sqrt{\Phi_Q{}^2 + \Phi_{S_t}{}^2 + \Phi_{S_u}{}^2}$	No direct constraint on evaporation	(6.5)
case MV-S_u	Q, S_t, E_a	$\Phi_{MV-S_u} = \sqrt{\Phi_Q{}^2 + \Phi_{S_t}{}^2 + \Phi_{E_a}{}^2}$	No specific constraint on surface soil moisture	(6.6)
case MV-S_t	Q, S_u, E_a	$\Phi_{MV-S_t} = \sqrt{\Phi_Q{}^2 + \Phi_{S_u}{}^2 + \Phi_{E_a}{}^2}$	No direct constraint on deep subsurface processes	(6.7)
case MV-Q	S_t, S_u, E_a	$\Phi_{MV-Q} = \sqrt{\Phi_{S_t}{}^2 + \Phi_{S_u}{}^2 + \Phi_{E_a}{}^2}$	Only satellite-based variables with no direct constraint on streamflow	(6.8)

6.4 Post-calibration model evaluation

The predictive skill of the model is evaluated by assessing the transferability of the global parameters across temporal periods and spatial scales obtained by the abovementioned calibration strategies. First, the temporal transferability is evaluated following a split-sample test that consists in assessing performances for a period that is different from the calibration period (Klemes, 1986). Secondly, spatial scale transferability is evaluated by using different grid cell (i.e. pixel) sizes as modelling resolution (Kumar *et al.*, 2013; Samaniego *et al.*, 2010). The global parameters of the model for all calibration cases are obtained for a resolution of 0.25° (~28 km, i.e. 619 pixels in the basin) and the same parameters are used to run the model without recalibration at four different finer scales: 0.125° (~14 km, i.e. 2,320 pixels), 0.0625° (~7 km, i.e. 8,974 pixels), 0.03125° (~3.5 km, i.e. 35,231 pixels) and 0.015625° (~1.75 km, i.e. 139,494 pixels). The evaluation data for model parameter transferability are streamflow for streamflow, using the Kling-Gupta efficiency (E_{KG}), fine scale T_s data to evaluate E_a and S_u, using r_s, while no high-resolution data is available for S_t evaluation. T_s is used has proxy data for the evaluation of S_u and E_a because past studies found significant negative correlation between T_s and S_u (Kumar *et al.*, 2013; Lakshmi *et al.*, 2003; Wang *et al.*, 2007), and a control of T_s over E_a (Boni *et al.*, 2001; Lakshmi, 2000).

Following Biondi *et al.* (2012), supplemental skill metrics different from those used in model calibration are computed for a thorough model evaluation because every metric has its own limitations (Fowler *et al.*, 2018b; Knoben *et al.*, 2019; Santos *et al.*, 2018; Schaefli and Gupta, 2007). In addition to E_{NS}, E_{NSlog}, E_{SP}, E_{RMS}, and r_s, the E_{KG} is reported for model evaluation.

6.5 Results and Discussions

The following section presents and discusses the results of model performances for different variables used in the calibration procedure. The results refer to the evaluation period when analysing the results if not clearly specified. However, both calibration and evaluation results are presented in figures. Hereafter, the SRS datasets are called reference data, as they are not direct observations. Additional results on model performances for each of the four climatic zones in the VRB are provided in the supplementary information of Dembélé *et al.* (2020b).

6.5.1 Model performance for streamflow

The model performance for streamflow (Q) at 11 gauging stations is given in Figure 6.1. For the calibration period, the mean E_{KG} is 0.67 (E_{NS} = 0.71, E_{NSlog} = 0.72) for the model calibration with only Q data (i.e. case Q). The performance of Q in the calibration period decreases when multiple variables are used to constrain the parameters search. The mean E_{KG} is 0.55 (E_{NS} = 0.57, E_{NSlog} = 0.66) for the multivariate calibration (i.e. case MV), corresponding to a decrease of 18% compared to case Q.

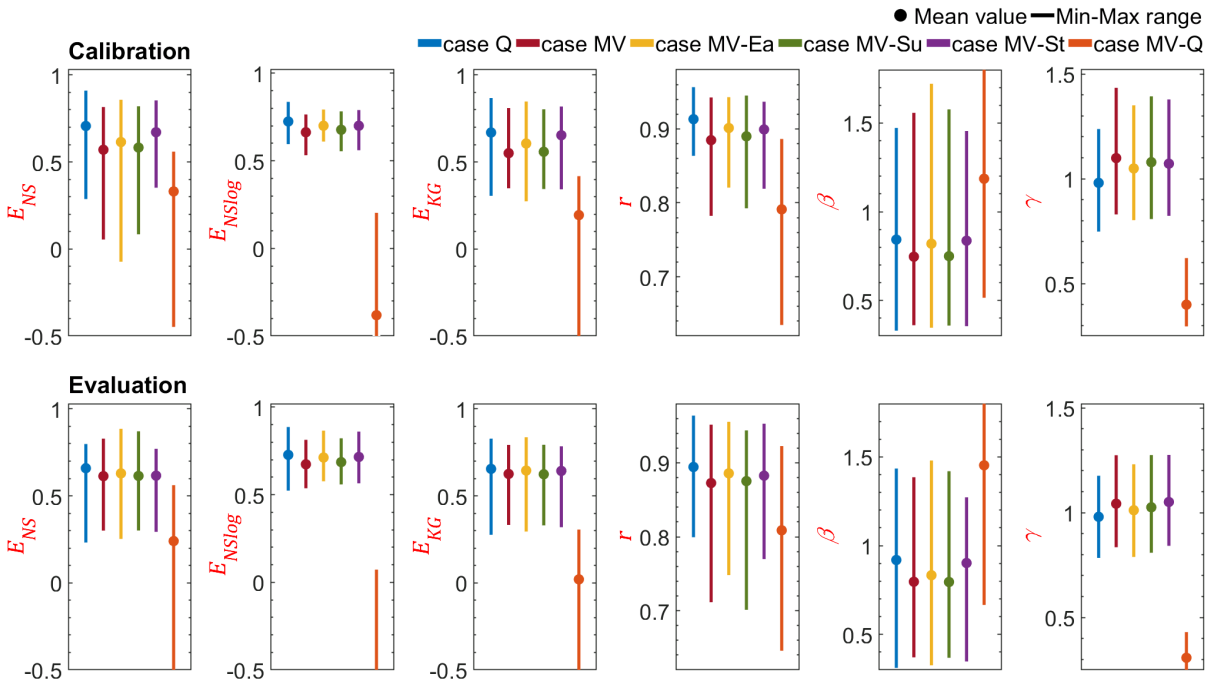


Figure 6.1. Statistics of model performance for streamflow. The best score of all the metrics is 1. The dots give the mean score and the bars represent the min-max range of values for 11 streamflow gauges. The colours correspond to the model calibration cases.

Regarding the other multivariate calibration cases, the best performance with respect to Q is obtained in case MV-S_t with a mean E_{KG} of 0.65 (E_{NS} = 0.67, E_{NSlog} = 0.70) that represents a slight decrease of 3% compared to case Q, while the weakest performance is given by case MV-Q with a mean E_{KG} of 0.19 (E_{NS} = 0.33, E_{NSlog} = -0.38). Differences in measurements scales between Q data and satellite products (i.e. river dimensions vs. pixel size) can justify a low performance for case MV-Q (see section 6.5.5). In general, all calibration cases give a good timing of Q with a mean correlation coefficient of $r > 0.79$, but they underestimate it with a mean bias of $\beta < 0.85$, except case MV-Q that shows overestimation with a positive bias (β = 1.19). They all show a higher variability of Q than the observed data, with mean $\gamma > 1.04$,

except case Q (γ = 0.98) and case MV-Q (γ = 0.40) that produce a lower variability. A subset of the hydrographs of three stations from different climatic zones are depicted in Figure 6.2.

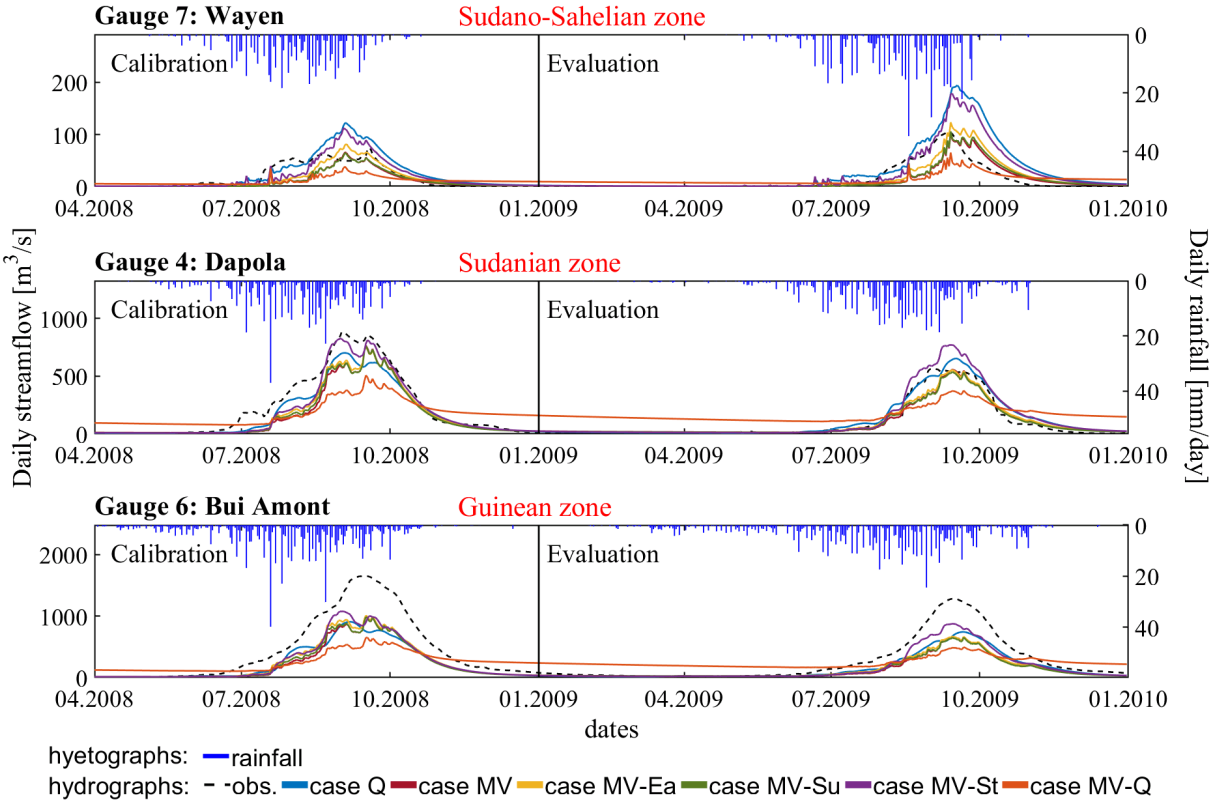


Figure 6.2. Hydrographs at selected stations in different climatic zones for all model calibration cases. Only a subset of the simulation period (2003-2012) is shown for visualization.

During the evaluation period, as compared to the calibration period, the model performance for the mean E_{KG} decreases by 2% for case Q (from 0.67 to 0.66) and case MV-S_t (from 0.65 to 0.64) and 90% for case MV-Q (from 0.19 to 0.02), while it increases by 7% for case MV-E_a (from 0.61 to 0.65), 12% for case MV-S_u (from 0.56 to 0.62) and 14% for case MV(from 0.55 to 0.63). Considering the mean E_{KG}, case MV performs less well than case Q by 11% on average, which means 18% less during the calibration and 4% less during the evaluation period. The deterioration of streamflow performance in a multivariate calibration setting is also reported in previous studies (Bai *et al.*, 2018b; Livneh and Lettenmaier, 2012; Poméon *et al.*, 2018; Rakovec *et al.*, 2016a). However, this is largely an artefact of Type I error (i.e. falsely accepting poor models, Beven, 2010) induced by the Q-only calibration, resulting in inconsistency in the representation of processes (Gupta *et al.*, 2012; Hrachowitz *et al.*, 2014). In addition, the performance of Q slightly increases when E_a (+7%) or S_t (+11%) are left out of the multivariate calibration with case MV during the evaluation period. Therefore, as shown in Figure 6.1, the combinations Q+S_t+S_u (i.e. case MV-E_a) and Q+S_u+E_a (i.e. case MV-S_t) are the

best for streamflow prediction, while Q+S_t+E_a (i.e. case MV-S_u) performs similar to Q+S_t+S_u+E_a (i.e. case MV).

6.5.2 Model performance for terrestrial water storage

The statistics for the monthly terrestrial water storage (S_t) anomalies are given in Figure 6.3. A similar trend in skill scores (i.e. E_{RMS} and r) among all calibration cases is observed in the calibration and evaluation periods with weaker scores during evaluation.

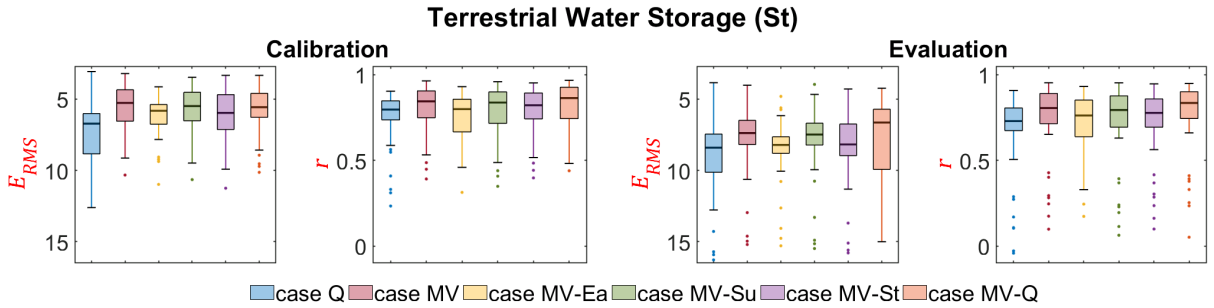


Figure 6.3. Statistics of model performance for terrestrial water storage (St). The y-axis is reversed for ERMS. The number of elements per boxplot (n = 52) corresponds to the number of gird cells for GRACE data in the study area. The colours correspond to the model calibration cases.

The evaluation period is characterized by a substantial improvement from model case Q (median E_{RMS} = 8.41 cm, r = 0.73) to case MV (E_{RMS} = 7.38 cm, r = 0.81). In general, all multivariate calibration cases reproduce the variability in S_t better than case Q. Previous studies also reported improvement of S_t prediction in multivariate settings (Chen *et al.*, 2017; Livneh and Lettenmaier, 2012; Werth *et al.*, 2009). The lowest performance increase is observed when E_a (E_{RMS} = 8.21 cm, r = 0.76) or S_t (E_{RMS} = 8.18 cm, r = 0.78) are removed from the multivariate setting. The best prediction is obtained with case MV-Q, yielding median E_{RMS} of 6.65 cm and r of 0.84.

Figure 6.4a shows the climatology of the basin-averaged S_t for all models. The temporal dynamics of the normalized GRACE-derived S_t is well reproduced by all models (r > 0.89) with different degrees of underestimation from September to March, which is a period with little or no rainfall, and slight overestimation from April to June, which is the beginning of the rainy season. All models fit well the period July-August, which is the wettest period of the rainy season.

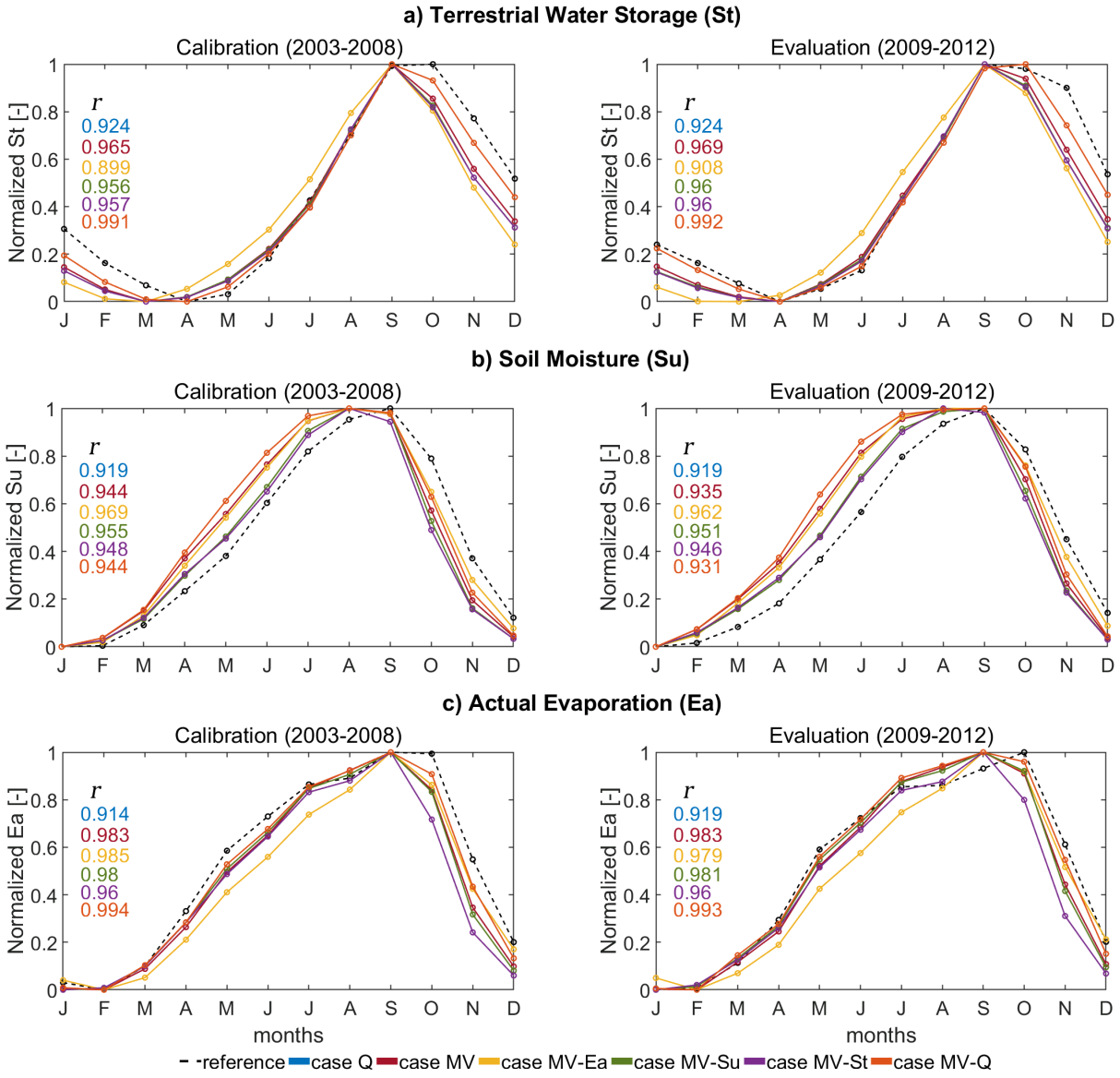


Figure 6.4. Climatology of terrestrial water storage (a), soil moisture (b) and actual evaporation (c) with the Pearson correlation coefficient (r) indicating the performance of all model calibration cases.

In general, case MV-E_a shows the highest deviation from the satellite signal ($r = 0.90$) followed by case Q ($r = 0.92$). Removing spatial patterns of E_a from the calibration leads to S_t overestimation during the rainy season and an underestimation during the dry season. The same trend can be observed when only Q is used for model calibration (i.e. case Q). The S_t simulation improves in the multivariate calibration including Q (i.e. case MV, $r = 0.97$), but the best match is obtained when Q is left out (i.e. case MV-Q, $r = 0.99$). When GRACE-derived S_t is excluded from the parameter estimation (i.e. case MV-S_t), the model still performs well for S_t climatology with $r = 0.96$. Consequently, E_a is the most critical variable for predicting the S_t signal in the

proposed multivariate calibration setting, while S_u is less critical, probably because the GRACE-derived S_t signal already accounts for S_u (Li *et al.*, 2012a).

6.5.3 Model performance for soil moisture

The climatology of the basin-averaged reference (i.e. ESA CCI) soil moisture (S_u) and different modelled S_u are depicted in Figure 6.4b. All calibration cases give a good performance ($r >$ 0.91), with a good representation of S_u seasonality during both calibration and evaluation periods.

All simulations overestimate the reference S_u during the rising limb (February-August) corresponding to the increasing rainfall period, and underestimate it during the recession limb (September-January). Simulations that show the highest deviation from the reference during the rising limb, on the contrary, show the lowest deviation during the recession and vice versa. The overall best performance is obtained with case MV-E_a ($r = 0.96$), with a better match when the basin is not water limited (i.e. September-January). Case MV-S_t and case MV-S_u show similar performances ($r \approx 0.95$) with a consistent deviation from the reference S_u for all months, while case MV and case MV-Q have similar performances ($r \approx 0.94$), but with a better fit of the reference S_u in the recession limb. It can be inferred that Q is the most critical variable for S_u reproduction during the rising limb, while S_t and satellite S_u improve the simulation during the recession limb. Case Q outperforms all multivariate calibration cases when soil water content increases and it underperforms them when the maximum water content is reached and starts decreasing. The overall lowest performance is given by case Q ($r = 0.92$), followed by case MV-Q ($r = 0.93$), Suggesting that Q alone is not sufficient for predicting the temporal dynamics of S_u, but it remains useful in the multivariate calibration setting. Surprisingly, E_a does not bring substantial information to the multivariate prediction of S_u. Contrastingly, Poméon *et al.* (2018) obtained a slight improvement in S_u simulation (+7%) when using absolute values of satellite E_a in their multivariate calibration. The model performance in reproducing spatial patterns is measured with E_{SP} and its components (i.e. r_s, γ, and α) summarized in Figure 6.5a.

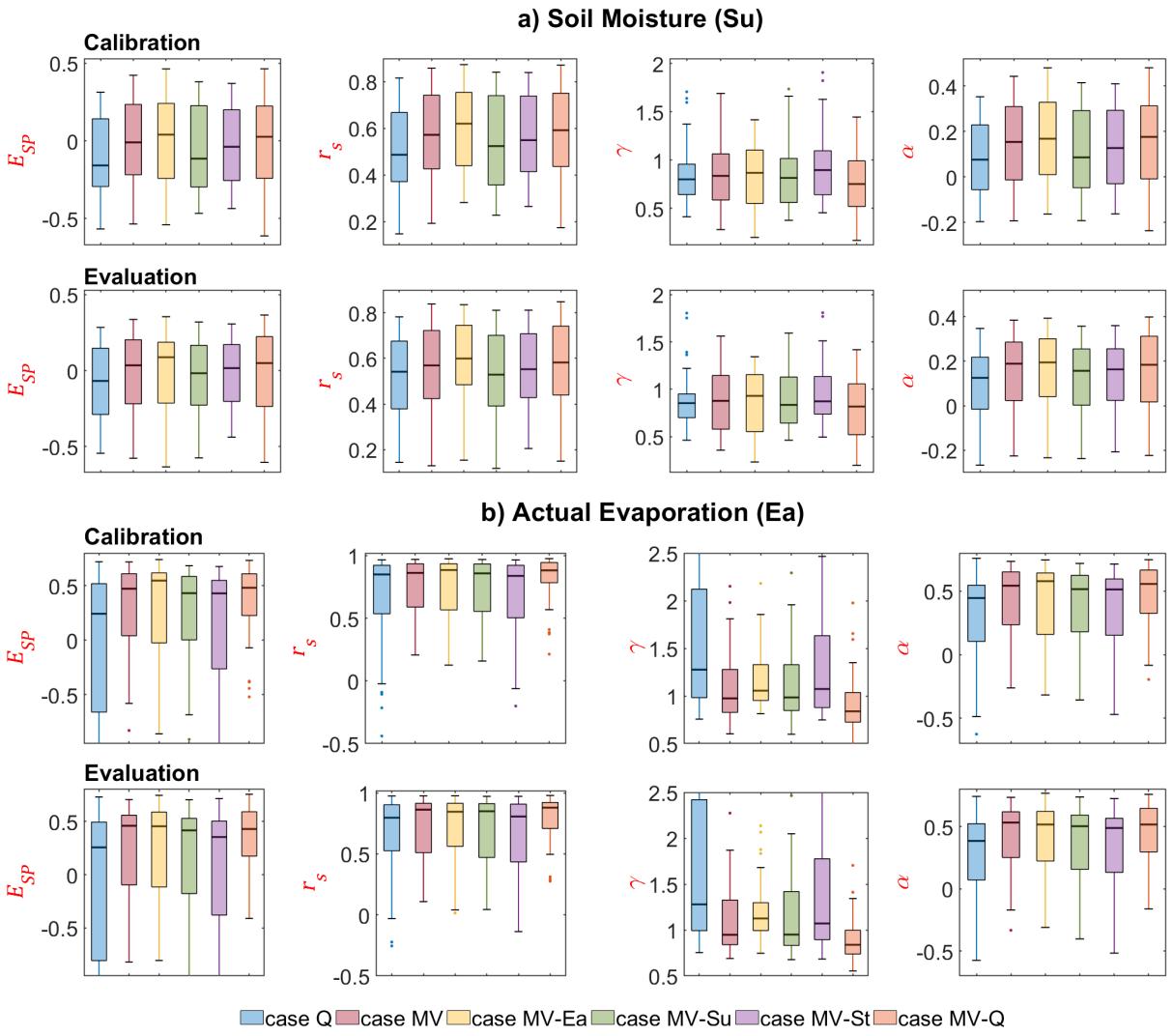


Figure 6.5. Spatial statistics of model performance for soil moisture (a) and actual evaporation (b). The best score is 1 for all the metrics. The number of elements per boxplot corresponds to the number of months in the calibration period (n = 72) or in the evaluation period (n = 48). The colours correspond to the model calibration cases.

The evaluation and calibration periods depict similar trends when comparing the different efficiency metrics (i.e. E_{SP}, r_s, γ, and α). The lowest performance is given by the Q-only calibration (i.e. case Q) with median E_{SP} = -0.07 (r_s = 0.54, γ = 0.85, and α = 0.13), followed by case MV-S_u with median E_{SP} = -0.02 (r_s = 0.53, γ = 0.84, and α = 0.16). Case MV shows better performances with median E_{SP} = 0.03 (r_s = 0.57, γ = 0.88, and α = 0.19). The best performance even with a low score is given by case MV-E_a with median E_{SP} = 0.09 (r_s = 0.60, γ = 0.93, and α = 0.19), followed by case MV-Q with median E_{SP} = 0.05 (r_s = 0.58, γ = 0.82, and α = 0.18) and case MV-S_t with median E_{SP} = 0.02 (r_s = 0.55, γ = 0.87, and α = 0.16). Removing E_a or Q from the multivariate setting results in better performances, while the

removal of S_t or satellite S_u results in lower performances. Consequently, satellite S_u and S_t are the most important variables for improving the spatial patterns of modelled S_u. Better simulation of S_u in multivariate settings is also reported by Lopez *et al.* (2017) using E_a+S_u calibration, and by Poméon *et al.* (2018) with Q+E_a calibration.

The long-term monthly average (2003-2012) of S_u is illustrated in Figure 6.6a. See section 6.5.5.2 for S_u comparison with T_s. Although the spatial patterns of modelled S_u in the multivariate cases are still different from the reference S_u (Figure 6.6a), they are better than the Q-only case.

6.5.4 Model performance for evaporation

The climatology of basin-averaged reference (i.e. GLEAM) actual evaporation (E_a) and different modelled E_a are depicted in Figure 6.4c. All calibration cases give a good performance ($r > 0.91$), reproducing well E_a seasonality during both the calibration and the evaluation periods.

The best performance is obtained with case MV-Q ($r = 0.99$). In general, all simulations tend to underestimate E_a. During the rising limb (February-August), corresponding to the increasing rainfall period, the highest deviation is given by case MV-E_a, although the overall performance is good ($r = 0.98$). Mismatches are more prominent during the recession limb (September-January), where all modelled E_a decrease faster than the reference. The highest deviation is observed when only Q data is used for model calibration ($r = 0.92$), followed by case MV-S_t ($r = 0.96$). It can be inferred that the model case MV-S_t is missing adequate information on the available water amount to be evaporated, which can be obtained from the satellite S_t signal. During the recession period, little to no rainfall occurs in the basin, but a part of antecedent rainfalls is stored in reservoirs and lakes, which represents a major source of land evaporation. It can be argue that Q alone is not sufficient for modelling E_a, while S_t brings additional and useful information for simulating E_a, which supports the research hypothesis. The performance of case MV-S_u is similar to case MV ($r \approx 0.98$), meaning that S_u is not critical for predicting the temporal dynamics of E_a. Moreover, satellite E_a improves the modelled E_a during water accumulation in the basin (i.e. February-August) and is no longer critical when the basin is not water limited (i.e. September-January). This result suggests that the model can mainly rely on GRACE-derived S_t to reproduce E_a. Similar results on the good estimation of E_a with GRACE-derived S_t are found in literature (e.g. Bai *et al.*, 2018b; Livneh and Lettenmaier, 2012; Rakovec

et al., 2016a). Poméon *et al.* (2018) also obtained a higher model performance for E_a in their multivariate setting (i.e. Q+E_a) with mHM in West Africa.

Figure 6.5b gives the spatial pattern efficiency of E_a for all model calibration cases. In general, the performance decreases from the calibration to the evaluation period, and the modelled E_a with all model calibration cases has higher spatial pattern efficiency scores ($E_{SP} > 0.25$) compared to modelled S_u ($E_{SP} < 0.1$). All multivariate calibration cases outperform the Q-only calibration, giving the lowest performance with median $E_{SP} = 0.28$. The Q-only calibration gives a good spatial correlation ($r_s = 0.8$) but overestimates the variability ($\gamma = 1.28$) and struggles to match the spatial location of grid cells ($\alpha = 0.39$) of E_a. The best spatial patterns matching is given by case MV with median $E_{SP} = 0.46$ ($r_s = 0.86$, $\gamma = 0.95$, and $\alpha = 0.53$). Removing Q from the multivariate setting (i.e. case MV-Q) results in underestimation of the spatial variability of E_a, with median $E_{SP} = 0.43$ ($r_s = 0.88$, $\gamma = 0.84$ and $\alpha = 0.52$). In contrast, the spatial variability of E_a is overestimated for case MV-E_a with median $E_{SP} = 0.45$ ($r_s = 0.85$, $\gamma = 1.13$ and $\alpha = 0.52$), while case MV-S_u yields a lower spatial location score with median $E_{SP} = 0.42$ ($r_s = 0.85$, $\gamma = 0.95$ and $\alpha = 0.50$). The spatial pattern performance of E_a is more sensitive to the removal of S_t, as shown by case MV-S_t with median $E_{SP} = 0.35$ ($r_s = 0.81$, $\gamma = 1.07$ and $\alpha = 0.49$). These results indicate that spatial patterns of S_u can improve the spatial patterns of E_a and S_t is critical for reproducing both the temporal and spatial dynamics of E_a. Demirel *et al.* (2018) similarly reported better spatial pattern performance for E_a when using a multivariate setting (i.e. Q+E_a) compared to the Q-only calibration.

Figure 6.6b illustrates the long-term (2003-2012) monthly average of E_a. See section 6.5.5.2 for E_a comparison with T_s. The southern region of the basin, with a sub-humid climate, is where the multivariate calibration cases show more differences in spatial patterns compared to case Q. Besides the south-north differences, it is interesting to see strong differences in the west-E_ast variability of the spatial pattern. As the southern part is sub-humid ($E_a \geq 70\%$), small variations in E_a are not well represented when the model is calibrated using only Q compared to the semi-arid northern part. These findings are in agreement with the study of Rakovec *et al.* (2016a), which revealed a more pronounced sensitivity in E_a estimation in humid catchments in Europe through a multivariate calibration setting (Q+S_t). Similar results are obtained by Bai *et al.* (2016) when testing different E_p formulas in China. Contrastingly, Bai *et al.* (2018b) found that their multivariate calibration setting (Q+S_t) benefitted more to E_a simulation in dry catchments than wet catchments in China.

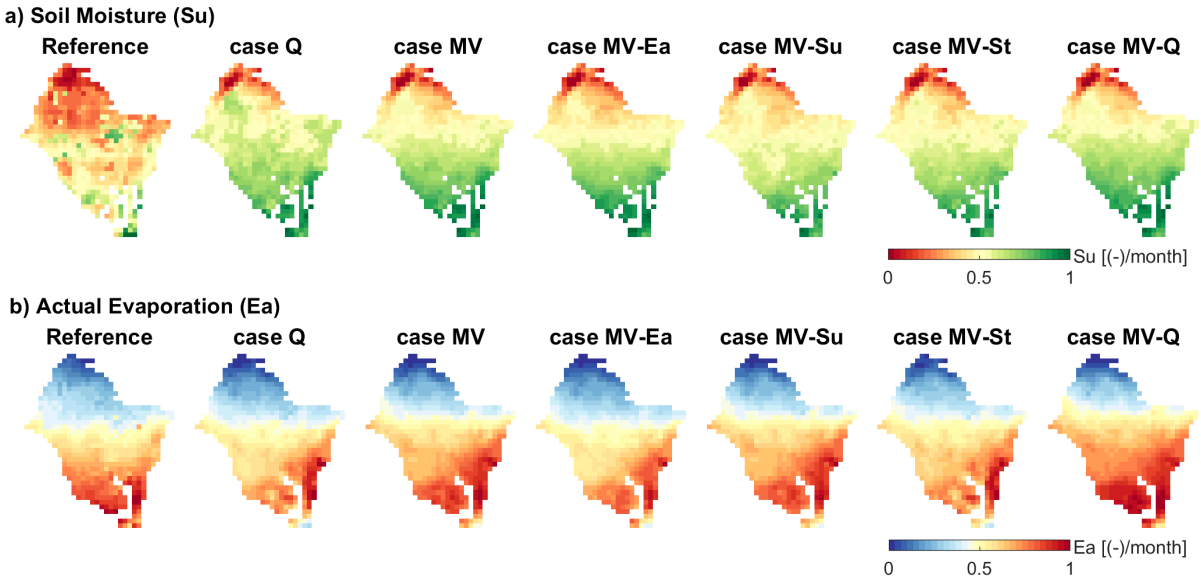


Figure 6.6. Long-term monthly average of soil moisture (a) and actual evaporation (b) for all model calibration cases over the simulation period (2003-2012). The reference map represents the satellite product (ESA CCI for Su and GLEAM for Ea). Masked pixels are gaps in satellite measurements or lake areas not modelled in mHM. The values are normalized for better emphasizing on patterns and using a unique colour scale.

6.5.5 Parameter transferability across spatial scales

6.5.5.1 Streamflow evaluation across spatial scales

The model performance of streamflow in terms of scale transferability of the global parameters is given in Figure 6.7. The differences in model performance among calibration cases are conserved across spatial scales, with a median coefficient of variation of 1.6% for E_{KG}, 0.1% for r, and 6.4% for β.

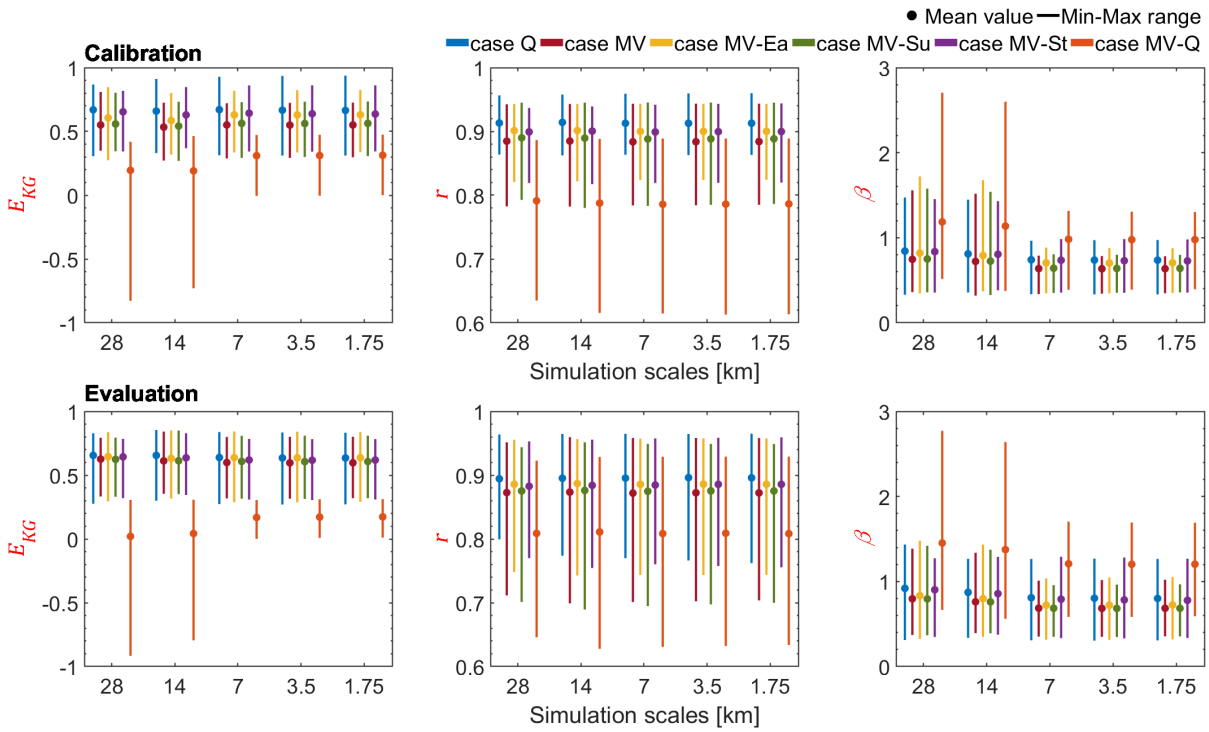


Figure 6.7**.** Statistics for model parameter transferability across spatial scales for streamflow performance for all calibration cases. The dots give the mean score and the bars represent the min-max range of all values for 11 streamflow gauges. The colours correspond to the model calibration cases.

6.5.5.2 Spatial pattern evaluation across spatial scales

Long-term monthly maps of S_u (Figure 6.8a) and E_a (Figure 6.8b) are plotted along with T_s maps at various spatial resolutions. Here, only the coarsest and finest resolutions (i.e. 28 and 1.75 km) are shown.

The patterns of S_u is consistent with the patterns of T_s because the expectation is that the higher the T_s, the lower the S_u and vice versa (Figure 6.8a). For semi-arid regions, E_a largely depends on water availability (i.e. rainfall) and is dominant for open water storages. In the VRB, E_a depicts an opposite pattern to T_s, which is shown in Figure 6.8b. The reproduction of both S_u and E_a in the multivariate calibration cases and across spatial scales show more plausible patterns with T_s, which are well preserved across scales with higher consistency, than their representation with case *Q*.

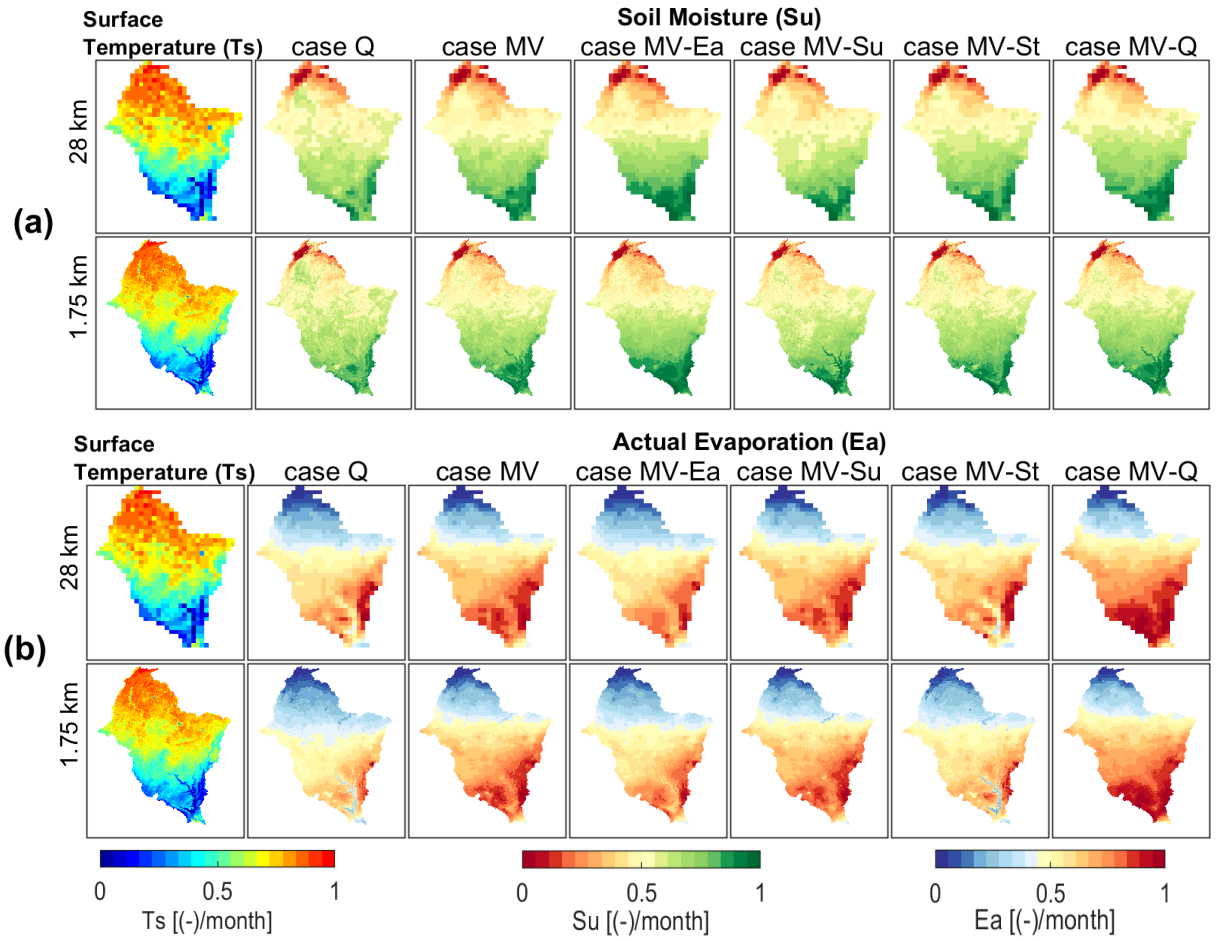


Figure 6.8. Long-term monthly average of land surface temperature compared to soil moisture (a) and actual evaporation (b) for all model calibration cases over the simulation period (2003-2012) at various spatial resolutions. The values are normalized for better emphasizing on patterns and using a unique colour scale.

6.5.6 Benefit of spatial patterns and data types in multivariate calibration

6.5.6.1 Analysis of the Lake Volta region

Evidence of the benefit of multivariate calibration with SRS data is exemplified on Figure 6.9 by zooming-in on the Volta Lake region in the southern part of the VRB (cf. Figure 2.2). Notwithstanding that mHM does not have a lake module, it is nicely noticeable that the model represents the heterogeneity in spatial patterns with the multivariate calibration cases better than the Q-only calibration case. As it should be expected from the T_s patterns, case MV shows higher S_u and E_a over the Lake Volta but with lower E_a in its surroundings. This improvement in spatial patterns is not observed with case Q, confirming the limitations of the Q-only calibration and emphasizing the importance of patterns of SRS data for model calibration. Moreover, it can be inferred from the results that S_t is the most important variable for representing the lake area while Q is the less critical variable.

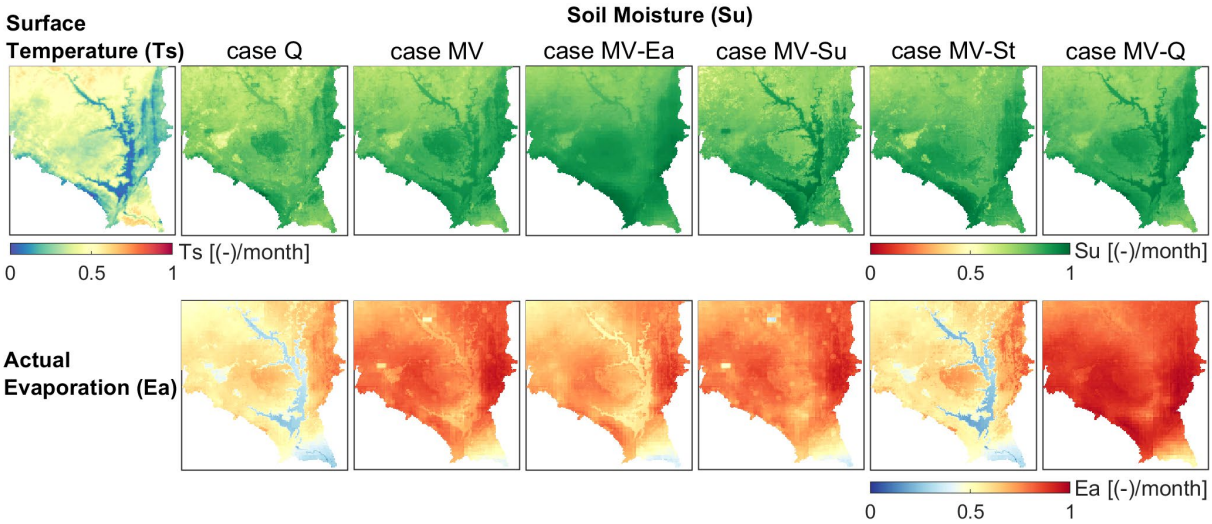


Figure 6.9. Comparison of the spatial patterns of soil moisture (top row) and actual evaporation (bottom row) with land surface temperature (first map from left) in the Volta Lake region. The Ts map used as benchmark shows the Lake Volta depicted in dark blue with the lowest temperature in the region. The ability of the mHM model to highlight the lake area is assessed with the patterns of Su and Ea for all model calibration cases.

6.5.6.2 Comparison of the multivariate calibration cases to the benchmark

Figure 6.10 gives the gain or loss in model performance with different multivariate model calibration cases compared to the *Q*-only calibration case (i.e. the benchmark).

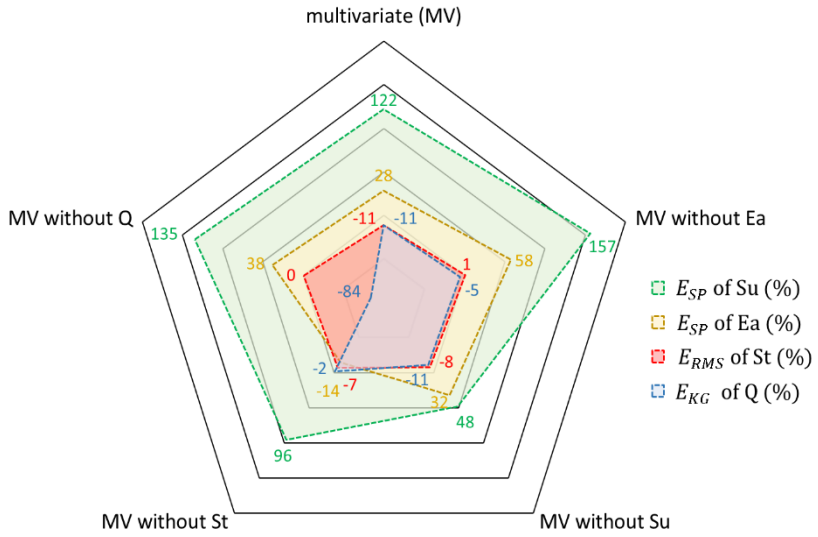


Figure 6.10. Relative difference in performance of multivariate (MV) calibration cases compared to Q-only calibration case. For every MV case, the values on the line from each vertex to the centre of the polygon give the relative difference in performance with the Q-only calibration case, for all variables (i.e. Q, St, Su and Ea).

In general, the multivariate calibration cases show higher model predictive skill for many components of the hydrological system (i.e. Q, S_t, S_u and E_a) when compared to Q-only calibration. The decrease in the model performance for Q is usually counterbalanced with an increase in performance for S_t, S_u and E_a, which might simply be the result of a solution to the artefacts caused by the Q-only calibration. These results reveal that the simulation of spatial patterns of S_u benefits most from the multivariate settings, followed by the simulation of E_a and S_t, while the decrease in Q performance varies widely depending on the multivariate calibration cases. A summary of the importance of each variable in predicting the others (i.e. Q, S_t, S_u and E_a) is given in Table 6.2.

Table 6.2. Importance of different variables in predicting others in a multivariate calibration setting. Degree of importance: Low (+), Moderate (++), High (+++) and Very High (++++)

		Predictands					
		Temporal dynamics				Spatial patterns	
		Q	S_t	E_a	S_u	E_a	S_u
Predictors	Q	++++	+	++	+++	+	+
	S_t	++	++	++++	+++	++++	++
	E_a	++	++++	+++	+	++	+
	S_u	++	+	+	+++	+++	+++

The most determinant variable for streamflow prediction in a multivariate setting is streamflow itself, similarly for S_u, while it is E_a for S_t, and S_t for E_a. Zeng and Cai (2016) also found that S_t controls the temporal variability of E_a. Surprisingly, E_a is the less critical variable for S_u prediction. However, it is worth stressing that only the spatial patterns of satellite E_a is exploited here. Moreover, E_a calculation in the model setup might be a reason of the limited contribution of satellite E_a in S_u prediction. The E_p calculation (cf. Section 4.5 of Chapter 4) is done with time-variant and gridded leaf area index data that imposes heterogeneity on modelled E_a (Birhanu *et al.*, 2019). Consequently, additional contribution from the satellite E_a in S_u prediction is expectedly limited in case the leaf area index data is in agreement with the satellite E_a. Moreover, not explicitly weighting the components of the multivariate objective function might have led to implicit weighting, which led to the artefact that some variables are not very good predictors for themselves.

6.6 Summary and Outlook

This work is a follow up on several recent studies on multi-objective calibration and spatial pattern improvement in hydrological modelling. The proposed multivariate calibration approach is a step forward in improving the realism of hydrological model predictions (Clark *et al.*, 2015; Rakovec *et al.*, 2016a) because not only a reliable temporal dynamic is sought for in the modelling objective but also plausible spatial patterns of several hydrological processes simultaneously. A key element of this study is the assessment of the plausibility of spatial patterns of soil moisture and evaporation with independent data of land surface temperature not used during the model calibration. With respect to the obtained performances, it can be concluded that spatial patterns of satellite data are a highly relevant and robust feature that can be used in multivariate calibration to improve the overall representation of the hydrological system even with trade-offs among the variables, which thereby confirms the research hypothesis.

A rigorous comparison of the proposed bias-insensitive metric with other spatial pattern metrics is left for future works. Further investigations can focus on setting a threshold for the acceptability of the modelled spatial patterns, which was not required here as the goal was to check the increase or the decrease of spatial pattern performance rather than determining whether the patterns are good or bad in an absolute sense, when switching between streamflow-only and multivariate calibration cases.

Our methodology lacks in-situ data for model evaluation, except streamflow. However, in-situ measurements of soil moisture, evaporation and terrestrial water storage at large scale are rather rare (Vereecken *et al.*, 2008), and are also subject to uncertainties due to the non-uniformity of the data collection in space (Stisen *et al.*, 2011). As this study focuses on spatial pattern assessment, satellite data remain the only possible option for the large Volta River basin in West Africa, where ground measurements are a luxury.

The presented multivariate calibration reveals trade-offs among the objective functions for streamflow and for satellite data. However, trade-offs cannot be avoided as they originate from errors in input data, model structure and lack of knowledge of the hydrological system (Bergström *et al.*, 2002; Gupta *et al.*, 1998). Moreover, it was a deliberate choice to equally weight the components of the multivariate objective function (Eq. 6.1) because no prior knowledge on the importance of each variable was available, and it was an objective of this study to know their contributions in the calibration procedure. In such situation, the default choice is to weight them equally (Bergström *et al.*, 2002; Stisen *et al.*, 2018). Weights are

sometimes assigned to objective function components by iterative optimization testing different weights, which is however computationally demanding. It is also possible to transform the components of the multivariate objective functions to solve differences in their magnitudes (Madsen, 2003), but the effects of such transformations on the calibration procedure are unknown and they are not required if the metrics are dimensionless or of the same order of magnitude (Bergström *et al.*, 2002).

The climatic inputs influence somehow the spatial variability of the hydrological processes due to the aridity gradient in the VRB. The detailed results are valid for the VRB but they can be generalized to regions with similar hydroclimatic characteristics. However, the applicability of the proposed multivariate calibration framework is, in principle, universal, as long as a DHM is used and spatial datasets are available. Further research can explore the applicability of the presented multivariate calibration strategies in different hydroclimatic regions with different spatial data sources, and different DHMs to understand how the model structure interacts with the performance of different calibration strategies. Choosing an adequate hydrological model (Addor and Melsen, 2019) is key to any good experiment. The MPR scheme used in mHM might have facilitated to some extent the reproduction of spatial patterns, but the MPR scheme can be similarly implemented with other models as demonstrated by previous studies (e.g. VIC and PCR-GLOBWB models, Mizukami *et al.*, 2017; Samaniego *et al.*, 2017). A sensitivity analysis to identify the model parameters that influence the representation of spatial patterns is a recommended outlook.

Future methodological developments could in particular focus on improved formulation of the multi-objective functions inspired by previous findings on following topics: fitting of low flows and system signatures (Fowler *et al.*, 2018b; Krause *et al.*, 2005; Pushpalatha *et al.*, 2012), gauge measurement weighting (Madsen, 2003) or sub-period calibration (Gharari *et al.*, 2013). Additional key questions to address in this context include the model structural deficiencies (Gupta *et al.*, 1998; Gupta *et al.*, 2012), and the uncertainties of modelling datasets (i.e. input, calibration and evaluation data), which can lead to erroneous model rejection (Beven, 2010; 2018; 2019a).

The above efforts in model improvement are particularly important for prediction in a changing environment (Fowler *et al.*, 2018a), and they can set avenues for prediction in ungauged basins solely from space.

6.7 Conclusion

This study presents a calibration approach using multiple data sources simultaneously, with the specificity of integrating only spatial patterns of satellite remote sensing data in the parameter estimation procedure. A bias-insensitive and multicomponent metric is proposed for spatial pattern matching. The study is carried out in the Volta River basin in West Africa. Results reveal the benefit of the multivariate calibration setting over the traditional calibration using only streamflow data. The main findings are as follows:

- Streamflow is a necessary variable but alone it is not sufficient for reliably reproducing other hydrological fluxes and states;
- Spatial patterns of satellite data, without the absolute values, can be incorporated in the calibration procedure with bias-insensitive metrics;
- Multivariate calibration based on streamflow and satellite data can improve the overall representation of the hydrological system, and thereby increase the model predictive skill;
- The reduction in streamflow performance in multivariate setting is largely compensated by the gain in performance for other hydrological processes (i.e. terrestrial water storage, soil moisture and evaporation).

It is advocated to adopt a multivariate calibration procedure focusing on spatial patterns in distributed hydrological models because it is a robust approach for addressing equifinality, reducing uncertainties and enhancing the predictive skill of hydrological models in a changing environment.

Chapter 7

Impacts of Climate Change on Hydrological Processes in the Volta River Basin*

The past is certain, the future obscure.

Thales

Not everything that is faced can be changed.

But nothing can be changed until it is faced.

James A. Baldwin

* This chapter is in preparation for publication.

Abstract

A comprehensive evaluation of the impacts of climate change on water resources of the Volta River basin located in West Africa is proposed in this study. A large ensemble of twelve global climate models (GCM) from CMIP5 that are dynamically downscaled by five regional climate models (RCM) from CORDEX-Africa is used. In total, 43 RCM-GCM combinations are considered under three representative concentration pathways (RCP2.6, RCP4.5 and RCP8.5). The reliability of the climate datasets is first evaluated with satellite and reanalysis reference datasets. Subsequently, a multivariate bias correction method is applied to the climate datasets before using them as input to the fully distributed Hydrologic Model (mHM) for hydrological projections. The analyses cover the twenty-first century and are carried out for the historical period 1991-2020, with projections over three future periods 2021-2050, 2051-2080 and 2071-2100.

Results reveal contrasting changes in the seasonality of precipitation depending on the RCPs and the future projection periods, while a clear increase in the seasonality of temperature is expected. Although temperature and potential evaporation increase under all RCPs, the intensification of the complete hydrological cycle during the twenty-first century is only expected under the RCP8.5 scenario. In this case, an increase is foreseen for the long-term annual estimates of precipitation (+6.2%), average temperature (+9.5%) and potential evaporation (+5.0%). These changes in climatic variables will lead to changes in actual evaporation (+4.2%), surface runoff (+42%), streamflow (+84%), groundwater recharge (+37%), soil moisture (+2.3%) and terrestrial water storage (+3.2%). As shown in this study, floods and droughts will be recurrent under these conditions, thereby weakening the water-energy-food security nexus and amplifying the vulnerability of the local population to climate change. These findings could serve as a guideline for decision makers, and contribute to the elaboration of adaptation and mitigation strategies to cope with dramatic consequences of climate change, and strengthen the regional socio-economic development. However, variabilities between climate models highlight uncertainties in the projections. While there is a 100% agreement between the RCM-GCM models for projections of temperature, only an agreement of 62% on the direction of change is obtained for precipitation, which underscores the complexity of modelling climate and hydrological systems in West Africa, and sparks a new call to both the scientific community and policymakers for further efforts in investigating climate change in the region.

7.1 Introduction

Climate warming is expected to occur at a faster rate in West Africa as compared to the global average during the twenty-first century (Todzo *et al.*, 2020). Anthropogenic greenhouse gas emissions have led to an unprecedented increase in surface air temperature, which has resulted in the intensification of the hydrological cycle (Sylla *et al.*, 2016). Therefore, recurrent floods and droughts could persist in the future because precipitation is expected to decrease in frequency but increase in intensity, whereas runoff variability is tightly linked to changes in rainfall (Aich *et al.*, 2014; Dosio *et al.*, 2020; Roudier *et al.*, 2014b). In the face of climate change and variability, West Africa is particularly vulnerable because of its high reliance on rainfed agriculture and limited institutional capacities to cope with climate change and variability (Karambiri *et al.*, 2011; Sultan and Gaetani, 2016; Yira *et al.*, 2017). Climate change and anthropogenic pressures pose an increasing stress on water resources (Sood *et al.*, 2013). Freshwater shortages with a decline in basin-scale irrigation potential could have dire consequences for sustainable agriculture (Sylla *et al.*, 2018b). Consequently, global warming is a serious thread for water and food security in West Africa.

In the transboundary Volta River basin (VRB) located in West Africa, water resources are fundamental for agriculture, hydroelectricity, fisheries and ecosystem services (Williams *et al.*, 2016). Most of the agriculture is rainfed but many regions rely on irrigated agriculture (Roudier *et al.*, 2014b). Hydropower is a major source of electricity production with the potential to unlock more access to energy in the region (Kling *et al.*, 2016; Stanzel *et al.*, 2018). Future water resources developments in the VRB focus primarily on hydroelectricity and irrigation (McCartney *et al.*, 2012). Nevertheless, severe impacts of climate change on water resources in the VRB will impede future socio-economic development (Sood *et al.*, 2013). Knowledge on the future evolution of the hydrological cycle in the VRB is capital to increase the adaptive capacities of the riparian countries to the regional consequences of global warming. However, there is little knowledge on the impacts of climate change on water resources in West Africa in general (e.g., Kasei, 2010; Oyerinde *et al.*, 2016; Yira *et al.*, 2017), and only a few studies focused on the VRB (Jung *et al.*, 2012; Okafor *et al.*, 2019; Roudier *et al.*, 2014b). Therefore, it is urgent to quantify the future changes in hydrological processes because of the fragility of the region to extreme hydroclimatic events and its high reliability on water resources for economic development (Jin *et al.*, 2018; Stanzel *et al.*, 2018).

The limitations of climate change impacts studies on water resources in West Africa arise from the lack of hydrological and meteorological observations to drive models, in addition to

uncertainties related to climate projection data as well as hydrological models (Dembélé *et al.*, 2019; Oyerinde *et al.*, 2016; Sidibe *et al.*, 2020; Sylla *et al.*, 2018a). Currently, there is a high variability in climate projections with different storylines over West Africa (Dosio *et al.*, 2020), which makes it a scientifically particular and interesting region, and underlines the necessity to further investigate the impacts of climate change on hydrological processes. Despite substantial progress in improving climate projections data and the number of studies investigating climate change in West Africa (e.g., Diallo *et al.*, 2016; Kebe *et al.*, 2017; Mahé *et al.*, 2013; Nikiema *et al.*, 2017), uncertainties still exist in the climate projections data (Eyring *et al.*, 2019; Sidibe *et al.*, 2020; Sylla *et al.*, 2016). However, knowledge on the impacts of climate change on water resources needs to be updated continuously.

This study assesses the impacts of climate change on water resources in the VRB, and the implications for water resources management. The goal is to analyse changes in patterns of precipitation and temperatures, and their associated repercussions on various components of the hydrological cycle (i.e., streamflow, surface runoff, potential and actual evaporation, groundwater recharge, soil moisture and terrestrial water storage) in the twenty-first century. Uncertainties in the climate projections are addressed by employing a large ensemble of twelve Global Climate Models (GCMs) downscaled by five Regional Climate Models (RCMs) under three Representative Concentration Pathways (RCPs; Moss *et al.*, 2010; Van Vuuren *et al.*, 2011). The RCMs are obtained from the COordinated Regional-climate Downscaling Experiment (CORDEX) for Africa (Giorgi *et al.*, 2009). CORDEX generates high-resolution historical and future climate projections for regional applications, by downscaling GCMs participating in the Coupled Model Intercomparison Project (CMIP5; Taylor *et al.*, 2012). As not all GCMs are downscaled by the RCMs, there are 21 RCM-GCM combinations available in the historical period (1951-2005), while 43 RCM-GCM combinations are available for the future projections (2006-2100) under various emission scenarios (RCP2.6, RCP4.5 and RCP8.5). Several RCPs are used in this study to consider the possibility of different futures. In fact, only considering the highest RCP8.5 scenario as the "business as usual" scenario in climate change studies is increasingly criticized because the assumption of the heavy use of coal in RCP8.5 is unrealistic (Hausfather and Peters, 2020; Ritchie and Dowlatabadi, 2017). However, the current emissions are found to be in line with the RCP8.5 scenario (Peters *et al.*, 2013), and there are suggestions for giving RCP8.5 a high priority (O'Neill *et al.*, 2016).

The RCM-GCM datasets are first evaluated by comparing them to the best satellite and reanalysis datasets of rainfall and temperature for hydrological modelling in the VRB (Dembélé

et al., 2020c). Subsequently, a multivariate bias correction is applied to the climate projection datasets using the Rank Resampling for Distributions and Dependences (R2D2) method (Vrac and Thao, 2020). Finally, the bias corrected climate projection datasets are used as input in the fully distributed mesoscale Hydrologic Model (mHM) to assess the impact of climate change on multiple hydrological fluxes and state variables. Although the performance of a model during past and present conditions does not guarantee its reliability for future projections (Stanzel *et al.*, 2018), having a well performing model that simulates realistic hydrological processes is a prerequisite for any sound impact study. The mHM model is thoroughly calibrated to reproduce plausible spatiotemporal patterns of hydrological processes in the VRB, following a robust approach based on multivariate and spatial patterns of satellite data (Dembélé *et al.*, 2020b). The following research questions are investigated:

- What are the changes in rainfall and temperature during the twenty-first century in the VRB?
- How will hydrological processes evolve under a changing climate?
- How uncertain are the projections in terms of multi-model variability?
- What are the implications of a changing climate for floods and droughts?

Overall, this study seeks to explore the implications of climate change for water resources management, and to provide knowledge that can serve to deploy adaptation and mitigation strategies to limit the negative impacts of global warming on the socioeconomic development of the VRB.

7.2 Overview of the methodology

The methodology adopted for the assessment of climate change impacts on hydrological processes is summarized in Figure 7.1. The main steps consist of the bias correction of the RCM and GCM datasets, the modelling of hydrological processes based on the climate projection datasets, and the analysis of the future changes in the modelled hydrological processes.

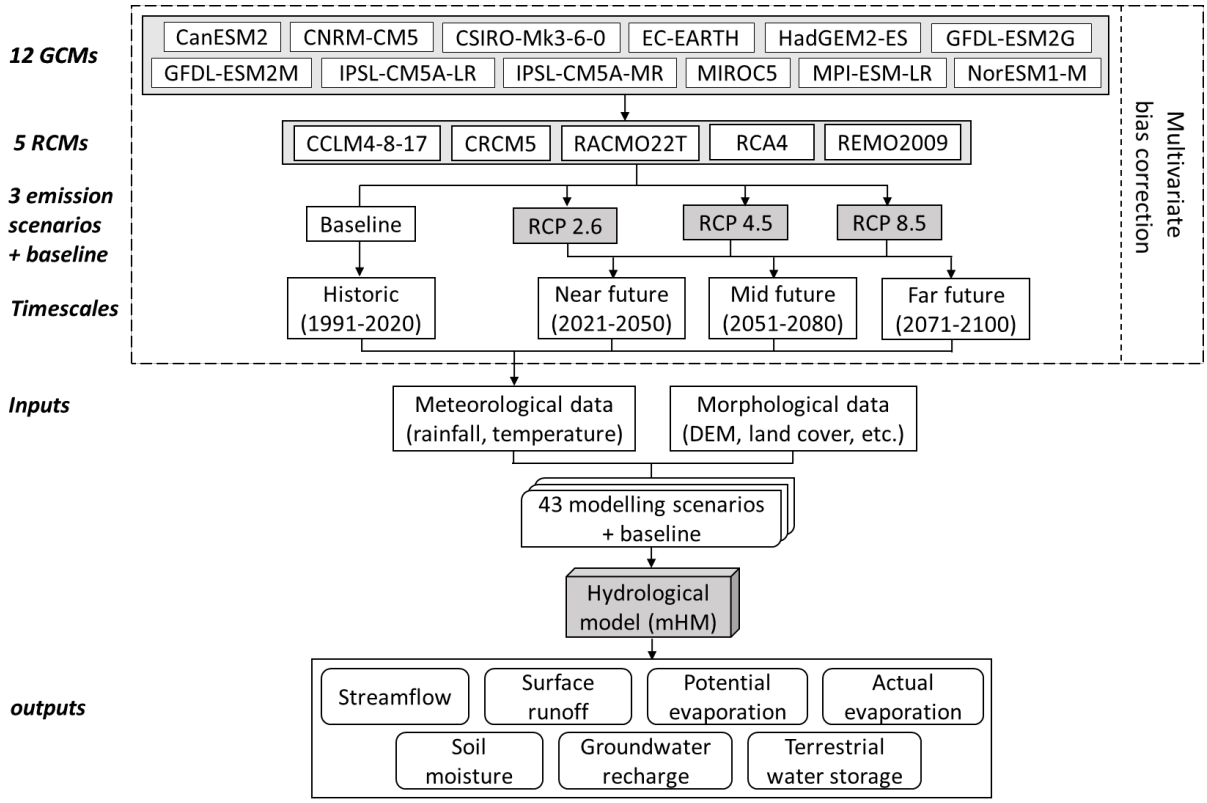


Figure 7.1. Overview of the procedure for assessing the impacts of climate change on hydrological processes

7.3 Climate datasets

Based on data availability on the Earth System Grid Federation (ESGF) platform, twelve GCMs from CMIP5 dynamically downscaled with five RCMs available from the CORDEX-Africa initiative are selected for this study (Table 7.1). Daily datasets of precipitation and minimum, maximum and average air temperature are obtained for the realization r1i1p1 over the historical period (1951-2005) and the future projections period (2006-2100). Three RCPs are used in this study, and they correspond to different future concentrations and emissions of greenhouse gases and air pollutants, and land-use change until 2100, relative to the pre-industrial times (Moss *et al.*, 2010; Van Vuuren *et al.*, 2011). The three RCPs are (i) RCP2.6 corresponding to a mitigation scenario with a very low radiative forcing level that peaks at ~3 W m^{-2} (~490 ppm CO_2 equivalent) before 2100 and then declines; (ii) RCP4.5 representing a medium stabilization scenario without overshoot pathway to 4.5 W m^{-2} (~650 ppm CO_2 equivalent) at stabilization after 2100; and (iii) RCP8.5 corresponding to a very high emission scenario with rising radiative forcing pathway leading to 8.5 W m^{-2} (~1370 ppm CO_2 equivalent) by 2100.

Table 7.1. List of the CORDEX-Africa Global Climate Models (GCMs) datasets downscaled by Regional Climate Models (RCMs) and their availability per Representative Concentration Pathways (RCP) marked with a cross (x).

RCMs	GCMs	RCP2.6	RCP4.5	RCP8.5
CCLM4-8-17	CNRM-CM5		x	x
	HadGEM2-ES		x	x
	MPI-ESM-LR		x	x
CRCM5	CanESM2		x	
	MPI-ESM-LR		x	
RACMO22T	EC-EARTH		x	x
RCA4	CanESM2		x	x
	CNRM-CM5		x	x
	CSIRO-Mk3-6-0		x	x
	EC-EARTH			x
	CM5A-MR		x	x
	MIROC5	x	x	x
	HadGEM2-ES	x	x	x
	MPI-ESM-LR	x	x	x
	NorESM1-M	x	x	x
	GFDL-ESM2M		x	x
REMO2009	CM5A-LR	x		x
	MIROC5	x		x
	HadGEM2-ES	x		x
	MPI-ESM-LR	x	x	x
	GFDL-ESM2G	x		
	TOTAL	9	16	18

7.4 Bias correction

Assessing the reliability of climate projections data in reproducing observations is a precondition to impact studies (Eyring *et al.*, 2019). However, observations can also be subject to uncertainties. To address these uncertainties, the climate projection datasets are evaluated with ten rainfall datasets (TAMSAT v3.0, CHIRPS v2.0, ARC v2.0, MSWEP v2.2, PERSIANN-CDR v1r1, PGF v3, ERA5, WFDEI-GPCC, WFDEI-CRU and MERRA-2) and six temperature datasets (PGF v3, ERA5, WFDEI, EWEMBI, MERRA-2 and JRA-55), which demonstrated the best performances for large scale hydrological modelling in the VRB (Dembélé *et al.*, 2020c). In the following, these datasets composed of both satellite and reanalysis products will be referred to as observations.

As discrepancies are observed between the observations and the climate projection datasets (Figure 7.2), a bias correction is applied before using the climate datasets in hydrological

modelling. The Rank Resampling for Distributions and Dependences (R2D2) method (Vrac and Thao, 2020) is adopted for a multivariate bias correction of climatic variables. R2D2 is a stochastic and analogue-based method that adjusts not only the univariate distributions of climatic variables relying on their ranks, but also the inter-variable and inter-site dependence structures (Vrac, 2018). Bias correction with the R2D2 approach is achieved in two steps. First, the marginal distributions of univariate time series are adjusted using any univariate bias correction method. Here, the cumulative distribution function transform (CDF-t) approach (Michelangeli *et al.*, 2009; Vrac *et al.*, 2012) is used to adjust the marginal properties of the univariate time series. Secondly, R2D2 is used to adjust the dependence structure between several variables independent of their marginal distribution (i.e., empirical copula function). R2D2 showed a good performance in comparison to other bias correction methods (François *et al.*, 2020).

In this study, precipitation and temperature datasets are corrected simultaneously to preserve the temporal and spatial dependences between the climatic variables. The bias correction is done using only the WFDEI as reference data, because it has both rainfall and temperature data over a long period (1979-2016), and previously demonstrated good performances in the VRB (Dembélé *et al.*, 2020a; Dembélé *et al.*, 2020b; Dembélé *et al.*, 2020c). Moreover, using a single observation dataset as reference avoids losing the natural variability of the time series, which might not be the case with a multi-dataset ensemble mean used as the reference. The period 1981-2005 is taken as the reference period for the bias correction of the climate projection datasets, whose time series are divided into several 25-year periods over the period 1951-2100 to correspond to the length of the reference period. The multivariate bias correction is applied by grouping the data per calendar month in each sub-period of 25 years, which allows preserving temporality in the corrected data.

7.5 Hydrological Modelling

The bias corrected climate projection datasets are used as forcing to the fully distributed mesoscale Hydrologic Model (mHM) to assess the changes in hydrological processes as a results of climate change. The mHM model simulates dominant hydrological processes with seamless spatiotemporal patterns in the modelling domain (Kumar *et al.*, 2013; Samaniego *et al.*, 2010; Samaniego *et al.*, 2017). Dembélé *et al.* (2020a) and Dembélé *et al.* (2020b) demonstrated the ability of the mHM model to reproduce reliable spatiotemporal patterns of various hydrological processes after a robust model calibration with multiple variables and

spatial patterns of satellite data in the VRB. The model parameters obtained from the study of Dembélé *et al.* (2020b) are used to run the mHM model in this study. Dembélé *et al.* (2020b) provide details on the setup and the performance of mHM for hydrological modelling in the VRB. The model is run over the entire availability period (1951-2100) of the RCM-GCM datasets, with a model warm-up period of 10 years (1951-1960). The 30-year period (1991-2020) is chosen as the baseline period (i.e., historical run) for the hydrological modelling. While projections are assessed for the near term future (2021-2050), the long-term future (2051-2080) and the late-century (2071-2100) under various emission scenarios. In total, 21 RCM-GCM combinations are available for the historical runs, while for future projections, 9 RCM-GCM combinations are available for the RCP2.6, 16 for RCP4.5 and 18 for RCP8.5 (Table 7.1). Although land use and land cover (LULC) scenarios are not used in this study, the temporal dynamic of LULC is accounted for by using different maps over the modelling period. Based on high resolution LULC data available between 1992 and 2015 from the European Space Agency Climate Change Initiative (ESA, 2017), LULC maps of 1992, 2005 and 2015 are used for the periods 1951-1990, 1991-2020 and 2021-2100, respectively.

The realism of the hydrological simulations is checked with the Budyko framework (Budyko, 1974), which is a tool to estimate mean annual water availability as a function of aridity. The supply-demand framework is valid for large catchments under steady state, considering long-term water and energy balance (Donohue *et al.*, 2010). The exercise consists in verifying if the ratio of the long-term mean annual potential evaporation to precipitation (aridity index) and the ratio of actual evaporation to precipitation (evaporative index) are in the limits of acceptability and do not exceed the energy and water limits (Sposito, 2017). An environment is considered water-limited when the catchment average atmospheric evaporative demand exceeds the water supply (i.e., precipitation is lower than potential evaporation), while the opposite applies to an energy-limited environment (Donohue *et al.*, 2011; McVicar *et al.*, 2012). The Budyko framework is frequently used in hydrological modelling (Greve *et al.*, 2020; Wang *et al.*, 2016).

Uncertainties in the results are assessed in terms of variability between climate models using the second order coefficient of variation (V_2, cf. Eq. 4.14 in Section 4.7 of Chapter 4) (Kvålseth, 2017).

7.6 Results

When reporting on the changes (future minus reference) in hydroclimatic processes in the following, the trend of the changes is determined only by the majority of the models agreeing on the same direction of change. For instance, a variable is said to be increasing in the future only if the majority of the RCM-GCM models in the large ensemble shows an increasing trend, and vice versa. Therefore, the magnitude of the change is obtained by averaging only models agreeing on the same trend. This approach is similar to the weighted multi-model mean where the major agreement is weighted by one, while the minor agreement is given a weight of zero. In this case, at least 50% of models present in the ensemble are always considered when reporting on future projections. The percentage of model agreement is provided to elucidate the robustness of the projections. However, the figures show the projections of all the RCM-GCM combinations.

7.6.1 Multivariate bias correction

The raw RCM-GCM datasets are evaluated by comparing their cumulative distribution functions to those of the observations over the period 1983-2005 corresponding to the concomitant availability period of all the observation datasets (Figure 7.2). The distribution of most of the RCM-GCM datasets presented discrepancies with the observations, with larger gaps for temperature than precipitation. The multivariate bias correction with the R2D2 method visually performed well by adjusting the distributions of the RCM-GCM datasets to the WFDEI reference dataset for all the climatic variables. Therefore, the corrected RCM-GCM datasets are expected to provide reliable hydrological simulations in the VRB.

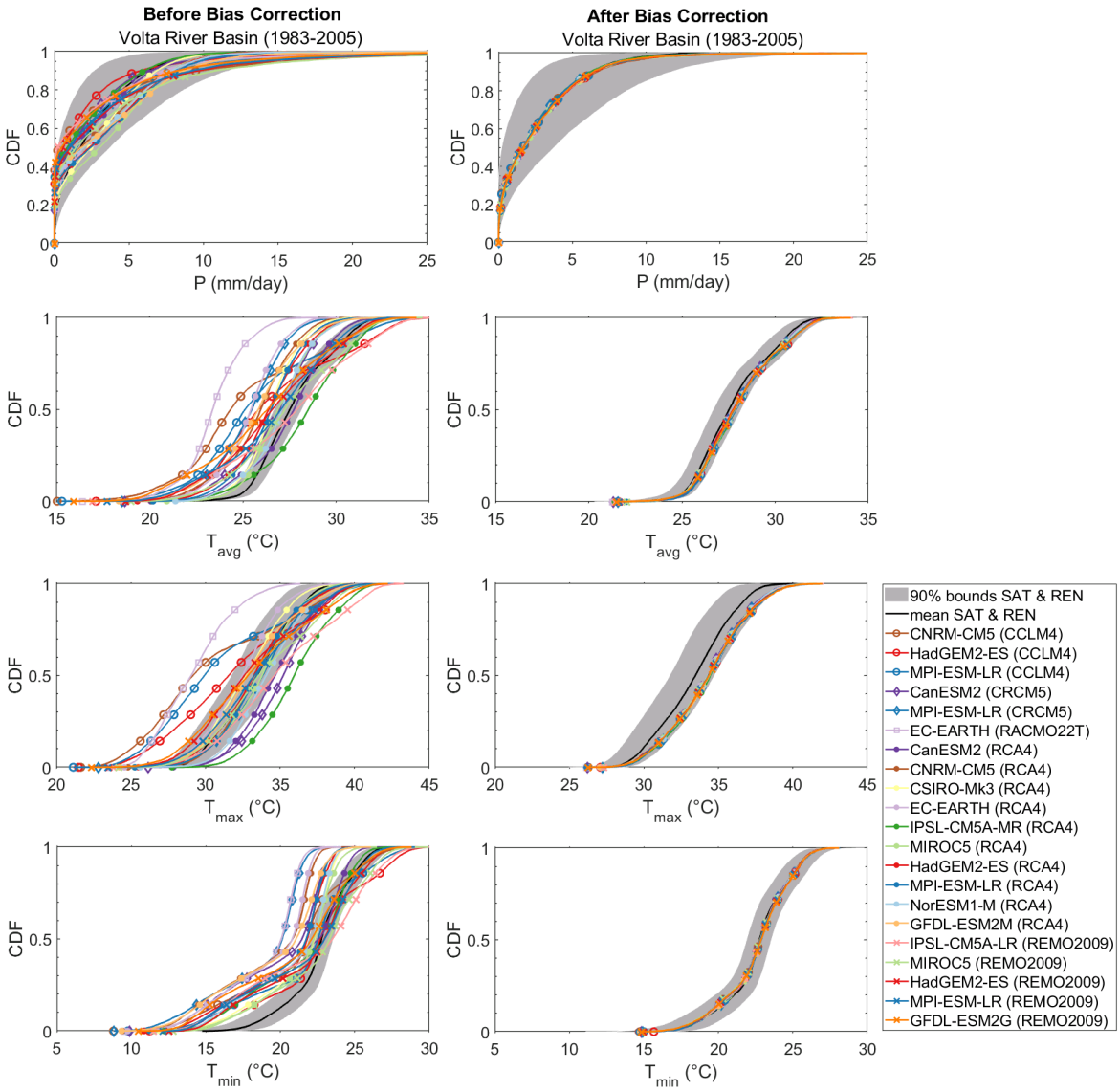


Figure 7.2. Cumulative distribution functions (CDF) of daily precipitation (P) and average, maximum and minimum daily air temperature (T_{avg}, T_{max} and T_{min} = 19.3 °C) before and after multivariate bias correction of various RCM-GCM datasets. The black line and grey-shaded area represent the mean and the 90% confidence interval of the best satellite and reanalysis datasets of rainfall (10 datasets) and temperature (6 datasets) in the Volta basin.

7.6.2 Plausibility of hydrological simulations

The plausibility of the hydrological simulations using various RCM-GCM datasets as inputs to the mHM model under various RCPs and long-term periods is shown with the Budyko framework in Figure 7.3.

All the RCM-GCM datasets provide plausible hydrological simulations, at least in terms of water and energy balance, as they respect the water and energy limits imposed within the Budyko framework. On average, the evaporative index is between 0.86 and 0.97, while the

aridity index is between 2.2 and 4.4, which corresponds to expected values for sub-humid and semi-arid environments such as the VRB (Gunkel and Lange, 2017).

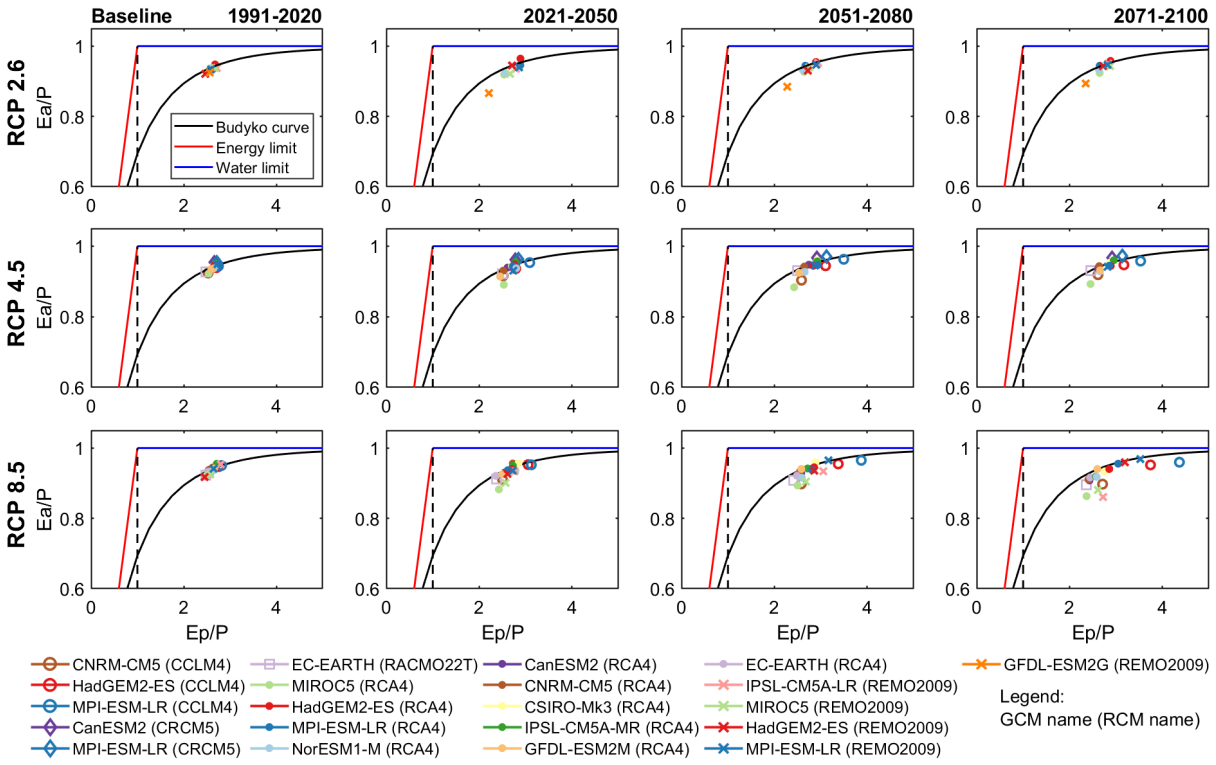


Figure 7.3. Plausibility of hydrological processes with the Budyko framework representing the evaporative index, i.e. ratio of actual evaporation to precipitation (E_a/P), as a function of the aridity index, i.e. ratio of potential evaporation to precipitation (E_p/P)

7.6.3 Seasonal changes in climate

7.6.3.1 Precipitation

The annual cycle of precipitation in the VRB shows a unimodal shape for all the RCPs (Figure 7.4). Monthly precipitation continuously increases from 2 mm/month in January to 208 mm/month in August and decreases to 3 mm/month in December. Uncertainties in the rainfall estimates are lower during the wet months (July-September) with an average variation of 4%, as compared to the dry months (November-March) with 43% of variation. The future evolution of the seasonality of precipitation depicts a contrasting trend across RCPs and across periods. A decline in the seasonality of precipitation by -17% is expected until the end of the twenty-first century for RCP2.6, while an increase of +31% up to the mid-century followed by a decrease of -16% is observed for RCP4.5, and an average increase of +42% is recorded over the entire century for RCP8.5. High inter-model variabilities ranging from 82% to 98% are observed in future seasonal projections of precipitation between the RCM-GCM datasets.

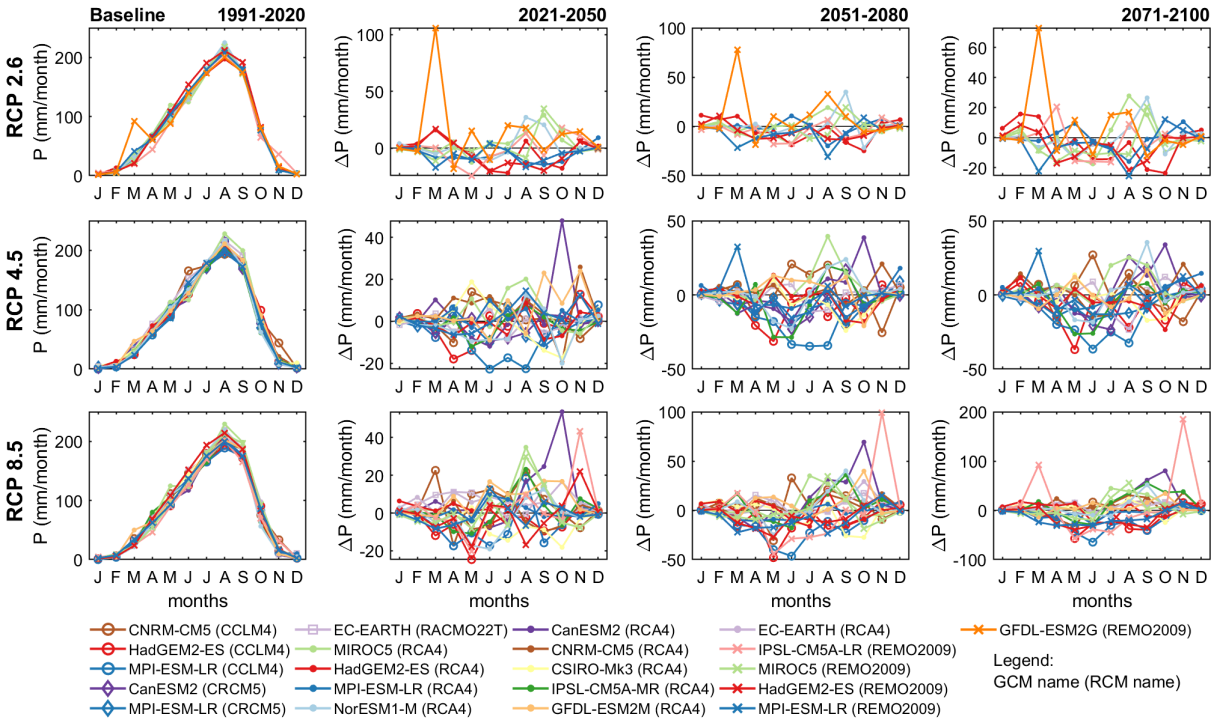


Figure 7.4. Annual cycle of precipitation (P) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.3.2 Temperature

Average, maximum and minimum temperature present a bimodal annual cycle in the VRB (Figure 7.5 to Figure 7.7), with a second mode less marked for the minimum temperature. Monthly temperature increases from January (T_{avg} = 26.5 °C, T_{max} = 34.7 °C, T_{min} = 19.3 °C) and marks its first peak between March and April (T_{avg} = 31.7 °C, T_{max} = 37.8 °C, T_{min} = 26.1 °C). Then, it decreases to its lowest record during the rainy season in August (T_{avg} = 26.4 °C, T_{max} = 30.8 °C, T_{min} = 22.8 °C), before rising to its second peak between October and November (T_{avg} = 28.4 °C, T_{max} = 35.7 °C, T_{min} = 23.1 °C) and decreases up to January. Low inter-model (i.e., RCM-GCMs) variabilities ranging from 0.3% to 0.9% are observed for monthly temperatures. Future changes in average, maximum and minimum monthly temperature agree to the same increasing trend across all RCPs over the twenty-first century. Temperature keeps increasing in time from 2021 to 2100 and with increasing radiative forcing level from RCP2.6 to RCP8.5. Generally, monthly average temperature increases by 3% to 14% (i.e., +1.0°C to +4.0 °C), maximum temperature increases by 3% to 11% (i.e., +1.1°C to +3.9 °C), and minimum temperature increases by 4% to 18% (i.e., +1.0°C to +4.1 °C) over the twenty-first

century. Medium variabilities in future seasonal projections between RCM-GCM datasets range from 24% to 35% for T_{avg}, from 28% to 42% for T_{max}, and from 21% to 35% for T_{min}.

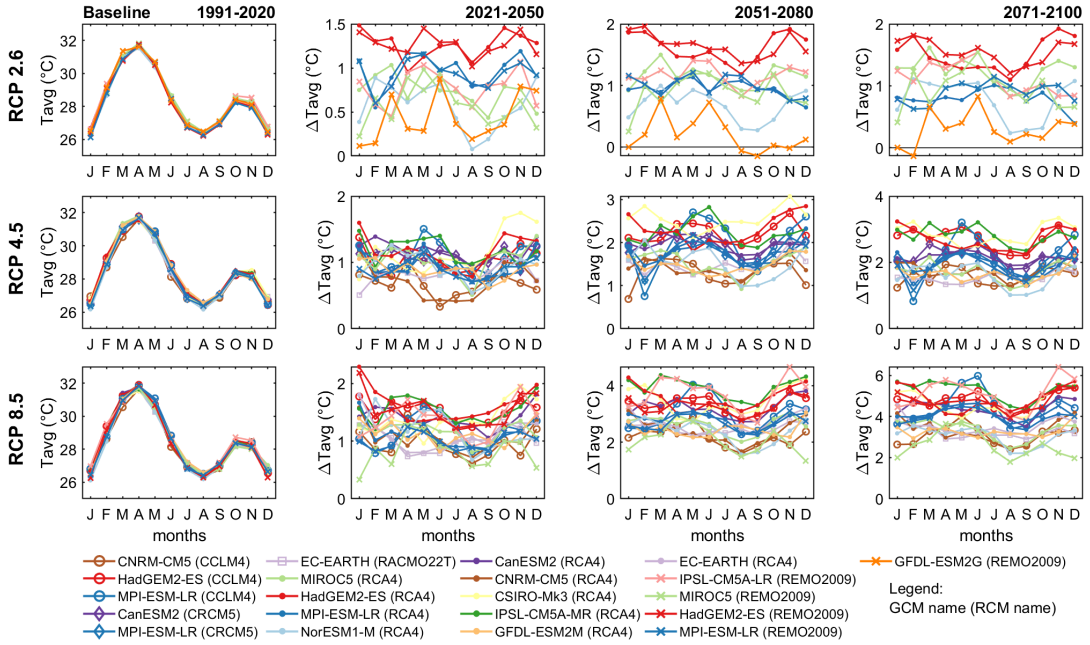


Figure 7.5. Annual cycle of average air temperature (Tavg) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

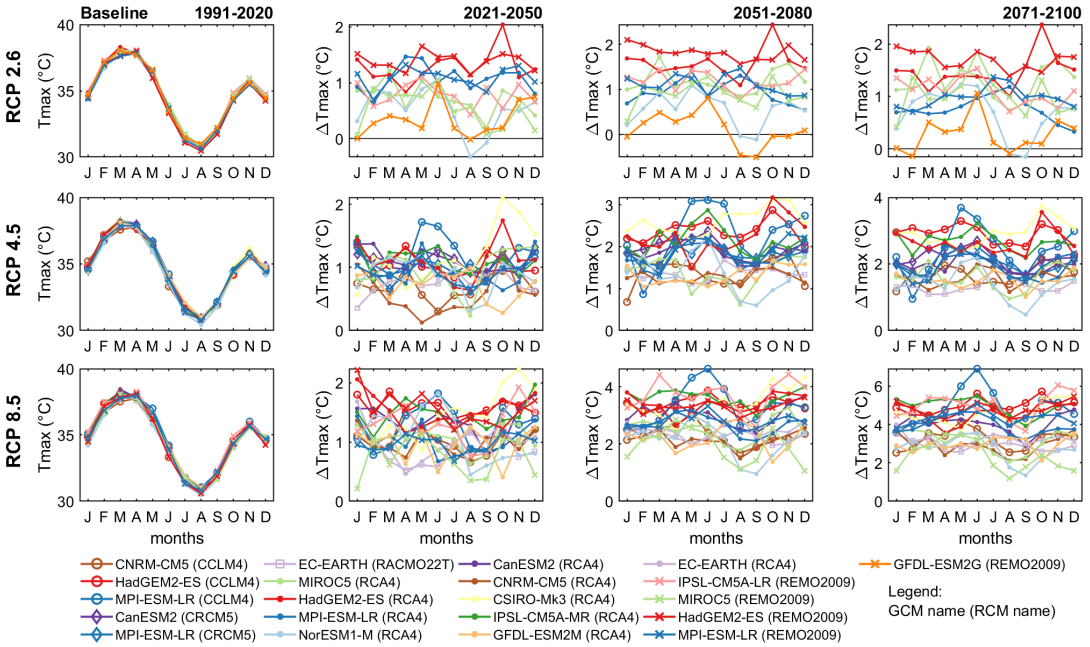


Figure 7.6. Annual cycle of maximum air temperature (Tmax) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

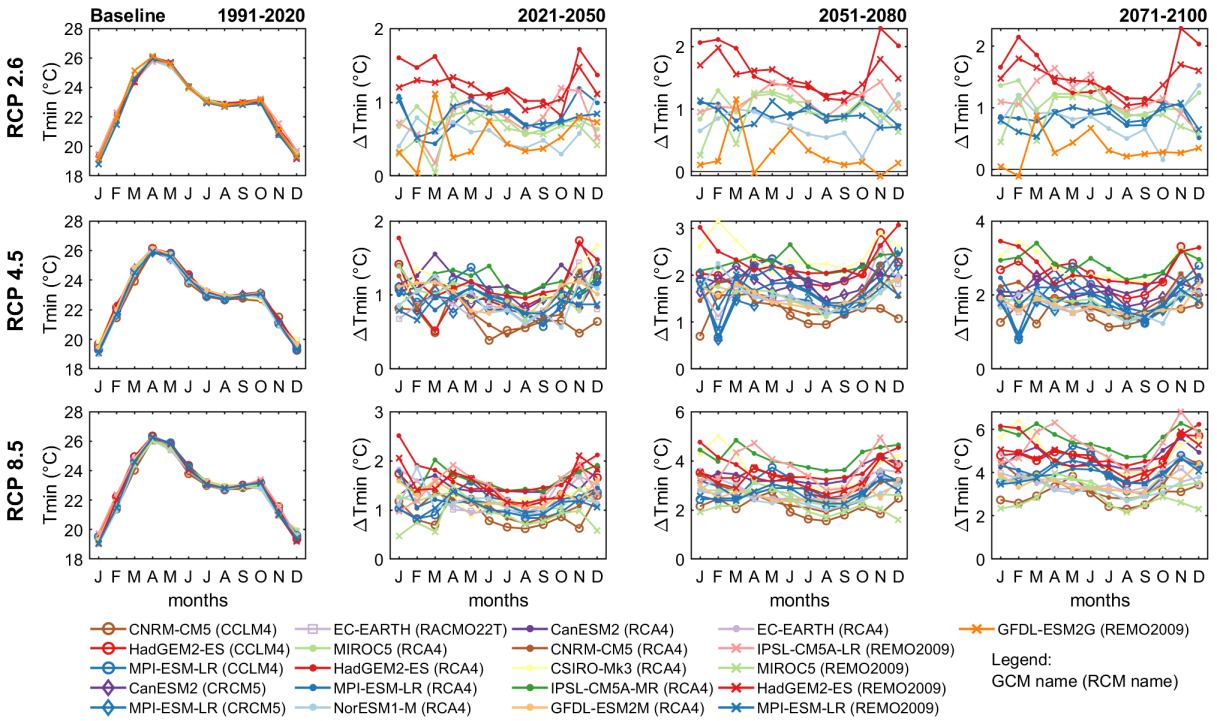


Figure 7.7. Annual cycle of minimum air temperature (Tmin) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.3.3 Potential evaporation

As expected, potential evaporation (E_p) has a bimodal annual cycle similar to that of temperature (Figure 7.8). E_p increases from 207 mm/month in January and peaks at 252 mm/month in March. Then, it decreases to its minimum at 185 mm/month during the rainy season in August, followed by a rise to 221 mm/month in October, and finally decreases to 202 mm/month in December. The inter-model variability is low and ranges from 0.6% during the dry season to 1.2% in the rainy season (June-September). Future projections of monthly E_p shows a concise increase across RCPs and for different periods in the twenty-first century. From 2021 to 2100, monthly E_p is expected to increase by +2.8% for RCP2.6, from +2.3% to 4.6% for RCP4.5, and from 2.6% to 7.8% for RCP8.5, as compared to the historical period 1991-2020. However, the magnitude of the changes in seasonal E_p varies highly across RCM-GCM models with 40% to 76% of variability.

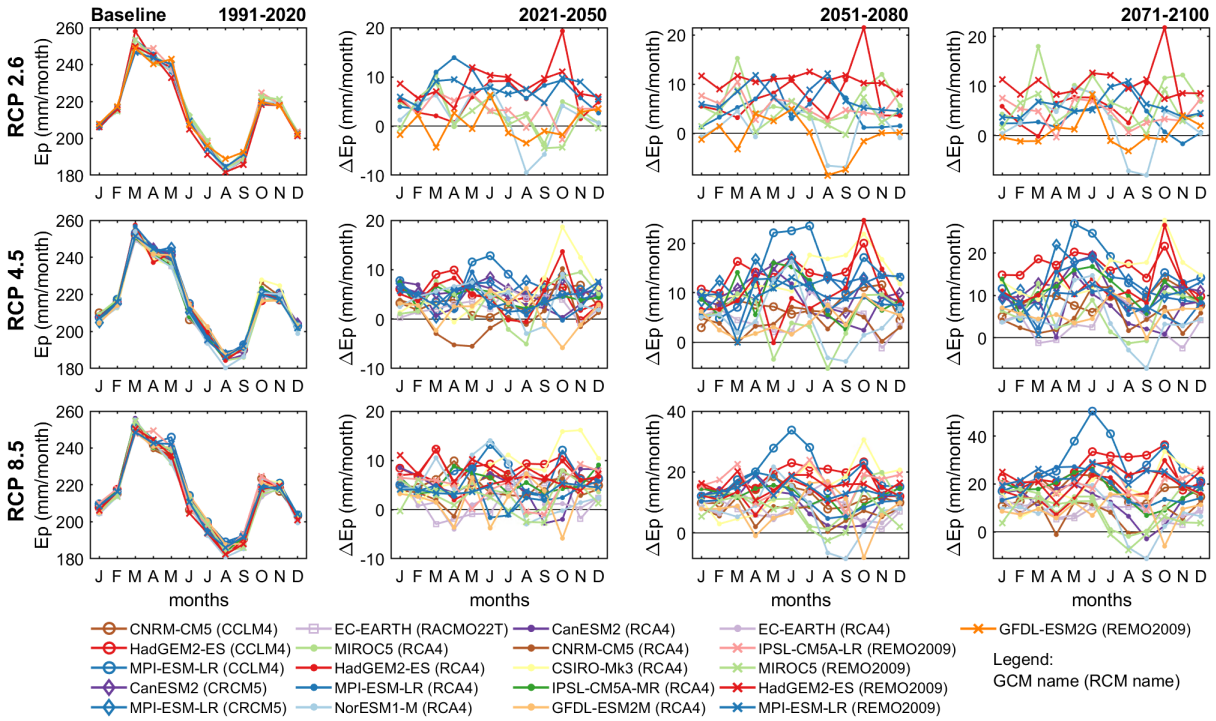


Figure 7.8. Annual cycle of potential evaporation (E_p) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.4 Seasonal changes in hydrology

7.6.4.1 Actual evaporation

Actual evaporation (E_a) follows a unimodal annual cycle similarly to precipitation (Figure 7.9). Monthly E_a increases from 24 mm/month in January to 140 mm/month in September and then decreases to 33 mm/month in December. The inter-model variabilities of monthly E_a is about 1.2% during the rainy season and increases to 8.5% in the dry season. Future projections of monthly E_a denote a contrasting trend between the RCPs. A decline of monthly E_a is expected under RCP2.6 (-7.6% to -6.3%) and RCP4.5 (-8% to -4.4%) between 2021 and 2100, while an increase is expected under RCP8.5 (+6.1% to +13%), with high inter-model variabilities varying between 88% and 97%.

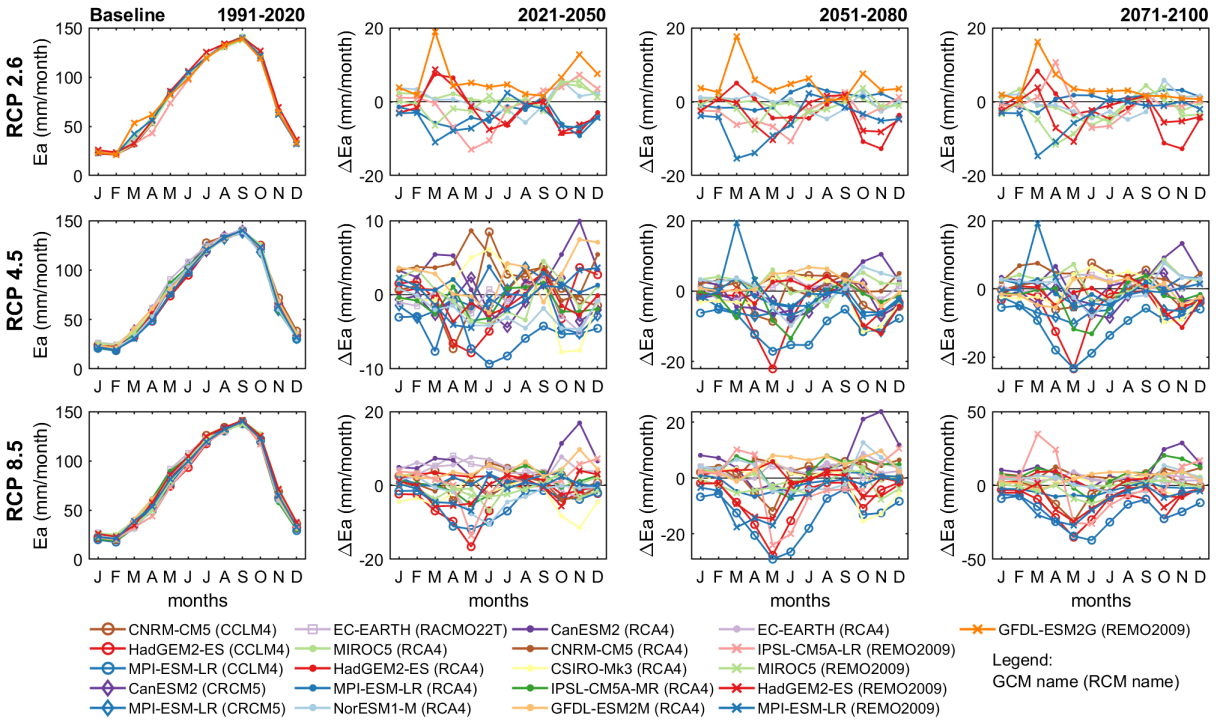


Figure 7.9. Annual cycle of actual evaporation (E_a) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.4.2 Surface runoff

The annual cycle of surface runoff (Q_{run}) presents a unimodal shape with steeper slopes as compared to that of precipitation (Figure 7.10). Monthly Q_{run} increases from 0.9 mm/month in February to 20 mm/month in September, and decreases to 1.2 mm/month in January. The inter-model variabilities range from 20% in the rainy season to 33% in the dry season. Future projections show an increase of monthly Q_{run} until the mid-century (2021-2050) by +43% for RCP2.6, +37% for RCP4.5 and +35% for RCP8.5. While monthly Q_{run} keeps increasing until the end of the twenty-first century for RCP8.5 (+59% to +120%), it decreases for RCP4.5 (-24% to -23%) and RCP2.6 (-24% to -20%), with inter-model variabilities ranging from 90% to 97%.

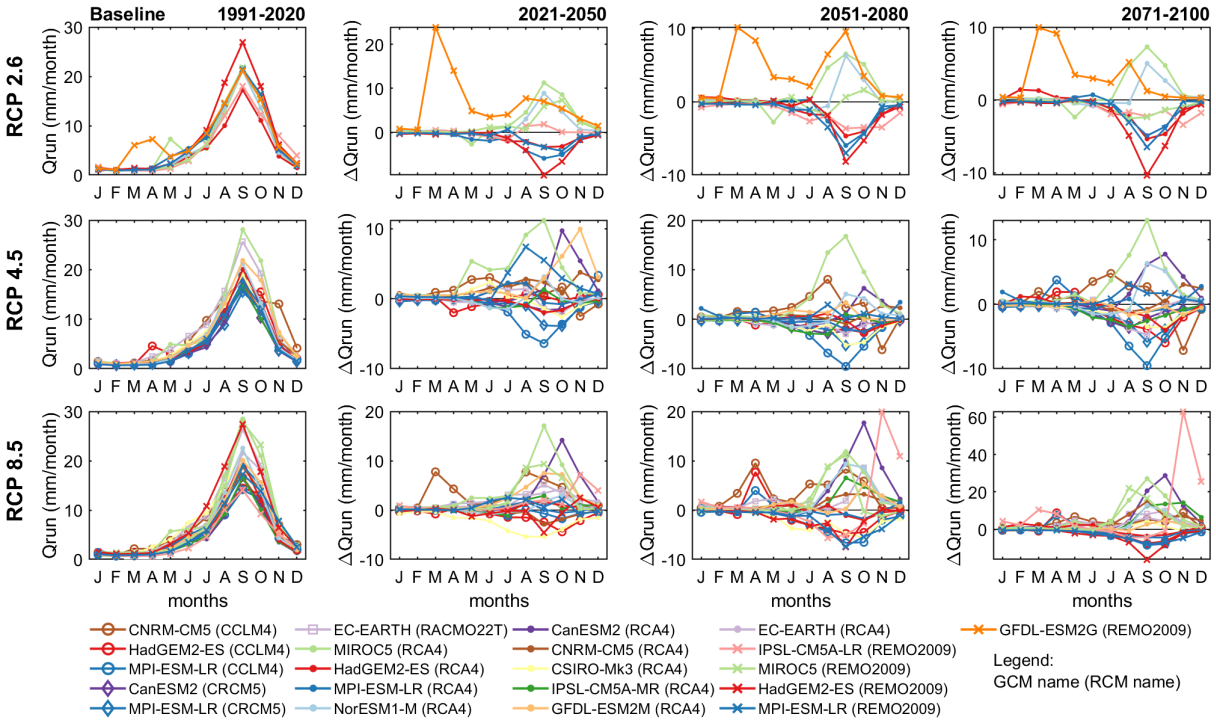


Figure 7.10. Annual cycle of surface runoff (Q_{run}) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.4.3 Streamflow

Streamflow (Q) presents an annual cycle similar to surface runoff (Figure 7.11). Average streamflow in the VRB is calculated based on 11 gauging stations (Figure 2.5, Chapter 2), therefore it does not include flows within the Lower Volta sub-basin. Monthly Q increases from 13 m^3/s in February to 338 m^3/s in September, and decreases to 15 m^3/s in January, with inter-model variabilities varying between 24% in the rainy season and 34% in the dry season. The future projections of monthly Q are similar to Q_{run}. An increase in monthly Q is observed from 2021 to 2050 with variation by +53% for RCP2.6, +41% for RCP4.5 and +44% for RCP8.5. From the mid-century, monthly Q keeps increasing for RCP8.5 (+98% to +154%), while it decreases for RCP4.5 (-26% to -25%) and RCP2.6 (-30% to -25%), with high inter-model variabilities ranging between 91% and 97%.

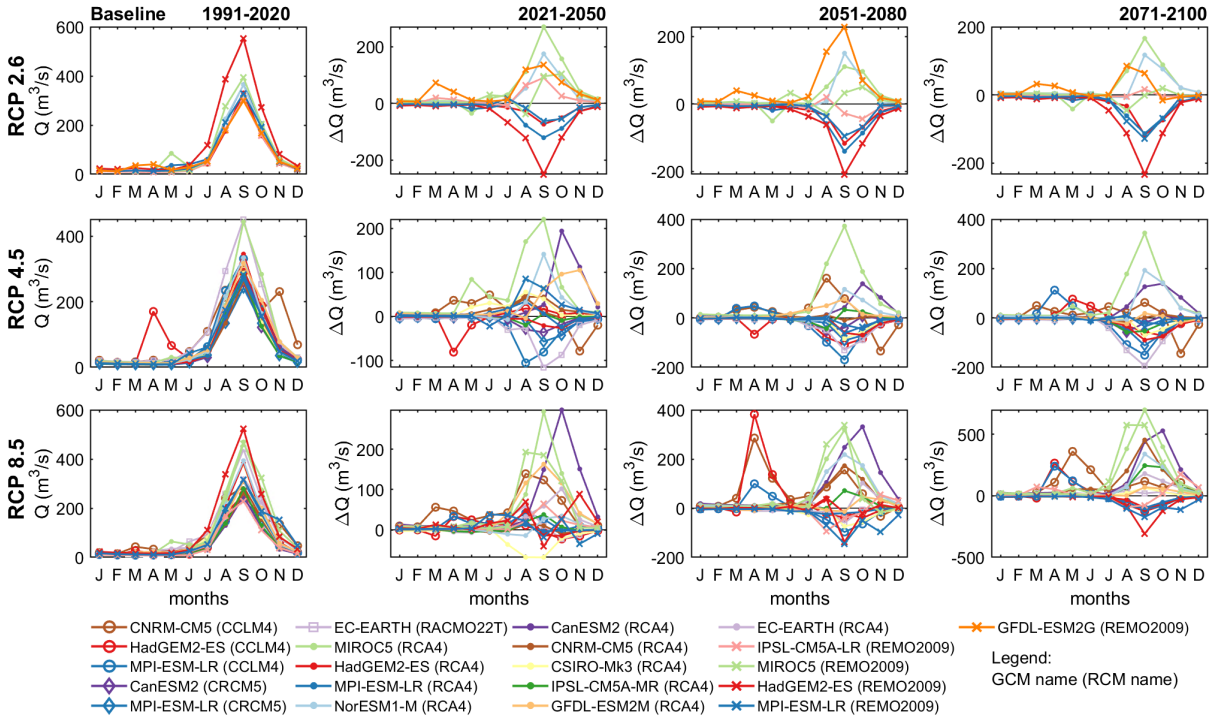


Figure 7.11. Annual cycle of streamflow (Q) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.4.4 Groundwater recharge

The annual cycle of groundwater recharge (R_r) is similar to those of surface runoff and streamflow (Figure 7.12). Monthly R_r increases from 0.02 mm/month in February to 3.7 mm/month in September, and decreases to 0.04 mm/month in January, with inter-model variabilities ranging from 18% in the rainy season to 59% in the dry season. A low groundwater recharge is observed in the VRB because it is mainly generated from precipitation (Williams et al, 2016), while actual evaporation only accounts for 93% of annual precipitation and surface runoff represents 7% of precipitation, based on the RCM-GCM datasets. Future projections of monthly R_r show similar trends to those of Q_{run} and Q. An increase in monthly R_r is expected in the period 2021-2050 by +63% for RCP2.6, +56% for RCP4.5 and +63% for RCP8.5. During the periods 2051-2100, monthly R_r keeps increasing for RCP8.5 (+107% to +307%), while it decreases for RCP4.5 (-28% to -27%) and RCP2.6 (-29% to -25%), with inter-model variabilities varying between 84% and 98%.

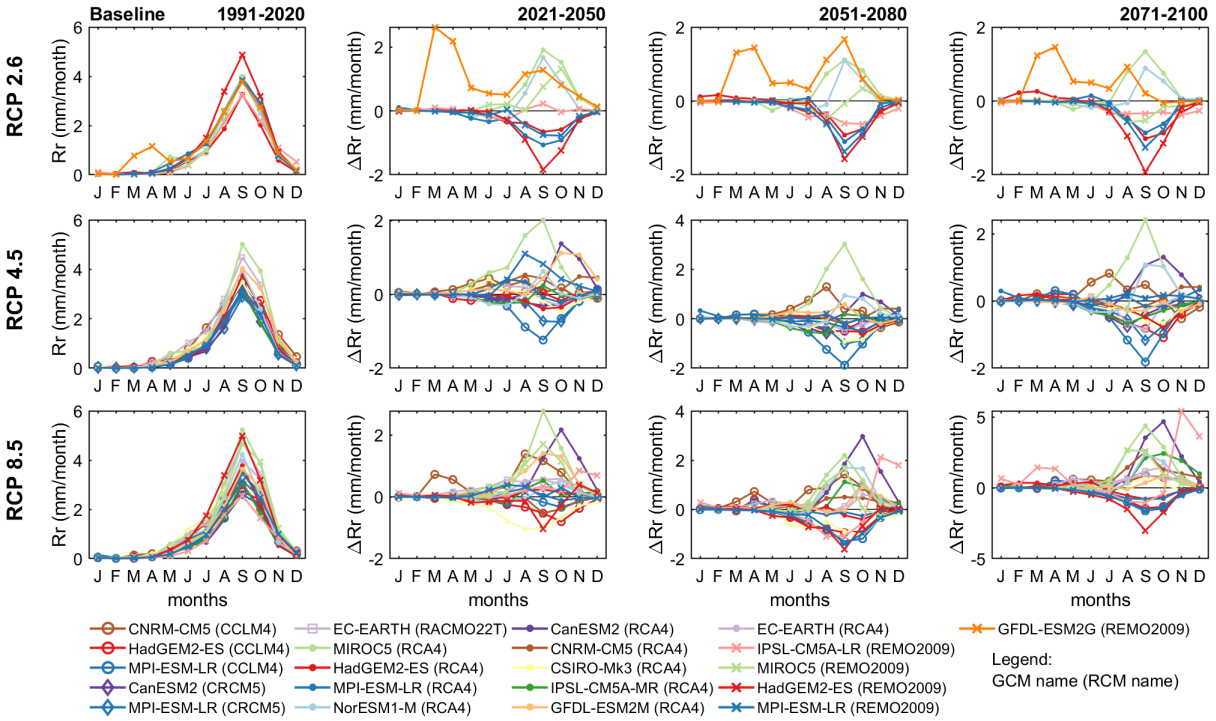


Figure 7.12. Annual cycle of groundwater recharge (R_r) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.4.5 Soil moisture

The root-zone soil moisture (S_u) follows a unimodal annual cycle with a few months forward shift as compared to precipitation (Figure 7.13). Monthly S_u increases from 0.50 mm/mm in April to 0.71 mm/mm in September, and decreases to 0.50 mm/mm in March, with inter-model variabilities ranging from 2.3% in the rainy season to 1.7% in the dry season. Future projections show a decrease of monthly S_u from 2021 to 2100 for RCP2.6 (-3.8% to -3.4%) and RCP4.5 (-4.0% to -2.1%). For RCP8.5, monthly S_u is expected to increase by +2.1% in 2021-2050, then decrease by -4.9% in 2051-2080, and increase again by +3.7% in 2071-2100. The overall inter-model variabilities vary between 84% and 95%.

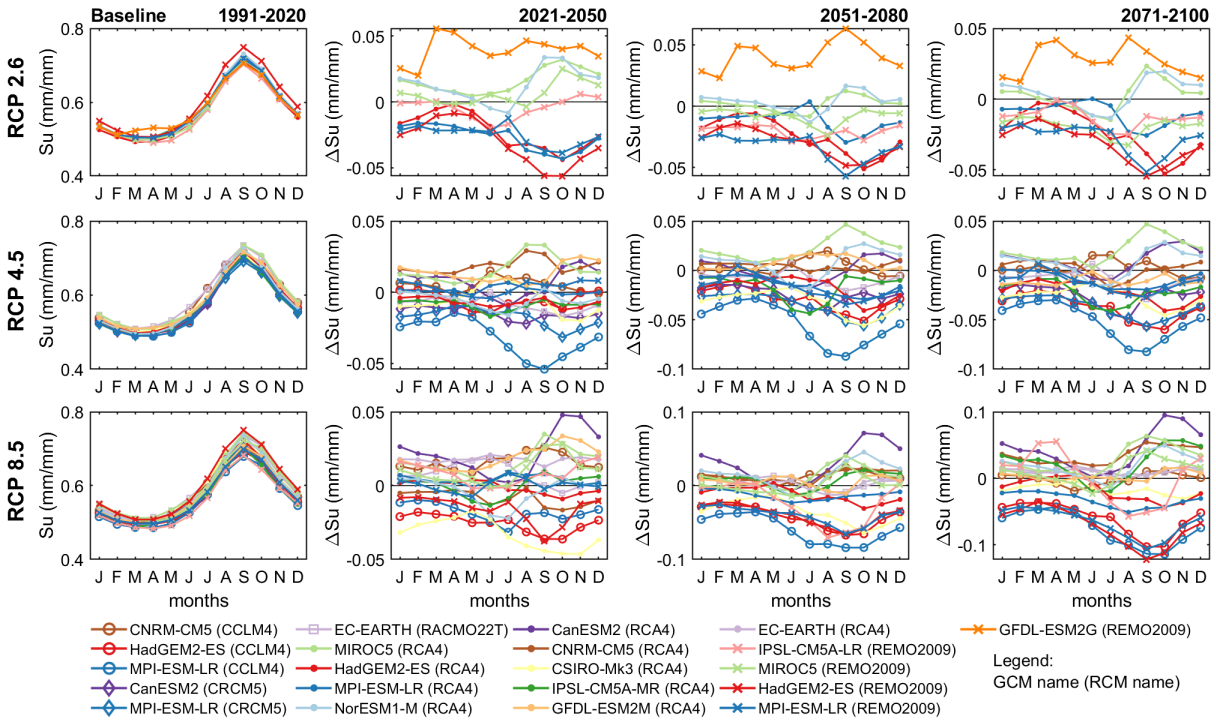


Figure 7.13. Annual cycle of root-zone soil moisture (S_u) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.4.6 Terrestrial water storage

Terrestrial water storage (S_t) is the sum of all water stored below and above land. The annual cycle of S_t is similar to that of S_u (Figure 7.14). Monthly S_t increases from 424 mm in April to 615 mm in September, and decreases to 421 mm in March, with inter-model variabilities varying between 2.6% in the rainy season to 2.1% in the dry season. Future projections of S_t are similar to those of S_u over the twenty-first century. A decline of monthly S_t is expected from 2021 to 2100 for RCP2.6 (-4.2% to -3.7%) and RCP4.5 (-4.3% to -2.3%). For RCP8.5, monthly S_t is expected to increase by +2.7% in 2021-2050, then decrease by -5.4% in 2051-2080, and increase again by +5.1% in 2071-2100. The overall inter-model variabilities vary between 89% and 96%.

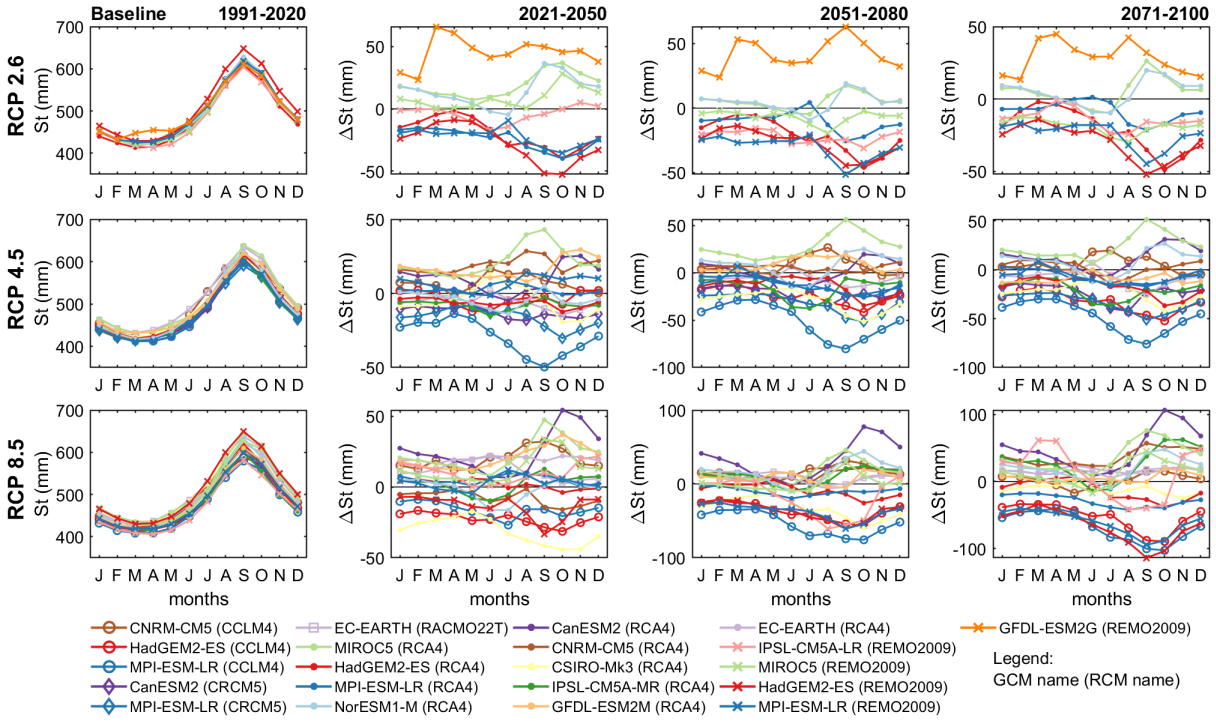


Figure 7.14. Annual cycle of terrestrial water storage (S_t) for the reference period (1991-2020) under different RCPs (first column) and projections of changes for future periods (2021-2100) as compared to the reference period for various RCM-GCM combinations

7.6.5 Annual mean changes in climatic and hydrological processes

Considering all RCPs over the historical period 1991-2020 in the VRB, the multi-model ensemble mean of long-term annual estimates of climatic variables is as follows: T_{max} = 34.7 °C, T_{min} = 22.8°C, T_{avg} = 28.4 °C, E_p = 2579 mm/year, P = 999 mm/year.

The future projections of climatic and hydrological processes over the twenty-first century (2021-2100) as compared to the historical period (1991-2020) show contrasting trends between RCPs and between projections periods (Table 7.2). All climatic variables (precipitation, temperature and potential evaporation) are expected to increase over the twenty-first century for all RCPs and projection periods, at the exception of precipitation that is expected to decrease under RCP2.6 and RCP4.5, and increase under RCP8.5. The average annual changes in climatic variables over 2021-2100 are amplified with increasing radiative forcing levels, and vary between RCP2.6 (T_{max} = +2.8%, T_{min} = +4.1%, T_{avg} = +3.4%, E_p = +2.4%, P = -4.2%), RCP4.5 (T_{max} = +4.6%, T_{min} = +7.1%, T_{avg} = +5.7%, E_p = +3.4%, P = -3.0%) and RCP8.5 (T_{max} = +7.5%, T_{min} = +12%, T_{avg} = +9.5%, E_p = +5.0%, P = +6.2%). More importantly, there is a 100% agreement among the RCM-GCM models concerning the increase in temperature and 98.8% of agreement for the increase of potential evaporation across all the RCPs and projection

periods. Therefore, there is a high confidence in future climate warming of the VRB region in West Africa during the twenty-first century. These findings align with previous studies in West Africa (e.g., Dosio *et al.*, 2020; Jin *et al.*, 2018; Todzo *et al.*, 2020). In agreement with Dosio *et al.* (2019), a lower RCM-GCM agreement of 62% is found for future changes in precipitation, thereby highlighting the complexity of modelling climate in West Africa (Fitzpatrick *et al.*, 2020; Panthou *et al.*, 2012) . The relatively low RCM-GCM agreement for precipitation can be justified by the difficult representation of the West African monsoon in most of the climate models (Akinsanola *et al.*, 2020; Philippon *et al.*, 2010; Xue *et al.*, 2010).

In contrast to the climatic variables, the hydrological processes (streamflow, surface runoff, actual evaporation, groundwater recharge, soil moisture and terrestrial water storage) decrease under RCP2.6 and RCP4.5, and only increase under the RCP8.5 scenario during the period 2021-2100 (Table 7.2). The multi-model ensemble mean of long-term annual estimates of hydrological processes for all RCPs over the historical period is as follows: E_a = 930 mm/year, Q_{run} = 73 mm/year, Q = 973 m^3/s, R_r = 18 mm/year, S_u = 0.58 mm/mm and S_t = 491 mm. The average annual changes in hydrological variables over 2021-2100 vary between RCP2.6 (E_a = -2.9%, Q_{run} = -2.7%, Q = -3.1%, R_r = -21%, S_u = -3.7%, S_t = -4.0%), RCP4.5 (E_a = -2.5%, Q_{run} = -11%, Q = -2.6%, R_r = -13%, S_u = -3.4%, S_t = -3.7%) and RCP8.5 (E_a = +4.2%, Q_{run} = +42%, Q = +84%, R_r = +37%, S_u = +2.3%, S_t = +3.2%). It can be concluded from these findings that an intensification of the hydrological cycle in the VRB is to be expected only under the RCP8.5 scenario. Moreover, it is noteworthy that surface runoff and streamflow are expected to increase while precipitation decreases under the RCP2.6 and RCP4.5 during the period 2021-2050. This paradoxical phenomenon of rainfall-runoff correlation is commonly referred to as the "Sahelian paradox" (Mahé and Paturel, 2009). Although the Sahelian paradox is not completely understood yet, several possible reasons include changes in soil properties and vegetation cover that can lead to a reduction in the soil water holding capacity (Gal *et al.*, 2017). The persistence of the Sahelian paradox under future climate scenarios can be justified by the fact that rainfall is expected to decrease in frequency but increase in intensity (Todzo *et al.*, 2020), thereby leading to a rapid saturation of the soil, which limits infiltration and increases runoff. Moreover, the parameters of the mHM model are constant from the historic to the future modelling periods, so that the runoff generation mechanisms do not change over time. The necessity for hydrological model recalibration under climate change should be further investigated in the future.

Table 7.2. Summary of annual projections of climatic and hydrological variables under different representative concentration pathways (RCPs) from the historical period (1991-2020) to different future periods (2021-2100). The trend of the changes is determined by the majority of the models agreeing on the same direction of change, which is expressed by the percentage of RCM-GCM agreement. V_2 is the second-order coefficient of variation.

Annual data	Baseline 1991-2020		**RCP2.6**					
			Future projections (%)			% RCM-GCM agreement		
	Ensemble mean	V2 (%)	2021-2050	2051-2080	2071-2100	2021-2050	2051-2080	2071-2100
Tmax (°C)	34.7	0.3	+2.5	+3.0	+3.0	100.0	100.0	100.0
Tmin (°C)	22.8	0.5	+3.5	+4.4	+4.4	100.0	100.0	100.0
Tavg (°C)	28.4	0.4	+2.9	+3.6	+3.6	100.0	100.0	100.0
Ep (mm/year)	2577	0.4	+2.0	+2.7	+2.4	100.0	88.9	100.0
P (mm/year)	1005	2.7	-4.7	-4.4	-3.4	55.6	66.7	77.8
Ea (mm/year)	933	2.0	-2.9	-3.1	-2.6	66.7	77.8	77.8
Q (mm/year)	77	14.1	+34.9	-23.0	-20.1	55.6	55.6	66.7
Rr (mm/year)	12.4	13.7	-20.2	-20.6	-21.2	55.6	66.7	66.7
Su (mm/mm)	0.58	1.4	-3.8	-3.7	-3.6	55.6	66.7	66.7
St (mm)	494	1.6	-4.0	-4.0	-4.0	55.6	66.7	66.7
Q (m3/s)	1025	21.9	+47.8	-29.6	-27.6	55.6	55.6	55.6

Annual data	Baseline 1991-2020		**RCP4.5**					
			Changes in variable (%)			% RCM-GCM agreement		
	Ensemble mean	V2 (%)	2021-2050	2051-2080	2071-2100	2021-2050	2051-2080	2071-2100
Tmax (°C)	34.7	0.3	+2.8	+5.2	+5.9	100.0	100.0	100.0
Tmin (°C)	22.8	0.4	+4.3	+7.9	+9.0	100.0	100.0	100.0
Tavg (°C)	28.4	0.3	+3.5	+6.4	+7.3	100.0	100.0	100.0
Ep (mm/year)	2585	0.4	+1.9	+4.0	+4.4	100.0	100.0	100.0
P (mm/year)	991	3.1	-2.4	-4.9	-1.7	56.3	62.5	50.0
Ea (mm/year)	927	2.1	-1.6	-4.3	-1.7	68.8	62.5	50.0
Q (mm/year)	69	18.7	+7.2	-18.7	-20.9	50.0	56.3	56.3
Rr (mm/year)	11.2	17.6	+5.2	-20.2	-22.8	50.0	56.3	56.3
Su (mm/mm)	0.57	1.7	-2.0	-4.1	-4.2	56.3	68.8	68.8
St (mm)	488	2.0	-2.1	-4.6	-4.4	56.3	62.5	68.8
Q (m3/s)	909	23.1	+34.0	-20.7	-21.1	62.5	56.3	56.3

Annual data	Baseline 1991-2020		**RCP8.5**					
			Changes in variable (%)			% RCM-GCM agreement		
	Ensemble mean	V2 (%)	2021-2050	2051-2080	2071-2100	2021-2050	2051-2080	2071-2100
Tmax (°C)	34.7	0.4	+3.4	+8.1	+11.1	100.0	100.0	100.0
Tmin (°C)	22.9	0.4	+5.5	+13.0	+17.9	100.0	100.0	100.0
Tavg (°C)	28.4	0.4	+4.3	+10.2	+14.0	100.0	100.0	100.0
Ep (mm/year)	2576	0.5	+2.2	+5.5	+7.3	100.0	100.0	100.0
P (mm/year)	999	3.9	+4.6	+5.5	+8.6	61.1	55.6	72.2
Ea (mm/year)	930	2.6	+3.1	+3.7	+5.9	55.6	55.6	61.1
Q (mm/year)	74	21.2	+27.4	+31.6	+65.6	72.2	66.7	72.2
Rr (mm/year)	11.9	20.9	+24.6	+29.2	+58.0	72.2	61.1	66.7
Su (mm/mm)	0.58	2.2	+1.9	+1.7	+3.2	61.1	55.6	55.6
St (mm)	491	2.7	+2.6	+2.6	+4.3	66.7	55.6	61.1
Q (m3/s)	985	25.3	+37.4	+76.8	+137.5	83.3	72.2	77.8

The percentage of agreement of 62% for hydrological projections between the climate models is similar to that of precipitation, which supports that precipitation is the key driver of hydrological processes in the VRB. Therefore, the improvement of precipitation representation in climate models would ultimately enhance the reliability of the assessment of climate change impacts on water resources.

7.7 Implications of climate change for water resources management

The future evolutions of hydrological processes in the VRB will have implications for the recurrence of extreme monthly hydroclimatic events in the region. The increase of surface runoff and streamflow in the near future (2021-2050) is confirmed among all the RCPs, thereby foreseeing an increase in the likelihood of floods in the VRB, as also reported by Jin *et al.* (2018). However, the decline in root-zone soil moisture is a warning for the occurrence of agricultural droughts, which will have dire consequences on the livelihoods of the rural population that relies on agriculture. The socioeconomic development of the region might also be hampered under these conditions.

Therefore, it is crucial to deploy sound water resources management practices that can help to cope with these intolerable impacts of climate change in the VRB. Rainfall water harvesting systems should be deployed in the region. Managed aquifer recharge techniques can also help slacken surface runoff and increase groundwater recharge, which is a non-negligible source of water for irrigation in some regions (McCartney *et al.*, 2012; Williams *et al.*, 2016). Moreover, there is a high potential for expanding agriculture as potential evaporation is expected to increase over the twenty-first century, thereby setting conditions to grow more crops if water availability is improved. Increasing water availability by the construction of infrastructures (e.g. dams) is vital for off-season agriculture as well as hydropower production. However, environmental and social consequences should be well thought upfront. These solutions among many others, accompanied by innovative initiatives like farmer-led irrigation, can help balance the water-energy-food nexus in the VRB, thereby providing a solid foundation for a sustainable socioeconomic development. However, such progress is only possible under a strong collaboration between climate and water resources scientists, development practitioners and policymakers. Finally, a better collaboration among the six riparian states for the management of their common water resources is key to bolster resilience and foster regional development.

7.8 Discussions

The findings of this study provide a comprehensive overview of the impacts of climate change on water resources in the VRB over the twenty-first century. The analyses are done for the entire basin, so the results might differ for sub-basins or climatic zones. As such, the findings are valid for the development of regional adaptation and mitigation strategies of climate change.

The large ensemble of RCM and GCM datasets along with the three RCPs are the basis of the robustness of this study. However, uncertainties exist in the climate projection datasets, mainly for precipitation, as demonstrated by the results. These uncertainties are propagated in the hydrological modelling as well. However, uncertainties in this study are quantified in terms of inter-model variabilities, which provide more insights into the findings. Moreover, only using the set of RCM-GCM projections that agree on the direction of future changes to predict the evolution of climatic and hydrological processes has both advantages and drawbacks. On the one hand, the advantages of a multi-model mean based on the majority of agreement are that models with contradictory projections are not mixed in the analysis, which avoids ending up with a multi-model mean that might not align with any of the models in the ensemble, and avoids dampening the magnitude of the change. On the other hand, adopting a multi-model mean based on the majority of agreement among RCM-GCM projections imposes the risk of falling in a "democracy-based decision" where the majority is not necessary right. However, conserving the trends and magnitudes of future projections as adopted in this study should be prioritized by avoiding mixing all RCM-GCM projections independently from their change direction in a multi-model mean, because the later will be influenced by the magnitude of each model, thereby leading to biases towards potential outliers. In fact, Rodrigues *et al.* (2014) showed that the multi-model mean is not necessarily the best approach.

The R2D2 method used for multivariate bias correction assumes stationarity in the inter-variable relationships, which might not hold over very long periods. However, predicting non-stationarity of biases under climate change is not straightforward. With advances in multivariate bias correction methods (François *et al.*, 2020), the added value of methods that consider non-stationarity in climate (Robin *et al.*, 2019) should be investigated in climate impact studies on water resources.

Although multiple RCMs, GCMs and RCPs are used in this study, a single hydrological model is used for the hydrological projections. Therefore, the results are also subject to potential deficiencies of the mHM model. However, the mHM model used in this study has been thoroughly calibrated to provide realistic simulations of hydrological state variables and fluxes

in the VRB, which is evaluated with the Budyko framework, and further described in previous studies (Dembélé *et al.*, 2020a; Dembélé *et al.*, 2020b; Dembélé *et al.*, 2020c). The calculation of potential evaporation based only on temperature data could lead to a strong increase in actual evaporation without accounting for changes in other variables that influence potential evaporation. However, the increase in temperature and potential evaporation that does not systematically lead to an increase in actual evaporation (Table 7.2). The method of Hargreaves and Samani (1985) was found reliable in estimating potential evaporation in the VRB (Dembélé *et al.*, 2020a; Dembélé *et al.*, 2020b) and similar semi-arid regions (Bai *et al.*, 2016; Er-Raki *et al.*, 2010; Gao *et al.*, 2017). The plausibility of the simulated potential and actual evaporation is further assessed in this study with the Budyko framework.

Furthermore, the hydrological projections focus on changes in climate and do not explicitly account for land use land cover change or changes in water management practices in the VRB. Although, land use and land cover changes play an important role in the production of hydrological processes, the primary focus in this study is climate change. However, land use changes are assumed to be accounted for to some extent in the RCPs, as their development is based on assumptions regarding future evolution of land use and land cover (Van Vuuren *et al.*, 2011). Future studies of climate change impacts on water resources should consider using different hydrological models, which might unveil the impacts of the choice of hydrological models in impact studies. Combined assessment of the impacts of climate change, land use land cover change, and water management practices could also provide new insights on the evolution of water resources in future studies. An update of impact studies should be continuously undertaken, considering advances in climate sciences such as the development of the new Shared Socioeconomic Pathways (SSPs) for instance (O'Neill *et al.*, 2014; O'Neill *et al.*, 2017; Riahi *et al.*, 2017).

Finally, more work is needed to improve the modelling of the West Africa monsoon and its representation in convection-permitting climate models (Berthou *et al.*, 2019; Kendon *et al.*, 2017; Kendon *et al.*, 2019). The conjunction of all these efforts will improve the accuracy of climate projections among the RCMs and GCMs, thereby improving the reliability of hydrological projections and fostering the adoption of the findings of impact studies by decision makers.

7.9 Conclusion

A large ensemble of twelve GCMs from CMIP5 and five RCMs from CORDEX-Africa is used to investigate the impacts of climate change on water resources in the Volta River basin under three RCPs. The climate projection datasets are used to force the fully distributed mesoscale Hydrologic Model (mHM) over the twenty first century. Changes in hydrological processes over the period 2021-2100 are estimated relatively to the historical period 1991-2020. The results reveal a contrasting intensification of the hydrological cycle depending on RCPs. The key findings can be summarized as follows:

- Climate warming is confirmed in the Volta basin as all RCM-GCM projections predict an increase in minimum, maximum and average surface air temperature under all RCPs. Similarly, potential evaporation is projected to increase.
- Precipitation is projected to decrease under RCP2.6 and RCP4.5, while an increase is foreseen under RCP8.5. Compared to temperature, there are more uncertainties in the trend of the changes in precipitation as there is only an agreement of 62% between the RCM-GCM projections.
- Actual evaporation, groundwater recharge, soil moisture and terrestrial water storage decline under RCP2.6 and RCP4.5, while they increase under RCP8.5.
- Only surface runoff and streamflow keep increasing under all RCPs and all over the twenty first century.
- A clear intensification of the entire hydrological cycle is foreseen only under RCP8.5, as all fluxes and state variables are expected to increase.

The changes in the hydrological cycle will have implications for future floods and droughts in the Volta basin, thereby amplifying the vulnerability of the local population to climate change. Sound water management practices are therefore required to bolster resilience and foster socioeconomic development. These finding can contribute to the elaboration of regional adaptation and mitigation strategies of climate change. However, significant inter-model variabilities highlight the complexity and uncertainties related to the assessment of climate change impacts on water resources. Therefore, more work is required to improve climate modelling in West Africa. A strong collaboration between climate and water resources scientists, practitioners and policymakers is key for advancing knowledge and development.

Chapter 8

Water Accounting for Sustainable Water Resources Management in the Volta River Basin*

What gets measured gets managed.

Peter Drucker

What we know is a drop,

what we don't know is an ocean.

Isaac Newton

* This chapter is in preparation for publication.

Abstract

Water management critically relies on quantified and reliable information on water resources and uses, which help decision makers to understand and deploy sound policies to cope with water scarcity and sustain water security. Water accounting frameworks are useful tools for reporting on water resources. This study demonstrates the possibility to comprehensively report on current and future conditions of water resources using the Water Accounting Plus (WA+) framework combined with hydrological modelling. The fully distributed mesoscale Hydrologic Model (mHM) is used to predict water fluxes, stocks and flows for the historical period 1999-2020 and the near term future period 2021-2050. Data of climate change projections are obtained from a large ensemble of nine global climate models (GCMs) and four regional climate models (RCMs) under the representative concentration pathway RCP8.5. Hydrological processes derived from mHM are used to feed the WA+ framework and provide an evidence-informed reporting on the state and trends of water resources in the Volta River basin located in West Africa.

The results show that the long-term net inflow in the basin over the period 1991-2020 is 388.6 km^3/year (936 mm/year) with 96% attributable to rainfall (374.6 km^3/year or 901 mm/year), and is projected to increase by +5% in the near future. However, only 8% of the net inflow is exploitable as blue water stored in reservoirs, lakes, streams, and aquifers. The remainder of the net inflow is depleted through landscape evaporation known as green water consumption, which dominantly occurs at 55% over lands that are not managed by human. The available water for various water uses in the basin is 14.6 km^3/year or 35 mm/year, of which 79% are actually utilized, while the remainder 21% are utilizable but are not consumed. The non-utilizable water is 8.3 km^3/year and represents 27% of the exploitable water, while the non-recoverable water ascribed to groundwater recharge and water pollution constitutes 28% of the exploitable water. Only 42% of water use is beneficial for the intended purposes, with agriculture representing 35% of the beneficial water consumption. Future projections show an increase of +20% in the exploitable water fraction, while the available water fraction is expected to decrease by -5%. These findings show that climate change could disproportionally affect the exploitable and the available water, thereby calling for adaptation measures. Finally, integrating climate change scenarios in water accounting should be further investigated as it is a way forward in improving water governance in transboundary basins.

8.1 Introduction

Chronic water scarcity is a big threat to prosperity as it could cost some regions up to 6% of their gross domestic product, thereby exposing some countries to fragility and instability (Damania, 2020; World Bank Group, 2016). Water scarcity is considered as an imbalance between water supply and demand, with an excess of water demand over available supply (Steduto *et al.*, 2012). Water scarcity results from constraints on the water availability in quantity and quality, on the accessibility by means of reliable supply or from the lack of infrastructure due to financial and technical restrictions. As such, water scarcity can be addressed with appropriate water management policies by deploying strategies for supply enhancement or demand management.

With 9 billion people to feed by 2050, it is critically becoming important to improve water use efficiency and crop water productivity (Godfray *et al.*, 2010; Guillou and Matheron, 2014). However, climate change and other development issues including urbanization, land degradation and deforestation are expected to exacerbate water scarcity, with implications for food insecurity and conflicts between those who share the resources (Zeitoun *et al.*, 2016).

Information on the state of water resources including inflows, outflows and water uses in a river basin is essential for sustainable water management. However, data unavailability and inaccessibility limit water resources management in many regions around the world (Karimi *et al.*, 2013b). In transboundary basins, where water is a vital source for regional socio-economic development for the riparian countries, sound and transparent management of natural resources is key for geopolitical stability. The absence of good information systems hamper long-term water management and development planning, which requires adequate tools for measuring, reporting and monitoring water resources. In this context, key tools for water resources assessment are water accounting frameworks.

Water accounting is the systematic organization and presentation of information on the status and trends in water supply, demand, accessibility and use in time and space within specified domains (Batchelor *et al.*, 2016; Steduto *et al.*, 2012). Water accounting serves as a basis for evidence-informed decision-making and should be used for any policy development and strategy to cope with water scarcity. Existing water accounting systems include the International Water Management Institute (IWMI) water accounting framework (Molden and Sakthivadivel, 1999), the System of Environmental and Economic Accounting for Water (SEEAW) of the UN Statistics Division (DESA, 2012), the Australian Water Accounting Conceptual Framework (Merz, 2006), the UNEP ’s Water Footprint, Neutrality, and Efficiency

(WaFNE) (Morrison *et al.*, 2010), and the Water-use accounts framework of the CPWF (Kirby *et al.*, 2010). However, none of the water accounting frameworks resulting from these initiatives has been adopted as a general standard (Dost *et al.*, 2013). Reasons for this include the fact that they present results without differentiating between managed, manageable and unmanageable water flows, their outputs are usually too complex for decision making, and input data are often not available or are based on expensive long-term monitoring activities (Karimi, 2014).

More recently, the Water Accounting Plus (WA+) framework was developed to address the shortcomings of previous water accounting frameworks (Karimi *et al.*, 2013a). WA+ provides estimates of manageable and unmanageable water flows, stocks, consumption among users, and interactions with land use. Land use is grouped into four major clusters that differ in terms of water management, namely, protected land use, utilized land use, managed land use and managed water use. The results of WA+ are presented with standardized accounting sheets accompanied with a set of indicators that summarize complex hydrological processes in a more accessible format to different water professionals. The core of the WA+ methodology is based on water balance calculation, spatial analysis, remote sensing, geographic information system, and spatial modelling. WA+ is a valuable tool for water resources planning and development, especially in data scarce regions, ungauged locations and transboundary basins because it primarily relies on remotely sensed data. Because of the usually coarse spatial resolution of satellite data, WA+ based on remote sensing data is more suited for large river basins and for regional studies. However, there are difficulties in closing the water balance when using only remote sensing data (FAO and IHE Delft, 2020).

Alternatively, hydrological models can be used to provide input data to the WA+ framework because they can close the water balance via simulations, and the sources of uncertainties in the components of the water cycle can be tracked as opposed to using various sources of remote sensing data.

In spite of the limitations of earth observation data, they are still a valuable source of information for water resources management, and they have a good potential for improving large scale hydrological modelling (Dembélé *et al.*, 2020a; Dembélé *et al.*, 2020b), which would ultimately enhance the information needed for WA+. Moreover, WA+ is usually done for the past or current periods, while decision-making for sustainable water resources management is becoming increasingly necessary as climate change is affecting water resources. Developing WA+ with climate change scenarios can help decision makers to elaborate and implement adaptation and mitigation strategies.

This study provides a comprehensive analysis of the current and future states of water resources using the Water Accounting Plus (WA+) framework in the Volta River basin (VRB) located in West Africa. The VRB is a transboundary basin covering about 415,600 km^2 shared among six countries, namely, Benin, Burkina Faso, Côte d'Ivoire, Ghana, Mali and Togo. Burkina Faso and Ghana alone share 82.5% of the basin total area (Table 2.1 in Chapter 2). The population of the VRB is essentially rural and represents 70% of the basin total population. The annual population growth rate is 2.5%, which implies a doubling of the population every 28 years (Rodgers *et al.*, 2006). In 2010, 23.8 million people were living in the VRB and the population is projected to reach 38.4 million in 2030 (Williams *et al.*, 2016). Water resources in the VRB play an important role in the socio-economic development, especially in agriculture, hydropower production, aquaculture and domestic water supply. They provide additional livelihood for the rural populations that are mostly active in the agricultural sector (van de Giesen *et al.*, 2001).

Water demand in the VRB is projected to increase by more than 1000% between 2000 and 2025 (Biney, 2010), which poses challenges for transboundary water resources management. First, the rainfall in the VRB is erratic and with high spatiotemporal and inter-annual variabilities, which is expected to be exacerbated under climate change (Nicholson *et al.*, 2018b). Secondly, countries in the VRB have different national priorities in terms of water use. The upstream consumptive use of water in Burkina Faso is essentially dominated by agriculture. As Burkina Faso occupies the direst part of the VRB, the priority is the construction of small and medium reservoirs to develop irrigated agriculture. While the downstream priority in Ghana is the production of hydroelectricity from the Lake Volta's Akosombo dam (De Condappa and Lemoalle, 2009). In spite of progress in water governance, the diverging water consumption priorities and water management differences remain sources of tension between both states (Biney, 2010). An independent and unbiased assessment of the spatiotemporal availability of water and various uses could potentially alleviate these tensions.

The proposed methodology is based on the use of climate change projection datasets selected based on previous findings on the impacts of climate change on water resources in the VRB (Chapter 7). The selected climate projection datasets comprise nine global climate models (GCMs) and four regional climate models (RCMs) under the RCP8.5 scenario. Climate projection datasets are used to force the fully distributed mesoscale Hydrological Model (mHM), and thus to simulate hydrological variables necessary for WA+. The mHM model previously demonstrated very good performances in the VRB (Dembélé *et al.*, 2020b). Finally,

the WA+ framework is used to provide a comprehensive view of water resources for the historical period (1991-2020) and the near term future (2021-2050). Future land use and land cover scenarios are not integrated in this study because the goal is firstly to understand the impacts of climate change on water resources, which might provide insights on future land use practices that can help cope with water scarcity and improve water security (Cook and Bakker, 2012). Besides the application of an existing methodology to a highly relevant case study, the major contribution to the current state-of-the-art is the combined use of climate change scenarios and hydrological modelling to feed the WA+ framework and to assess future water accounts in the VRB, which is lacking in previous studies (e.g., Delavar *et al.*, 2020; Hunink *et al.*, 2019; Karimi *et al.*, 2013b). The following research questions are addressed:

- What is the state and the evolution of water resources under climate change in the VRB?
- How does water consumption relate to different land use and management practices?

The current research aims at bringing quantified information on water resources using an independent and standardized framework for reporting on water resources in the VRB.

8.2 Water Accounting Plus (WA+)

WA+ is a standardized reporting framework that summarizes and displays water conditions and management practices in river basins (Karimi *et al.*, 2013a). It was developed based on the water accounting framework of the International Water Management Institute (Molden, 1997). Beyond the quantification of water volumes, WA+ explicitly considers land use interactions with the water cycle. Therefore, WA+ differentiates between available, utilizable, manageable and reserved water flows and stocks among many other components of the water cycle. WA+ results are presented in volume of water and the water accounts are usually done for annual or longer periods because WA+ is meant for long-term planning. Definitions of the WA+ terminology are provided in the Appendix 28. More information and updates on WA+ can be accessed at https://www.wateraccounting.org/index.html (last accessed 07.07.2020).

8.3 Land use and land cover in WA+

Land use and land cover (LULC) is an important factor in WA+ because it determines whether the water is manageable or non-manageable. Four clusters are used to group land use and land cover classes and they differ in terms of water management, namely, the Protected Land Use (PLU), the Utilized Land Use (ULU), the Managed Land Use (MLU) and the Managed Water

Use (MWU) (Karimi *et al.*, 2013a). Table 8.1 provides a description of each WA+ LULC group with associated examples.

Table 8.1. Four land use and land cover groups used in Water Accounting Plus (WA+), adapted from Karimi *et al.* (2013a)

WA+ LULC groups	Description	Examples
Protected Land Use (PLU)	Areas where no changes in land and/or water management are possible or advisable because they are protected by National Governments or Internationals NGO's.	Tropical rainforests, wetlands, mountainous vegetation, national parks, RAMSAR sites, etc.
Utilized Land Use (ULU)	Land where vegetation is not managed on a regular basis and the human influence is limited. Water flow is essentially natural.	Forests, woodlands, shrublands, grasslands, lakes, natural pastures, savannas, deserts, etc.
Modified Land Use (MLU)	Areas where vegetation and/or soils are managed, but all water supply is natural (rainfall). Water is not diverted but land use affects the vertical soil water balance.	Rainfed agriculture, biofuel crops, timber plantation, built-up areas, urban encroachment, etc.
Managed Water Use (MWU)	Areas where water flows are regulated by humans. All sectors that withdraw water from surface water and/or groundwater.	Irrigation schemes, urban water supply, diversion dams, canals, ditches, weirs, industrial extractions, storage for hydropower, etc.

Although LULC scenarios are not used in this study, the temporal dynamic of LULC is considered by using different land cover maps over the study period. Based on the availability of high resolution LULC data from the European Space Agency Climate Change Initiative (ESA, 2017), LULC data of 2005 and 2015 are used for the historical period (1991-2020) and for the future period (2021-2050), respectively. The ESACCI-LC-L4-LCCS v2.0.7 data with a high spatial resolution of 300 m is used in this study. For simplicity, the ESA maps for the VRB are first reclassified into ten basic and major LULC groups, namely, water bodies, bare areas, urban areas, rainfed croplands, irrigated croplands, grassland, shrubland, evergreen forest, deciduous forest and wetlands (Table 8.2; Figure 8.1). The final LULC maps for WA+ is obtained by crossing and overlapping the basic LULC map from ESA CCI with other spatial maps on various land status and uses. The maps of the World Database on Protected Areas (WDPA, 2016) and the Global Reservoir and Dam Database (GRanD; Lehner *et al.*, 2011;

Mulligan *et al.*, 2020) are used to reclassify the basic LULC data and distinguish between protected versus non-protected lands and identify managed water bodies.

Table 8.2. Proportions of land use and land cover classes in the Volta River Basin per WA+ LULC classes

WA+ class	LULC	1991-2020			2021-2050		
		Area (km^2)	Area (%)		Area (km^2)	Area (%)	
PLU	Protected Water Bodies	16	0.004	10.73	28	0.01	10.68
	Protected Bare areas	1	0.0002		2	0.0004	
	Protected Grasslands	6720	1.62		6885	1.66	
	Protected Shrublands	17933	4.31		17432	4.19	
	Protected Evergreen forest	488	0.12		487	0.12	
	Protected Deciduous forest	19313	4.65		19460	4.68	
	Protected Wetlands	142	0.03		83	0.02	
ULU	Water Bodies	1004	0.24	53.76	1017	0.24	53.36
	Bare areas	57	0.01		59	0.01	
	Grasslands	89142	21.45		90762	21.84	
	Shrublands	57705	13.88		52919	12.73	
	Evergreen forest	952	0.23		1009	0.24	
	Deciduous forest	74366	17.89		75802	18.24	
	Wetlands	196	0.05		194	0.05	
MLU	Urban areas	407	0.10	33.61	721	0.17	34.00
	Rainfed croplands	139283	33.51		140582	33.83	
MWU	Managed Water Bodies	6185	1.49	1.90	6465	1.56	1.96
	Irrigated croplands	1693	0.41		1697	0.41	

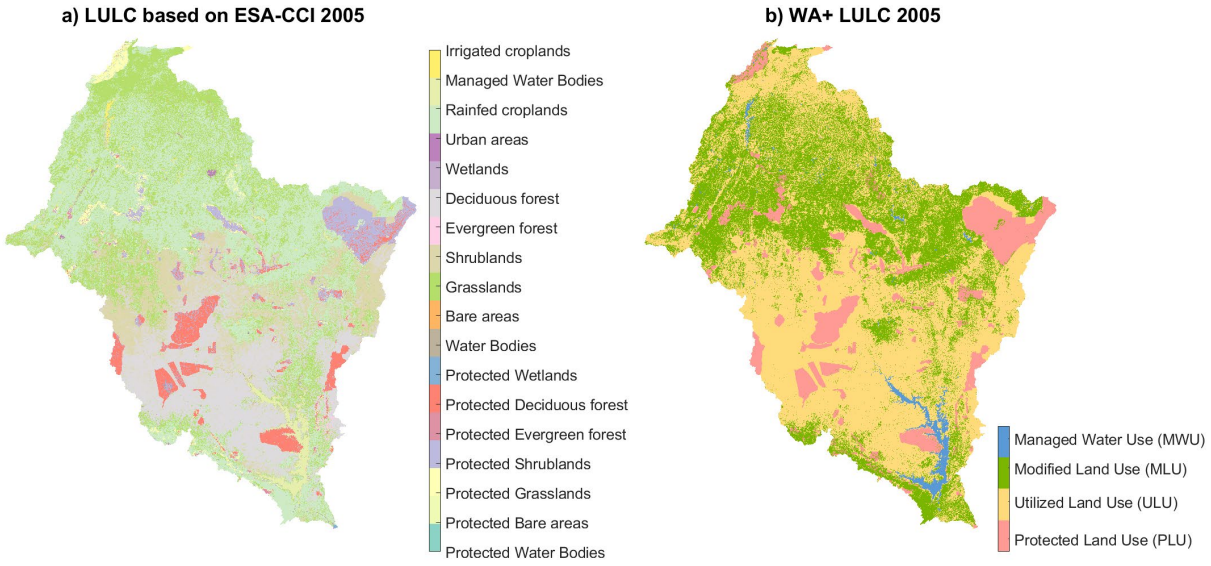


Figure 8.1. Land use and land cover in the Volta River basin using the ESA CCI data (a), and regrouped into WA+ classes (b).

8.4 WA+ sheets and performance indicators

8.4.1 Overview

In the following, the term total terrestrial evaporation is used in replacement of the debated "evapotranspiration" term (Miralles *et al.*, 2020; Savenije, 2004), which is however used in the terminology of WA+. To avoid changing the WA+ terminology, the abbreviation "ET" is conserved but is defined as total terrestrial evaporation in this study.

WA+ is still under development with currently eight sheets to describe water conditions. These sheets are named: resource base, total evaporation, utilized flow, agricultural services, surface water, groundwater, ecosystem services and sustainability (Table 8.3). Each sheet has a set of indicators that are used to summarize the overall water resources situation. The first four sheets are considered the main sheets for reporting on water resources. However, this study focuses on the two most important sheets (i.e., the resource base sheet and the total evaporation sheet) because the other sheets require information that are not available for future predictions (e.g., biomass production, agriculture, etc.). Examples of the other sheets can be found in the literature (e.g., FAO and IHE Delft, 2019).

Table 8.3. Water Accounting Plus (WA+) reporting sheets. Adapted from Bastiaanssen *et al.* (2015).

Number	WA+ Sheets	Purpose
1	Resource Base Sheet	Provide an overview on over-exploitation, unmanageable, manageable, exploitable, reserved, utilized and utilizable flows at river basin scale.
2	Evapotranspiration Sheet	Quantify beneficial and non-beneficial water consumption for all land use classes, and sectors including agriculture, environment, economy, energy and leisure.
3	Agricultural Services Sheet	Assess agricultural production (kg/ha) and the related water productivity (kg/m³).
4	Utilized Flow Sheet	Identify manmade and natural withdrawals from surface water and groundwater, with a distinction between consumed and non-consumed water.
5	Surface Water Sheet	Quantify the natural and actual river flow and determine the surface water availability and utilizable withdrawals.
6	Groundwater Sheet	Assess the role of groundwater in renewable water resources to support safe groundwater withdrawals.
7	Ecosystem Services Sheet	Express the regulating role of vegetation in the exchanges between land and atmosphere and estimate the reduction of greenhouse gas emission to prevent biodiversity degradation.
8	Sustainability Sheet	Quantify reliability, resilience and vulnerability of water resources and asses various land and water changes in a spatial context.

8.4.2 Resource base sheet

The WA+ resource base sheet provides information on inflows and outflows of water volumes in a river basin and relates them to various processes (Figure 8.3). The net inflow to the basin is obtained by adjusting the gross inflow with the change in total water storage. A part of the net inflow is consumed as landscape ET, representing the total evaporation from precipitation and considered as green water consumption (Falkenmark and Rockström, 2006). The remaining water is the exploitable water, i.e. the non-evaporated water, which is considered as blue water (Falkenmark and Rockström, 2006).

The exploitable water corresponds to the sum of the available water for various water uses, the non-utilizable outflow and the reserved outflow (i.e., downstream commitment, environmental flows, navigational flow, etc.). The non-utilizable outflow is the water that is difficult or impossible to tap during peak flows or inundations (Shilpakar *et al.*, 2011), and is considered

as 30% of annual surface runoff in the VRB. The reserved outflow is estimated as 28% of the mean annual runoff in the VRB (Smakhtin *et al.*, 2004).

The available water is partitioned into utilized flow and utilizable outflow. A part of the utilized flow is consumed as incremental ET, which can be natural or manmade, while the other part is the non-recoverable flow. The non-recoverable flow comprises groundwater recharge and polluted water estimated as 15% of annual runoff in the VRB (Mekonnen and Hoekstra, 2015). The utilizable outflow is the water that is not depleted and represents additional water that could be utilized. The non-consumed water is the sum of the utilizable flow, the non-utilizable outflow and the reserved flow. The total outflow is the sum of the non-consumed water and the non-recoverable flow. It represents the amount of water that physically leaves the basin through surface and subsurface water systems.

Because of moisture recycling, a part of the terrestrial evaporation contributes to the generation of precipitation. In the VRB, it is estimated that 4% of the regional terrestrial evaporation is recycled and contributes to the total precipitation in that same region (Van der Ent *et al.*, 2010).

The depleted water is the net inflow minus the total outflow and minus the recycled ET. Blue and Green ET separation is achieved with the Budyko framework (Budyko, 1974; McVicar *et al.*, 2012). The WA+ resource base sheet has a set of standard indicators described as follows (Karimi *et al.*, 2013a; Karimi *et al.*, 2013b):

The Exploitable Water fraction (*EWF*) represents the part of the net inflow that is not depleted through landscape ET.

$$EWF = \frac{Exploitable\ Water}{Net\ Inflow} \tag{8.1}$$

The Storage Change Fraction (*SCF*) defines the portion of storage change in the exploitable water, and expresses the degree of dependency on the total freshwater storage change, which include surface and subsurface water storages.

$$SCF = \frac{Total\ Storage\ Change}{Exploitable\ Water} \tag{8.2}$$

The Available Water Fraction (*AWF*) describes the proportion of the exploitable water that can be withdrawn from the basin.

$$AWF = \frac{Available\ Water}{Exploitable\ Water} \tag{8.3}$$

The Basin Closure Fraction (*BCF*) defines the extent of the depletion of the available water. A closed basin occurs when all available water is depleted.

$$BCF = \frac{Utilized\ Flow}{Available\ Water} \tag{8.4}$$

The Reserved Flow Fraction (*RFF*) indicates the degree of commitment to downstream flow requirements.

$$RFF = \frac{Reserved\ Outflow}{Total\ Outflow} \tag{8.5}$$

8.4.3 Total evaporation sheet

The WA+ total evaporation sheet (Figure 8.4) presents water depletion and describes parts of water consumption that are managed, manageable and non-manageable based on the LULC types (Figure 4). It presents the breakdown of the total evaporation (*ET*) into soil and water evaporation (*E*) and vegetation transpiration (*T*), and then differentiates between beneficial and non-beneficial water consumptions. Beneficial and non-beneficial ET is determined based on a value judgement that has to be adjusted depending on case studies.

In this study, transpiration is assumed to be beneficial as it reflects the amount of water transferred to the atmosphere from plants through stomata in the leaves, thereby indicating plant growth (e.g., crops). However, transpiration can be non-beneficial in case of undesirable vegetation such as weed infestation in croplands, alien invasive species and floating vegetation in water bodies. Evaporation from soil and water as well as from wet surfaces such as leaves, roads and building is here considered non-beneficial. However, evaporation from natural surfaces and from interception can be beneficial, for instance in case of plant temperature regulation, natural lakes, wetlands, water bodies exploited for fishing, aquatic birds, water sports and leisure (Karimi *et al.*, 2013a).

Table 8.4. Beneficial total evaporation (ET) fraction and repartition per sector retained for the VRB.

WA+ LULC	LULC	Beneficial ET fraction (%)				Beneficial ET per sector (%)				
		Transpiration	Water E	Soil E	Interception E	Agriculture	Environment	Economy	Energy	Leisure
PLU	Protected Water Bodies	100	100	0	0	0	85	0	0	15
	Protected Bare areas	100	0	30	0	0	85	0	0	15
	Protected Grasslands	100	0	30	0	0	85	0	0	15
	Protected Shrublands	100	0	30	0	0	85	0	0	15
	Protected Evergreen forest	100	0	30	0	0	85	0	0	15
	Protected Deciduous forest	100	0	30	0	0	85	0	0	15
	Protected Wetlands	100	100	100	0	0	85	0	0	15
ULU	Water Bodies	0	50	0	0	35	40	5	0	20
	Bare areas	0	0	0	0	0	100	0	0	0
	Grasslands	50	0	0	0	5	95	0	0	0
	Shrublands	70	0	0	0	5	85	0	10	0
	Evergreen forest	100	0	0	0	5	90	0	0	5
	Deciduous forest	100	0	0	0	5	90	0	0	5
	Wetlands	100	50	0	0	5	80	5	0	10
MLU	Urban areas	70	0	0	0	0	0	35	0	65
	Rainfed croplands	100	0	0	0	90	0	10	0	0
MWU	Managed Water Bodies	0	100	0	0	35	5	30	20	10
	Irrigated croplands	100	0	0	0	90	0	10	0	0

The repartition of the beneficial water consumption for agriculture, environment, energy, economy and leisure is based on assumptions that can be adapted to cases studies. Table 8.4 provides details on the decisions made here for beneficial ET fraction and its reparation per sector in the VRB.

The following set of indicators are used to summarize the information in the WA+ total evaporation sheet (Karimi *et al.*, 2013a; Karimi *et al.*, 2013b):

The Transpiration Fraction (*TF*) describes the part of total evaporation (*ET*) that is produced by plants.

$$TF = \frac{Transpiration}{Total\ Evaporation} \tag{8.6}$$

The Beneficial Fraction (*BF*) represents the proportion of total evaporation that occurred as beneficial evaporation and beneficial transpiration.

$$BF = \frac{Beneficial\ ET}{Total\ Evaporation} \tag{8.7}$$

The Managed Fraction (*MF*) indicates the proportion of total evaporation that occurred by manipulation of land use and water management.

$$MF = \frac{Managed\ ET}{Total\ Evaporation} \tag{8.8}$$

The Agricultural ET Fraction (*AEF*) corresponds to the part of total evaporation related to agricultural production.

$$AEF = \frac{Agricultural\ ET}{Total\ Evaporation} \tag{8.9}$$

The Irrigated ET Fraction (*IEF*) describes the portion of agricultural *ET* that is attributable irrigation.

$$IEF = \frac{Irrigated\ Agricultural\ ET}{Agriculture\ ET} \tag{8.10}$$

8.5 Hydrological modelling for climate change projections

The fully distributed mesoscale Hydrologic Model (mHM) is used to simulate the hydrological variables required for WA+. The mHM model is a conceptual model that has the specificity of producing a seamless representation of hydrological processes across scales (Kumar *et al.*, 2013; Samaniego *et al.*, 2010; Samaniego *et al.*, 2017), including evaporation, interception, transpiration, runoff, streamflow and groundwater recharge. The model has shown good performances in modelling the hydrological cycle in the VRB, and its configuration for this study is comprehensively described in the work of Dembélé *et al.* (2020b).

The mHM model is thoroughly constrained to produce an accurate representation of hydrological processes in the VRB. Particularly, the representation of spatial patterns of various processes including evaporation and soil moisture are improved by calibrating the mHM model with multiple satellite remote sensing datasets (Dembélé *et al.*, 2020b). Moreover, the overall water balance of the model is simulated by calibrating the model with the GRACE data to better represent the terrestrial water storage.

Based on the findings on climate change impacts assessment on water resources in the VRB (Chapter 7), a large ensemble of eleven combinations of GCM-RCM datasets, comprised of nine GCMs and four RCMs, is selected and used to force the mHM model (Table 8.5). The hydrological model outputs are subsequently used for the WA+ analyses. The mHM model is run at a daily time step with a spatial discretization of 0.25° (~28 km), which corresponds to 619 grid cells in the VRB.

Table 8.5. List of global climate models (GCMs) and regional climate models (RCMs) used for WA+

RCMs	GCMs
CCLM4-8-17	CNRM-CM5
RACMO22T	EC-EARTH
RCA4	CanESM2
	EC-EARTH
	CM5A-MR
	MIROC5
	HadGEM2-ES
	MPI-ESM-LR
	GFDL-ESM2M
REMO2009	CM5A-LR
	MIROC5

8.6 Results

For simplicity, the results are presented for the 30-year mean and multi-model mean of the RCM-GCM datasets. This allows comparing the long-term conditions of water resources from the historic period (1991-2020) to those of the future period (2021-2050).

8.6.1 Consistency of the WA+ LULC

A first important check is made here to show that the water fluxes simulated with mHM model for the LULC classes show a consistent pattern in the Budyko space (Figure 8.2). The Budyko space (Budyko, 1974) is formed by the aridity index (ratio of the long-term mean annual potential evaporation, E_p, to precipitation, P), and the evaporative index (ratio of actual evaporation, E_a, to precipitation, P).

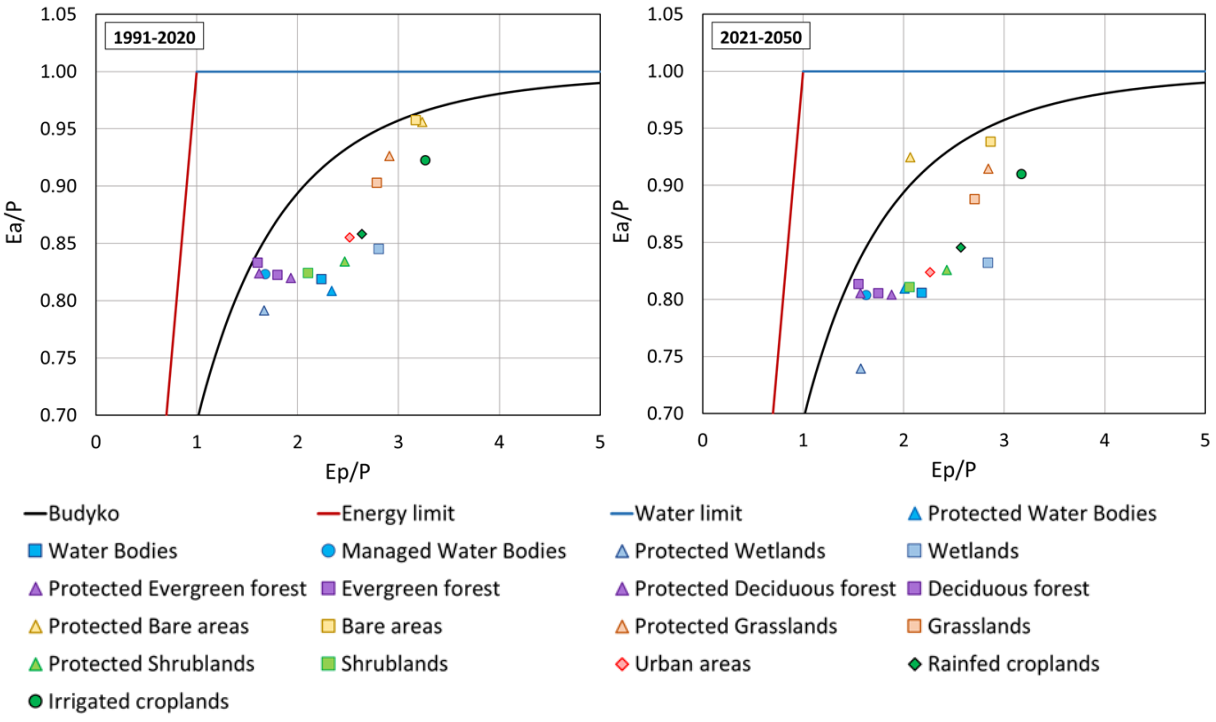


Figure 8.2. WA+ LULC repartition in the Budyko framework where the evaporative index (actual evaporation divided by precipitation) is plotted as a function of the aridity index (potential evaporation divided by precipitation). The y-axis is truncated at 0.7 for a better display.

The simulated long-term values for the different LULC classes (period 1991-2020 and 2021-2050) plot well in the physically possible space below the energy and water limits (Donohue *et al.*, 2011; McVicar *et al.*, 2012), and close to the theoretical curve postulated by Budyko. The evaporative index is between 0.73 and 0.97, and the aridity index is between 1.5 and 3.5, which are expected values for sub-humid and semi-arid environments such as the VRB (Gunkel

and Lange, 2017). The consistency of the simulated water fluxes for the retained LULC classes is further demonstrated by the fact that the irrigated croplands have a higher evaporative index than the rainfed croplands, and forests have a lower aridity index than the other LULC classes.

8.6.2 Sheet 1: Resource base

The WA+ resource base sheet gives an overview of the water repartition into flows, stocks and fluxes (Figure 8.3). For the period 1991-2020, the long-term annual average net inflow in the VRB was 388.6 km^3/year (935 mm/year) of which 96% is attributable to precipitation resulting from advection (374.6 km^3/year or 901 mm/year). The landscape ET accounts for 92% of the net inflow and occurs mainly through the ULU (55%) (for abbreviations see Table 8.1) and the MLU (32%). The landscape ET is the green water consumption that is ascribed to rainfall. In the MLU, rainfed croplands represent about 33% of the basin area, which justifies the high portion of the landscape ET in that LULC class. The ULU is dominated by grasslands and shrublands, which represent together 35% (146,848 km^2) of the basin area (Table 8.2).

Only 30.7 km^3/year of water in the VRB were exploitable and corresponded to 8% of the net inflow. The exploitable water refers to the blue water, which is water stored in reservoirs, lakes, streams, and aquifers. The non-utilizable water originating from peak flows or inundations and the reserved outflow for environmental and downstream requirements together constituted about 52% of the exploitable water. The available water for various uses in the basin was 14.6 km^3/year, with 79% that were utilized, while the remainder 21% were utilizable but were not consumed. The non-recoverable flow due to groundwater recharge and water pollution was 8.6 km^3/year and corresponds to 74% of the utilized flow. The remainder 26% of the utilized flow represents the incremental ET (3.03 km^3/year), which refers to additional evaporation from blue water and is not related to rainfall. Natural processes mainly generated the incremental ET, while only 7% are ascribed to manmade activities. Moreover, the MWU group consumed only 7% of the utilized flow, while it encompasses the irrigated croplands and the managed water bodies.

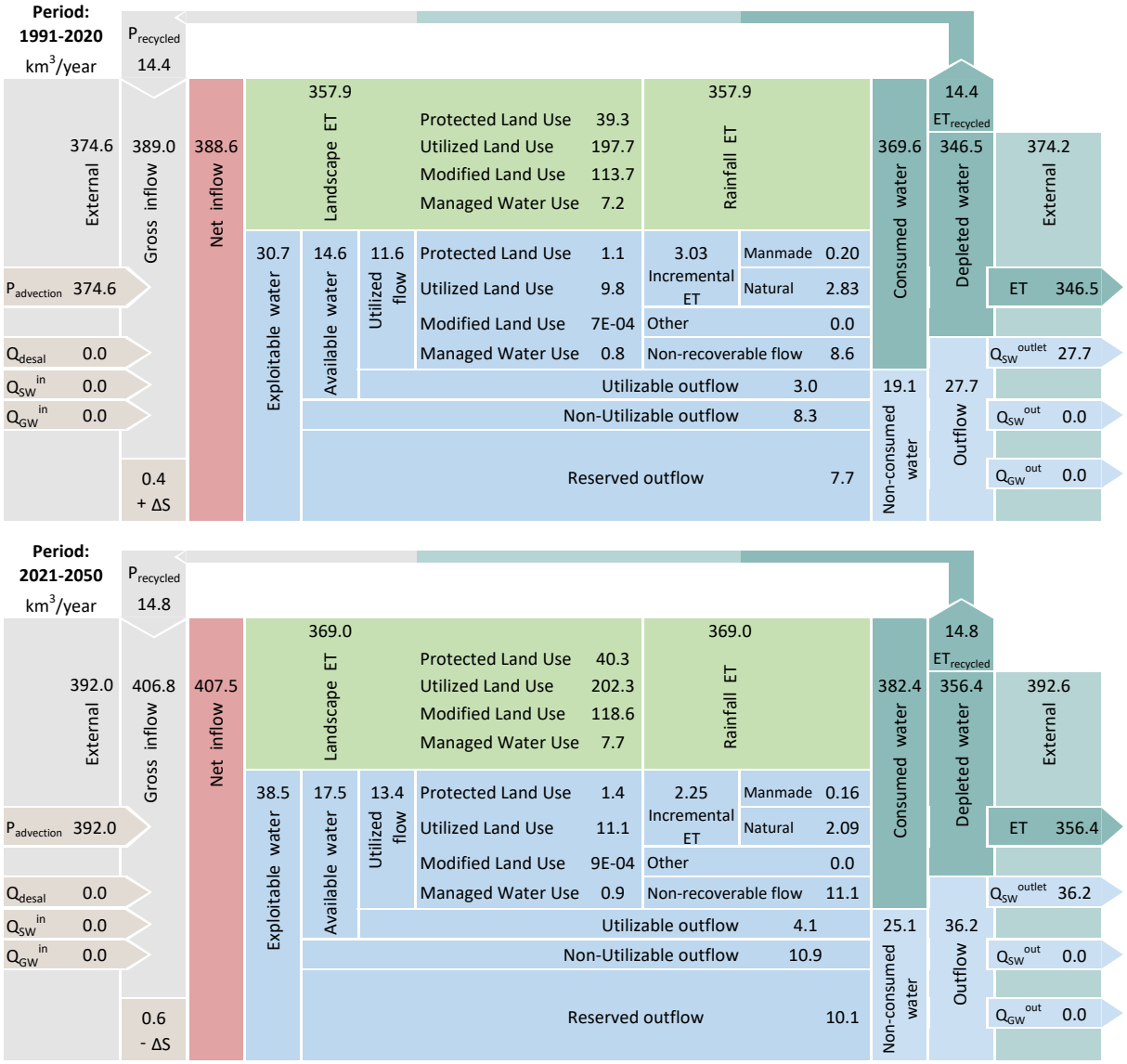


Figure 8.3. WA+ resource base sheets for the historical (1991-2020) and future (2021-2050) periods.

The total depleted water was 346.5 km^3/year and the outflow at the basin outlet was 27.7 km^3/year, corresponding to 89% and 7% of the net inflow. A part of the consumed water representing 4% of the net inflow is recycled and falls back in the basin as precipitation.

The evolution of the water resources over the future period 2021-2050 shows an increase in most of the water accounts in the VRB (Figure 8.3). An increase in net inflow by +5% (18.9 km^3/year) relative to the historical period is expected. As a consequence of the increase in net inflow, other water accounts are projected to increase, including landscape ET (+3% or 11 km^3/year), exploitable water (+25% or 7.8 km^3/year) and available water (+19% or 2.8 km^3/year). However, a decrease is projected for the incremental ET (-26% or -0.8 km^3/year).

The projected increase in some water accounts are not favourable for the VRB. This includes the utilizable outflow (+15% or 1.8 km^3/year), the non-utilizable outflow (+31% or 2.6 km^3/year) and the non-recoverable flow (+30% or 2.5 km^3/year).

8.6.3 Sheet 2: Total evaporation

The WA+ total evaporation sheet summarizes water consumption and provides the breakdown of total evaporation (*ET*) into transpiration and evaporation from soil, water bodies and interception (Figure 8.4). Over the period from 1991 to 2020, the long-term annual ET was 361 km^3/year (869 mm/year) of which 11% were non-manageable because occurring in the PLU, 55% were manageable from the ULU and 34% were managed from the MLU and MWU. Transpiration was 177.1 km^3/year and alone accounted for 49% of total ET, followed by soil evaporation (26.4% or 95.2 km^3/year), interception evaporation (24% or 86.4 km^3/year), while water evaporation was the lowest (0.6% or 2.2 km^3/year).

From the total water consumed in the VRB during the period 1991 to 2020, only 42% were beneficial. The total beneficial consumption was 151.1 km^3/year, with 57% attributable to the environment, 34.6% to agriculture, 3.6% to the economy, 3.6% to leisure and 1.2% to energy production. The non-beneficial water consumption represents 58% of the total water depleted. The large portion of non-beneficial water consumption is ascribed to interception and soil evaporation that occurred at 60% in the ULU and 29% in the MLU.

The projected water accounts over the period 2021-2050 shows an increase of the overall water consumption in the VRB. The total ET is projected to increase by +3% (10.3 km^3/year), which could be expected as the net inflow is projected to increase by +5% over the same period (Figure 8.3). By maintaining the current land and water management practices, the beneficial water consumption could increase by +5% as a result of the +5% increase in transpiration. In the same momentum, the beneficial water consumption is projected to increase by +7% for agriculture and +4% for the environment.

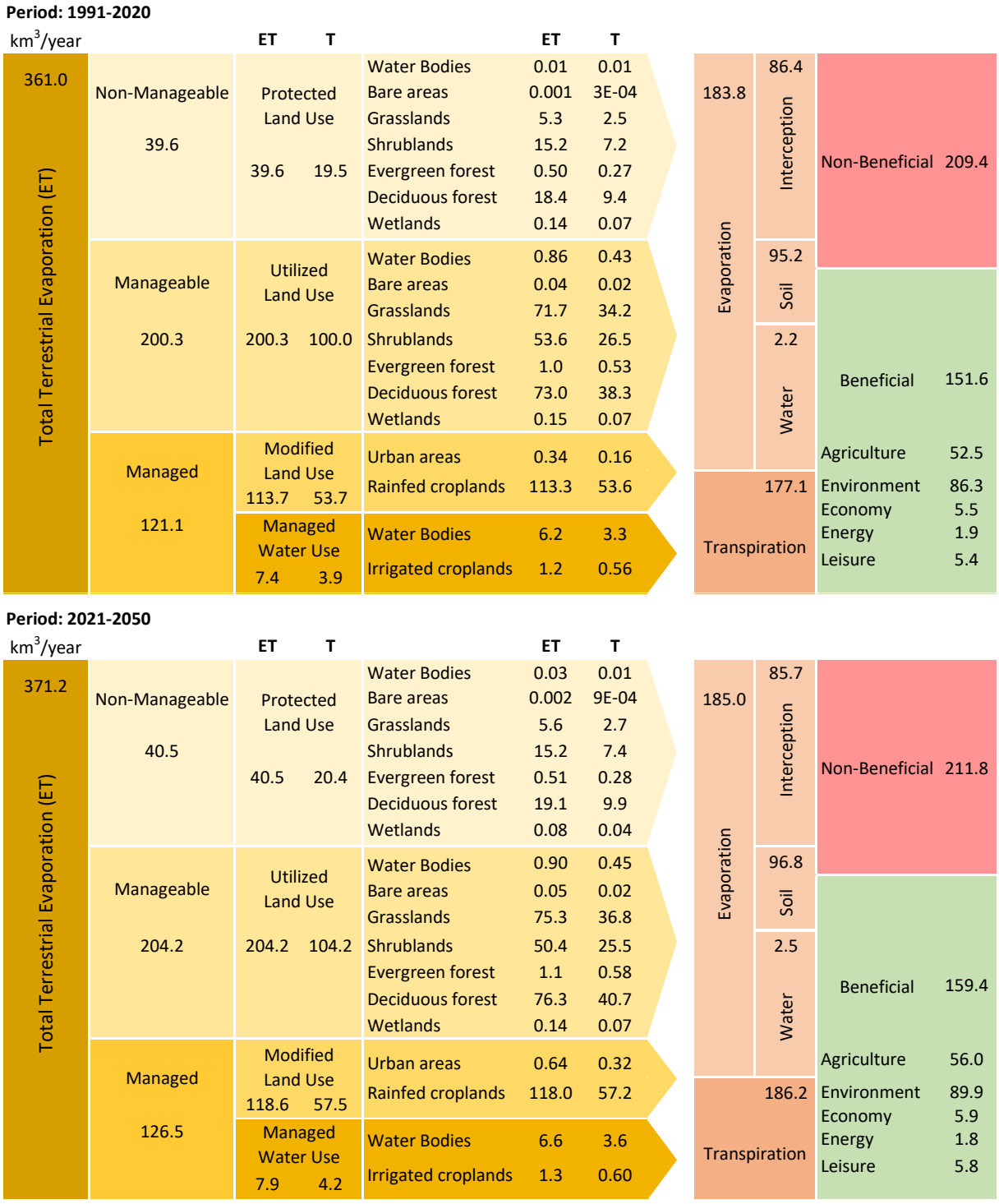


Figure 8.4. WA+ total evaporation sheets for the historical (1991-2020) and future (2021-2050) periods

The contribution of each WA+ LULC to total ET and its components as well as to the beneficial fraction and the water consumption in different sectors is shown in Figure 8.5. Most of the depleted water occurs in the ULU (55%), followed by the MLU (32%), the PLU (11%), and the MWU (2%). The MLU accounts for 92% of the water depleted for agriculture and 98% for the economy. The ULU is accountable for 78% of the water consumed by the environment, 98%

for energy production and 36% for leisure. The PLU contributes at 61% of the water depleted for leisure and at 22% for the environment.

The beneficial water consumption mainly occurs in the ULU (48.7%) due to the forests, followed the MLU (36%) because of rainfed croplands, and the PLU (14.8%) because of protected vegetation species, forests and wetlands. Only 0.5% of the beneficial water consumption occurs in the MWU, which encompasses the irrigated croplands and the managed water bodies.

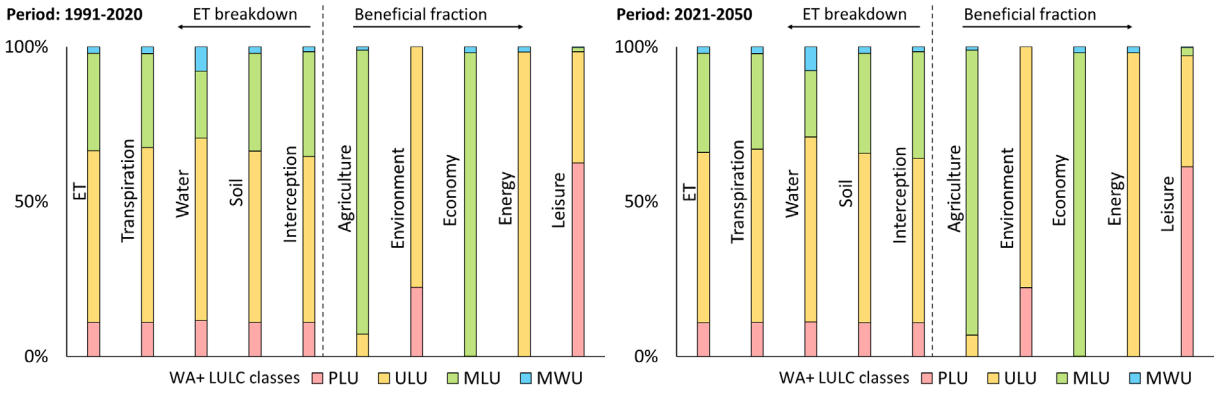


Figure 8.5. Total evaporation (ET) breakdown and beneficial fraction from WA+ LULC classes for each activity sector

8.6.4 WA+ performance indicators

A set of indicators are used to better understand the water accounts summarized in the WA+ sheets. They are calculated for the historical and future periods along with their relative changes (Table 8.6). The indicators of the resource base sheet show that the exploitable water fraction (*EWF*) is 0.08 over the period 1991-2020 with an expected increase of +20% in the near future between 2020 and 2050. The low *EWF* indicates that a small portion of the net inflow can be exploited in the VRB. Moreover, the available water fraction (*AWF*) was 0.48 and is projected to decrease by -5%. The projected decrease in *AWF* while the EWF is expected to increase is justified by that fact that the exploitable water could increase by +25% while the available water would only increase by +19% (Figure 8.3). The *AWF* shows that less than half of the exploitable water is actually available for withdrawals from the blue water storages. The downstream flow requirements are met as the reserved flow fraction (*RFF*) is 0.41, indicating that a large portion of the basin outflow is dedicated to environmental flows. Therefore, the *AWF* indicates that a huge amount of water is being lost through inundations during peak flows as the commitment for the reserved flow is satisfied. The storage change fraction (*SCF*) of 0.01 indicates a low

contribution of the freshwater storage to the exploitable water. The basin closure fraction (*BCF*) is 0.79 and is projected to decrease by -3%, showing that although a large portion of the available water is utilized, there is still a substantial portion that can be used.

Table 8.6. Change in WA+ indicators from the historical (1991-2020) to the future (2021-2050) periods

Indicators	1991-2020	2020-2050	Change (%)
	Sheet 1: Resource Base		
Exploitable Water Fraction (EWF)	0.079	0.094	19.6
Storage Change Fraction (SCF)	0.014	-0.016	-219.1
Available Water Fraction (AWF)	0.477	0.454	-4.9
Basin Closure Fraction (BCF)	0.794	0.766	-3.4
Reserved Flow Fraction (RFF)	0.406	0.404	-0.5
	Sheet 2: Total Evaporation		
Transpiration Fraction (TF)	0.491	0.502	2.2
Beneficial Fraction (BF)	0.420	0.429	2.2
Managed Fraction (MF)	0.335	0.341	1.6
Agriculture ET Fraction (AEF)	0.145	0.151	3.6
Irrigated ET Fraction (IEF)	0.024	0.023	-2.8

The performance indicators of the total evaporation sheet only show very small changes for future projections of water accounts (Table 8.6). All the performance indicators are projected to slightly increase between +1.6% and 3.6%, except the irrigated ET fraction (*IEF*) that could decrease by -2.8%. The transpiration fraction (*TF*) is 0.49 and indicates that transpiration is a major process in water depletion in the VRB, which can be explained by the large presence of vegetated lands (rainfed croplands, irrigated croplands, grasslands, shrublands and forests) covering about 98% of the basin area. However, the beneficial fraction (*BF*) of the depleted water is only 0.42, which can be justified by the low land and water management practices as the managed fraction (*MF*) is 0.34. Although agriculture occupies 34% of the basin area, the agricultural ET fraction (*AEF*) is only 0.15, while the contribution of irrigated agriculture is very low with an irrigated ET fraction of 0.02. These results suggest that there are possibilities for improving land and water management to increase the benefits of water consumption in the VRB.

8.7 Possible adaptation measures

Possible adaptation measures are discussed here to illustrate how the findings from this study can contribute to identifying required actions for water management in the VRB. The presented WA+ report based on independent datasets obtained from global databases provides a common ground for the assessment of water resources, which is very important in transboundary water management. Therefore, the uniformity of the data collection and the analysis techniques increases the acceptability of the results by the stakeholders in the VRB.

Based on the results of the resource base sheet, the socio-economic development in the VRB can largely beneficiate from the available water resources by developing irrigated agriculture and hydropower generation, which are the top priorities of the upstream and downstream countries (Burkina Faso and Ghana). However, mutual agreement between the countries on the timing of water storage and release is key for a peaceful upstream-downstream cohabitation.

To benefit from the projected increase in the net inflow over the period from 2021 to 2050, it is essential for authorities in the VRB to invest in water storage infrastructure with low evaporative losses, expand agriculture lands and deploy strategies to reduce water pollution. A conjunction of these strategies among many others can help reduce the non-consumed water, which is projected to increase by +32% or 6 km^3/year. However, the reserved flow for downstream requirements should be conserved. The landscape ET in the ULU can be made more beneficial by increasing the proportion of livelihood-generating lands. A large portion of lands in the ULU group (grasslands and shrublands) can be converted to MLU or MWU with adequate land and water management practices. However, care should be taken to avoid biodiversity degradation and deforestation. Nature-based solutions such as managed aquifer recharge systems and blue-green infrastructure such as storm water control systems can help achieve better water consumption in the VRB (Keesstra *et al.*, 2018; Nesshöver *et al.*, 2017).

From the total evaporation sheet, it appears that higher benefits can be obtained from the overall water consumption in the VRB by adopting appropriate land and water management policies that can improve water productivity and water use efficiency. Initiatives that can support an efficient water consumption include practices that seek for producing more food with less water. Then, agriculture lands should be expanded and small scale initiatives such as farmer led irrigation be encouraged and promoted (de Bont *et al.*, 2019; Lefore *et al.*, 2019; Woodhouse *et al.*, 2017). With more water, the production of hydroelectricity becomes essential with a high potential to unlock more access to green energy (Gyamfi *et al.*, 2018; Kling *et al.*, 2016).

8.8 Discussions

Despite the classical sources of uncertainty associated with climate change impact projections (Sylla *et al.*, 2018b), the key uncertainties in the presented methodology are associated with the identification of the land use classes (LULC), with some used global data sets and with expert knowledge.

LULC is the backbone of the WA+ framework. Therefore, the reliability of the results highly depends on the accuracy of the LULC data. However, the ESA LULC data has the advantage of being available at a high resolution and being subject to thorough quality check. At the very large scale of the VRB, having high resolution LULC data such as the ESA data is satisfactory for the application of WA+. Moreover, to bring confidence into the analyses, the Budyko framework is used to check the consistency of the LULC classification and the proportions of water and energy fluxes.

Some calculations of the WA+ framework are done using global datasets, such as the level of water pollution in the basin and the requirements for reserved flow such as the environmental flow and navigational flow. However, when available and reliable, regional or local datasets should be used for future studies.

The estimation of some components of the WA+ requires value judgement, which makes it a non-deterministic but flexible framework that can be adjusted for each case study. However, the value judgment requires expert knowledge so that results based on value judgement need to be interpreted with cautions. This concerns the estimation of the beneficial and non-beneficial fractions of water use, and the repartition of water depletion per sector including agriculture, environment, economy, energy and leisure.

The water accounts presented in this study are for long-term averages, while annual variations might occur. However, the long-term analysis is best suited for comparing trends over distinct periods and for attributing differences to climate change, which is the main goal of this study.

As this study aims at demonstrating the applicability of climate change scenarios with the WA+ framework, only the RCP8.5 scenario is used. Moreover, the analyses are done for the near future because the results for the far future might be less realistic and not very useful for water management as assumptions underlying the WA+ framework might rapidly evolve in the future.

8.9 Conclusion

This study demonstrate the applicability of the WA+ framework with climate change scenarios and hydrological modelling. The fully distributed mesoscale Hydrologic Model (mHM) is forced with climate change projections of a large ensemble of four RCMs and nine GCMs. The projected hydrological processes from mHM are used to inform the WA+ framework. The results show a clear projected increase in the exploitable water fraction and a decrease in the available water fraction over the period from 2021 to 2050. These findings can be translated into a clear need for adaptation measures to increase the water storage capacity in the Volta River basin to facilitate a good exploitation of the projected increase in net inflow, mainly for agriculture and hydropower generation.

Future studies should investigate the use of different RCPs, climate models and hydrological models. Scenarios of LULC can also be used to check how decisions on land use management can affect the water accounts. The conjunction of these efforts will move water governance forward and strengthen water security in the Volta River basin.

Chapter 9

Concluding Remarks

I think and think for months and years.
Ninety-nine times, the conclusion is false.
The hundredth time I am right.
Albert Einstein

It is the mark of an educated mind to rest satisfied
with the degree of precision which the nature
of the subject admits and not to seek exactness
where only an approximation is possible.
Aristotle

9.1 Overview

This PhD thesis demonstrates the applicability of the Water Accounting Plus (WA+) framework combined with the fully distributed mesoscale Hydrologic Model (mHM) and a large ensemble of climate change scenarios to comprehensively report on the current and future states of water resources in the Volta River basin (VRB). Achieving the main research objective required efforts to first address key hydrological challenges that are posed by large-scale hydrological modelling in data scarce and predominantly semi-arid environments such as the VRB. Therefore, the methodological framework adopted in this thesis provides answers to the identified challenges whose results serve as the basis for answering the main research questions. The key findings and the future research directions are discussed in the following.

9.2 Addressing the challenge of missing streamflow data

The proposed framework for gap-filling time series of streamflow data with the Direct Sampling (DS) method gives satisfactory performances. The method is tested for various missing data scenarios and it allows an accurate simulation of the missing streamflow values by using data patterns from predictor variables in the VRB (Dembélé *et al.*, 2019). In fact, the results show that the statistical content of the target variable is preserved, the probability distribution of the simulation matches accurately the reference data, and the shape of the hydrograph shows a good timing with a strong preservation of the annual seasonality. DS is a data-driven method and therefore it requires historical records that are sufficiently informative for the simulation of missing values, which is its main limitation. The results highlight that some conditions are necessary for obtaining a better prediction of the missing streamflow data. These conditions include the proximity and the good correlation of the target and the predictor streamflow stations, and the absence of missing values in the predictor variable. The findings suggest that DS is a promising tool for missing data simulation in environmental sciences and the proposed gap-filling framework could be transferred to other hydrological applications.

9.3 Addressing the challenge of reliable meteorological data

The evaluation of the adequacy of gridded meteorological datasets for large-scale hydrological modelling reveals that there is no single rainfall or temperature dataset that consistently ranks first in reproducing the spatiotemporal variability of all hydrological processes (Dembélé *et al.*, 2020c). The satellite-based and reanalysis rainfall datasets have contrasting performances across the four climatic zones present in the VRB, which suggests cautiousness in performance

generalizability to different spatial domains. A dataset that is best in reproducing the temporal dynamics might not be the best for the spatial patterns. In addition, there are more uncertainties in representing the spatial patterns of hydrological processes than their temporal dynamics. Finally, some global rainfall datasets performed less than regional datasets, which highlights the importance and the necessity of regional evaluation studies for satellite and reanalysis meteorological datasets. A final ranking of the seventeen gridded rainfall datasets is provided, which allows the selection of the most convenient product for simulating plausible hydrological flux and state variables with the mHM model in the VRB. It is noteworthy that these findings can be subject to uncertainties related to potential model structural deficiencies as well as errors in the observational datasets used for the model evaluation. However, mHM is chosen for its ability to produce seamless spatial patterns across scales. The findings are expected to draw the users' attention on the reliability of the gridded rainfall datasets for large-scale hydrological modelling and trigger further efforts for the improvement of these datasets. Finally, as no rainfall dataset consistently ranks first in reproducing several hydrological fluxes and state variables, choosing the most suitable rainfall dataset is not sufficient for improving process representation in models. Therefore, multivariate calibration of hydrological models stands as a promising way for improving hydrological simulations.

9.4 Addressing the challenge of accurate process representation

The novel multivariate calibration framework exploiting spatial patterns and simultaneously incorporating streamflow and satellite datasets shows that there are benefits in using satellite datasets, when suitably integrated in a robust model parametrization scheme (Dembélé *et al.*, 2020a; Dembélé *et al.*, 2020b). The relatively small decrease in the model performance for streamflow is greatly counterbalance by the increase in the model performance for other hydrological fluxes and state variables (i.e., actual evaporation, soil moisture and terrestrial water storage). The decrease in the performance of streamflow in the multivariate calibration strategies can be justified by the fact that there is an artefact caused by the non-uniqueness of model parameters when adopting the traditional streamflow-only calibration. Therefore, these findings unveil the pitfalls of the streamflow-only calibration and suggest the adoption of multivariate calibration strategies in hydrological modelling. Furthermore, it was found that spatial patterns of satellite data are a highly relevant and robust feature that can be used in multivariate calibration to improve the overall representation of the hydrological system. However, some trade-offs among the objective functions for streamflow and for satellite data

cannot be avoided as they originate from errors in the input data, the model structure and the lack of knowledge of the hydrological system.

9.5 Impacts of climate change on water resources

The assessment of the impacts of climate change on water resources in the VRB reveal contrasting changes in the seasonality of precipitation depending on the representative concentration pathways (RCPs) and the future projection periods (2021-2100), while a clear increase in the seasonality of temperature is expected. Although temperature and potential evaporation increase under all RCPs, a clear intensification of the hydrological cycle during the twenty-first century is only expected under the RCP8.5 scenario. Consequently, an increase is expected for the long-term annual estimates of precipitation, average temperature and potential evaporation. These changes in climatic variables subsequently lead to an increase in actual evaporation, surface runoff, streamflow, groundwater recharge, soil moisture and terrestrial water storage. In this context, floods and droughts are expected to be recurrent, which can weaken the water-energy-food security nexus and amplify the vulnerability of the local population to climate change. These findings could serve as a guideline for decision makers, and contribute to the elaboration of adaptation and mitigation strategies to cope with the dramatic consequences of climate change, and strengthen the regional socio-economic development. However, uncertainties exist in the projections because of the strong variabilities between the climate models. While there is a strong agreement between the climate models for the projection of temperature, there is a medium agreement on the direction of future changes for precipitation. These results underscore the complexity of modelling climate and hydrological systems in West Africa, and sparks a new call for further efforts in investigating climate change in the region.

9.6 Water accounting for sustainable water management

By addressing key challenges encountered in large-scale hydrological modelling, it was possible to comprehensively report on water resources in the VRB, using the Water Accounting Plus (WA+) framework combined with hydrological modelling and climate change scenarios. The results show that the net inflow in the VRB is projected to increase by +5% in the near future (2021-2050), but only 8% of the net inflow is exploitable as blue water stored in surface and subsurface reservoirs, while the remainder is depleted through landscape evaporation. The landscape evaporation, also referred to as green water consumption, dominantly occurs over

lands that are not managed by human, and that are essentially composed of grasslands and shrublands. It is estimated that only 42% of the water use is beneficial for the intended purposes, and agriculture represents 35% of the beneficial water consumption. The exploitable water fraction is expected to increase by +20% over the period from 2021-2050, while the available water fraction is expected to decrease by -5%. The disproportionate and contrasting changes in the exploitable and the available water fractions highlight the need for sound adaptation measures. These measures include the construction of infrastructure for water storage with low evaporative losses to better exploit the projected increase in the net inflow, which can help to boost agriculture and hydropower production in the VRB. It is noteworthy that the reliability of the findings highly depends on the accuracy of the land use and land cover (LULC) maps. However, the consistency of the LULC classification and the proportions of water and energy fluxes is checked with the Budyko framework, which showed satisfactory results.

9.7 Lessons learnt and future research directions

Overall, this PhD thesis presents scientific methodological developments for large-scale hydrological modelling and their applications in the Volta River Basin (VRB). The chapters of the thesis explore various aspect of hydrological data processing and analysis, spatially explicit model set-up and application, evolution of water resources under climate change and water resources assessment with standardized tools. The organization of the chapters shows a systematic and incremental advancement of knowledge and complexity from data preparation to hydrological model calibration and evaluation in a large river basin for improved water resources assessment using readily accessible datasets.

The key messages from this thesis are centered on using good quality data and robust methods to obtain more reliable information on water resources. For instance, global satellite data are increasingly available and they offer a unique opportunity to observe the water cycle. Although these global satellite data are relevant, they are usually not direct estimates of environmental processes, so that they have uncertainties. Therefore, the use of these datasets requires a careful selection and evaluation before using them to derive information that can be used for policy-making. Hydrological models also suffer from uncertainty in their structures and their parametrizations, so that they are not perfect tools but can still be improved for water resources assessment. The WA+ tool provides a unique opportunity to summarize complex processes in more readable formats through the water accounting sheets. It is a relevant tool for water resources management, but it still requires simplification of the sheets and clarification of some

underlying assumptions such as water consumption classification as beneficial or not. There are also uncertainties in the global and regional climate models, such that future water predictions should always be interpreted with cautious and followed with uncertainty analysis.

Future research directions should focus on challenges identified in the chapters of this thesis:

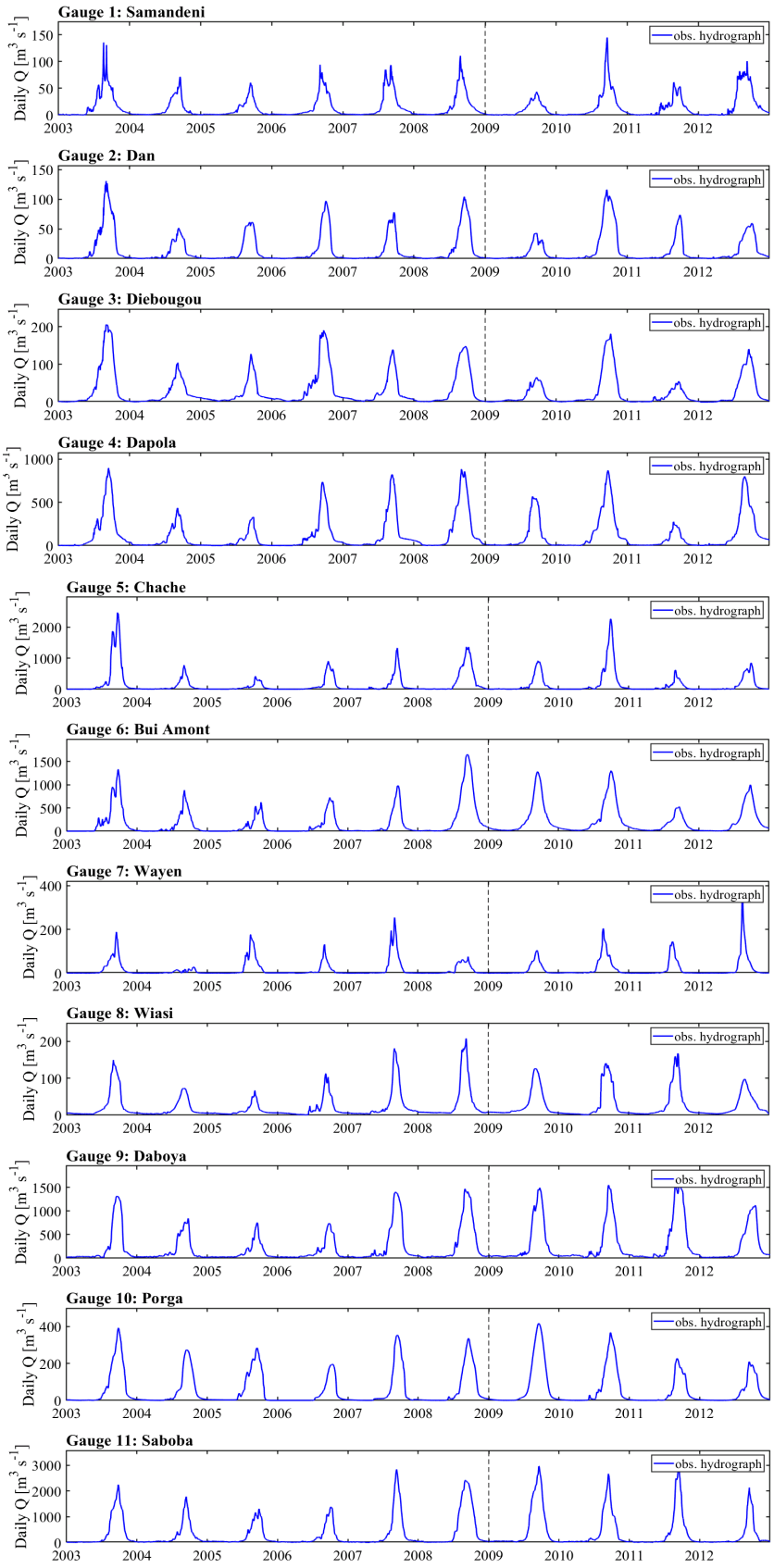
- Further development of the proposed gap-filling framework (Chapter 3) should investigate the simultaneous use of multiple predictor variables to inform the simulation of missing values. Additional research on the gap-filling of streamflow time series can consider the use of exogenous auxiliary variable (e.g., rainfall, evaporation), which can complement the predictor variables when they have gaps in their time series, and yield better estimates of the missing values. Although DS relies on a small set of parameters that are manually tuned, it would be more interesting and efficient to use a simple calibration technique that automatically adjust the parameters. The propagation of the uncertainties associated with the gap-filled data to subsequent applications (e.g. hydrological model calibration) should be investigated.
- The proposed methodology for the evaluation of gridded meteorological datasets (Chapter 4) should be applied to different regions and with different hydrological models, which might give more insights in the global suitability of these datasets for large-scale hydrological modelling. A detailed analysis of parameter variability as a function of input data can build the basis of future research and resolve potential structural deficiencies of the mHM model. Moreover, the transferability of the model's global parameters across different input datasets can be tested, which could shed light on the necessity of model recalibration when using different meteorological forcing.
- The developed multivariate calibration framework based on a new spatial bias-insensitive metric (Chapter 5 and Chapter 6) can be used with other spatially distributed hydrological models and can be applied to different climatic regions, which might reveal the advantages and the drawbacks of the approach. Further efforts in improving process representation in hydrological models should focus on fitting high and low flows, sub-period model calibration and calibration based on system signatures. The conjunctions of these future efforts could improve the prediction of floods and droughts, and improve model simulations for climate change and land use change impact studies.
- Uncertainties are ubiquitous in studies of climate change impacts on water resources (Chapter 7). Therefore, future studies in the VRB should consider different hydrological

models, which might unveil the adequacy of these models for the region. Moreover, new insights on the future evolution of water resources could be obtained by assessing the combined impacts of climate change, land use land cover change and water management practices. Environmental change impact studies should be continuously undertaken in the region, considering advances in climate and hydrological sciences.

- Integrating climate change scenarios in the WA+ framework is a way forward in improving water governance in transboundary basins, and as such it should be further investigated in future studies. Moreover, future LULC change scenarios can be developed to assess changes in the water accounts based on socio-economic developments. The subsequent findings could provide better insights for sound policy development and facilitate sustainable water management.

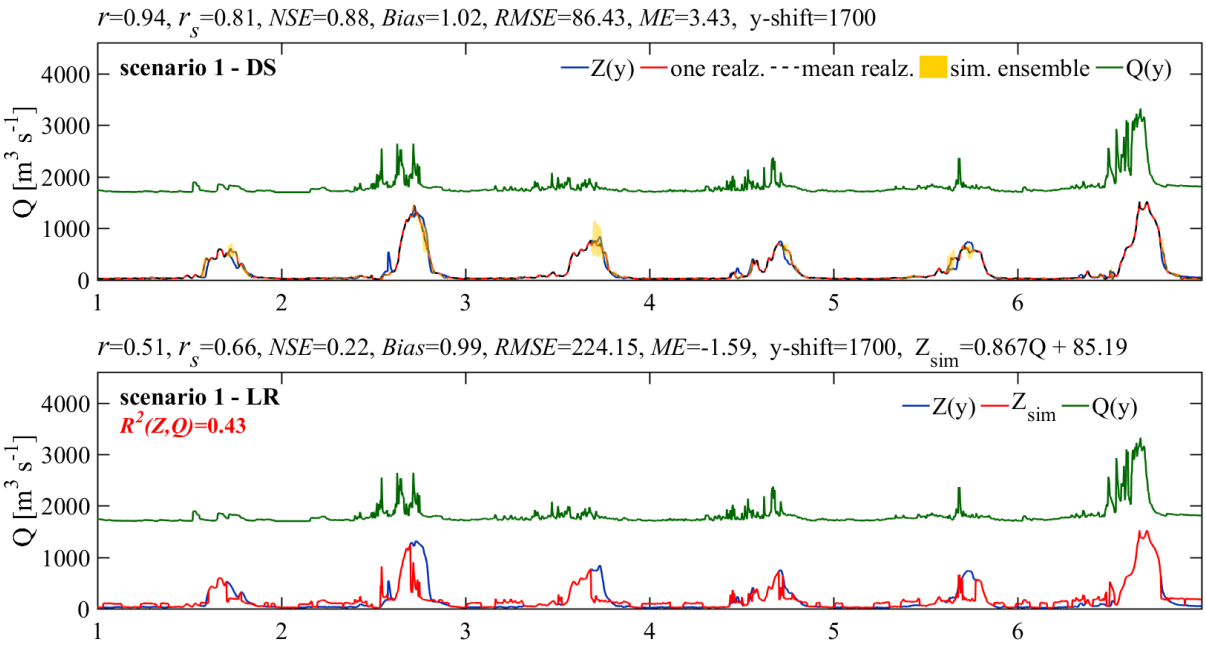
Appendices

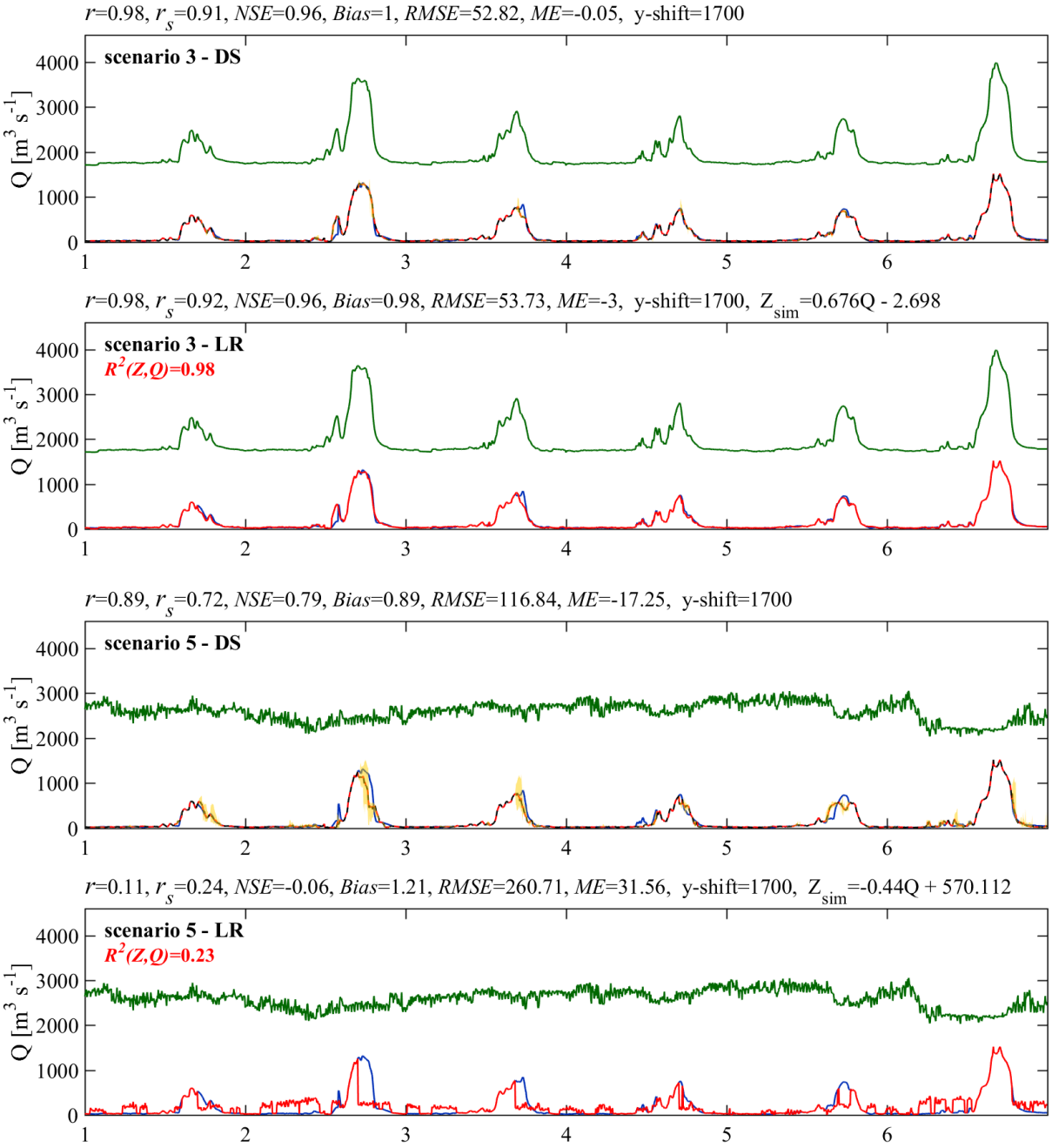
Appendix 1. Hydrographs at different streamflow gauging stations in the Volta basin

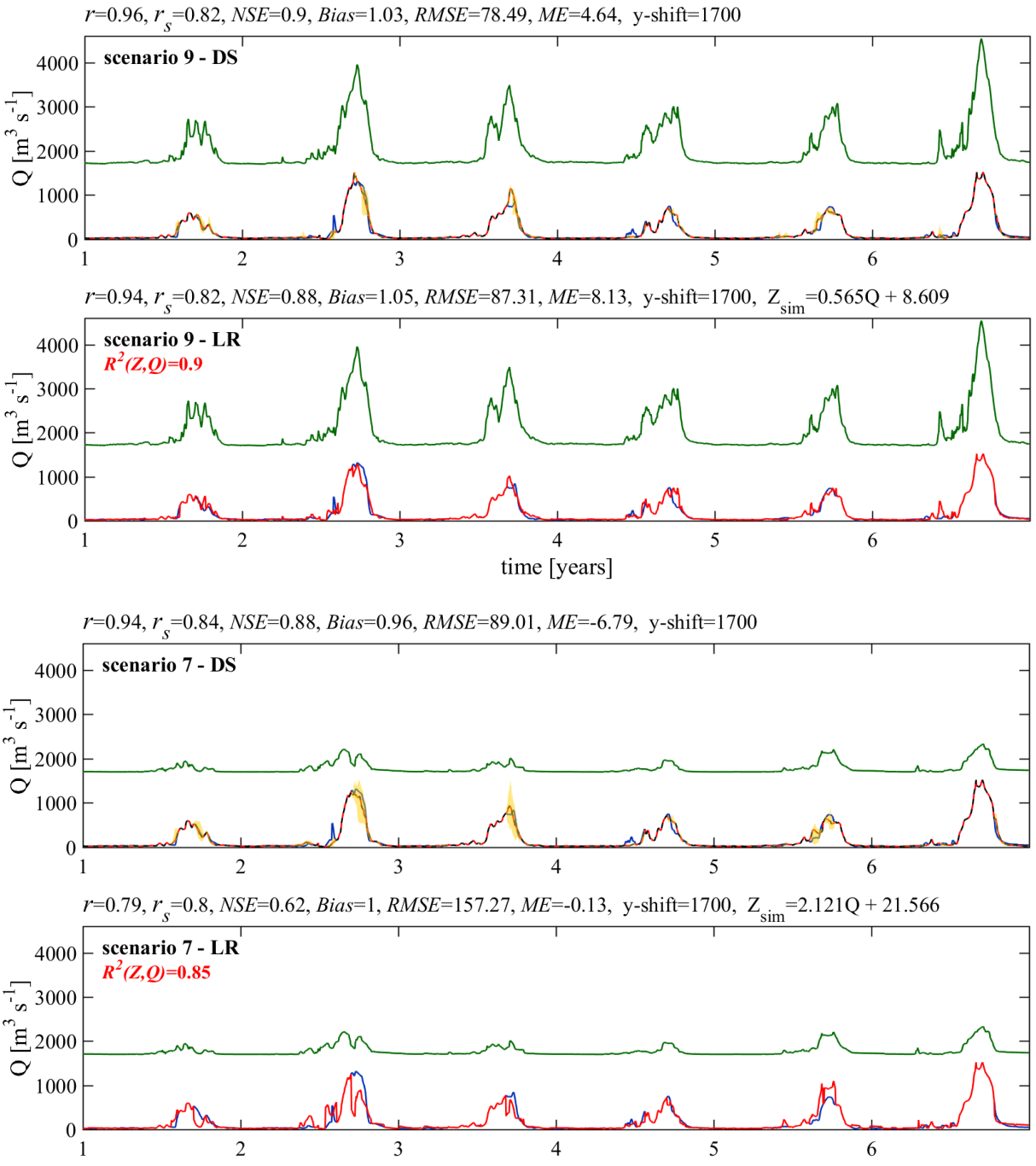


Appendix 2. Comparison of the Direct Sampling (DS) and the linear regression (LR) methods for gap-filling of streamflow time series Q(y) is shifted on the y-axis (y-shift) for display purpose.

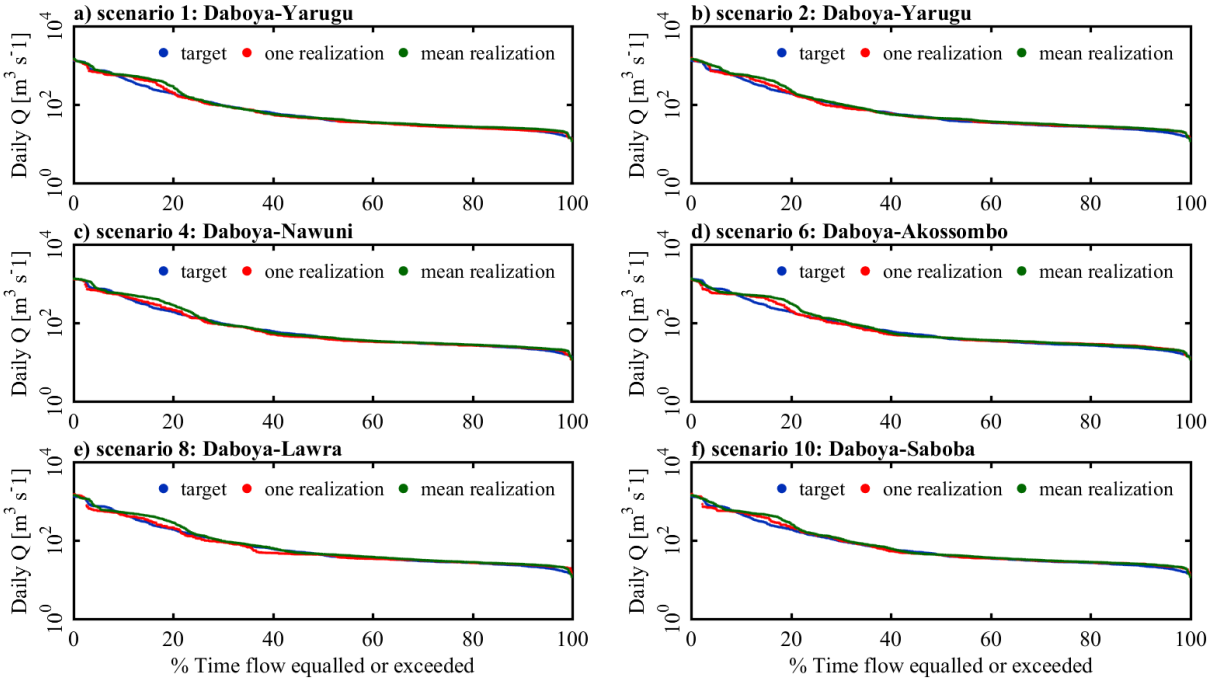
The supplementary materials contain the results of the comparison of the Direct Sampling method with the simple linear regression method, and flow duration curves of the gap filled time series. Simple linear regression (LR) is used to infill gaps in the target variable (Z) by using the corresponding predictor variable (Q) for scenarios described in Table 3.3. Only scenarios where the predictor is fully informed (i.e. without gaps) are presented because the LR method cannot be applied when the predictor variable also contains gaps at concomitant time steps with the target variable. This is, in exchange, possible with the DS method, which is one of its advantages over other gap-filling methods. The average scores of the statistical indicators are given alongside the plots with the regression equation for the LR method. Z_{sim} is the gap-filled time series of the target obtained by the LR method.



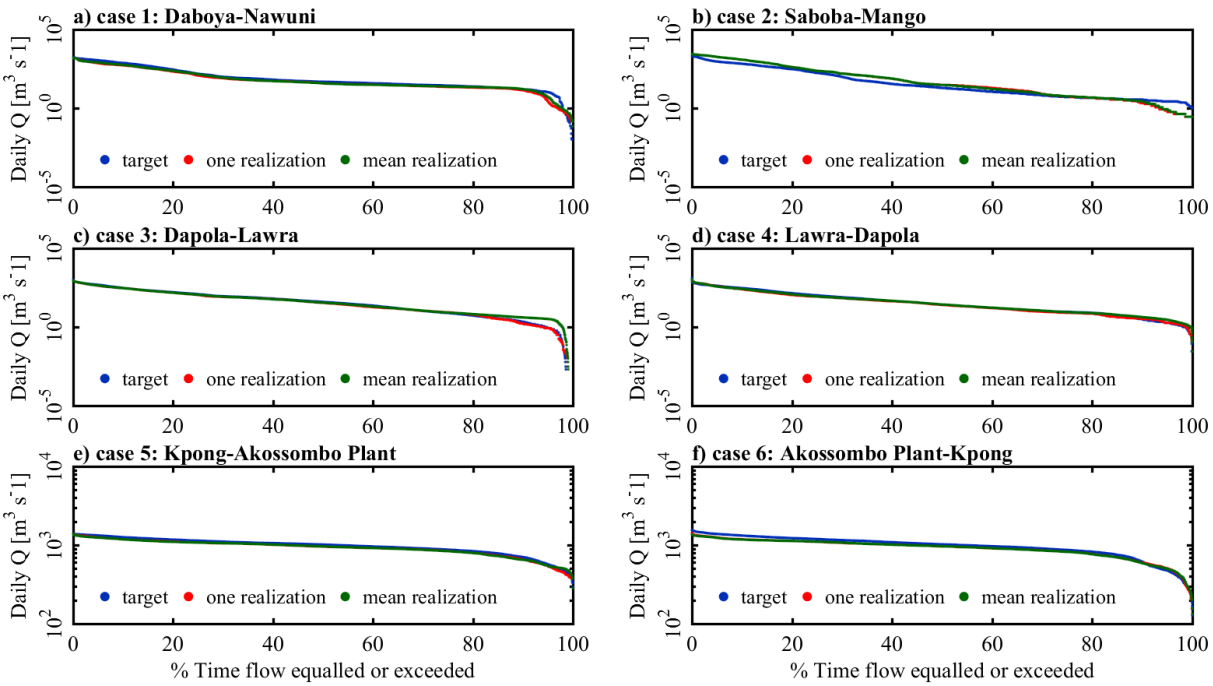

r=0.98, r_s=0.91, NSE=0.96, Bias=1, RMSE=52.82, ME=-0.05, y-shift=1700
scenario 3 - DS
Q [m^3 s^-1]
r=0.98, r_s=0.92, NSE=0.96, Bias=0.98, RMSE=53.73, ME=-3, y-shift=1700, Z_sim=0.676Q - 2.698
scenario 3 - LR
R^2(Z,Q)=0.98
r=0.89, r_s=0.72, NSE=0.79, Bias=0.89, RMSE=116.84, ME=-17.25, y-shift=1700
scenario 5 - DS
r=0.11, r_s=0.24, NSE=-0.06, Bias=1.21, RMSE=260.71, ME=31.56, y-shift=1700, Z_sim=-0.44Q + 570.112
scenario 5 - LR
R^2(Z,Q)=0.23


r=0.96, r_s=0.82, NSE=0.9, Bias=1.03, RMSE=78.49, ME=4.64, y-shift=1700
scenario 9 - DS
Q [m^3 s^-1]
r=0.94, r_s=0.82, NSE=0.88, Bias=1.05, RMSE=87.31, ME=8.13, y-shift=1700, Z_sim=0.565Q + 8.609
scenario 9 - LR
R^2(Z,Q)=0.9
time [years]
r=0.94, r_s=0.84, NSE=0.88, Bias=0.96, RMSE=89.01, ME=-6.79, y-shift=1700
scenario 7 - DS
r=0.79, r_s=0.8, NSE=0.62, Bias=1, RMSE=157.27, ME=-0.13, y-shift=1700, Z_sim=2.121Q + 21.566
scenario 7 - LR
R^2(Z,Q)=0.85

Appendix 3. Flow duration curves of the gap-filling scenarios with the Direct Sampling method



Appendix 4. Flow duration curves for different cases on the real gaps reconstruction with the Direct Sampling method



Appendix 5. Meteorological data access portals

Datasets	Name/ website
TAMSAT v3.0	Tropical Applications of Meteorology using SATellite (TAMSAT), African Rainfall Climatology and Time-series (TARCAT) https://www.tamsat.org.uk/data/archive
CHIRPS v2.0	Climate Hazards group InfraRed Precipitation with Stations (CHIRPS) V2.0 http://chg.ucsb.edu/data/chirps/
ARC v2.0	Africa Rainfall Estimate Climatology (ARC 2.0) https://www.cpc.ncep.noaa.gov/products/international/data.shtml
RFE v2.0	Climate Prediction Centre (CPC) African Rainfall Estimate (RFE) https://www.cpc.ncep.noaa.gov/products/international/data.shtml
MSWEP v2.2	Multi-Source Weighted-Ensemble Precipitation (MSWEP) V2.2 http://www.gloh2o.org/
GSMaP-std v6	Global Satellite Mapping of Precipitation (GSMaP) Moving Vector with Kalman (MVK) Standard V6 https://sharaku.eorc.jaxa.jp/GSMaP/
PERSIANN-CDR v1r1	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Climate Data Record (CDR) V1R1 http://chrsdata.eng.uci.edu/
CMORPH-CRT v1.0	Climate Prediction Centre (CPC) MORPHing technique (CMORPH) bias corrected (CRT) V1.0 www.cpc.ncep.noaa.gov
TRMM 3B42 v7	TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42 V7 https://mirador.gsfc.nasa.gov/
TRMM 3B42 RT v7	TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42 Real Time V7 https://mirador.gsfc.nasa.gov/
WFDEI-CRU	WATCH Forcing Data ERA-Interim (WFDEI) corrected using Climatic Research Unit (CRU) dataset www.eu-watch.org
WFDEI-GPCC	WATCH Forcing Data ERA-Interim (WFDEI) corrected using Global Precipitation Climatology Centre (GPCC) dataset ftp://rfdata:forceDATA@ftp.iiasa.ac.at/
PGF v3	Princeton University global meteorological forcing (PGF) http://hydrology.princeton.edu/data/pgf/
ERA5	European Centre for Medium-range Weather Forecasts ReAnalysis 5 (ERA5) hourly data on single levels https://cds.climate.copernicus.eu/
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications 2 (rainfall: M2T1NXFLX_V5.12.4; temperature: M2SDNXSLV_V5.12.4) https://disc.gsfc.nasa.gov/datasets/
EWEMBI v1.1	EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI) http://doi.org/10.5880/pik.2016.004
JRA-55	Japanese 55 year ReAnalysis (JRA-55); rainfall: fcst_phy2m125; temperature: anl_surf125 https://jra.kishou.go.jp/JRA-55/index_en.html

Appendix 6. Hydrological modelling data access portals

Products	Data access portals
GMTED 2010	https://topotools.cr.usgs.gov/
SoilGrids	https://www.isric.org/explore/soilgrids
GLiM v1.0	https://doi.pangaea.de/10.1594/PANGAEA.788537
Globcover 2009	http://due.esrin.esa.int/page_globcover.php
GIMMS	http://cliveg.bu.edu/modismisr/lai3g-fpar3g.html
GRACE TellUS v5.0	https://grace.jpl.nasa.gov/
ESA CCI SM v4.2	https://www.esa-soilmoisture-cci.org/
GLEAM v3.2a	https://www.gleam.eu/
MYD11A2 v6	https://lpdaac.usgs.gov/products/myd11a2v006/

Appendix 7. Comparison of *SPAEF* to E_{SP}

The following presents an experiment to compare *SPAEF* (Koch et al., 2018; Demirel et al., 2018) and E_{SP} (spatial pattern efficiency metric, Eq. 4.8). The experiment is done as follows:

- An observed variable (OBS) is defined as a 3x3 matrix containing a random permutation of values between 1 and 9.
- A tentative to reproduce OBS is done by producing 500 random permutations of the cells in OBS.
- Each permutation of OBS corresponds to a simulated variable (SIM). The best simulation is considered as the one having the highest *SPAEF* or E_{SP}, and the worst simulation has the lowest values of *SPAEF* or E_{SP}. The simulations are numbered from #1 to #500.

Figure A5 shows the results of the experiment. A summary of the results is given in Table A5.

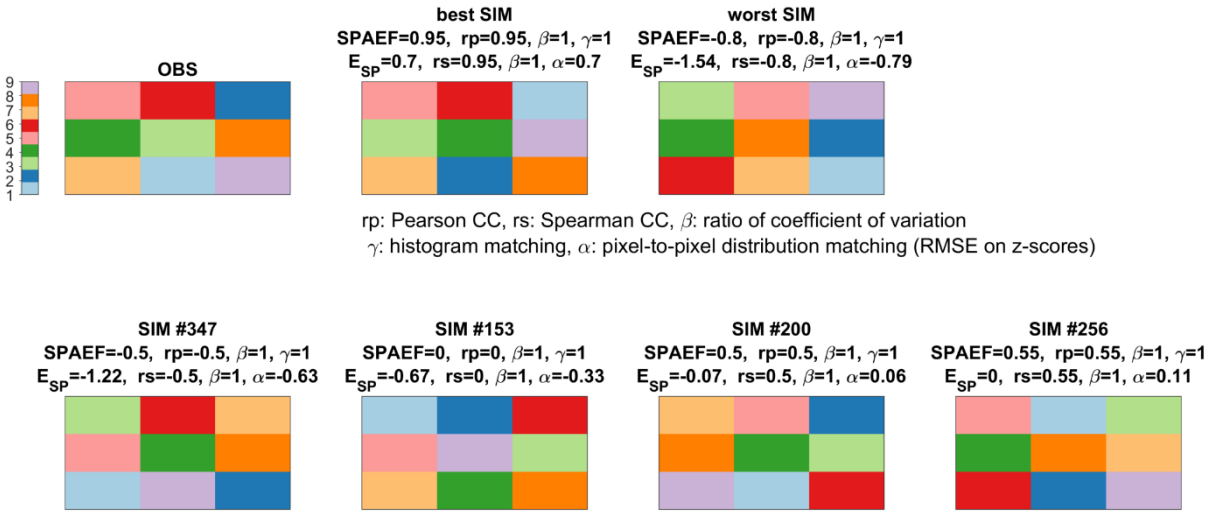


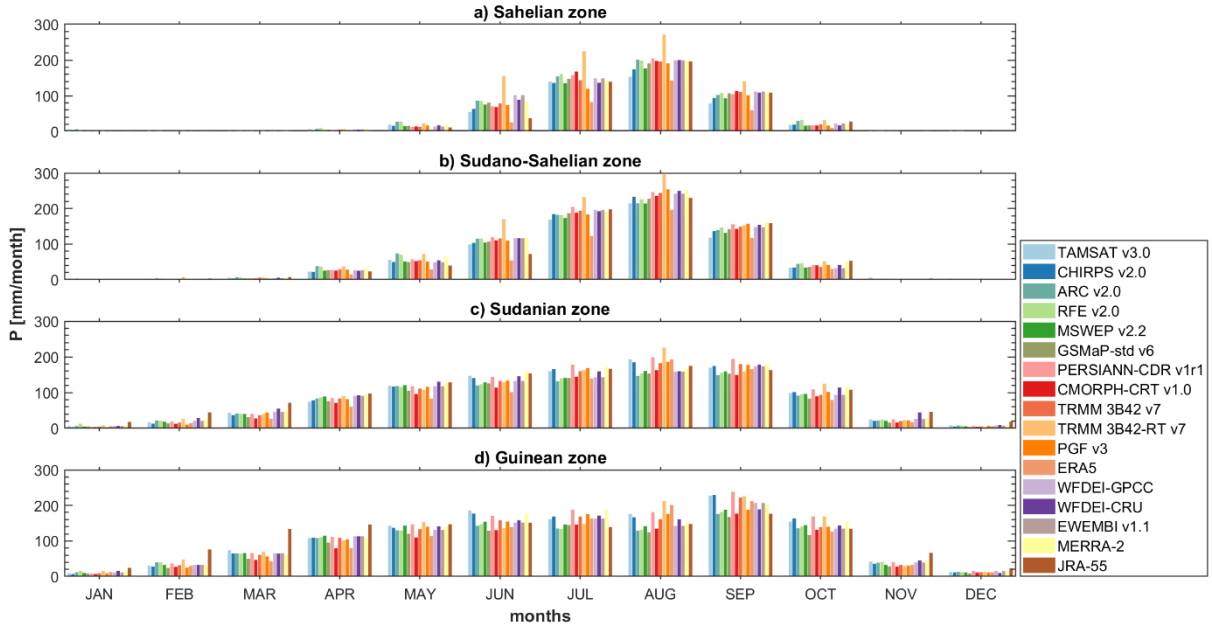
Figure A5 Comparison of *SPAEF* and E_{SP}. The best and worst simulations (SIM) of the observed variable (OBS) are presented with the corresponding *SPAEF* and E_{SP} scores along with their components. Selected simulations #347, #153, #200 and #256 show the difference between *SPAEF* and E_{SP} for different simulated spatial patterns.

Knowing that the optimal value for both E_{SP} and *SPAEF* is 1, we can conclude that E_{SP} is more discriminant than *SPAEF* because for 33% of pixels being located correctly (Figure A5 and Table A5), the *SPAEF*=0.95 while E_{SP} =0.7 for the best SIM (or *SPAEF*=0.55 and E_{SP} =0 for SIM#256). Because of the histogram match, *SPAEF* overestimate the efficiency of spatial patterns when the spatial variability is high (i.e. β).

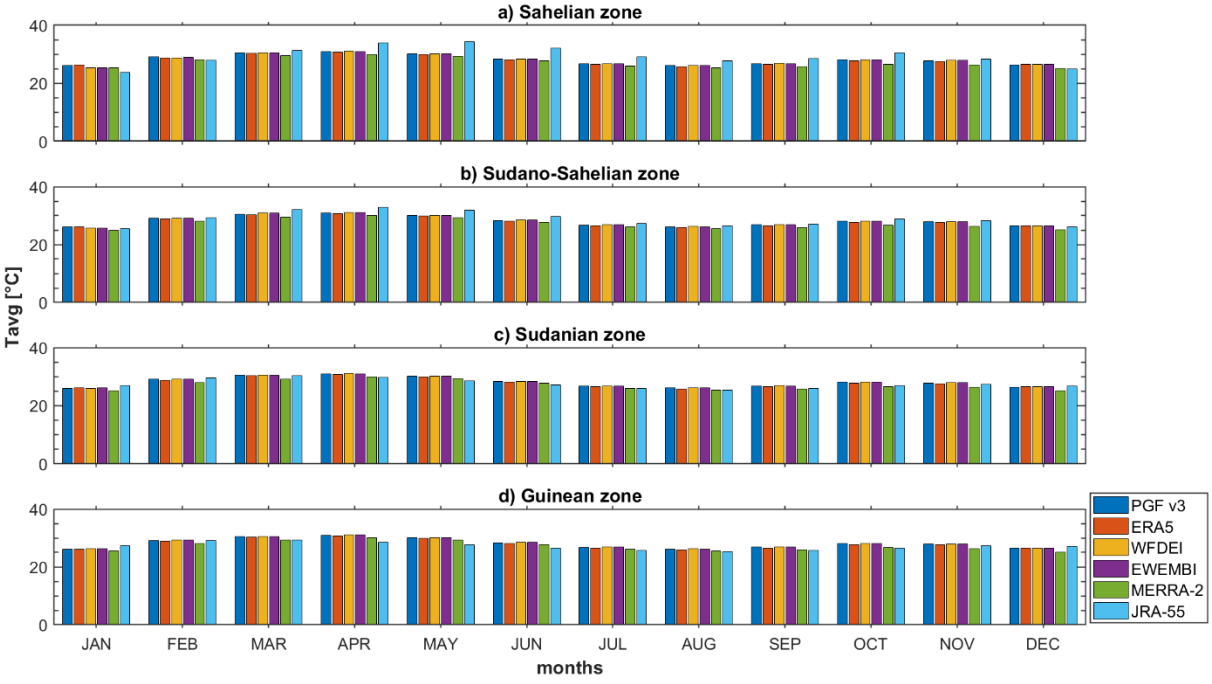
Table A5 Results of the comparison of *SPAEF* to E_{SP} for selected simulations

SIM	**# color matches**	**% pixel location match**	**Color match**	***SPAEF***	E_{SP}	r_p **or** r_s	α
Worst	1	11%	dark green	-0.8	-1.54	-0.8	-0.79
#347	1	11%	red	-0.5	-0.74	-0.5	-0.63
#153	1	11%	yellow	0	-0.67	0	-0.33
#200	2	22%	dark blue, light blue	0.5	-0.07	0.5	0.06
#256	3	33%	pink, dark green, purple	0.55	0	0.55	0.11
Best	3	33%	red, pink, yellow	0.95	0.7	0.95	0.7

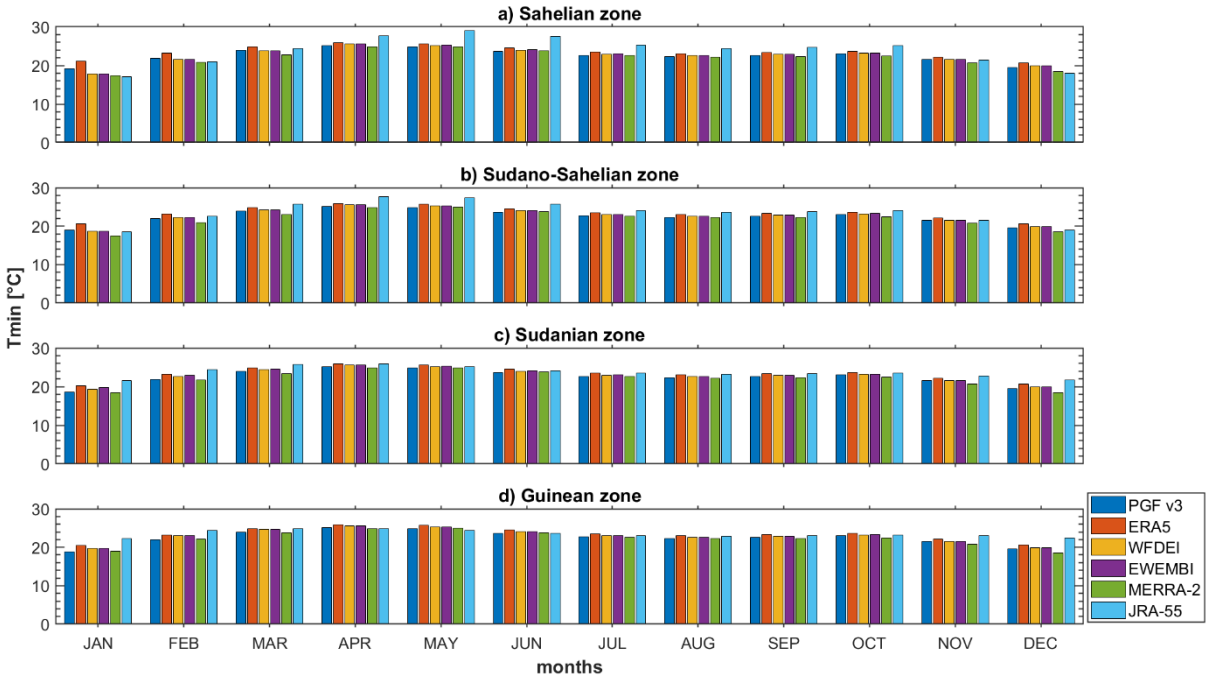
Appendix 8. Climatology of mean monthly rainfall totals over the period 2003-2012, averaged over four climatic zones (a, b, c and d). . The coloured bars represent 17 rainfall datasets.



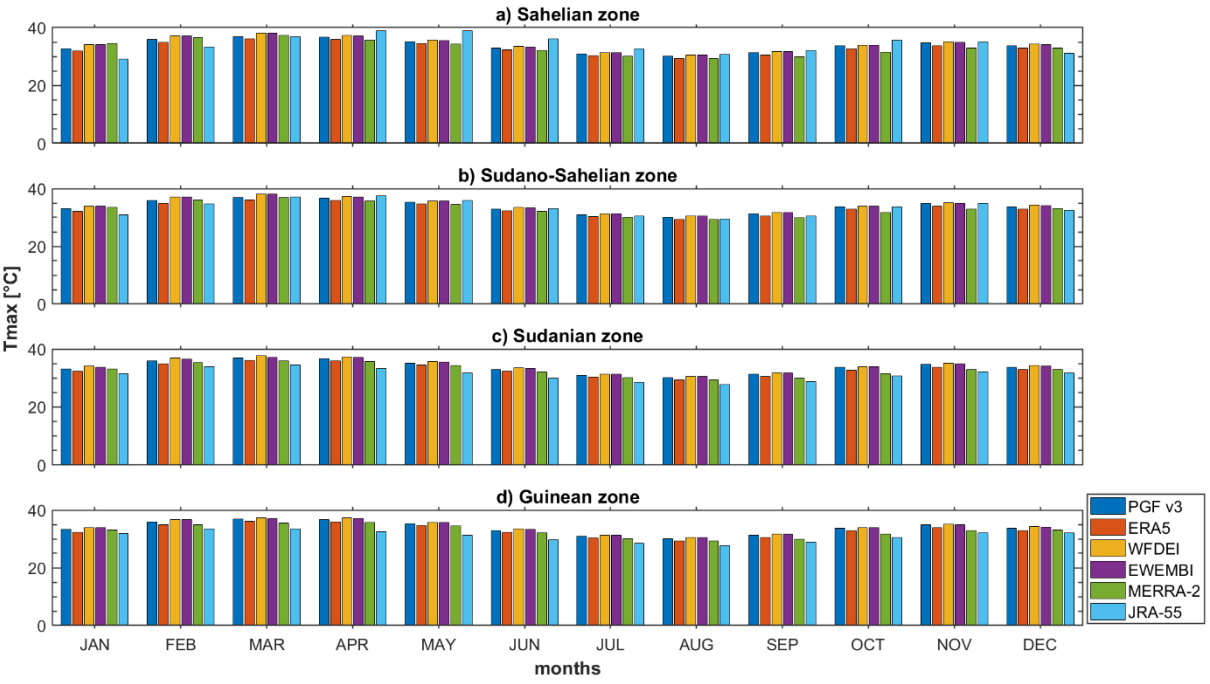
Appendix 9. Climatology of mean monthly average air temperature over the period 2003-2012, averaged over four climatic zones (a, b, c and d). The coloured bars represent 6 temperature datasets.



Appendix 10. Climatology of mean monthly minimum air temperature over the period 2003-2012, averaged over four climatic zones (a, b, c and d). The coloured bars represent 6 temperature datasets.



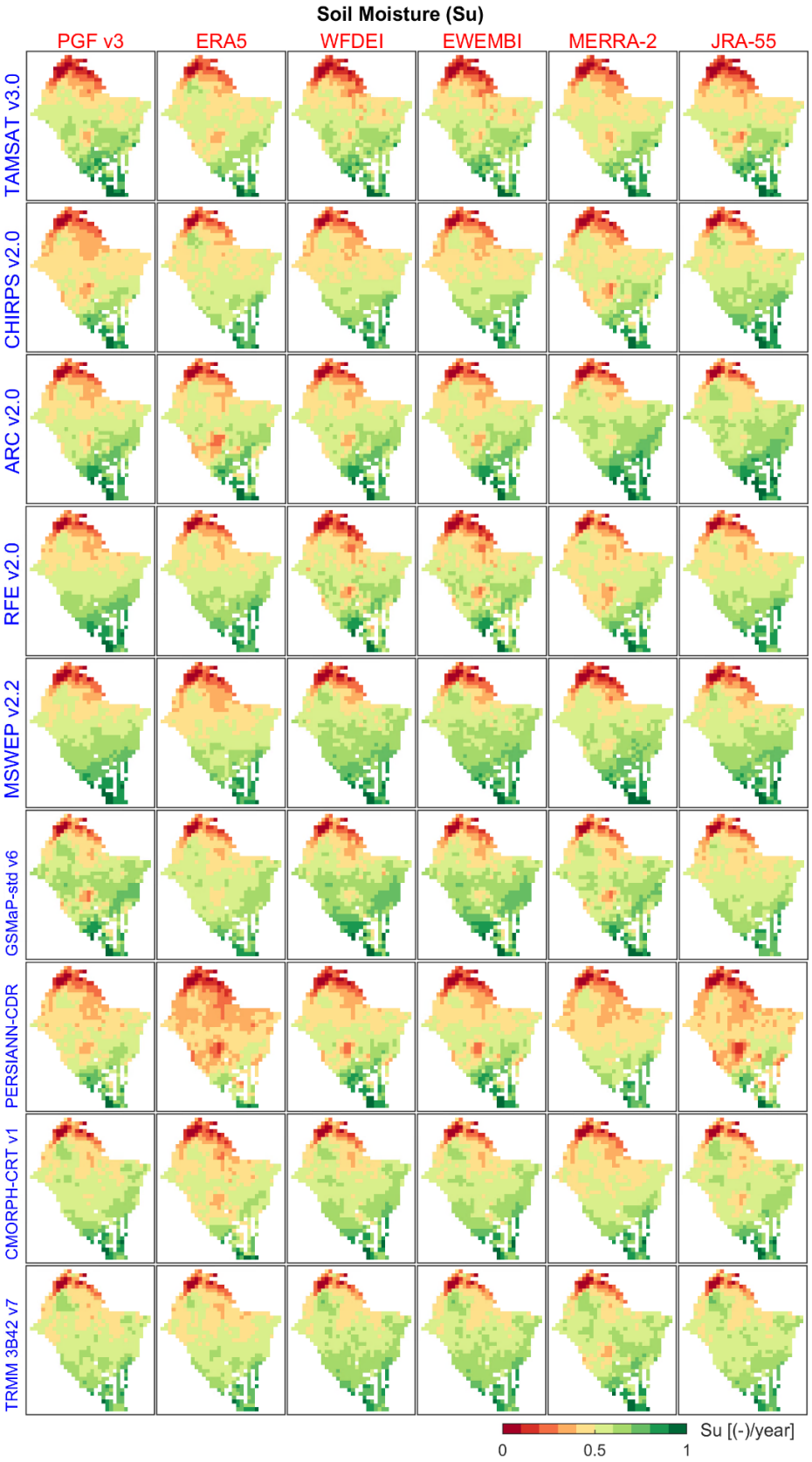
Appendix 11. Climatology of mean monthly maximum air temperature over the period 2003-2012, averaged over four climatic zones (a, b, c and d). The coloured bars represent 6 temperature datasets.

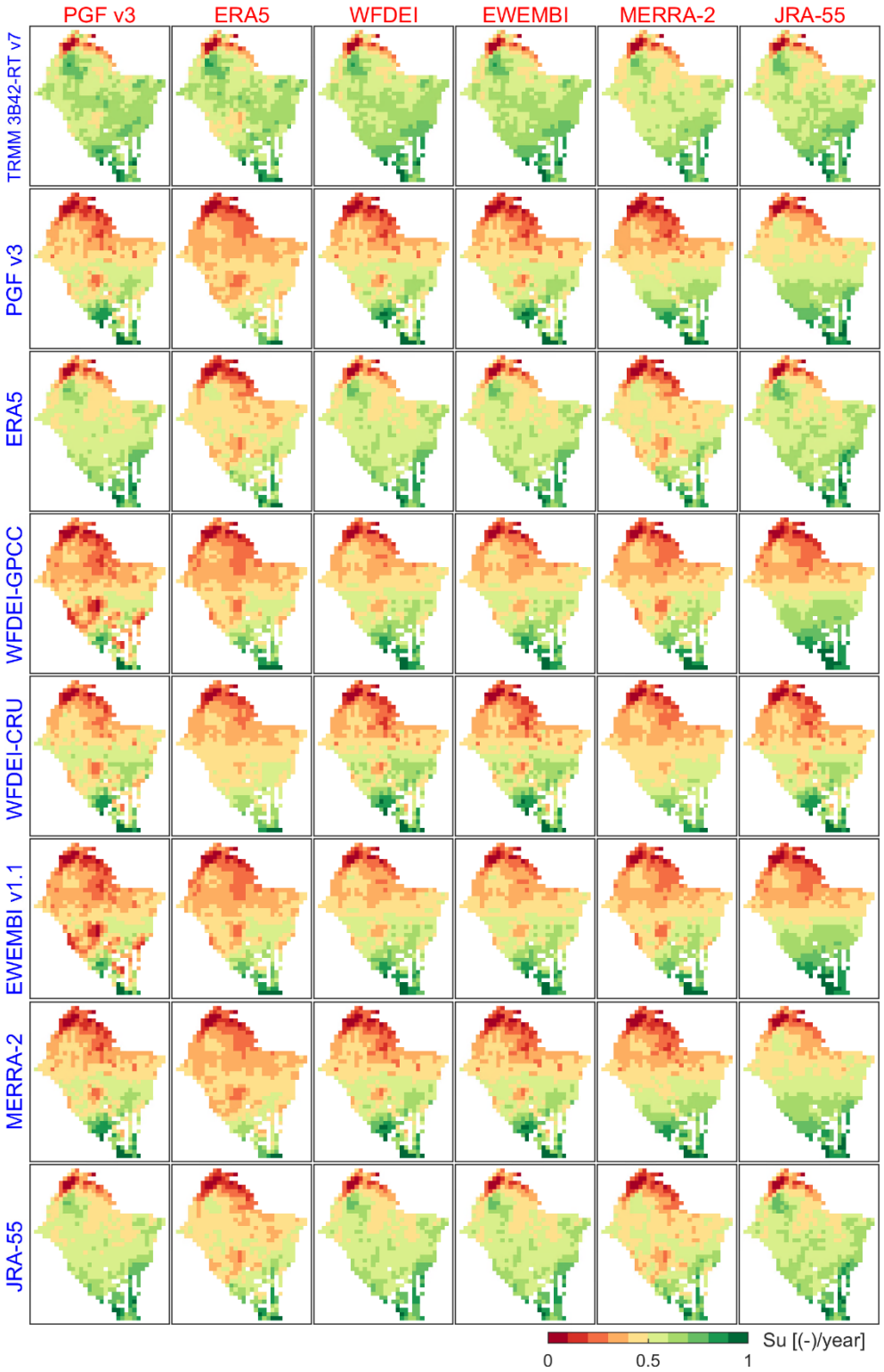


Appendix 12. Best rainfall-temperature dataset combinations for simulating the spatial patterns and the temporal dynamics of streamflow (Q), terrestrial water storage (S_t), soil moisture (S_u) and actual evaporation (E_a) in the Volta River basin.

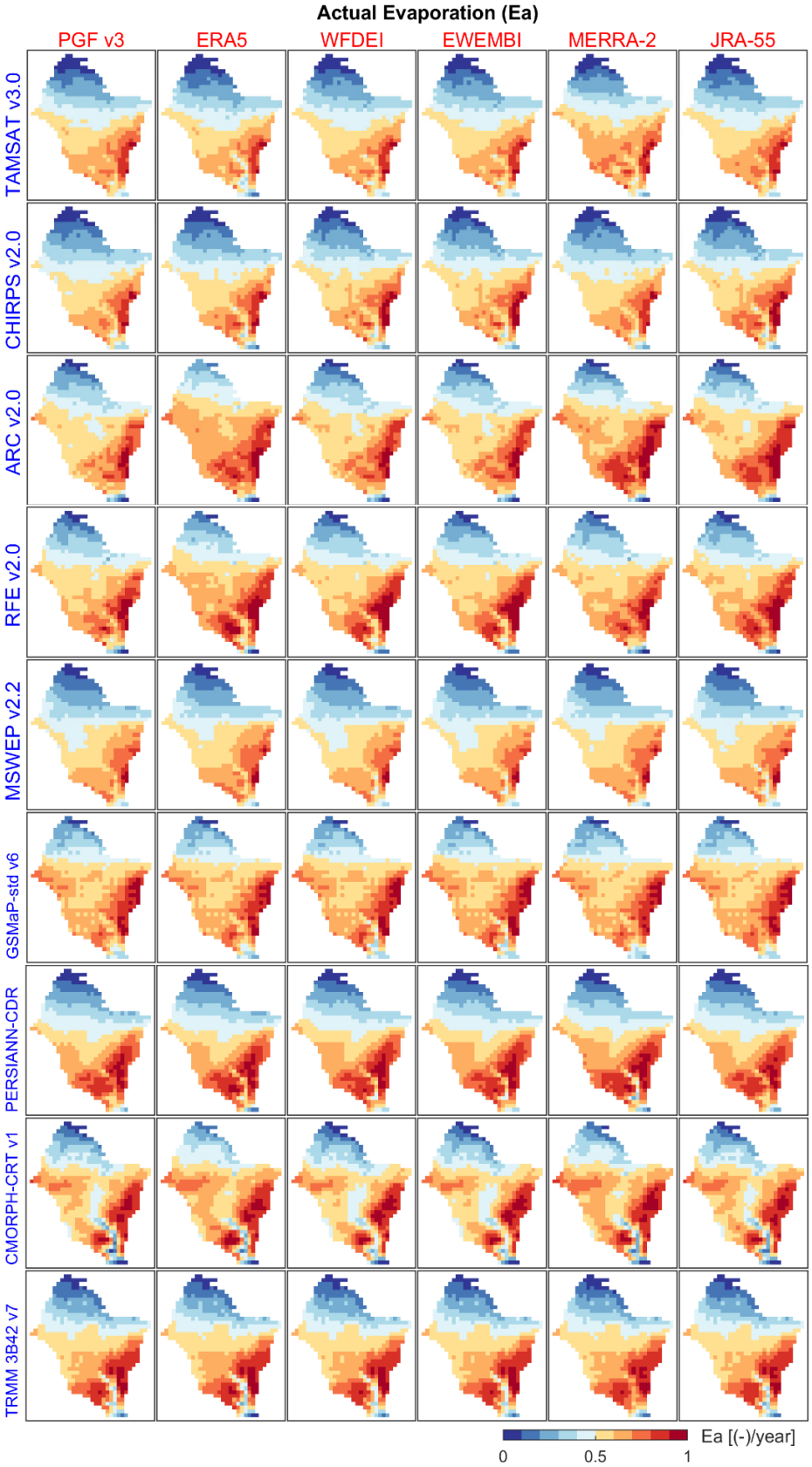
		Top 5 rainfall datasets for a given temperature dataset				
		1	2	3	4	5
		Temporal dynamics of *Q*				
Temperature datasets	JRA-55	TAMSAT	GSMaP-std	CHIRPS	MERRA-2	CMORPH-CRT
	MERRA-2	TAMSAT	GSMaP-std	PERSIANN-CDR	CHIRPS	CMORPH-CRT
	EWEMBI	TAMSAT	GSMaP-std	CHIRPS	MERRA-2	ARC
	WFDEI	TAMSAT	GSMaP-std	CHIRPS	MERRA-2	ARC
	ERA5	CHIRPS	TAMSAT	CMORPH-CRT	GSMaP-std	ARC
	PGF v3	TAMSAT	GSMaP-std	PERSIANN-CDR	CHIRPS	MERRA-2
		Temporal dynamics of *St*				
	JRA-55	CMORPH-CRT	ARC	RFE	TAMSAT	CHIRPS
	MERRA-2	ARC	RFE	TRMM 3B42-RT	WFDEI-CRU	CHIRPS
	EWEMBI	ARC	CHIRPS	TRMM 3B42-RT	GSMaP-std	ERA5
	WFDEI	ARC	CHIRPS	TRMM 3B42-RT	GSMaP-std	ERA5
	ERA5	ARC	RFE	CMORPH-CRT	TAMSAT	TRMM 3B42-RT
	PGF v3	CMORPH-CRT	RFE	ARC	CHIRPS	TAMSAT
		Temporal dynamics of *Su*				
	JRA-55	WFDEI-GPCC	EWEMBI	MERRA-2	PGF	TAMSAT
	MERRA-2	MERRA-2	WFDEI-GPCC	EWEMBI	PERSIANN-CDR	ERA5
	EWEMBI	WFDEI-CRU	PGF	PERSIANN-CDR	MERRA-2	GSMaP-std
	WFDEI	WFDEI-CRU	PGF	PERSIANN-CDR	MERRA-2	GSMaP-std
	ERA5	WFDEI-GPCC	EWEMBI	MERRA-2	PGF	ERA5
	PGF v3	WFDEI-GPCC	EWEMBI	WFDEI-CRU	MERRA-2	TAMSAT
		Temporal dynamics of *Ea*				
	JRA-55	ARC	RFE	TAMSAT	WFDEI-GPCC	EWEMBI
	MERRA-2	ARC	CMORPH-CRT	RFE	WFDEI-GPCC	EWEMBI
	EWEMBI	RFE	GSMaP-std	ARC	TAMSAT	PERSIANN-CDR
	WFDEI	RFE	GSMaP-std	ARC	TAMSAT	PERSIANN-CDR
	ERA5	ARC	RFE	WFDEI-GPCC	EWEMBI	GSMaP-std
	PGF v3	TAMSAT	WFDEI-GPCC	EWEMBI	WFDEI-CRU	ARC
		Spatial patterns of *Su*				
	JRA-55	WFDEI-GPCC	EWEMBI	TAMSAT	MSWEP	RFE
	MERRA-2	MSWEP	PGF	ARC	WFDEI-CRU	WFDEI-GPCC
	EWEMBI	TAMSAT	PERSIANN-CDR	MSWEP	WFDEI-GPCC	EWEMBI
	WFDEI	TAMSAT	PERSIANN-CDR	MSWEP	WFDEI-GPCC	EWEMBI
	ERA5	MSWEP	RFE	WFDEI-GPCC	EWEMBI	WFDEI-CRU
	PGF v3	RFE	TAMSAT	MSWEP	ARC	MERRA-2
		Spatial patterns of *Ea*				
	JRA-55	TAMSAT	WFDEI-GPCC	EWEMBI	MSWEP	MERRA-2
	MERRA-2	MSWEP	CHIRPS	MERRA-2	WFDEI-CRU	PGF
	EWEMBI	PERSIANN-CDR	TAMSAT	MSWEP	WFDEI-GPCC	EWEMBI
	WFDEI	PERSIANN-CDR	TAMSAT	MSWEP	WFDEI-GPCC	EWEMBI
	ERA5	MSWEP	MERRA-2	WFDEI-GPCC	EWEMBI	PERSIANN-CDR
	PGF v3	MSWEP	CHIRPS	TAMSAT	MERRA-2	PERSIANN-CDR

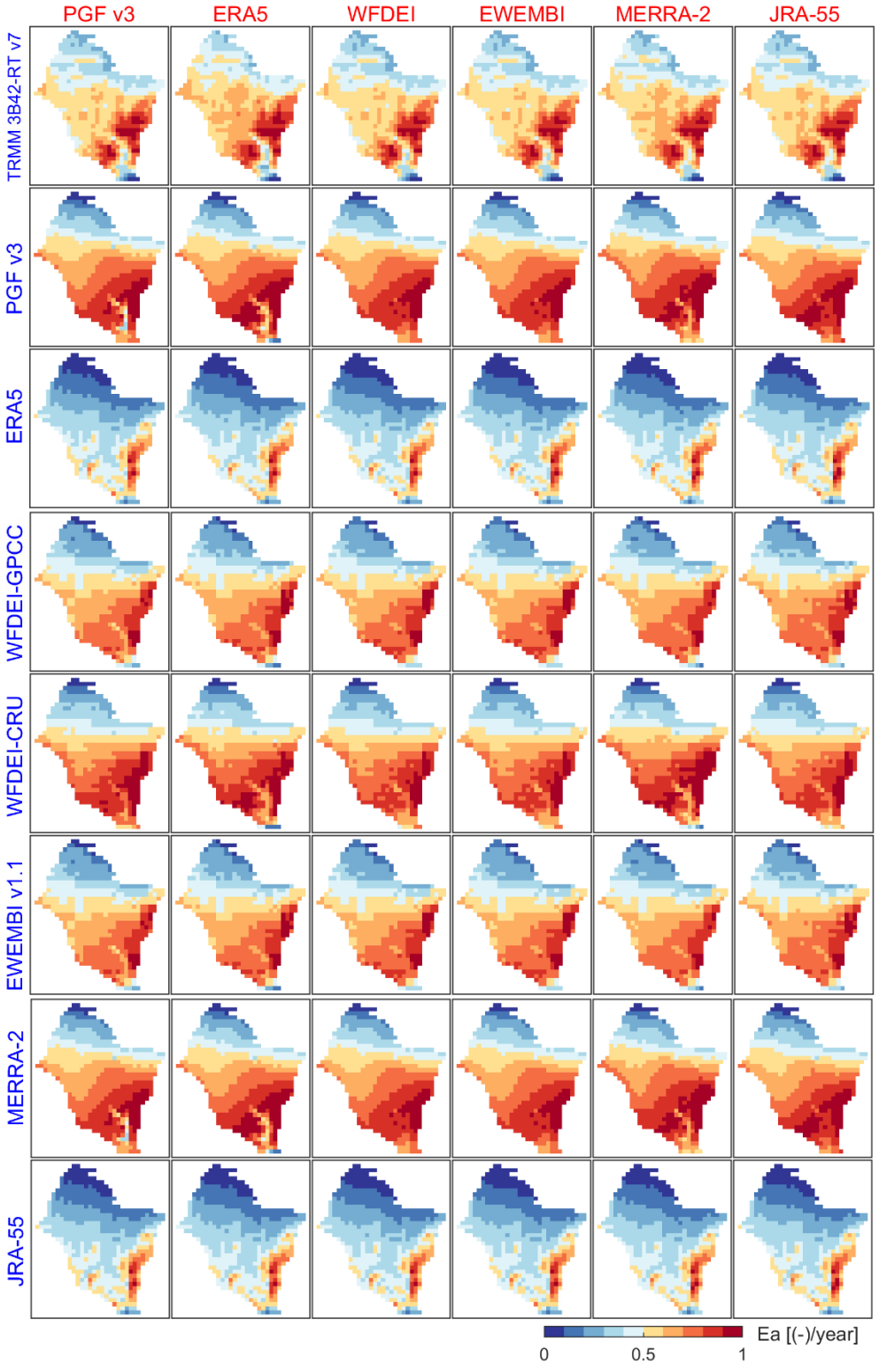
Appendix 13. Maps of long-term (2003-2012) annual average of soil moisture (S_u) obtained as outputs of hydrological modelling using different combinations of rainfall datasets (y-axis, blue font) and temperature datasets (x-axis, red font). The values are normalized for better emphasizing on patterns and using a unique color scale.

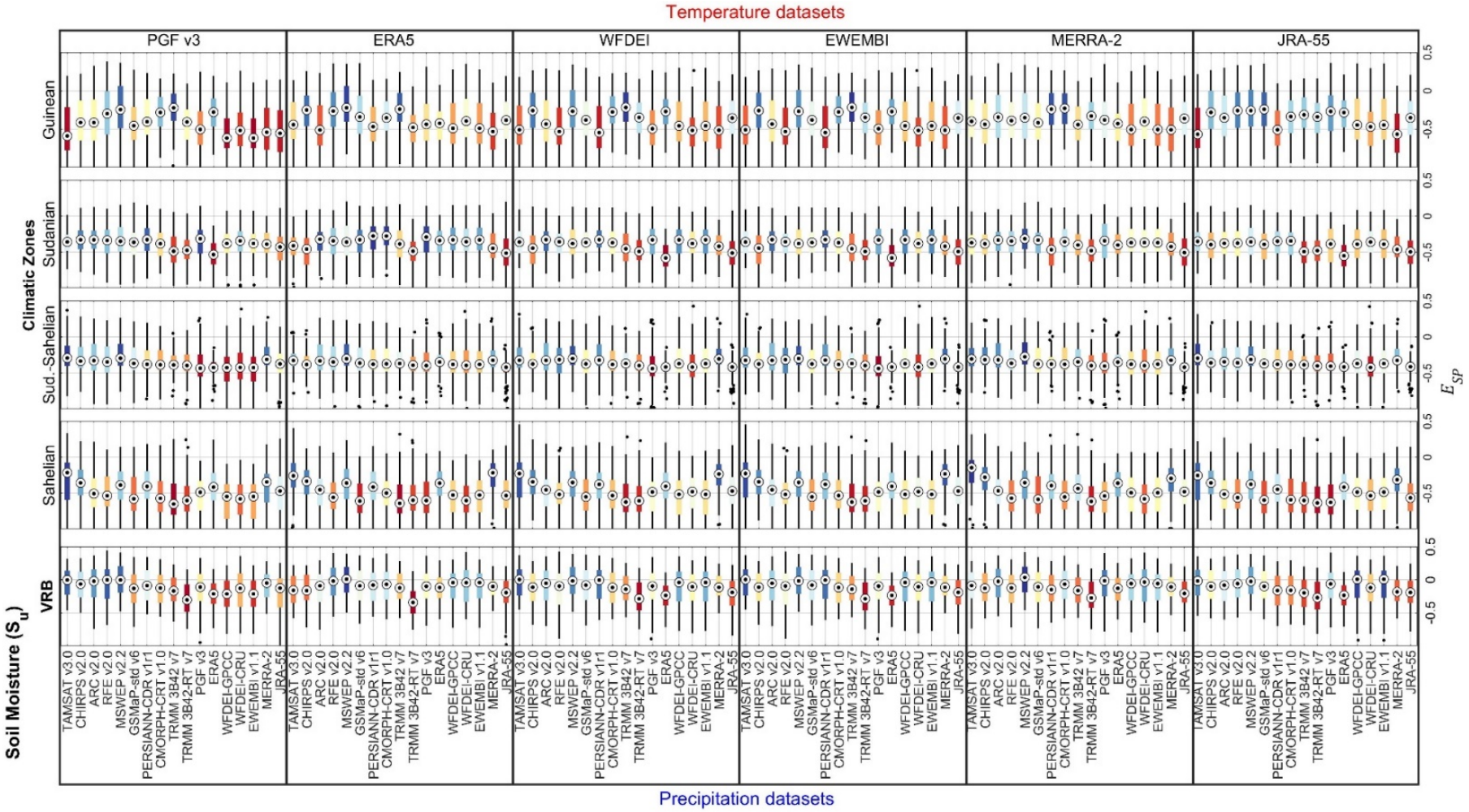



PGF v3
ERA5
WFDEI
EWEMBI
MERRA-2
JRA-55
TRMM 3B42-RT v7
PGF v3
ERA5
WFDEI-GPCC
WFDEI-CRU
EWEMBI v1.1
MERRA-2
JRA-55
Su [(-)/year]
0
0.5
1

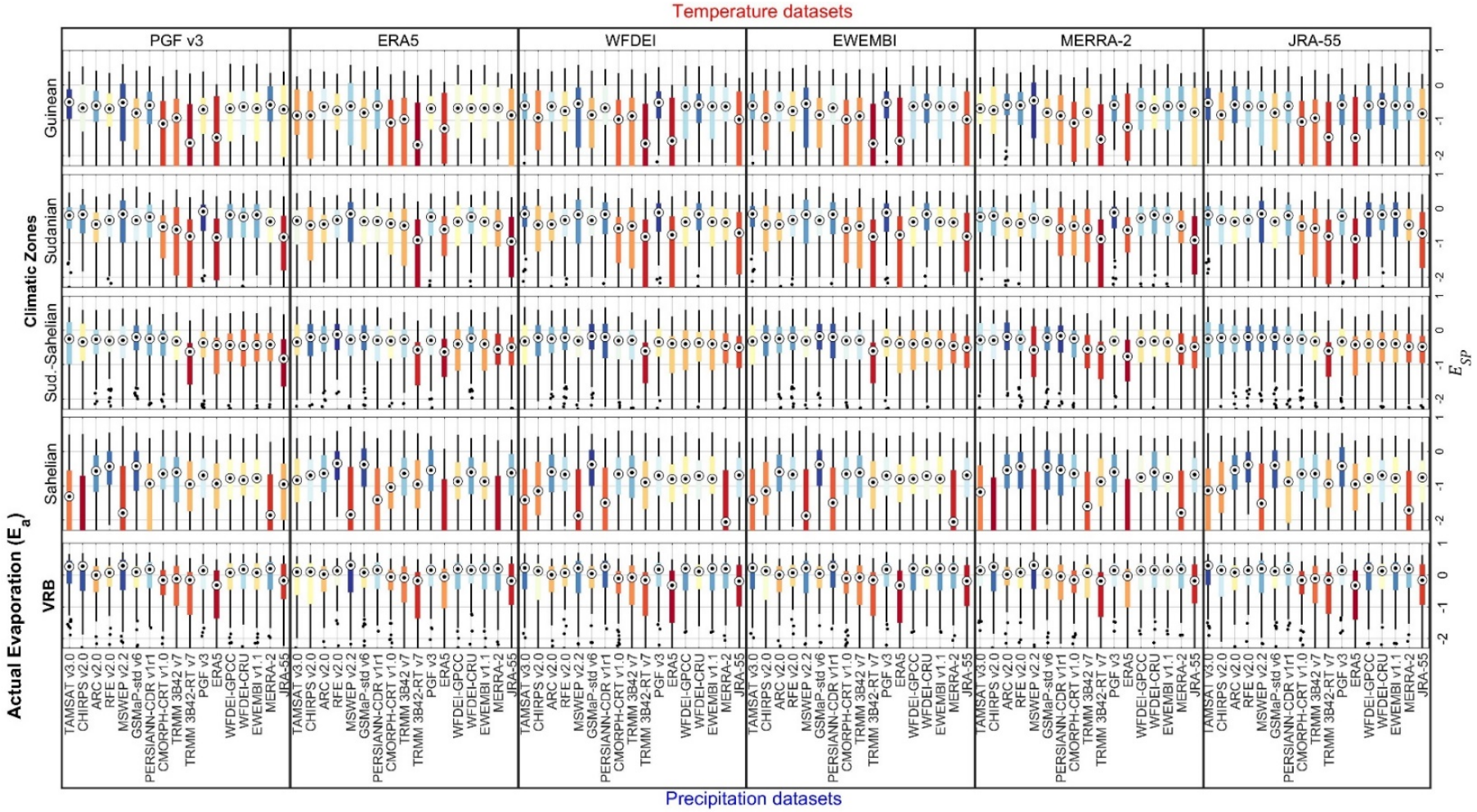
Appendix 14. Maps of long-term (2003-2012) annual average of actual evaporation (E_a) obtained as outputs of hydrological modelling using different combinations of rainfall datasets (y-axis, blue font) and temperature datasets (x-axis, red font). The values are normalized for better emphasizing on patterns and using a unique color scale.




PGF v3
ERA5
WFDEI
EWEMBI
MERRA-2
JRA-55
TRMM 3B42-RT v7
PGF v3
ERA5
WFDEI-GPCC
WFDEI-CRU
EWEMBI v1.1
MERRA-2
JRA-55
0
0.5
1
Ea [(-)/year]



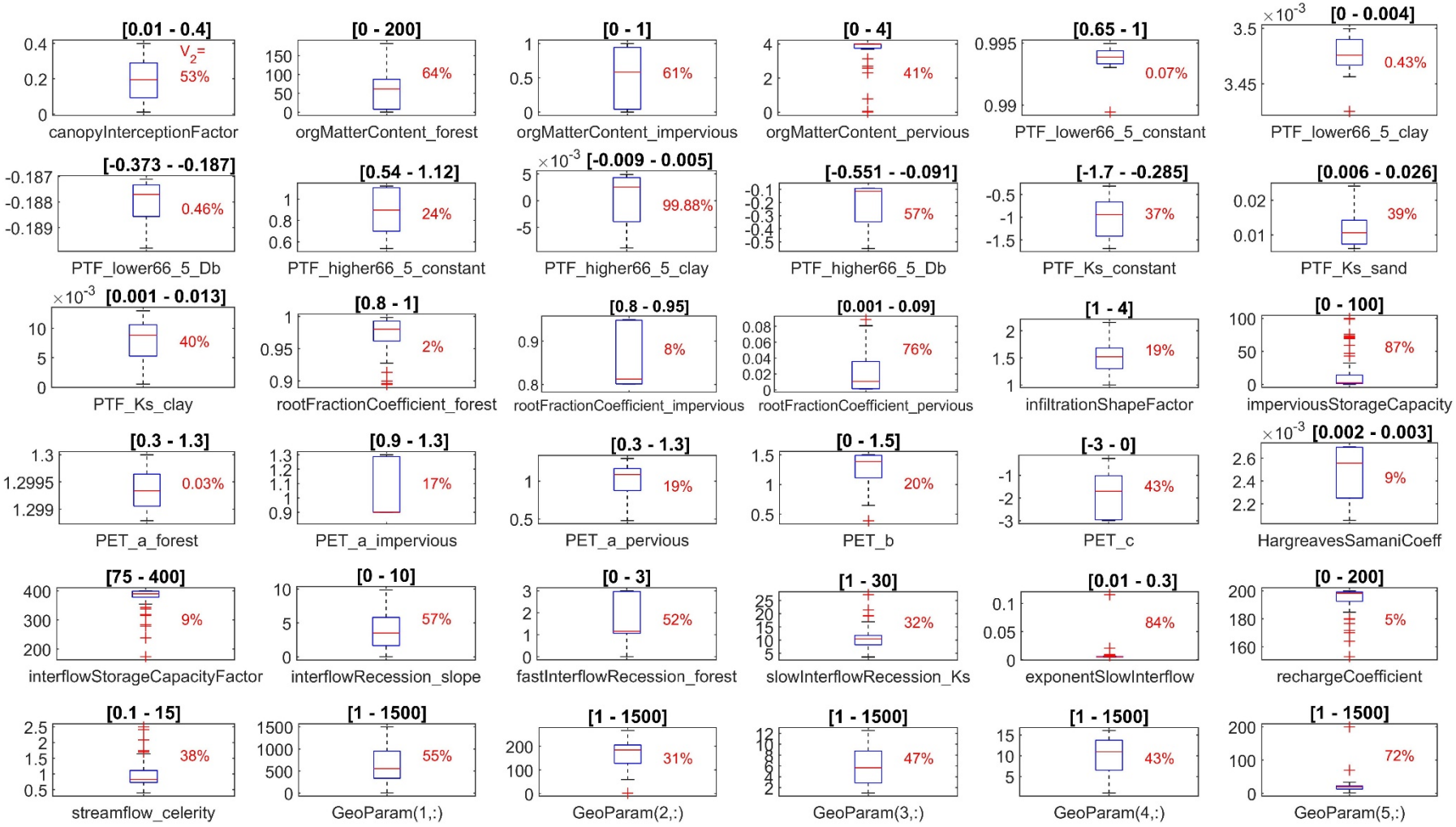
Appendix 15. Spatial pattern efficiency (E_{SP}) of soil moisture (S_u) over the entire simulation period (2003-2012) for the Volta River basin (VRB) and its climatic zones, using different combinations of precipitation and temperature datasets used as input for hydrological modelling. Each boxplot has 120 values corresponding to the number of months. The boxplots are coloured from the best (blue) to the worst performance (red) based on the median value.



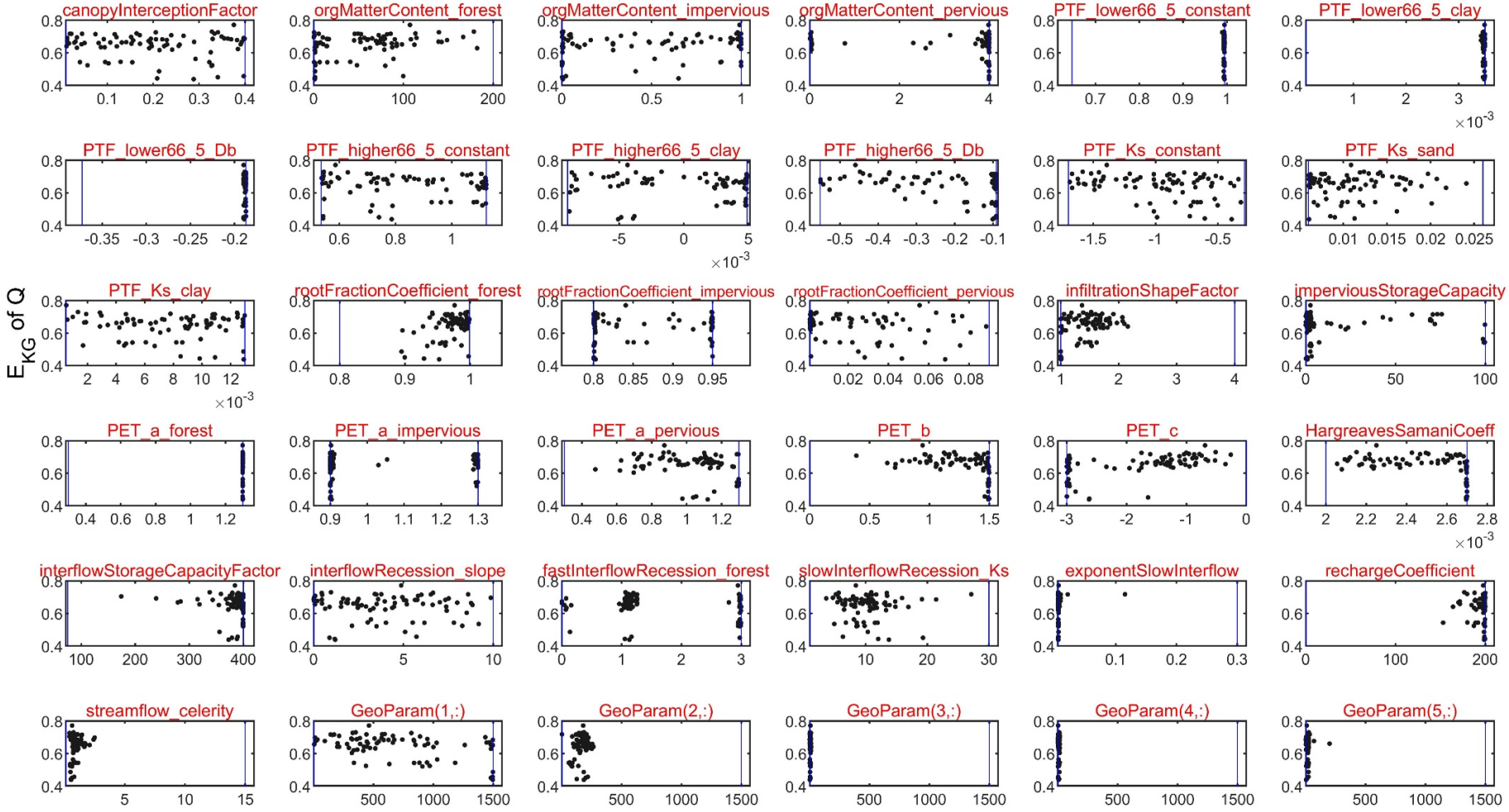
Appendix 16. Spatial pattern efficiency (E_{SP}) of actual evaporation (E_a) over the entire simulation period (2003-2012) for the Volta River basin (VRB) and its climatic zones, using different combinations of precipitation and temperature datasets used as input for hydrological modelling. Each boxplot has 120 values corresponding to the number of months. The boxplots are coloured from the best (blue) to the worst performance (red) based on the median value.

Appendix 17. mHM global parameters. The description of the model parameters can be found in the work of Samaniego et al. (2010).

Parameters	Lower bound	Upper bound	Initial value
Interception			
canopyInterceptionFactor	0.01	0.40	0.278
Soil Moisture			
orgMatterContent_forest	0.00	200.00	105.84
orgMatterContent_impervious	0.00	1.00	0.755
orgMatterContent_pervious	0.00	4.00	2.967
PTF_lower66_5_constant	0.65	1.00	0.939
PTF_lower66_5_clay	0.000	0.004	0.003
PTF_lower66_5_Db	-0.373	-0.187	-0.204
PTF_higher66_5_constant	0.536	1.123	0.924
PTF_higher66_5_clay	-0.009	0.005	-0.001
PTF_higher66_5_Db	-0.551	-0.091	-0.107
PTF_Ks_constant	-1.700	-0.285	-0.424
PTF_Ks_sand	0.0060	0.0260	0.0064
PTF_Ks_clay	0.001	0.013	0.005
rootFractionCoefficient_forest	0.800	0.999	0.864
rootFractionCoefficient_impervious	0.800	0.950	0.912
rootFractionCoefficient_pervious	0.001	0.090	0.022
infiltrationShapeFactor	1.000	4.000	3.945
Direct sealed area runoff			
imperviousStorageCapacity	0.00	100.00	5.340
Potential evapotranspiration			
PET_a_forest	0.30	1.30	0.5282
PET_a_impervious	0.90	1.30	0.9743
PET_a_pervious	0.30	1.30	0.5088
PET_b	0.00	1.50	1.4993
PET_c	-3.00	0.00	-1.0330
HargreavesSamaniCoeff	0.0020	0.0027	0.0023
Interflow			
interflowStorageCapacityFactor	75.00	400.00	313.85
interflowRecession_slope	0.00	10.00	5.51
fastInterflowRecession_forest	0.00	3.00	2.52
slowInterflowRecession_Ks	1.00	30.00	22.21
exponentSlowInterflow	0.005	0.300	0.283
Percolation			
rechargeCoefficient	0.00	200.00	183.51
Routing			
streamflow_celerity	0.10	15.00	12.47
Geology			
GeoParam1	1.00	1500.00	1037.16
GeoParam2	1.00	1500.00	709.99
GeoParam3	1.00	1500.00	381.09
GeoParam4	1.00	1500.00	525.30
GeoParam5	1.00	1500.00	783.83



Appendix 18 Distribution of mHM global parameters after model calibration with different combinations of precipitation and temperature datasets as inputs. Each boxplot has 102 elements corresponding to parameter values obtained with different input datasets. The initial parameters' ranges are provided in squared brackets. The second-order coefficient of variation (V_2) is given in percentage.

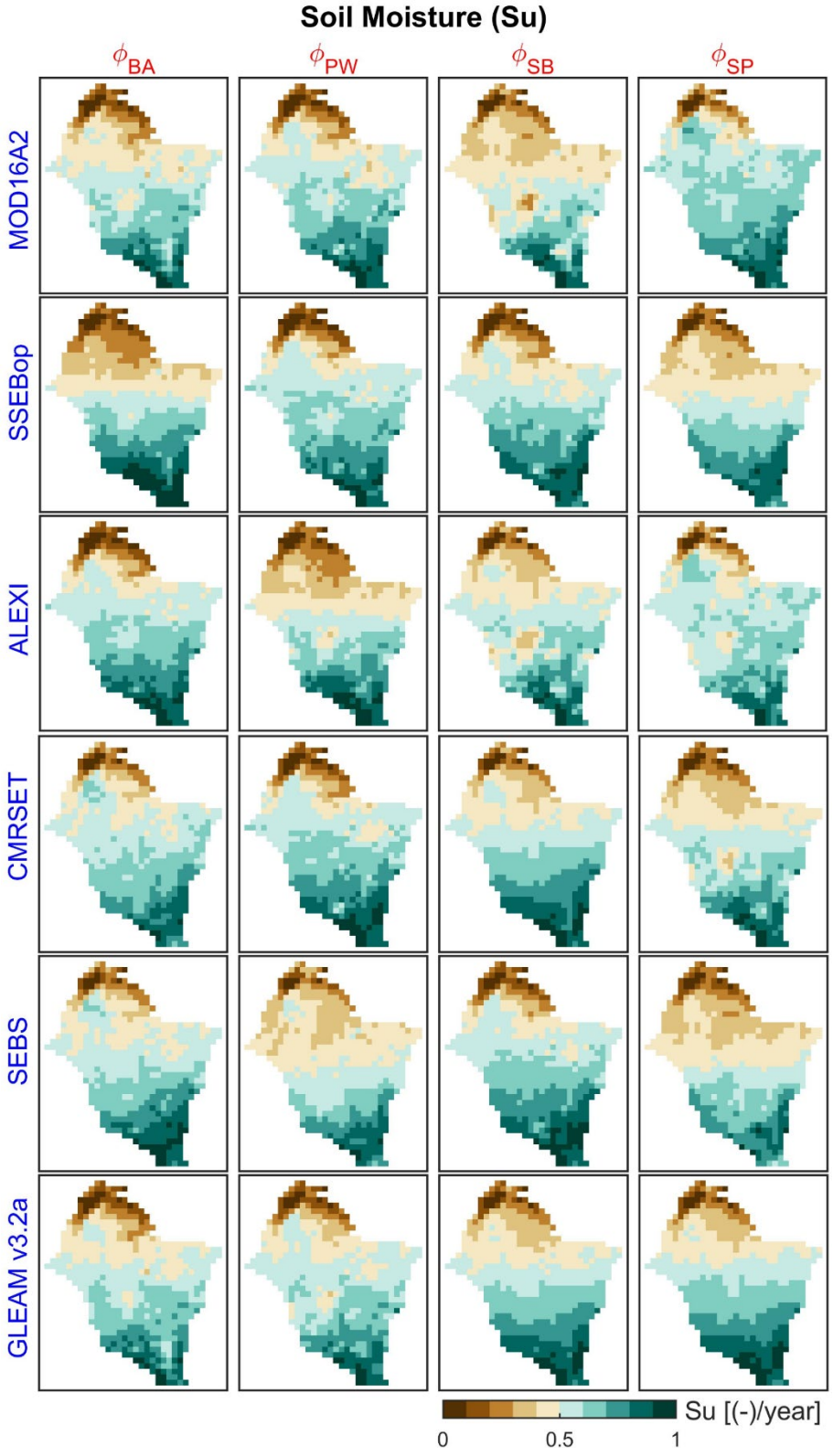


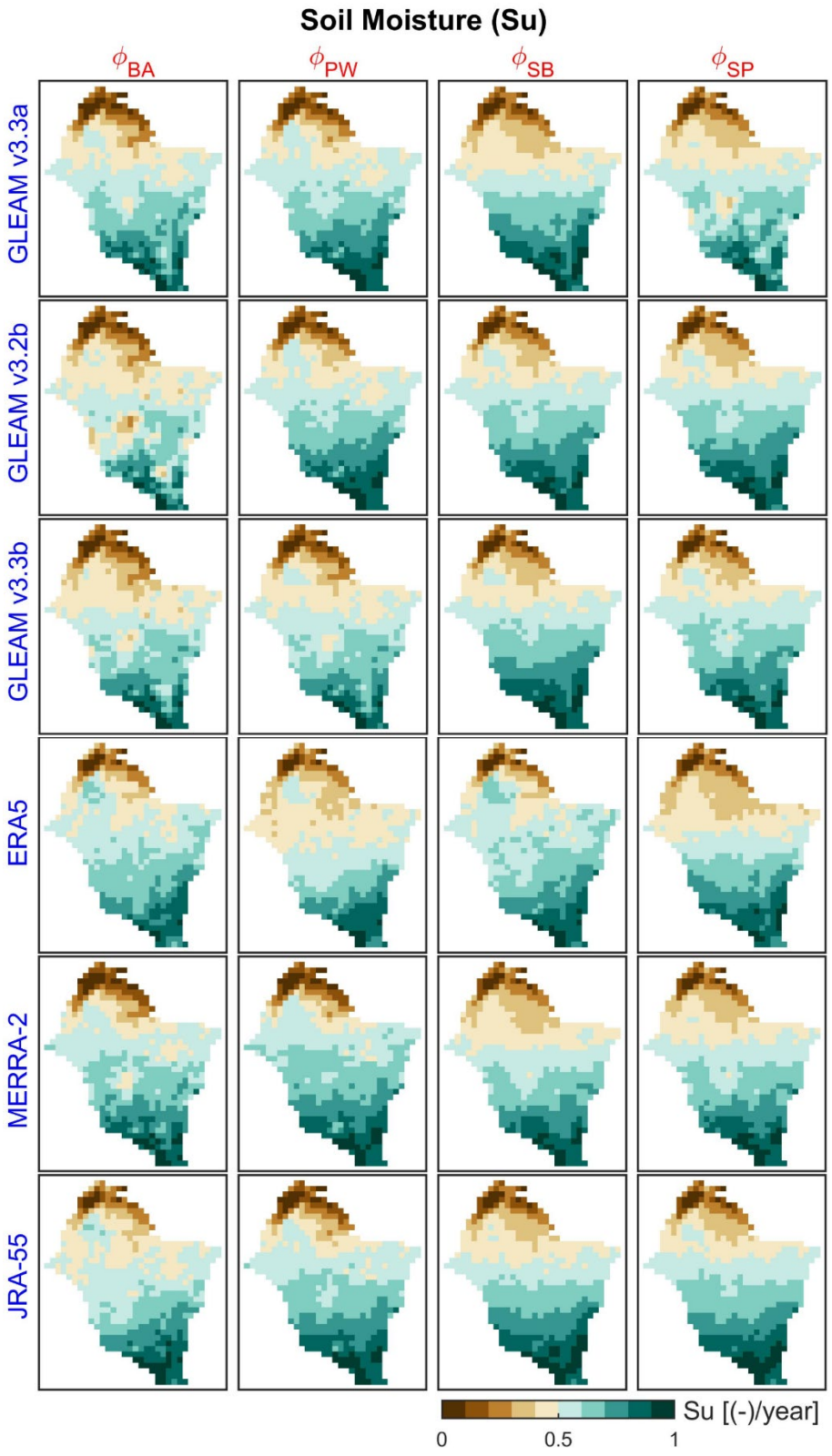
Appendix 19. Model performance for streamflow (Q) as a function of final global parameters (subplot titles) obtained after model calibration with different combinations of precipitation and temperature datasets as inputs. The x-axis give the parameter values for 102 different input datasets. The y-axis gives the Kling-Gupta efficiency (E_{KG}) of Q over the simulation period (2003-2012). The vertical lines (in blue) show the lower and upper bounds of the parameter ranges.

Appendix 20. Evaporation data access portals

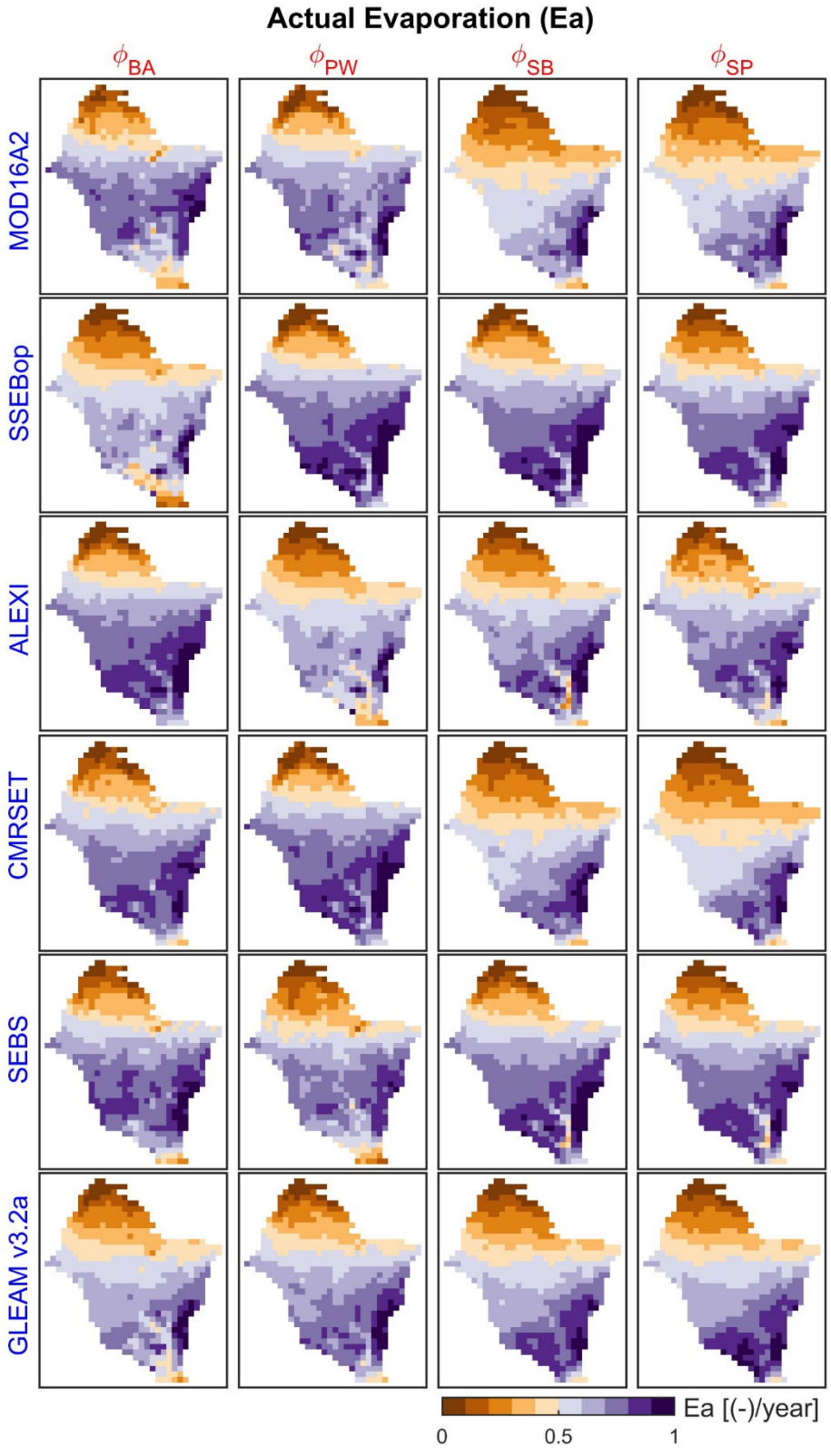
Datasets	Name/ Data portal
MOD16A2	Moderate Resolution Imaging Spectroradiometer (MODIS) Global Terrestrial Evapotranspiration Algorithm version 5 http://files.ntsg.umt.edu/data/NTSG_Products/MOD16/MOD16A2_MONTHLY.MERRA_GMAO_1kmALB/
SSEBop	Operational Simplified Surface Energy Balance https://edcintl.cr.usgs.gov/downloads/sciweb1/shared/fews/web/
ALEXI	Atmosphere-Land Exchange Inverse ftp://ftp.wateraccounting.unesco-ihe.org/WaterAccounting/Data_Satellite/Evaporation/ALEXI/World/
CMRSET	CSIRO MODIS Reflectance Scaling EvapoTranspiration http://remote-sensing.nci.org.au/u39/public/data/wirada/cmrset/
SEBS	Surface Energy Balance System ftp://ftp.wateraccounting.unesco-ihe.org/WaterAccounting/Data_Satellite/Evaporation/SEBS/SEBS/
GLEAM v3.2a	Global Land Evaporation Amsterdam Model https://www.gleam.eu
GLEAM v3.3a	
GLEAM v3.2b	
GLEAM v3.3b	
ERA5	European Centre for Medium-range Weather Forecasts ReAnalysis 5 (ERA5) hourly data on single levels https://cds.climate.copernicus.eu/
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications 2 (Evaporation_land: M2TUNXLND_V5.12.4) https://disc.gsfc.nasa.gov/datasets/M2TUNXLND_V5.12.4/summary
JRA-55	Japanese 55 year ReAnalysis (JRA-55); evaporation: fcst_phy2m125 https://jra.kishou.go.jp/JRA-55/index_en.html

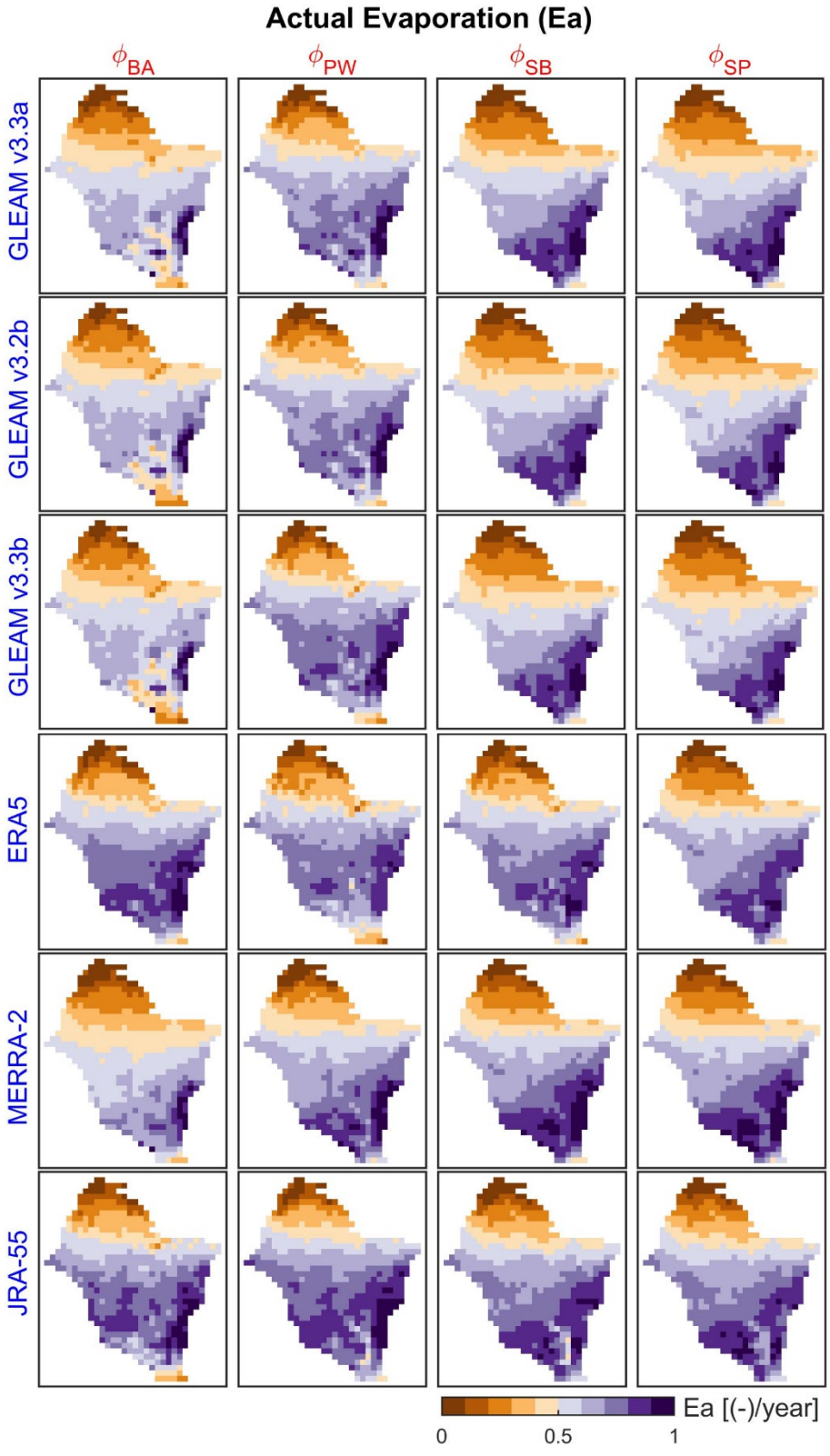
Appendix 21. Maps of long-term (2003-2012) annual average of modelled soil moisture (S_u) using different evaporation datasets (y-axis, blue font) and calibration strategies (x-axis, red font) to calibrate the mHM model. The values are normalized by their range to better emphasize on patterns and use a unique color scale.

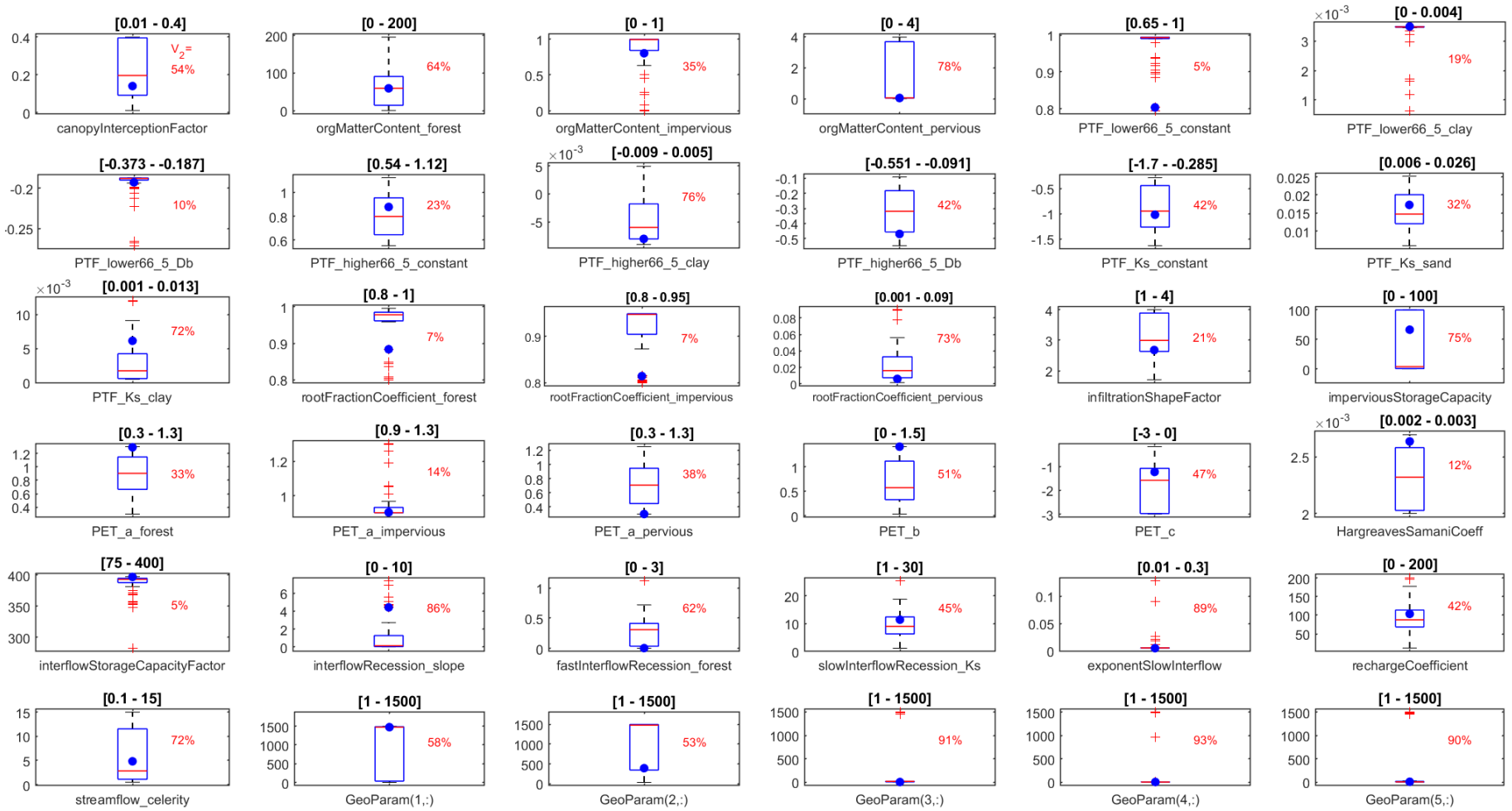



Soil Moisture (Su)
ϕ_{BA}
ϕ_{PW}
ϕ_{SB}
ϕ_{SP}
GLEAM v3.3a
GLEAM v3.2b
GLEAM v3.3b
ERA5
MERRA-2
JRA-55
Su [(-)/year]
0
0.5
1

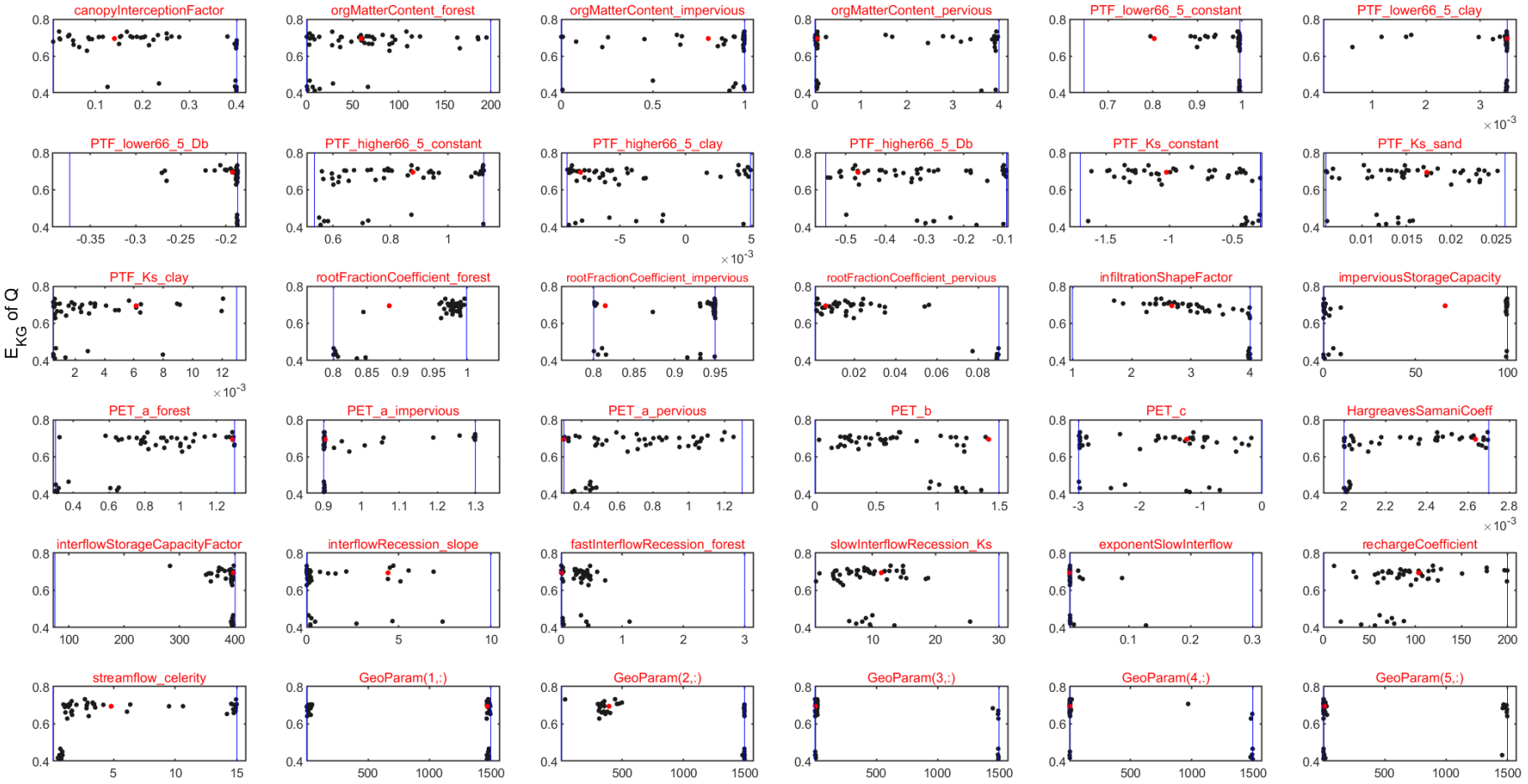
Appendix 22. Maps of long-term (2003-2012) annual average of modelled actual evaporation (E_a) using different evaporation datasets (y-axis, blue font) and calibration strategies (x-axis, red font) to calibrate the mHM model. The values are normalized by their range to better emphasize on patterns and use a unique color scale.



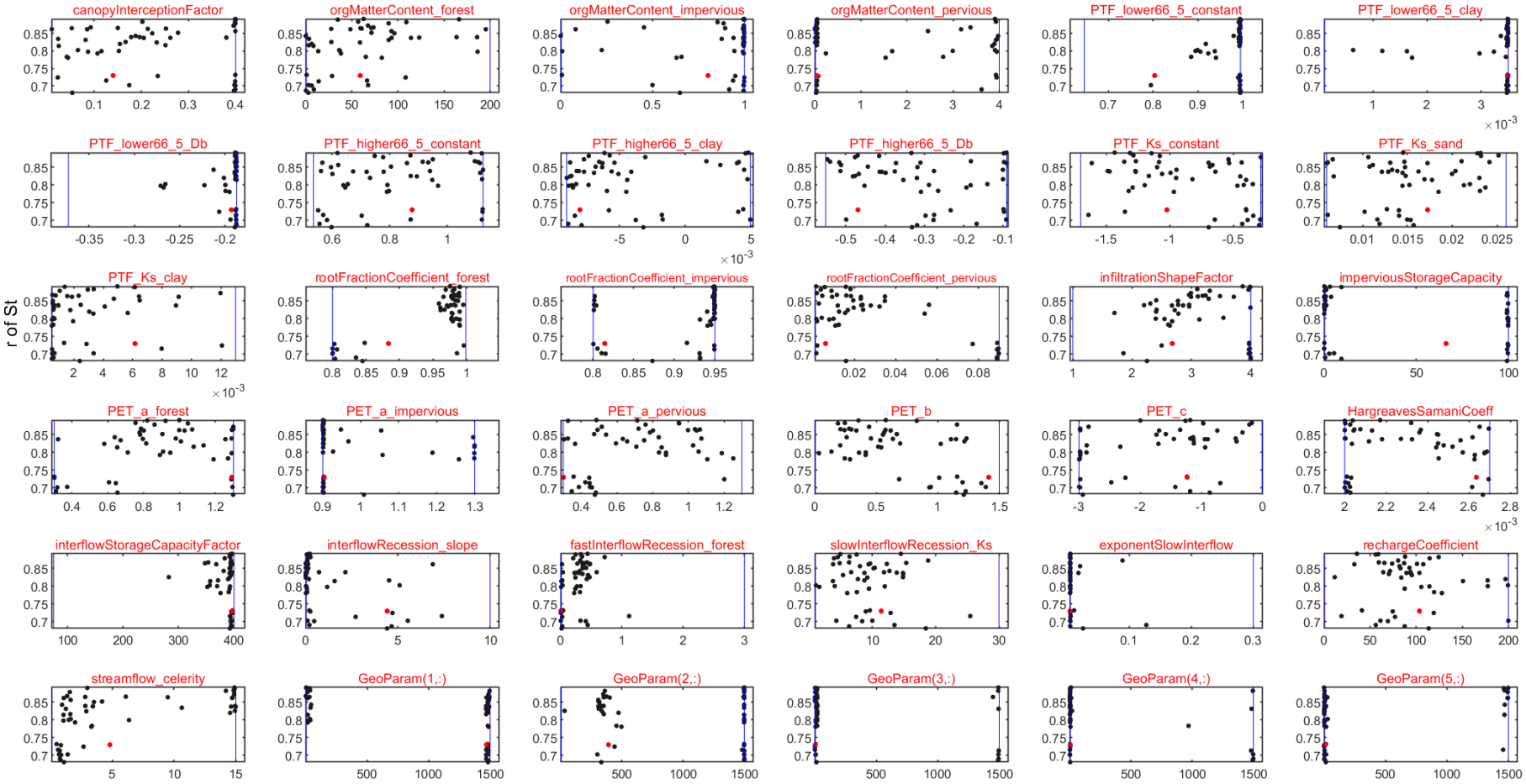

Actual Evaporation (Ea)
ϕ_{BA}
ϕ_{PW}
ϕ_{SB}
ϕ_{SP}
GLEAM v3.3a
GLEAM v3.2b
GLEAM v3.3b
ERA5
MERRA-2
JRA-55
Ea [(-)/year]
0
0.5
1



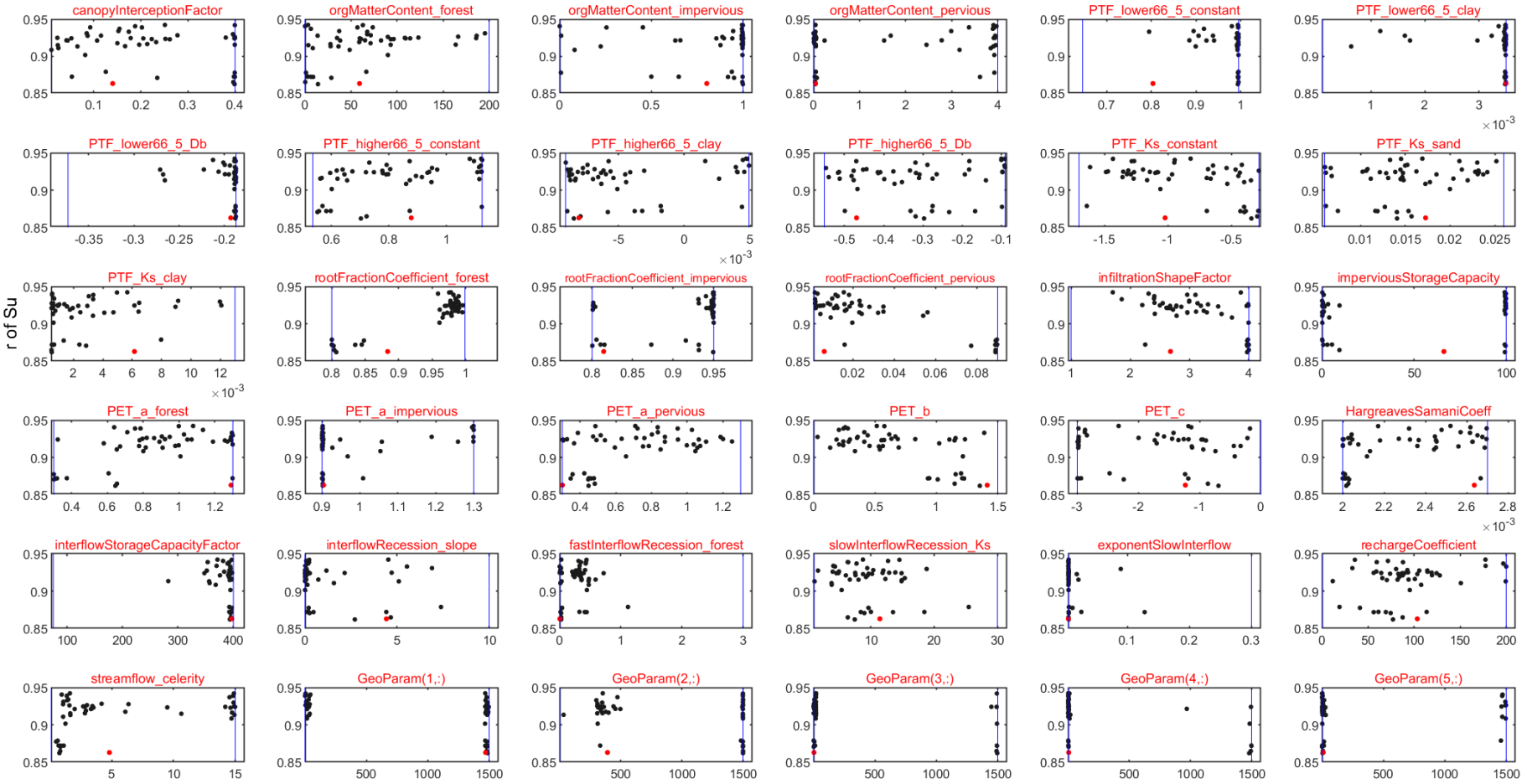
Appendix 23. Distribution of mHM global parameters for the 48 modelling scenarios with various evaporation datasets. Each boxplot has 48 elements corresponding to parameter values obtained with twelve evaporation datasets and four calibration strategies. The blue dots are the parameter values obtained with the Q-only calibration. The initial parameters' ranges are provided in squared brackets. The second-order coefficient of variation (V_2, cf. Kvålseth, 2017) is given in percentage.



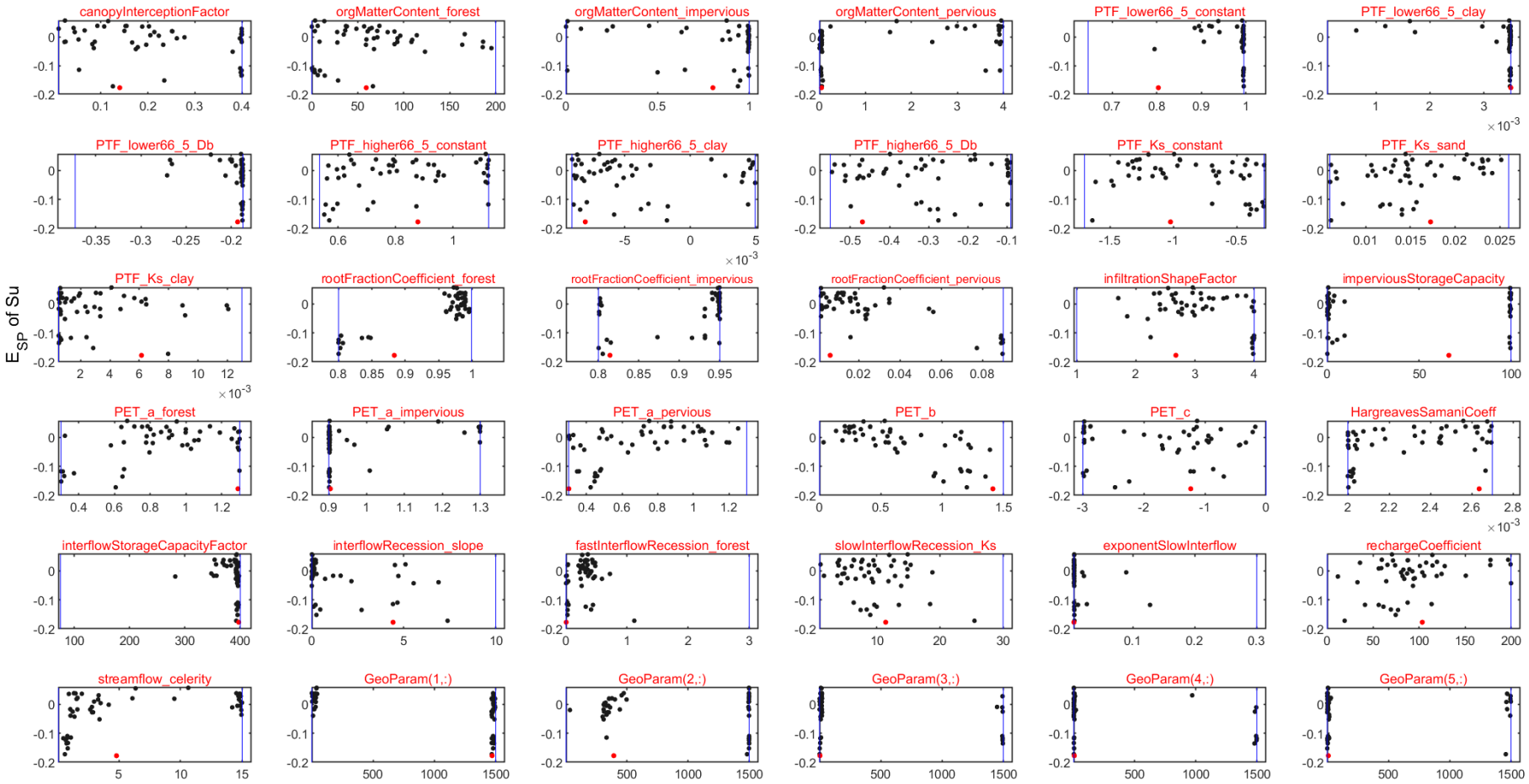
Appendix 24. Model performance for streamflow (Q) as a function of final global parameters (subplot titles) obtained after model calibration. The x-axis gives the parameter values for 48 modelling scenarios with various evaporation datasets (black dots), and the benchmark Q-only calibration (red dot). The y-axis gives the Kling-Gupta efficiency (E_{KG}) of Q over the simulation period (2003-2012). The vertical lines (in blue) show the lower and upper bounds of the parameter ranges.



Appendix 25. Model performance for terrestrial water storage (*St*) as a function of final global parameters (subplot titles) obtained after model calibration. The x-axis gives the parameter values for 48 modelling scenarios with various evaporation datasets (black dots), and the benchmark Q-only calibration (red dot). The y-axis gives the Pearson correlation coefficient (*r*) of *St* over the simulation period (2003-2012). The vertical lines (in blue) show the lower and upper bounds of the parameter ranges.



Appendix 26. Model performance for soil moisture (*Su*) as a function of final global parameters (subplot titles) obtained after model calibration. The x-axis gives the parameter values for 48 modelling scenarios with various evaporation datasets (black dots), and the benchmark Q-only calibration (red dot). The y-axis gives the Pearson correlation coefficient (*r*) of *Su* over the simulation period (2003-2012). The vertical lines (in blue) show the lower and upper bounds of the parameter ranges.



Appendix 27. Model performance for soil moisture (*Su*) as a function of final global parameters (subplot titles) obtained after model calibration. The x-axis gives the parameter values for 48 modelling scenarios with various evaporation datasets (black dots), and the benchmark Q-only calibration (red dot). The y-axis gives the spatial pattern efficiency (E_{SP}) of *Su* over the simulation period (2003-2012). The vertical lines (in blue) show the lower and upper bounds of the parameter ranges.

Appendix 28. WA+ definitions

These WA+ definitions are retrieved from https://www.wateraccounting.org/background.html (last accessed 07.07.2020)

Available water	The total exploitable water minus reserved flows. It represents the water that is available for use at the domain.
Beneficial consumption	Water consumed for the intended purpose.
Closed basin	A basin where utilizable flows are negligible.
Committed flow	Water that has been allocated for a special purpose such as an inter-basin transfer or cross–boundary flow. Often decrees and acts legally describe the agreed flows between one or more parties.
Consumptive use	Water withdrawn for an intended process that does not return back into the basin, except from atmospheric recycling. It relates to total evaporation, water pollution and water incorporated in products.
Environmental flow	Minimum river flow required for maintaining the biodiversity of riverine ecosystem including endangered fish species and the riparian corridor.
Gross inflow	The total amount of water that flows into the domain, this includes precipitation plus any inflow from surface or ground water sources and desalinized water.
Incremental ET	Enhancement of total evaporation (ET) due to rainfall by supplying additional water resources.
Landscape ET	Water from rainfall that evaporates locally from leaves, litter, soil, and via plants that extract moisture from the unsaturated zone.
Managed water use	Represents land use elements with anthropogenic regulation of withdrawals and water supplies. It includes water withdrawals for irrigation, aquaculture, domestic use and industries, among others.
Modified land use	Represents land use elements where vegetation is replaced with the intention to increase the utilization of land resources. Examples are plantation forests, pastures and rainfed crops, among others.
Net inflow	The gross inflow after correction of storage change (ΔS). It represents water available for landscape ET and exploitable water.
Non–beneficial consumption	Water consumed for purposes other than the use.
Non–recoverable flow	Non–consumed water that is lost to further use, by pollution that exceeds international standards, flows to saline groundwater sinks, deep aquifers that are not economically exploitable or flows to the sea.
Productivity of water	The physical mass of production or the economic value of production measured against consumptive use of water.
Protected land use	Environmentally sensitive land uses and natural ecosystem that cannot be modified due to protective measures.
Recoverable flow	Non–consumed water that can be captured and reused in streams, rivers, lakes, reservoirs and aquifers.

Reserved flow	Surface water that has been reserved to meet committed flows, navigational flows, and environmental flow.
Total water stocks	Water present in reservoirs, rivers, lakes and groundwater that can be used for withdrawals.
Utilized land use	Represents land use classes with a low to moderate utilization of natural resources, such as savannah, woodland and mixed pastures.
Utilized water	Part of available water that is depleted for uses.
Utilizable water	Water available for additional resources development.
Water governance	Political, social, economic and administrative systems that are in place, and which directly or indirectly affect the use, development and management of water resources and the delivery of water service delivery at different levels of society.
Water withdrawals	Water taken away from a water source, either natural (e.g. inundation, leakage) or artificial (i.e. pumping, diversions).

References

Abubakari, A., G. Nyarko, J. Yidana, G. Mahunu, F. Abagale, A. Quainoo, F. Chimsah, and V. Avornyo (2012), Comparative studies of soil characteristics in Shea parklands of Ghana, https://doi.org/10.5897/JSSEM11.145.

Acharya, S. C., R. Nathan, Q. J. Wang, C.-H. Su, and N. Eizenberg (2019), An evaluation of daily precipitation from a regional atmospheric reanalysis over Australia, *Hydrology and Earth System Sciences*, *23*(8), 3387-3403, https://doi.org/10.5194/hess-23-3387-2019.

Addor, N., and L. Melsen (2019), Legacy, rather than adequacy, drives the selection of hydrological models, *Water Resources Research*, *55*(1), 378-390, https://doi.org/10.1029/2018WR022958.

AghaKouchak, A., A. Farahmand, F. S. Melton, J. Teixeira, M. C. Anderson, B. D. Wardlow, and C. R. Hain (2015), Remote sensing of drought: Progress, challenges and opportunities, *Rev Geophys*, *53*(2), 452-480, https://doi.org/10.1002/2014rg000456.

Aich, V., S. Liersch, T. Vetter, S. Huang, J. Tecklenburg, P. Hoffmann, H. Koch, S. Fournet, V. Krysanova, and E. Müller (2014), Comparing impacts of climate change on streamflow in four large African river basins, *Hydrology & Earth System Sciences*, *18*(4), https://doi.org/10.5194/hess-18-1305-2014.

Akinsanola, A. A., W. Zhou, T. Zhou, and N. Keenlyside (2020), Amplification of synoptic to annual variability of West African summer monsoon rainfall under global warming, *npj Climate and Atmospheric Science*, *3*(1), 1-10, https://doi.org/10.1038/s41612-020-0125-1.

Alazzy, A. A., H. Lü, R. Chen, A. B. Ali, Y. Zhu, and J. Su (2017), Evaluation of satellite precipitation products and their potential influence on hydrological modeling over the Ganzi River Basin of the Tibetan Plateau, *Advances in Meteorology*, *2017*, https://doi.org/10.1155/2017/3695285.

Alemohammad, S., K. McColl, A. Konings, D. Entekhabi, and A. Stoffelen (2015), Characterization of precipitation product errors across the United States using multiplicative triple collocation, *Hydrology and Earth System Sciences*, https://doi.org/10.5194/hess-19-3489-2015.

Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, 326 pp, http://academic.uprm.edu/abe/backup2/tomas/fao%2056.pdf.

Amisigo, B. A., and N. C. van de Giesen (2005), Using a spatio-temporal dynamic state-space model with the EM algorithm to patch gaps in daily riverflow series, *Hydrology and Earth System Sciences*, *9*(3), 209-224, https://doi.org/10.5194/hess-9-209-2005.

Andah, W. E., N. van de Giesen, and C. A. Biney (2003), Water, climate, food, and environment in the Volta Basin, *Adaptation strategies to changing environments. Contribution to the ADAPT project*, http://www.weap21.org/downloads/ADAPTVolta.pdf.

Anderson, M. C., J. M. Norman, G. R. Diak, W. P. Kustas, and J. R. Mecikalski (1997), A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing, *Remote Sens Environ*, *60*(2), 195-216, https://doi.org/10.1016/S0034-4257(96)00215-5.

Anderson, M. C., J. M. Norman, J. R. Mecikalski, J. A. Otkin, and W. P. Kustas (2007), A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 1. Model formulation, *Journal of Geophysical Research: Atmospheres*, *112*(D10), https://doi.org/10.1029/2006JD007506.

Andreini, M., N. Giesen, A. Edig, M. Fosu, and W. Andah (2000), Volta Basin Water Balance. ZEF-Discussion Papers on Development Policy, No. 21, 37 pp, University of Bonn, Center for Development Research (ZEF), https://www.zef.de/uploads/tx_zefportal/Publications/zef-dp21-00.pdf.

Armstrong, R. N., J. W. Pomeroy, and L. W. Martz (2019), Spatial variability of mean daily estimates of actual evaporation from remotely sensed imagery and surface reference data, *Hydrology and Earth System Sciences*, *23*(12), 4891-4907, https://doi.org/10.5194/hess-23-4891-2019.

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, https://doi.org/10.1175/BAMS-D-13-00068.1.

Awange, J., V. Ferreira, E. Forootan, S. Andam-Akorful, N. Agutu, and X. He (2016), Uncertainties in remotely sensed precipitation data over Africa, *International Journal of Climatology*, *36*(1), 303-323, https://doi.org/10.1002/joc.4346.

Awange, J., K. Hu, and M. Khaki (2019), The newly merged satellite remotely sensed, gauge and reanalysis-based Multi-Source Weighted-Ensemble Precipitation: Evaluation over Australia and Africa (1981–2016), *Science of The Total Environment*, *670*, 448-465, https://doi.org/10.1016/j.scitotenv.2019.03.148.

Badgley, G., J. B. Fisher, C. Jiménez, K. P. Tu, and R. Vinukollu (2015), On uncertainty in global terrestrial evapotranspiration estimates from choice of input forcing datasets, *J Hydrometeorol*, *16*(4), 1449-1455, https://doi.org/10.1175/JHM-D-14-0040.1.

Bai, P., X. Liu, T. Yang, F. Li, K. Liang, S. Hu, and C. Liu (2016), Assessment of the influences of different potential evapotranspiration inputs on the performance of monthly hydrological models under different climatic conditions, *J Hydrometeorol*, *17*(8), 2259-2274, https://doi.org/10.1175/JHM-D-15-0202.1.

Bai, P., and X. Liu (2018), Intercomparison and evaluation of three global high-resolution evapotranspiration products across China, *Journal of hydrology*, *566*, 743-755, https://doi.org/10.1016/j.jhydrol.2018.09.065.

Bai, P., X. Liu, Y. Zhang, and C. Liu (2018a), Incorporating vegetation dynamics noticeably improved performance of hydrological model under vegetation greening, *Sci Total Environ*, *643*, 610-622, https://doi.org/10.1016/j.scitotenv.2018.06.233.

Bai, P., X. M. Liu, and C. M. Liu (2018b), Improving hydrological simulations by incorporating GRACE data for model calibration, *Journal of Hydrology*, *557*, 291-304, https://doi.org/10.1016/j.jhydrol.2017.12.025.

Baik, J., U. W. Liaqat, and M. Choi (2018), Assessment of satellite-and reanalysis-based evapotranspiration products with two blending approaches over the complex landscapes and climates of Australia, *Agricultural and Forest Meteorology*, *263*, 388-398, https://doi.org/10.1016/j.agrformet.2018.09.007.

Balsamo, G., et al. (2018), Satellite and In Situ Observations for Advancing Global Earth Surface Modelling: A Review, *Remote Sens-Basel*, *10*(12), 2038, https://doi.org/10.3390/rs10122038.

Bárdossy, A., and T. Das (2008), Influence of rainfall observation network on model calibration and application, *Hydrology and Earth System Sciences*, *12*(1), 77-89, https://doi.org/10.5194/hess-12-77-2008.

Bárdossy, A., and G. Pegram (2013), Interpolation of precipitation under topographic influence at different time scales, *Water Resources Research*, *49*(8), 4545-4565, https://doi.org/10.1002/wrcr.20307.

Bárdossy, A., and G. Pegram (2014), Infilling missing precipitation records–A comparison of a new copula-based method with other techniques, *Journal of hydrology*, *519*, 1162-1170, https://doi.org/10.1016/j.jhydrol.2014.08.025.

Baroni, G., B. Schalge, O. Rakovec, R. Kumar, L. Schüler, L. Samaniego, C. Simmer, and S. Attinger (2019), A Comprehensive Distributed Hydrological Modeling Intercomparison to Support Process Representation and Data Collection Strategies, *Water Resources Research*, *55*(2), 990-1010, https://doi.org/10.1029/2018wr023941.

Barrett, E. C., and D. W. Martin (1981), *Use of satellite data in rainfall monitoring*, edited, Academic press.

Barry, B., E. Obuobie, M. Andreini, W. Andah, and M. Pluquet (2005), The Volta River Basin, comprehensive assessment of water management in agriculture, comparative study of river basin development and management, *International Water Management Institute*, http://www.iwmi.cgiar.org/assessment/files_new/research_projects/river_basin_development_and_management/VoltaRiverBasin_Boubacar.pdf.

Bastiaanssen, W., L. T. Ha, and M. Fenn (2015), Water Accounting Plus (WA+) for Reporting Water Resources Conditions and Management: A Case Study in the Ca River Basin, Vietnam, 35 pp, Winrock International, https://www.wateraccounting.org/files/White_Paper_Water_Accounting_Winrock.pdf.

Batchelor, C., J. Hoogeveen, J.-M. Faurès, and L. Peiser (2016), Water accounting and auditing. A sourcebook*Rep. 9251093318*, http://www.fao.org/3/a-i5923e.pdf.

Beck, H. E., A. I. Van Dijk, V. Levizzani, J. Schellekens, D. Gonzalez Miralles, B. Martens, and A. De Roo (2017a), MSWEP: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data, *Hydrology and Earth System Sciences*, *21*(1), 589-615, https://doi.org/10.5194/hess-21-589-2017.

Beck, H. E., N. Vergopolan, M. Pan, V. Levizzani, A. I. J. M. van Dijk, G. P. Weedon, L. Brocca, F. Pappenberger, G. J. Huffman, and E. F. Wood (2017b), Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling, *Hydrology and Earth System Sciences*, *21*(12), 6201-6217, https://doi.org/10.5194/hess-21-6201-2017.

Beck, H. E., M. Pan, T. Roy, G. P. Weedon, F. Pappenberger, A. I. Van Dijk, G. J. Huffman, R. F. Adler, and E. F. Wood (2019a), Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS, *Hydrology and Earth System Sciences*, *23*(1), 207-224, https://doi.org/10.5194/hess-23-207-2019.

Beck, H. E., E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. Van Dijk, T. R. McVicar, and R. F. Adler (2019b), MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment, *Bulletin of the American Meteorological Society*, *100*(3), 473-500, https://doi.org/10.1175/BAMS-D-17-0138.1.

Becker, R., A. Koppa, S. Schulz, M. Usman, T. aus der Beek, and C. Schüth (2019), Spatially distributed model calibration of a highly managed hydrological system using remote sensing-derived ET data, *Journal of Hydrology*, https://doi.org/10.1016/j.jhydrol.2019.123944.

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, https://doi.org/10.1016/j.jhydrol.2010.11.043.

Benneh, G., G. Agyepong, and J. Allotey (1990), Land degradation in Ghana. Food Production and Rural Development Division, edited, Commonwealth Secretariat, Marlborough House. Pall Mall. London.

Bennis, S., F. Berrada, and N. Kang (1997), Improving single-variable and multivariable techniques for estimating missing hydrological data, *Journal of Hydrology*, *191*(1-4), 87-105, https://doi.org/10.1016/S0022-1694(96)03076-4.

Berendrecht, W. L., and F. C. van Geer (2016), A dynamic factor modeling framework for analyzing multiple groundwater head series simultaneously, *Journal of Hydrology*, *536*, 50-60, https://doi.org/10.1016/j.jhydrol.2016.02.028.

Berg, P., C. Donnelly, and D. Gustafsson (2018), Near-real-time adjusted reanalysis forcing data for hydrology, *Hydrology and Earth System Sciences*, *22*(2), 989-1000, https://doi.org/10.5194/hess-22-989-2018.

Bergström, S., G. Lindström, and A. Pettersson (2002), Multi-variable parameter estimation to increase confidence in hydrological modelling, *Hydrological processes*, *16*(2), 413-421, https://doi.org/10.1002/hyp.332.

Berthou, S., D. P. Rowell, E. J. Kendon, M. J. Roberts, R. A. Stratton, J. A. Crook, and C. Wilcox (2019), Improved climatological precipitation characteristics over West Africa at convection-permitting scales, *Clim Dynam*, *53*(3-4), 1-21, https://doi.org/10.1007/s00382-019-04759-4.

Beven, K., and J. Feyen (2002), The future of distributed modelling - Special issue, *Hydrological Processes*, *16*(2), 169-172, https://doi.org/10.1002/hyp.325.

Beven, K. (2006a), Searching for the Holy Grail of scientific hydrology: Q t=(S, R, Δt) A as closure, *Hydrology Earth System Sciences*, *10*(5), 609-618, https://doi.org/10.5194/hess-10-609-2006.

Beven, K. (2006b), A manifesto for the equifinality thesis, *Journal of Hydrology*, *320*(1-2), 18-36, https://doi.org/10.1016/j.jhydrol.2005.07.007.

Beven, K. (2010), Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models, *Hydrological Processes*, *24*(12), 1537-1547, https://doi.org/10.1002/hyp.7718.

Beven, K. (2011), *Rainfall-runoff modelling: the primer*, edited, John Wiley & Sons.

Beven, K. (2016), Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication, *Hydrological Sciences Journal*, *61*(9), 1652-1665, https://doi.org/10.1080/02626667.2015.1031761.

Beven, K. (2018), On hypothesis testing in hydrology: Why falsification of models is still a really good idea, *Wiley Interdisciplinary Reviews: Water*, *5*(3), e1278, https://doi.org/10.1002/wat2.1278.

Beven, K. (2019a), Towards a methodology for testing models as hypotheses in the inexact sciences, *Proceedings of the Royal Society A*, *475*(2224), 20180862, https://doi.org/10.1098/rspa.2018.0862.

Beven, K. (2019b), How to make advances in hydrological modelling, *Hydrology Research*, https://doi.org/10.2166/nh.2019.134.

Bhattacharya, T., D. Khare, and M. Arora (2019), A case study for the assessment of the suitability of gridded reanalysis weather data for hydrological simulation in Beas river basin of North Western Himalaya, *Applied Water Science*, *9*(4), 110, https://doi.org/10.1007/s13201-019-0993-x.

Bhattarai, N., K. Mallick, J. Stuart, B. D. Vishwakarma, R. Niraula, S. Sen, and M. Jain (2019), An automated multi-model evapotranspiration mapping framework using remotely sensed and reanalysis data, *Remote Sens Environ*, *229*, 69-92, https://doi.org/10.1016/j.rse.2019.04.026.

Bhuiyan, E., M. Abul, E. I. Nikolopoulos, E. N. Anagnostou, J. Polcher, C. Albergel, E. Dutra, G. Fink, A. Martínez-de la Torre, and S. Munier (2019), Assessment of precipitation error propagation in multi-model global water resource reanalysis, *Hydrology and Earth System Sciences*, *23*(4), 1973-1994, https://doi.org/10.5194/hess-23-1973-2019.

Biasutti, M., A. Voigt, W. R. Boos, P. Braconnot, J. C. Hargreaves, S. P. Harrison, S. M. Kang, B. E. Mapes, J. Scheff, and C. Schumacher (2018), Global energetics and local physics as drivers of past, present and future monsoons, *Nat Geosci*, *11*(6), 392, https://doi.org/10.1038/s41561-018-0137-1.

Biasutti, M. (2019), Rainfall trends in the African Sahel: Characteristics, processes, and causes, *Wiley Interdisciplinary Reviews: Climate Change*, e591, https://doi.org/10.1002/wcc.591.

Bichet, A., and A. Diedhiou (2018), West African Sahel has become wetter during the last 30 years, but dry spells are shorter and more frequent, *Climate Research*, *75*(2), 155-162, https://doi.org/10.3354/cr01515.

Biney, C. (2010), Connectivities and linkages within the Volta Basin, *The Global Dimensions of Change in River Basins*, *91*.

Biondi, D., G. Freni, V. Iacobellis, G. Mascaro, and A. Montanari (2012), Validation of hydrological models: Conceptual basis, methodological approaches and a proposal for a code of practice, *Phys Chem Earth*, *42-44*, 70-76, https://doi.org/10.1016/j.pce.2011.07.037.

Birhanu, D., H. Kim, and C. Jang (2019), Effectiveness of introducing crop coefficient and leaf area index to enhance evapotranspiration simulations in hydrologic models, *Hydrological Processes*, https://doi.org/10.1002/hyp.13464.
Bisselink, B., M. Zambrano-Bigiarini, P. Burek, and A. De Roo (2016), Assessing the role of uncertain precipitation estimates on the robustness of hydrological model parameters under highly variable climate conditions, *Journal of Hydrology: Regional Studies*, *8*, 112-129, https://doi.org/10.1016/j.ejrh.2016.09.003.
Bitew, M. M., and M. Gebremichael (2011), Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model, *Water Resources Research*, *47*(6), https://doi.org/10.1029/2010wr009917.
Blöschl, G., and E. Zehe (2005), On hydrological predictability, *Hydrological Processes: An International Journal*, *19*(19), 3923-3929, https://doi.org/10.1002/hyp.6075.
Blöschl, G., M. Sivapalan, H. Savenije, T. Wagener, and A. Viglione (2013), *Runoff prediction in ungauged basins: synthesis across processes, places and scales*, edited, 490 pp., Cambridge University Press.
Blöschl, G., et al. (2019), Twenty-three Unsolved Problems in Hydrology (UPH)–a community perspective, *Hydrological Sciences Journal*, *64*(10), 1141-1158, https://doi.org/10.1080/02626667.2019.1620507.
Boadi, S. A., and K. Owusu (2019), Impact of climate change and variability on hydropower in Ghana, *African Geographical Review*, *38*(1), 19-31, https://doi.org/10.1080/19376812.2017.1284598.
Boni, G., D. Entekhabi, and F. Castelli (2001), Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation, *Water resources research*, *37*(6), 1713-1722, https://doi.org/10.1029/2001WR900020.
Bontemps, S., P. Defourny, E. V. Bogaert, O. Arino, V. Kalogirou, and J. R. Perez (2011), GLOBCOVER 2009-Products description and validation report, 53 pp, http://due.esrin.esa.int/files/GLOBCOVER2009_Validation_Report_2.2.pdf.
Bosilovich, M. G., J. Chen, F. R. Robertson, and R. F. Adler (2008), Evaluation of global precipitation in reanalyses, *Journal of applied meteorology and climatology*, *47*(9), 2279-2299, https://doi.org/10.1175/2008jamc1921.1.
Breiman, L. (2001), Random forests, *Machine Learning*, *45*(1), 5-32, https://doi.org/10.1023/A:1010933404324.
Brocca, L., T. Moramarco, F. Melone, and W. Wagner (2013), A new method for rainfall estimation through soil moisture observations, *Geophysical Research Letters*, *40*(5), 853-858, https://doi.org/10.1002/grl.50173.
Brocca, L., L. Ciabatta, C. Massari, T. Moramarco, S. Hahn, S. Hasenauer, R. Kidd, W. Dorigo, W. Wagner, and V. Levizzani (2014), Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data, *Journal of Geophysical Research: Atmospheres*, *119*(9), 5128-5141, https://doi.org/10.1002/2014JD021489.
Brocca, L., P. Filippucci, S. Hahn, L. Ciabatta, C. Massari, S. Camici, L. Schüller, B. Bojkov, and W. Wagner (2019), SM2RAIN–ASCAT (2007–2018): global daily satellite rainfall data from ASCAT soil moisture observations, *Earth System Science Data*, *11*(4), 1583-1601, https://doi.org/10.5194/essd-11-1583-2019.
Budyko, M. (1974), Climate and life, *International Geophysics Series*, *18*.
Butler, D. (2014), Earth observation enters next phase, *Nature*, *508*(7495), 160-161, https://doi.org/10.1038/508160a.
Camici, S., L. Ciabatta, C. Massari, and L. Brocca (2018), How reliable are satellite precipitation estimates for driving hydrological models: A verification study over the Mediterranean area, *Journal of hydrology*, *563*, 950-961, https://doi.org/10.1016/j.jhydrol.2018.06.067.
Campozano, L., E. Sánchez, A. Aviles, and E. Samaniego (2015), Evaluation of infilling methods for time series of daily precipitation and temperature: The case of the Ecuadorian Andes, *Maskana*, *5*(1), 99-115, https://doi.org/10.18537/mskn.05.01.07.
Caroletti, G. N., R. Coscarelli, and T. Caloiero (2019), Validation of Satellite, Reanalysis and RCM Data of Monthly Rainfall in Calabria (Southern Italy), *Remote Sens-Basel*, *11*(13), 1625, https://doi.org/10.3390/rs11131625
Casse, C., M. Gosset, C. Peugeot, V. Pedinotti, A. Boone, B. Tanimoun, and B. Decharme (2015), Potential of satellite rainfall products to predict Niger River flood events in Niamey, *Atmos Res*, *163*, 162-176, https://doi.org/10.1016/j.atmosres.2015.01.010.
Cazenave, A., N. Champollion, J. Benveniste, and J. Chen (2016), *Remote Sensing and Water Resources*, edited, 337 pp., Springer, https://doi.org/10.1007/978-3-319-32449-4.
Ceperley, N. C., T. Mande, N. van de Giesen, S. Tyler, H. Yacouba, and M. B. Parlange (2017), Evaporation from cultivated and semi-wild Sudanian Savanna in west Africa, *Hydrology and Earth System Sciences*, *21*, 4149-4167, https://doi.org/10.5194/hess-21-4149-2017.
Chambers, J. M. (2017), *Graphical Methods for Data Analysis*, edited, 410 pp., Chapman and Hall/CRC, https://doi.org/10.1201/9781351072304.

Chen, J. M., and J. Liu (2020), Evolution of evapotranspiration models using thermal and shortwave remote sensing data, *Remote Sens Environ*, *237*, 111594, https://doi.org/10.1016/j.rse.2019.111594.
Chen, L., and L. Wang (2018), Recent advance in earth observation big data for hydrology, *Big Earth Data*, *2*(1), 86-107, https://doi.org/10.1080/20964471.2018.1435072.
Chen, X., D. Long, Y. Hong, C. Zeng, and D. Yan (2017), Improved modeling of snow and glacier melting by a progressive two-stage calibration strategy with GRACE and multisource data: How snow and glacier meltwater contributes to the runoff of the Upper Brahmaputra River basin?, *Water Resources Research*, *53*(3), 2431-2466, https://doi.org/10.1002/2016WR019656.
Chen, Y., J. Xia, S. Liang, J. Feng, J. B. Fisher, X. Li, X. Li, S. Liu, Z. Ma, and A. Miyata (2014), Comparison of satellite-based evapotranspiration models over terrestrial ecosystems in China, *Remote Sens Environ*, *140*, 279-293, https://doi.org/10.1016/j.rse.2013.08.045.
Ciabatta, L., L. Brocca, C. Massari, T. Moramarco, S. Gabellani, S. Puca, and W. Wagner (2016), Rainfall-runoff modelling by using SM2RAIN-derived and state-of-the-art satellite rainfall products over Italy, *International journal of applied earth observation and geoinformation*, *48*, 163-173, https://doi.org/10.1016/j.jag.2015.10.004.
Ciabatta, L., C. Massari, L. Brocca, A. Gruber, C. Reimer, S. Hahn, C. Paulik, W. Dorigo, R. Kidd, and W. Wagner (2018), SM2RAIN-CCI: A new global long-term rainfall data set derived from ESA CCI soil moisture, *Earth System Science Data*, *10*(1), 267, https://doi.org/10.5194/essd-10-267-2018.
Clark, M. P., D. E. Rupp, R. A. Woods, H. Tromp-van Meerveld, N. Peters, and J. Freer (2009), Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale, *Hydrological Processes: An International Journal*, *23*(2), 311-319, https://doi.org/10.1002/hyp.7154.
Clark, M. P., et al. (2015), Improving the representation of hydrologic processes in Earth System Models, *Water Resources Research*, *51*(8), 5929-5956, https://doi.org/10.1002/2015wr017096.
Clark, M. P., B. Schaefli, S. J. Schymanski, L. Samaniego, C. H. Luce, B. M. Jackson, J. E. Freer, J. R. Arnold, R. D. Moore, and E. Istanbulluoglu (2016), Improving the theoretical underpinnings of process-based hydrologic models, *Water Resources Research*, *52*(3), 2350-2365, https://doi.org/10.1002/2015WR017910.
Clark, M. P., M. F. P. Bierkens, L. Samaniego, R. A. Woods, R. Uijlenhoet, K. E. Bennett, V. R. N. Pauwels, X. Cai, A. W. Wood, and C. D. Peters-Lidard (2017), The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism, *Hydrology and Earth System Sciences*, *21*(7), 3427-3440, https://doi.org/10.5194/hess-21-3427-2017.
Coenders-Gerrits, M., B. Schilperoort, and C. Jiménez-Rodríguez (2020), Evaporative Processes on Vegetation: An Inside Look, in *Precipitation Partitioning by Vegetation*, edited by J. T. Van Stan, E. Gutmann and J. Friesen, pp. 35-48, Springer, Switzerland https://doi.org/10.1007/978-3-030-29702-2.
Contractor, S., M. Donat, L. V. Alexandre, M. Ziese, A. Meyer-Christoffer, U. Schneider, E. Rustemeier, A. Becker, I. Durre, and R. S. Vose (2020), Rainfall Estimates on a Gridded Network (REGEN)–a global land-based gridded dataset of daily precipitation from 1950 to 2016, *Hydrology and Earth System Sciences (HESS)*, *24*(2), 919-943, https://doi.org/10.5194/hess-24-919-2020.
Cook, C., and K. Bakker (2012), Water security: Debating an emerging paradigm, *Global environmental change*, *22*(1), 94-102, https://doi.org/10.1016/j.gloenvcha.2011.10.011.
Cook, K. H., and E. K. Vizy (2019), Contemporary climate change of the African monsoon systems, *Current Climate Change Reports*, *5*(3), 145-159, https://doi.org/10.1007/s40641-019-00130-1s.
Coulibaly, P., and C. K. Baldwin (2005), Nonstationary hydrological time series forecasting using nonlinear dynamic methods, *Journal of Hydrology*, *307*(1-4), 164-174, https://doi.org/10.1016/j.jhydrol.2004.10.008.
Cucchi, M., G. P. Weedon, A. Amici, N. Bellouin, S. Lange, H. M. Schmied, H. Hersbach, and C. Buontempo (2020), WFDE5: bias adjusted ERA5 reanalysis data for impact studies, *Earth System Science Data Discussions*, 1-32, https://doi.org/10.5194/essd-2020-28.
Cui, X., X. Guo, Y. Wang, X. Wang, W. Zhu, J. Shi, C. Lin, and X. Gao (2019), Application of remote sensing to water environmental processes under a changing climate, *Journal of Hydrology*, https://doi.org/10.1016/j.jhydrol.2019.04.078.
Cui, Y., X. Chen, J. Gao, B. Yan, G. Tang, and Y. Hong (2018), Global water cycle and remote sensing big data: overview, challenge, and opportunities, *Big Earth Data*, *2*(3), 282-297, https://doi.org/10.1080/20964471.2018.1548052.
Cunge, J. A. (1969), On The Subject Of A Flood Propagation Computation Method (Musklngum Method), *Journal of Hydraulic Research*, *7*(2), 205-230, https://doi.org/10.1080/00221686909500264.
da Motta Paca, V. H., G. E. Espinoza-Dávalos, T. M. Hessels, D. M. Moreira, G. F. Comair, and W. G. Bastiaanssen (2019), The spatial variability of actual evapotranspiration across the Amazon River Basin

based on remote sensing products validated with flux towers, *Ecological Processes*, *8*(1), 6, https://doi.org/10.1186/s13717-019-0158-8.
Damania, R. (2020), The economics of water scarcity and variability, *Oxford Review of Economic Policy*, *36*(1), 24-44, https://doi.org/10.1093/oxrep/grz027.
Danielson, J. J., and D. B. Gesch (2011), Global multi-resolution terrain elevation data 2010 (GMTED2010)*Rep. 2331-1258*, 34 pp, US Geological Survey, https://doi.org/10.3133/ofr20111073.
Dastorani, M. T., A. Moghadamnia, J. Piri, and M. Rico-Ramirez (2010), Application of ANN and ANFIS models for reconstructing missing flow data, *Environ Monit Assess*, *166*(1-4), 421-434, https://doi.org/10.1007/s10661-009-1012-8.
Dawdy, D. R., and T. O'Donnell (1965), Mathematical models of catchment behavior, *Journal of the Hydraulics Division*, *91*(4), 123-137.
Dawson, C. W., C. Harpham, R. L. Wilby, and Y. Chen (2002), Evaluation of artificial neural network techniques for flow forecasting in the River Yangtze, China, *Hydrology and Earth System Sciences*, *6*(4), 619-626, https://doi.org/10.5194/hess-6-619-2002.
de Bont, C., H. C. Komakech, and G. J. Veldwisch (2019), Neither modern nor traditional: Farmer-led irrigation development in Kilimanjaro Region, Tanzania, *World Development*, *116*, 15-27, https://doi.org/10.1016/j.worlddev.2018.11.018.
De Condappa, D., and J. Lemoalle (2009), Water atlas of the Volta Basin-Atlas de l'eau dans le basin de la Volta, 93 pp, http://hal.ird.fr/ird-00505116/document.
de Fraiture, C., G. N. Kouali, H. Sally, and P. Kabre (2014), Pirates or pioneers? Unplanned irrigation around small reservoirs in Burkina Faso, *Agricultural Water Management*, *131*, 212-220, https://doi.org/10.1016/j.agwat.2013.07.001.
Delavar, M., S. Morid, R. Morid, A. Farokhnia, F. Babaeian, R. Srinivasan, and P. Karimi (2020), Basin-wide water accounting based on modified SWAT model and WA+ framework for better policy making, *Journal of Hydrology*, 124762, https://doi.org/10.1016/j.jhydrol.2020.124762.
Dembélé, M., and S. J. Zwart (2016), Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa, *International Journal of Remote Sensing*, *37*(17), 3995-4014, https://doi.org/10.1080/01431161.2016.1207258.
Dembélé, M., F. Oriani, J. Tumbulto, G. Mariethoz, and B. Schaefli (2019), Gap-filling of daily streamflow time series using Direct Sampling in various hydroclimatic settings, *Journal of Hydrology*, *569*, 573-586, https://doi.org/10.1016/j.jhydrol.2018.11.076.
Dembélé, M., N. Ceperley, S. J. Zwart, G. Mariéthoz, and B. Schaefli (2020a), Potential of Satellite and Reanalysis Evaporation Datasets for Hydrological Modelling under Various Model Calibration Strategies, *Advances in Water Resources*, https://doi.org/10.1016/j.advwatres.2020.103667.
Dembélé, M., M. Hrachowitz, H. H. G. Savenije, G. Mariéthoz, and B. Schaefli (2020b), Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, *Water Resources Research*, *56*(1), https://doi.org/10.1029/2019wr026085.
Dembélé, M., B. Schaefli, N. van de Giesen, and G. Mariéthoz (2020c), Suitability of 17 rainfall and temperature gridded datasets for large-scale hydrological modelling in West Africa, *Hydrol. Earth Syst. Sci. Discuss.*, https://doi.org/10.5194/hess-2020-68.
Demirel, M. C., J. Mai, G. Mendiguren, J. Koch, L. Samaniego, and S. Stisen (2018), Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, *Hydrology and Earth System Sciences*, *22*(2), 1299-1315, https://doi.org/10.5194/hess-22-1299-2018.
DESA (2012), System of Environmental-Economic Accounting for Water, *Statistics Division, United Nations Department of Economic and Social Affairs, New York. UNESCO (2012f) Indicators, Retrieved June*, *6*, 2012, https://unstats.un.org/unsd/publication/seriesf/Seriesf_100e.pdf.
Dezfuli, A. (2017), Climate of western and central equatorial Africa, in *Oxford Research Encyclopedia of Climate Science*, edited, https://doi.org/10.1093/acrefore/9780190228620.013.511
Di Baldassarre, G., A. Montanari, H. Lins, D. Koutsoyiannis, L. Brandimarte, and G. Blöschl (2010), Flood fatalities in Africa: from diagnosis to mitigation, *Geophysical Research Letters*, *37*(22), https://doi.org/10.1029/2010GL045467.
Di Piazza, A., F. L. Conti, L. V. Noto, F. Viola, and G. La Loggia (2011), Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy, *International Journal of Applied Earth Observation and Geoinformation*, *13*(3), 396-408, https://doi.org/10.1016/j.jag.2011.01.005.
Diallo, I., F. Giorgi, A. Deme, M. Tall, L. Mariotti, and A. T. Gaye (2016), Projected changes of summer monsoon extremes and hydroclimatic regimes over West Africa for the twenty-first century, *Clim Dynam*, *47*(12), 3931-3954, https://doi.org/10.1007/s00382-016-3052-4.

Diaz, V., G. A. C. Perez, H. A. Van Lanen, D. Solomatine, and E. A. Varouchakis (2019), Characterisation of the dynamics of past droughts, *Science of The Total Environment*, 134588, https://doi.org/10.1016/j.scitotenv.2019.134588.

Dinku, T. (2019), Challenges with availability and quality of climate data in Africa, in *Extreme Hydrology and Climate Variability*, edited, pp. 71-80, Elsevier, https://doi.org/10.1016/B978-0-12-815998-9.00007-5.

Donat, M. G., A. L. Lowry, L. V. Alexander, P. A. O'Gorman, and N. Maher (2016), More extreme precipitation in the world's dry and wet regions, *Nature Climate Change*, *6*(5), 508-513, https://doi.org/10.1038/nclimate2941.

Donohue, R., M. Roderick, and T. R. McVicar (2010), Can dynamic vegetation information improve the accuracy of Budyko's hydrological model?, *Journal of hydrology*, *390*(1-2), 23-34, https://doi.org/10.1016/j.jhydrol.2010.06.025.

Donohue, R. J., M. L. Roderick, and T. R. McVicar (2011), Assessing the differences in sensitivities of runoff to changes in climatic conditions across a large basin, *Journal of Hydrology*, *406*(3-4), 234-244, https://doi.org/10.1016/j.jhydrol.2011.07.003.

Dorigo, W., et al. (2017), ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, *Remote Sens Environ*, *203*, 185-215, https://doi.org/10.1016/j.rse.2017.07.001.

Dosio, A., R. G. Jones, C. Jack, C. Lennard, G. Nikulin, and B. Hewitson (2019), What can we know about future precipitation in Africa? Robustness, significance and added value of projections from a large ensemble of regional climate models, *Clim Dynam*, *53*(9-10), 5833-5858, https://doi.org/10.1007/s00382-019-04900-3.

Dosio, A., A. G. Turner, A. T. Tamoffo, M. B. Sylla, C. Lennard, R. G. Jones, L. Terray, G. Nikulin, and B. Hewitson (2020), A tale of two futures: contrasting scenarios of future precipitation for West Africa from an ensemble of regional climate models, *Environmental Research Letters*, *15*(6), 064007, https://doi.org/10.1088/1748-9326/ab7fde.

Dost, R., E. B. Obando, W. G. M. Bastiaanssen, and J. Hoogeveen (2013), Background report: Water accounting+ in the Awash River Basin, 8 pp, FAO, Land and Water Division.

Duan, Z., Y. Tuo, J. Liu, H. Gao, X. Song, Z. Zhang, L. Yang, and D. F. Mekonnen (2019), Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia, *Journal of hydrology*, *569*, 612-626, https://doi.org/10.1016/j.jhydrol.2018.12.026.

Duethmann, D., J. Zimmer, A. Gafurov, A. Güntner, D. Kriegel, B. Merz, and S. Vorogushyn (2013), Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modelling, *Hydrology and Earth System Sciences*, *17*(7), 2415-2434, https://doi.org/10.5194/hess-17-2415-2013.

Dumedah, G., and P. Coulibaly (2011), Evaluation of statistical methods for infilling missing values in high-resolution soil moisture data, *Journal of Hydrology*, *400*(1-2), 95-102, https://doi.org/10.1016/j.jhydrol.2011.01.028.

Efron, B. (1992), Bootstrap methods: another look at the jackknife, in *Breakthroughs in statistics*, edited, pp. 569-593, Springer, https://doi.org/10.1007/978-1-4612-4380-9_41.

Efstratiadis, A., and D. Koutsoyiannis (2010), One decade of multi-objective calibration approaches in hydrological modelling: a review, *Hydrolog Sci J*, *55*(1), 58-78, https://doi.org/10.1080/02626660903526292.

Ehlers, L. B., T. O. Sonnenborg, and J. C. Refsgaard (2018), Observational and predictive uncertainties for multiple variables in a spatially distributed hydrological model, *Hydrological Processes*, https://doi.org/10.1002/hyp.13367.

Elshorbagy, A. A., U. S. Panu, and S. P. Simonovic (2000), Group-based estimation of missing hydrological data: I. Approach and general methodology, *Hydrolog Sci J*, *45*(6), 849-866, https://doi.org/10.1080/02626660009492388.

Enders, C. K. (2010), *Applied missing data analysis*, edited, 377 pp., Guilford press.

Er-Raki, S., A. Chehbouni, S. Khabba, V. Simonneaux, L. Jarlan, A. Ouldbba, J. Rodriguez, and R. Allen (2010), Assessment of reference evapotranspiration methods in semi-arid regions: can weather forecast data be used as alternate of ground meteorological parameters?, *Journal of Arid Environments*, *74*(12), 1587-1596, https://doi.org/10.1016/j.jaridenv.2010.07.002.

ESA (2017), Land Cover CCI Product User Guide Version 2. Tech. Rep., maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.

Essou, G. R., F. Sabarly, P. Lucas-Picher, F. Brissette, and A. Poulin (2016), Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling?, *J Hydrometeorol*, *17*(7), 1929-1950, https://doi.org/10.1175/JHM-D-15-0138.1.

Euser, T., H. Winsemius, M. Hrachowitz, F. Fenicia, S. Uhlenbrook, and H. Savenije (2013), A framework to assess the realism of model structures using hydrological signatures, *Hydrology Earth System Sciences*, *17*(5), 1893-1912, https://doi.org/10.5194/hess-17-1893-2013.

Eyring, V., P. M. Cox, G. M. Flato, P. J. Gleckler, G. Abramowitz, P. Caldwell, W. D. Collins, B. K. Gier, A. D. Hall, and F. M. Hoffman (2019), Taking climate model evaluation to the next level, *Nature Climate Change*, *9*(2), 102-110, https://doi.org/10.1038/s41558-018-0355-y.

Falck, A. S., V. Maggioni, J. Tomasella, D. A. Vila, and F. L. Diniz (2015), Propagation of satellite precipitation uncertainties through a distributed hydrologic model: A case study in the Tocantins–Araguaia basin in Brazil, *Journal of Hydrology*, *527*, 943-957, https://doi.org/10.1016/j.jhydrol.2015.05.042.

Falkenmark, M., and J. Rockström (2006), The new blue and green water paradigm: Breaking new ground for water resources planning and management, edited, American Society of Civil Engineers.

Fallah, A., S. O, and R. Orth (2020), Climate-dependent propagation of precipitation uncertainty into the water cycle, *Hydrol. Earth Syst. Sci.*, *24*(7), 3725-3735, https://doi.org/10.5194/hess-24-3725-2020.

Famiglietti, J. S., and M. Rodell (2013), Water in the balance, *Science*, *340*(6138), 1300-1301, https://doi.org/10.1126/science.1236460

FAO and IHE Delft (2019), Water Accounting in the Litani River Basin, Rome, http://www.fao.org/3/ca6679en/CA6679EN.pdf.

FAO and IHE Delft (2020), Water Accounting in the Jordan River Basin, Rome.

Fatichi, S., et al. (2016), An overview of current applications, challenges, and future trends in distributed process-based models in hydrology, *Journal of Hydrology*, *537*, 45-60, https://doi.org/10.1016/j.jhydrol.2016.03.026.

Feddes, R. A., P. Kowalik, K. Kolinskamalinka, and H. Zaradny (1976), Simulation of Field Water-Uptake by Plants Using a Soil-Water Dependent Root Extraction Function, *Journal of Hydrology*, *31*(1-2), 13-26, https://doi.org/10.1016/0022-1694(76)90017-2.

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

Feng, T., T. Su, R. Zhi, G. Tu, and F. Ji (2019), Assessment of actual evapotranspiration variability over global land derived from seven reanalysis datasets, *International Journal of Climatology*, *39*(6), 2919-2932, https://doi.org/10.1002/joc.5992.

Fenicia, F., J. J. McDonnell, and H. H. Savenije (2008), Learning from model improvement: On the contribution of complementary data to process understanding, *Water Resources Research*, *44*(6), https://doi.org/10.1029/2007WR006386.

Fisher, J. B., et al. (2017), The future of evapotranspiration: Global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources, *Water Resources Research*, *53*(4), 2618-2626, https://doi.org/10.1002/2016wr020175.

Fitzpatrick, R. G., D. J. Parker, J. H. Marsham, D. P. Rowell, F. M. Guichard, C. M. Taylor, K. H. Cook, E. K. Vizy, L. S. Jackson, and D. Finney (2020), What drives the intensification of mesoscale convective systems over the West African Sahel under climate change?, *Journal of Climate*, *33*(8), 3151-3172, https://doi.org/10.1175/JCLI-D-19-0380.1.

Fovet, O., L. Ruiz, M. Hrachowitz, M. Faucheux, and C. Gascuel-Odoux (2015), Hydrological hysteresis and its value for assessing process consistency in catchment conceptual models, *Hydrology Earth System Sciences*, *19*(1), 105-123, https://doi.org/10.5194/hess-19-105-2015.

Fowler, K., G. Coxon, J. Freer, M. Peel, T. Wagener, A. Western, R. Woods, and L. Zhang (2018a), Simulating Runoff Under Changing Climatic Conditions: A Framework for Model Improvement, *Water Resources Research*, *54*(12), 9812-9832, https://doi.org/10.1029/2018wr023989.

Fowler, K., M. Peel, A. Western, and L. Zhang (2018b), Improved Rainfall-Runoff Calibration for Drying Climate: Choice of Objective Function, *Water Resources Research*, *54*(5), 3392-3408, https://doi.org/10.1029/2017WR022466.

François, B., M. Vrac, A. J. Cannon, Y. Robin, and D. Allard (2020), Multivariate bias corrections of climate simulations: which benefits for which losses?, *Earth System Dynamics*, *11*(2), 537-562, https://doi.org/10.5194/esd-11-537-2020.

Frenken, K. (1997), *Irrigation potential in Africa: A basin approach*, edited, Food & Agriculture Org., http://www.fao.org/publications/card/en/c/744c0135-10ef-59f3-9346-624e7c89b90a/.

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

Gal, L., M. Grippa, P. Hiernaux, L. Pons, and L. Kergoat (2017), The paradoxical evolution of runoff in the pastoral Sahel: analysis of the hydrological changes over the Agoufou watershed (Mali) using the KINEROS-2 model, *Hydrology and Earth System Sciences*, *21*(9), 4591, https://doi.org/10.5194/hess-21-4591-2017.

Gao, F., G. Feng, Y. Ouyang, H. Wang, D. Fisher, A. Adeli, and J. Jenkins (2017), Evaluation of reference evapotranspiration methods in arid, semiarid, and humid regions, *JAWRA Journal of the American Water Resources Association*, *53*(4), 791-808, https://doi.org/10.1111/1752-1688.12530.

Gash, J. (1979), An analytical model of rainfall interception by forests, *Quarterly Journal of the Royal Meteorological Society*, *105*(443), 43-55, https://doi.org/10.1002/qj.49710544304.
Gaupp, F., J. Hall, S. Hochrainer-Stigler, and S. Dadson (2020), Changing risks of simultaneous global breadbasket failure, *Nature Climate Change*, *10*(1), 54-57, https://doi.org/10.1038/s41558-019-0600-z.
Gebremichael, M. (2010), Framework for satellite rainfall product evaluation, *Geophys Monogr Ser*, *191*, 265-275, https://doi.org/10.1029/2010gm000974.
Gelaro, R., W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, and R. Reichle (2017), The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *Journal of Climate*, *30*(14), 5419-5454, https://doi.org/10.1175/JCLI-D-16-0758.1.
Gharari, S., M. Hrachowitz, F. Fenicia, and H. H. G. Savenije (2013), An approach to identify time consistent model parameters: sub-period calibration, *Hydrology and Earth System Sciences*, *17*(1), 149-161, https://doi.org/10.5194/hess-17-149-2013.
Giorgi, F., C. Jones, and G. R. Asrar (2009), Addressing climate information needs at the regional level: the CORDEX framework, *World Meteorological Organization (WMO) Bulletin*, *58*(3), 175, https://public.wmo.int/en/bulletin/addressing-climate-information-needs-regional-level-cordex-framework.
Giustarini, L., O. Parisot, M. Ghoniem, R. Hostache, I. Trebs, and B. Otjacques (2016), A user-driven case-based reasoning tool for infilling missing values in daily mean river flow records, *Environmental Modelling & Software*, *82*, 308-320, https://doi.org/10.1016/j.envsoft.2016.04.013.
Godfray, H. C. J., J. R. Beddington, I. R. Crute, L. Haddad, D. Lawrence, J. F. Muir, J. Pretty, S. Robinson, S. M. Thomas, and C. Toulmin (2010), Food security: the challenge of feeding 9 billion people, *science*, *327*(5967), 812-818, https://doi.org/10.1126/science.1185383
Golian, S., M. Javadian, and A. Behrangi (2019), On the use of satellite, gauge, and reanalysis precipitation products for drought studies, *Environmental Research Letters*, *14*(7), 075005, https://doi.org/10.1088/1748-9326/ab2203.
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, https://doi.org/10.1002/qj.2130.
Gravetter, F. J., and L. B. Wallnau (2013), Chapter 5, z-Scores: Location of Scores and Standardized Distributions, in *Statistics for the behavioral sciences*, edited, pp. 137-162, Cengage Learning, Wadsworth, USA, http://www.paehub.com/download/Stats_Behavioral_Science.pdf.
Grayson, R., and G. Bloschl (2001), *Spatial patterns in catchment hydrology: observations and modelling*, edited, 416 pp., CUP Archive.
Greve, P., P. Burek, and Y. Wada (2020), Using the Budyko framework for calibrating a global hydrological model, *Water Resources Research*, e2019WR026280, https://doi.org/10.1029/2019WR026280.
Gruber, A., W. A. Dorigo, W. Crow, and W. Wagner (2017), Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, *Ieee T Geosci Remote*, *55*(12), 6780-6792, https://doi.org/10.1109/Tgrs.2017.2734070.
Gründemann, G. J., M. Werner, and T. I. Veldkamp (2018), The potential of global reanalysis datasets in identifying flood events in Southern Africa, *Hydrology and Earth System Sciences*, *22*(9), 4667-4683, https://doi.org/10.5194/hess-22-4667-2018.
Guardiano, F. B., and R. M. Srivastava (1993), Multivariate geostatistics: beyond bivariate moments, in *Geostatistics Troia'92*, edited, pp. 133-144, Springer.
Guerschman, J. P., A. I. Van Dijk, G. Mattersdorf, J. Beringer, L. B. Hutley, R. Leuning, R. C. Pipunic, and B. S. Sherman (2009), Scaling of potential evapotranspiration with MODIS data reproduces flux observations and catchment water balance observations across Australia, *Journal of Hydrology*, *369*(1-2), 107-119, https://doi.org/10.1016/j.jhydrol.2009.02.013.
Guillou, M., and G. Matheron (2014), *The World's challenge: feeding 9 Billion people*, edited, Springer, https://doi.org/10.1007/978-94-017-8569-3.
Gunkel, A., and J. Lange (2017), Water scarcity, data scarcity and the Budyko curve—An application in the Lower Jordan River Basin, *Journal of Hydrology: Regional Studies*, *12*, 136-149, https://doi.org/10.1016/j.ejrh.2017.04.004.
Gupta, A., and R. Govindaraju (2019), Propagation of structural uncertainty in watershed hydrologic models, *Journal of Hydrology*, *575*, 66-81, https://doi.org/10.1016/j.jhydrol.2019.05.026.
Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resources Research*, *34*(4), 751-763, https://doi.org/10.1029/97wr03495.

Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrological Processes: An International Journal*, *22*(18), 3802-3813, https://doi.org/10.1002/hyp.6989.
Gupta, H. V., M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye (2012), Towards a comprehensive assessment of model structural adequacy, *Water Resources Research*, *48*(8), https://doi.org/10.1029/2011WR011044.
Gyamfi, S., N. S. Derkyi, and E. Y. Asuamah (2018), The Potential and the Economics of Hydropower Investment in West Africa, in *Sustainable Hydropower in West Africa*, edited, pp. 95-107, Elsevier, https://doi.org/10.1016/B978-0-12-813016-2.00007-1.
Gyau-Boakye, P., and G. Schultz (1994), Filling gaps in runoff time series in West Africa, *Hydrological sciences journal*, *39*(6), 621-636, https://doi.org/10.1080/02626669409492784.
Harada, Y., H. Kamahori, C. Kobayashi, H. Endo, S. Kobayashi, Y. Ota, H. Onoda, K. Onogi, K. Miyaoka, and K. Takahashi (2016), The JRA-55 Reanalysis: Representation of atmospheric circulation and climate variability, *Journal of the Meteorological Society of Japan. Ser. II*, *94*(3), 269-302, https://doi.org/10.2151/jmsj.2016-015.
Hargreaves, G. H., and Z. A. Samani (1985), Reference crop evapotranspiration from temperature, *Applied engineering in agriculture*, *1*(2), 96-99, https://doi.org/10.13031/2013.26773.
Harrison, L., C. Funk, and P. Peterson (2019), Identifying changing precipitation extremes in Sub-Saharan Africa with gauge and satellite products, *Environmental Research Letters*, *14*(8), 085007, https://doi.org/10.1088/1748-9326/ab2cae.
Hartanto, I. M., J. Van Der Kwast, T. K. Alexandridis, W. Almeida, Y. Song, S. van Andel, and D. Solomatine (2017), Data assimilation of satellite-based actual evapotranspiration in a distributed hydrological model of a controlled water system, *International journal of applied earth observation and geoinformation*, *57*, 123-135, https://doi.org/10.1016/j.jag.2016.12.015.
Hartmann, J., and N. Moosdorf (2012), The new global lithological map database GLiM: A representation of rock properties at the Earth surface, *Geochem Geophy Geosy*, *13*(12), https://doi.org/10.1029/2012gc004370.
Harvey, C. L., H. Dixon, and J. Hannaford (2010), Developing best practice for infilling daily river flow data, *Role of Hydrology in Managing Consequences of a Changing Global Environment*, 816-823, https://doi.org/10.7558/bhs.2010.ic119.
Harvey, C. L., H. Dixon, and J. Hannaford (2012), An appraisal of the performance of data-infilling methods for application to daily mean river flow records in the UK, *Hydrology Research*, *43*(5), 618-636, https://doi.org/10.2166/nh.2012.110.
Hausfather, Z., and G. P. Peters (2020), Emissions–the 'business as usual'story is misleading, edited, Nature Publishing Group.
He, M., J. S. Kimball, Y. Yi, S. W. Running, K. Guan, A. Moreno, X. Wu, and M. Maneta (2019), Satellite data-driven modeling of field scale evapotranspiration in croplands using the MOD16 algorithm framework, *Remote Sens Environ*, *230*, 111201, https://doi.org/10.1016/j.rse.2019.05.020.
He, X., M. Pan, Z. Wei, E. F. Wood, and J. Sheffield (2020), A Global Drought and Flood Catalogue from 1950 to 2016, *Bulletin of the American Meteorological Society*(2020), https://doi.org/10.1175/BAMS-D-18-0269.1.
Hengl, T., et al. (2017), SoilGrids250m: Global gridded soil information based on machine learning, *PLoS One*, *12*(2), e0169748, https://doi.org/10.1371/journal.pone.0169748.
Herman, A., V. B. Kumar, P. A. Arkin, and J. V. Kousky (1997), Objectively determined 10-day African rainfall estimates created for famine early warning systems, *International Journal of Remote Sensing*, *18*(10), 2147-2159, https://doi.org/10.1080/014311697217800.
Herman, M. R., A. P. Nejadhashemi, M. Abouali, J. S. Hernandez-Suarez, F. Daneshvar, Z. Zhang, M. C. Anderson, A. M. Sadeghi, C. R. Hain, and A. Sharifi (2018), Evaluating the role of evapotranspiration remote sensing data in improving hydrological modeling predictability, *Journal of Hydrology*, *556*, 39-49, https://doi.org/10.1016/j.jhydrol.2017.11.009.
Hersbach, H., et al. (2018), Operational global reanalysis: progress, future directions and synergies with NWP, ERA Report Series 27, ECMWF, Reading, UK, European Centre for Medium Range Weather Forecasts, https://www.ecmwf.int/node/18765.
Hirsch, R. M. (1979), An evaluation of some record reconstruction techniques, *Water resources research*, *15*(6), 1781-1790, https://doi.org/10.1029/WR015i006p01781.
Hirsch, R. M. (1982), A comparison of four streamflow record extension techniques, *Water Resources Research*, *18*(4), 1081-1088, https://doi.org/10.1029/WR018i004p01081.
Hoekstra, A. Y., M. M. Mekonnen, A. K. Chapagain, R. E. Mathews, and B. D. Richter (2012), Global monthly water scarcity: blue water footprints versus blue water availability, *PloS one*, *7*(2), e32688, https://doi.org/10.1371/journal.pone.0032688.
Hrachowitz, M., et al. (2013), A decade of Predictions in Ungauged Basins (PUB)—a review, *Hydrological Sciences Journal*, *58*(6), 1198-1255, https://doi.org/10.1080/02626667.2013.803183.

Hrachowitz, M., O. Fovet, L. Ruiz, T. Euser, S. Gharari, R. Nijzink, J. Freer, H. Savenije, and C. Gascuel-Odoux (2014), Process consistency in models: The importance of system signatures, expert knowledge, and process complexity, *Water resources research*, *50*(9), 7445-7469, https://doi.org/10.1002/2014WR015484.

Hrachowitz, M., and M. P. Clark (2017), HESS Opinions: The complementary merits of competing modelling philosophies in hydrology, *Hydrology and Earth System Sciences*, *21*(8), 3953-3973, https://doi.org/10.5194/hess-21-3953-2017.

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, *J Hydrometeorol*, *8*(1), 38-55, https://doi.org/10.1175/JHM560.1.

Hunink, J., G. Simons, S. Suárez-Almiñana, A. Solera, J. Andreu, M. Giuliani, P. Zamberletti, M. Grillakis, A. Koutroulis, and I. Tsanis (2019), A simplified water accounting procedure to assess climate change impact on water resources for agriculture across different European river basins, *Water*, *11*(10), 1976, https://doi.org/10.3390/w11101976.

Huntington, T. G. (2006), Evidence for intensification of the global water cycle: review and synthesis, *Journal of Hydrology*, *319*(1-4), 83-95, https://doi.org/10.1016/j.jhydrol.2005.07.003.

ILEC (2017), World Lake Database, http://wldb.ilec.or.jp/Details/Lake/AFR-16, [accessed 08.23.2017].

Immerzeel, W. W., and P. Droogers (2008), Calibration of a distributed hydrological model based on satellite evapotranspiration, *Journal of Hydrology*, *349*(3-4), 411-424, https://doi.org/10.1016/j.jhydrol.2007.11.017.

Jiang, D., and K. Wang (2019), The Role of Satellite-Based Remote Sensing in Improving Simulated Streamflow: A Review, *Water*, *11*(8), 1615, https://doi.org/10.3390/w11081615.

Jiang, L., H. Wu, J. Tao, J. S. Kimball, L. Alfieri, and X. Chen (2020), Satellite-Based Evapotranspiration in Hydrological Model Calibration, *Remote Sens-Basel*, *12*(3), 428, https://doi.org/10.3390/rs12030428.

Jiao, Y., H. M. Lei, D. W. Yang, M. Y. Huang, D. F. Liu, and X. Yuan (2017), Impact of vegetation dynamics on hydrological processes in a semi-arid basin by using a land surface-hydrology coupled model, *Journal of Hydrology*, *551*, 116-131, https://doi.org/10.1016/j.jhydrol.2017.05.060.

Jimenez, C., C. Prigent, B. Mueller, S. I. Seneviratne, M. McCabe, E. F. Wood, W. Rossow, G. Balsamo, A. Betts, and P. Dirmeyer (2011), Global intercomparison of 12 land surface heat flux estimates, *Journal of Geophysical Research: Atmospheres*, *116*(D2), https://doi.org/10.1029/2010JD014545.

Jiménez, C., B. Martens, D. M. Miralles, J. B. Fisher, H. E. Beck, and D. Fernández-Prieto (2018), Exploring the merging of the global land evaporation WACMOS-ET products based on local tower measurements, *Hydrology and Earth System Sciences*, *22*(8), https://doi.org/10.5194/hess-22-4513-2018.

Jin, L., P. G. Whitehead, K. A. Addo, B. Amisigo, I. Macadam, T. Janes, J. Crossman, R. J. Nicholls, M. McCartney, and H. J. Rodda (2018), Modeling future flows of the Volta River system: Impacts of climate change and socio-economic changes, *Science of the Total Environment*, *637*, 1069-1080, https://doi.org/10.1016/j.scitotenv.2018.04.350.

Joiner, J., Y. Yoshida, M. Anderson, T. Holmes, C. Hain, R. Reichle, R. Koster, E. Middleton, and F.-W. Zeng (2018), Global relationships among traditional reflectance vegetation indices (NDVI and NDII), evapotranspiration (ET), and soil moisture variability on weekly timescales, *Remote Sens Environ*, *219*, 339-352, https://doi.org/10.1016/j.rse.2018.10.020.

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, *J Hydrometeorol*, *5*(3), 487-503, https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Jung, G., S. Wagner, and H. Kunstmann (2012), Joint climate–hydrology modeling: an impact study for the data-sparse environment of the Volta Basin in West Africa, *Hydrology Research*, *43*(3), 231-248, https://doi.org/10.2166/nh.2012.044.

Jung, H. C., A. Getirana, K. R. Arsenault, T. R. Holmes, and A. McNally (2019), Uncertainties in Evapotranspiration Estimates over West Africa, *Remote Sens-Basel*, *11*(8), 892, https://doi.org/10.3390/rs11080892

Karambiri, H., S. García Galiano, J. Giraldo, H. Yacouba, B. Ibrahim, B. Barbier, and J. Polcher (2011), Assessing the impact of climate variability and climate change on runoff in West Africa: the case of Senegal and Nakambe River basins, *Atmospheric Science Letters*, *12*(1), 109-115, https://doi.org/10.1002/asl.317.

Karimi, P., W. G. Bastiaanssen, and D. Molden (2013a), Water Accounting Plus (WA+)--a water accounting procedure for complex river basins based on satellite measurements, *Hydrology & Earth System Sciences*, *17*(7), https://doi.org/10.5194/hess-17-2459-2013.

Karimi, P., W. G. Bastiaanssen, D. Molden, and M. J. M. Cheema (2013b), Basin-wide water accounting based on remote sensing data: an application for the Indus Basin, *Hydrology & Earth System Sciences*, *17*(7), https://doi.org/10.5194/hess-17-2473-2013.

Karimi, P. (2014), Water Accounting Plus for Water Resources Reporting and River Basin Planning, https://www.wateraccounting.org/files/poolad_karimi_phd_thesis.pdf.
Kasei, R. A. (2010), *Modelling impacts of climate change on water resources in the Volta Basin, West Africa*, edited, Citeseer, http://hss.ulb.uni-bonn.de/2010/1977/1977.pdf.
Kebe, I., M. B. Sylla, J. A. Omotosho, P. M. Nikiema, P. Gibba, and F. Giorgi (2017), Impact of GCM boundary forcing on regional climate modeling of West African summer monsoon precipitation and circulation features, *Clim Dynam*, *48*(5-6), 1503-1516, https://doi.org/10.1007/s00382-016-3156-x.
Keesstra, S., J. Nunes, A. Novara, D. Finger, D. Avelar, Z. Kalantari, and A. Cerdà (2018), The superior effect of nature based solutions in land management for enhancing ecosystem services, *Science of the Total Environment*, *610*, 997-1009, https://doi.org/10.1016/j.scitotenv.2017.08.077.
Kendon, E. J., N. Ban, N. M. Roberts, H. J. Fowler, M. J. Roberts, S. C. Chan, J. P. Evans, G. Fosser, and J. M. Wilkinson (2017), Do convection-permitting regional climate models improve projections of future precipitation change?, *Bulletin of the American Meteorological Society*, *98*(1), 79-93, https://doi.org/10.1175/BAMS-D-15-0004.1.
Kendon, E. J., R. A. Stratton, S. Tucker, J. H. Marsham, S. Berthou, D. P. Rowell, and C. A. Senior (2019), Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale, *Nature communications*, *10*(1), 1-14, https://doi.org/10.1038/s41467-019-09776-9.
Khan, M. S., U. W. Liaqat, J. Baik, and M. Choi (2018), Stand-alone uncertainty characterization of GLEAM, GLDAS and MOD16 evapotranspiration products using an extended triple collocation approach, *Agricultural and Forest Meteorology*, *252*, 256-268, https://doi.org/10.1016/j.agrformet.2018.01.022.
Khu, S. T., and H. Madsen (2005), Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration, *Water Resources Research*, *41*(3), https://doi.org/10.1029/2004WR003041.
Kidd, C., and G. Huffman (2011), Global precipitation measurement, *Meteorol Appl*, *18*(3), 334-353, https://doi.org/10.1002/met.284.
Kidd, C., and V. Levizzani (2011), Status of satellite precipitation retrievals, *Hydrology and Earth System Sciences*, *15*(4), 1109-1116, https://doi.org/10.5194/hess-15-1109-2011.
Kidd, C., A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B. Kirschbaum (2017), So, how much of the Earth's surface is covered by rain gauges?, *Bull Am Meteorol Soc*, *98*(1), 69-78, https://doi.org/10.1175/BAMS-D-14-00283.1.
Kirby, M., M. Mainuddin, and J. Eastham (2010), Water-use accounts in CPWF basins: Models and Concepts, https://assets.publishing.service.gov.uk/media/57a08b1240f0b64974000950/CP_WP_1_Model_concepts_and_description.pdf.
Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resources Research*, *42*(3), https://doi.org/10.1029/2005wr004362.
Kirchner, J. W., and S. T. Allen (2020), Seasonal partitioning of precipitation between streamflow and evapotranspiration, inferred from end-member splitting analysis, *Hydrology and Earth System Sciences*, *24*(1), 17-39, https://doi.org/10.5194/hess-24-17-2020.
Klemas, V. (2014), Remote sensing of floods and flood-prone areas: an overview, *Journal of Coastal Research*, *31*(4), 1005-1013, https://doi.org/10.2112/JCOASTRES-D-14-00160.1.
Klemes, V. (1986), Operational Testing of Hydrological Simulation-Models, *Hydrolog Sci J*, *31*(1), 13-24, https://doi.org/10.1080/02626668609491024.
Kling, H., M. Fuchs, and M. Paulin (2012), Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, *Journal of Hydrology*, *424*, 264-277, https://doi.org/10.1016/j.jhydrol.2012.01.011.
Kling, H., P. Stanzel, and M. Fuchs (2016), Regional assessment of the hydropower potential of rivers in West Africa, *Energy Procedia*, *97*, 286-293, https://doi.org/10.1016/j.egypro.2016.10.002.
Knoben, W. J., J. E. Freer, and R. A. Woods (2019), Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, *Hydrology and Earth System Sciences*, *23*(10), 4323-4331, https://doi.org/10.5194/hess-23-4323-2019.
Knoche, M., C. Fischer, E. Pohl, P. Krause, and R. Merz (2014), Combined uncertainty of hydrological model complexity and satellite-based forcing data evaluated in two data-scarce semi-arid catchments in Ethiopia, *Journal of Hydrology*, *519*, 2049-2066, https://doi.org/10.1016/j.jhydrol.2014.10.003.
Ko, A., G. Mascaro, and E. R. Vivoni (2019), Strategies to Improve and Evaluate Physics-Based Hyperresolution Hydrologic Simulations at Regional Basin Scales, *Water Resources Research*, https://doi.org/10.1029/2018WR023521.
Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, and H. Endo (2015), The JRA-55 reanalysis: General specifications and basic characteristics, *Journal of the Meteorological Society of Japan. Ser. II*, *93*(1), 5-48, https://doi.org/10.2151/jmsj.2015-001.

Koch, J., M. C. Demirel, and S. Stisen (2018), The SPAtial EFficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models, *Geosci Model Dev*, *11*(5), 1873-1886, https://doi.org/10.5194/gmd-11-1873-2018.
Kölbel, J., C. Strong, C. Noe, and P. Reig (2018), Mapping Public Water Management by Harmonizing and Sharing Corporate Water Risk Information, www.wri.org/publication/mapping-public-water.
Koppa, A., M. Gebremichael, and W. W. Yeh (2019), Multivariate Calibration of Large Scale Hydrologic Models: The Necessity and Value of a Pareto Optimal Approach, *Advances in Water Resources*, https://doi.org/10.1016/j.advwatres.2019.06.005.
Koppa, A., and M. Gebremichael (2020), Improving the Applicability of Hydrologic Models for Food–Energy–Water Nexus Studies Using Remote Sensing Data, *Remote Sens-Basel*, *12*(4), 599, https://doi.org/10.3390/rs12040599
Krause, P., D. Boyle, and F. Bäse (2005), Comparison of different efficiency criteria for hydrological model assessment, *Advances in geosciences*, *5*, 89-97, https://doi.org/10.5194/adgeo-5-89-2005.
Kuczera, G., B. Renard, M. Thyer, and D. Kavetski (2010), There are no hydrological monsters, just models and observations with large uncertainties!, *Hydrological Sciences Journal*, *55*(6), 980-991, https://doi.org/10.1080/02626667.2010.504677.
Kumar, R., L. Samaniego, and S. Attinger (2013), Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, *Water Resources Research*, *49*(1), 360-379, https://doi.org/10.1029/2012wr012195.
Kunnath-Poovakka, A., D. Ryu, L. Renzullo, and B. George (2016), The efficacy of calibrating hydrologic model using remotely sensed evapotranspiration and soil moisture for streamflow prediction, *Journal of hydrology*, *535*, 509-524, https://doi.org/10.1016/j.jhydrol.2016.02.018.
Kvålseth, T. O. (2017), Coefficient of variation: the second-order alternative, *Journal of Applied Statistics*, *44*(3), 402-415, https://doi.org/10.1080/02664763.2016.1174195.
Laiti, L., S. Mallucci, S. Piccolroaz, A. Bellin, D. Zardi, A. Fiori, G. Nikulin, and B. Majone (2018), Testing the Hydrological Coherence of High-Resolution Gridded Precipitation and Temperature Data Sets, *Water Resources Research*, *54*(3), 1999-2016, https://doi.org/10.1002/2017WR021633.
Lakshmi, V. (2000), A simple surface temperature assimilation scheme for use in land surface models, *Water Resources Research*, *36*(12), 3687-3700, https://doi.org/10.1029/2000WR900204.
Lakshmi, V., T. J. Jackson, and D. Zehrfuhs (2003), Soil moisture–temperature relationships: results from two field experiments, *Hydrological processes*, *17*(15), 3041-3057, https://doi.org/10.1002/hyp.1275.
Lambert, A. (1969), A comprehensive rainfall/runoff model for an upland catchment area, *J. Instn. Water Engrs*, *23*(4), 231-238.
Landerer, F. W., and S. C. Swenson (2012), Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resources Research*, *48*(4), https://doi.org/10.1029/2011wr011453.
Lange, S. (2016), EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI), edited by G. D. Services.
Lauri, H., T. Räsänen, and M. Kummu (2014), Using reanalysis and remotely sensed temperature and precipitation data for hydrological modeling in monsoon climate: Mekong River case study, *J Hydrometeorol*, *15*(4), 1532-1545, https://doi.org/10.1175/Jhm-D-13-084.1.
Le Coz, C., and N. van de Giesen (2019), Comparison of rainfall products over sub-Sahara Africa, *J Hydrometeorol*(2019), https://doi.org/10.1175/JHM-D-18-0256.1.
Ledesma, J. L., and M. N. Futter (2017), Gridded climate data products are an alternative to instrumental measurements as inputs to rainfall–runoff models, *Hydrological Processes*, *31*(18), 3283-3293, https://doi.org/10.1002/hyp.11269.
Lefore, N., M. A. Giordano, C. Ringler, and J. Barron (2019), Sustainable and equitable growth in farmer-led irrigation in sub-Saharan Africa: what will it take?, *Water Alternatives*, http://www.water-alternatives.org/index.php/alldoc/articles/vol12/v12issue1/484-a12-1-10/file.
Lehner, B., C. R. Liermann, C. Revenga, C. Vörösmarty, B. Fekete, P. Crouzet, P. Döll, M. Endejan, K. Frenken, and J. Magome (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, https://doi.org/10.1890/100125.
Lepot, M., J. B. Aubin, and F. H. L. R. Clemens (2017), Interpolation in Time Series: An Introductive Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment, *Water*, *9*(10), 796, https://doi.org/10.3390/w9100796.
Leroux, D. J., T. Pellarin, T. Vischel, J.-M. Cohard, T. Gascon, F. Gibon, A. Mialon, S. Galle, C. Peugeot, and L. Seguis (2016), Assimilation of SMOS soil moisture into a distributed hydrological model and impacts on the water cycle variables over the Ouémé catchment in Benin, *Hydrology and Earth System Sciences*, *20*(7), 2827-2840, https://doi.org/10.5194/hess-20-2827-2016.

Lettenmaier, D. P., D. Alsdorf, J. Dozier, G. J. Huffman, M. Pan, and E. F. Wood (2015), Inroads of remote sensing into hydrologic science during the WRR era, *Water Resources Research*, *51*(9), 7309-7342, https://doi.org/10.1002/2015wr017616.

Li, B. L., M. Rodell, B. F. Zaitchik, R. H. Reichle, R. D. Koster, and T. M. van Dam (2012a), Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central Europe, *Journal of Hydrology*, *446*, 103-115, https://doi.org/10.1016/j.jhydrol.2012.04.035.

Li, L., C. S. Ngongondo, C.-Y. Xu, and L. Gong (2012b), Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in southern Africa, *Hydrology Research*, *44*(5), 770-788, https://doi.org/10.2166/nh.2012.175.

Li, X.-H., Q. Zhang, and C.-Y. Xu (2012c), Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin, *Journal of Hydrology*, *426*, 28-38, https://doi.org/10.1016/j.jhydrol.2012.01.013.

Li, Y., S. Grimaldi, V. R. Pauwels, and J. P. Walker (2018), Hydrologic model calibration using remotely sensed soil moisture and discharge measurements: The impact on predictions at gauged and ungauged locations, *Journal of hydrology*, *557*, 897-909, https://doi.org/10.1016/j.jhydrol.2018.01.013.

Liou, Y.-A., and S. K. Kar (2014), Evapotranspiration estimation with remote sensing and various surface energy balance algorithms—A review, *Energies*, *7*(5), 2821-2849, https://doi.org/10.3390/en7052821.

Liu, W., L. Wang, J. Zhou, Y. Li, F. Sun, G. Fu, X. Li, and Y.-F. Sang (2016), A worldwide evaluation of basin-scale evapotranspiration estimates against the water balance method, *Journal of Hydrology*, *538*, 82-95, https://doi.org/10.1016/j.jhydrol.2016.04.006.

Liu, X., T. Yang, K. Hsu, C. Liu, and S. Sorooshian (2017), Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau, *Hydrology and Earth System Sciences*, *21*(1), 169-181, https://doi.org/10.5194/hess-21-169-2017.

Liu, Y., et al. (2012a), Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities, *Hydrology and Earth System Sciences*, *16*(10), 3863-3887, https://doi.org/10.5194/hess-16-3863-2012.

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 (2012b), Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sens Environ*, *123*, 280-297, https://doi.org/10.1016/j.rse.2012.03.014.

Livneh, B., and D. P. Lettenmaier (2012), Multi-criteria parameter estimation for the Unified Land Model, *Hydrology and Earth System Sciences*, *16*(8), 3029-3048, https://doi.org/10.5194/hess-16-3029-2012.

Long, D., L. Longuevergne, and B. R. Scanlon (2014), Uncertainty in evapotranspiration from land surface modeling, remote sensing, and GRACE satellites, *Water Resources Research*, *50*(2), 1131-1151, https://doi.org/10.1002/2013WR014581.

López, O., R. Houborg, and M. F. McCabe (2017), Evaluating the hydrological consistency of evaporation products using satellite-based gravity and rainfall data, *Hydrology and Earth System Sciences*, *21*(1), 323-343, https://doi.org/10.5194/hess-21-323-2017.

Lopez, P. L., E. H. Sutanudjaja, J. Schellekens, G. Sterk, and M. F. P. 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, https://doi.org/10.5194/hess-21-3125-2017.

Lorenz, C., and H. Kunstmann (2012), The hydrological cycle in three state-of-the-art reanalyses: Intercomparison and performance analysis, *J Hydrometeorol*, *13*(5), 1397-1420, https://doi.org/10.1175/Jhm-D-11-088.1.

Ma, Q., L. Xiong, D. Liu, C.-Y. Xu, and S. Guo (2018), Evaluating the Temporal Dynamics of Uncertainty Contribution from Satellite Precipitation Input in Rainfall-Runoff Modeling Using the Variance Decomposition Method, *Remote Sens-Basel*, *10*(12), 1876, https://doi.org/10.3390/rs10121876.

Madsen, H. (2003), Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Advances in Water Resources*, *26*(2), 205-216, https://doi.org/10.1016/S0309-1708(02)00092-1.

Maggioni, V., P. C. Meyers, and M. D. Robinson (2016), A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era, *J Hydrometeorol*, *17*(4), 1101-1117, https://doi.org/10.1175/Jhm-D-15-0190.1.

Maggioni, V., and C. Massari (2018), On the performance of satellite precipitation products in riverine flood modeling: A review, *Journal of hydrology*, *558*, 214-224, https://doi.org/10.1016/j.jhydrol.2018.01.039.

Mahé, G., and J.-E. Paturel (2009), 1896–2006 Sahelian annual rainfall variability and runoff increase of Sahelian Rivers, *Comptes Rendus Geoscience*, *341*(7), 538-546, https://doi.org/10.1016/j.crte.2009.05.002.

Mahé, G., G. Lienou, L. Descroix, F. Bamba, J.-E. Paturel, A. Laraque, M. Meddi, H. Habaieb, O. Adeaga, and C. Dieulin (2013), The rivers of Africa: witness of climate change and human impact on the environment, *Hydrological Processes*, *27*(15), 2105-2114, https://doi.org/10.1002/hyp.9813.

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*(18), 10,619-610,644, https://doi.org/10.1002/2014jd021927.
Maidment, R. I., R. P. Allan, and E. Black (2015), Recent observed and simulated changes in precipitation over Africa, *Geophysical Research Letters*, *42*(19), 8155-8164, https://doi.org/10.1002/2015gl065765.
Maidment, R. I., D. Grimes, E. Black, E. Tarnavsky, M. Young, H. Greatrex, R. P. Allan, T. Stein, E. Nkonde, and S. Senkunda (2017), A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa, *Scientific data*, *4*, 170063, https://doi.org/10.1038/sdata.2017.63.
Maloney, E. D., A. Gettelman, Y. Ming, J. D. Neelin, D. Barrie, A. Mariotti, C.-C. Chen, D. R. Coleman, Y.-H. Kuo, and B. Singh (2019), Process-Oriented Evaluation of Climate and Weather Forecasting Models, *Bulletin of the American Meteorological Society*, *100*(9), 1665-1686, https://doi.org/10.1175/BAMS-D-18-0042.1.
Mando, A. (1997), The impact of termites and mulch on the water balance of crusted Sahelian soil, *Soil technology*, *11*(2), 121-138, https://doi.org/10.1016/S0933-3630(97)00003-2.
Mao, Y., and K. Wang (2017), Comparison of evapotranspiration estimates based on the surface water balance, modified Penman-Monteith model, and reanalysis data sets for continental China, *Journal of Geophysical Research: Atmospheres*, *122*(6), 3228-3244, https://doi.org/10.1002/2016JD026065.
Mariethoz, G., P. Renard, and J. Straubhaar (2010), The direct sampling method to perform multiple-point geostatistical simulations, *Water Resources Research*, *46*(11), https://doi.org/10.1029/2008WR007621.
Marra, F., E. I. Nikolopoulos, E. N. Anagnostou, A. Bardossy, and E. Morin (2019), Precipitation Frequency Analysis from Remotely Sensed Datasets: A Focused Review, *Journal of Hydrology*, *574*, 699-705, https://doi.org/10.1016/j.jhydrol.2019.04.081.
Martens, B., D. G. Miralles, H. Lievens, R. van der Schalie, R. A. M. de Jeu, D. Fernández-Prieto, H. E. Beck, W. A. Dorigo, and N. E. C. Verhoest (2017), GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geosci Model Dev*, *10*(5), 1903-1925, https://doi.org/10.5194/gmd-10-1903-2017.
Martens, B., W. Waegeman, W. A. Dorigo, N. E. Verhoest, and D. G. Miralles (2018), Terrestrial evaporation response to modes of climate variability, *NPJ Climate and Atmospheric Science*, *1*(1), 1-7, https://doi.org/10.1038/s41612-018-0053-5.
Marthews, T. R., E. M. Blyth, A. Martínez-de la Torre, and T. I. Veldkamp (2020), A global-scale evaluation of extreme event uncertainty in the eartH2Observe project, *Hydrology and Earth System Sciences*, *24*(1), 75-92, https://doi.org/10.5194/hess-24-75-2020.
Martin, N. (2006), *Development of a water balance for the Atankwidi catchment, West Africa: A case study of groundwater recharge in a semi-arid climate*, edited, Cuvillier, https://cuvillier.de/uploads/preview/public_file/3980/3865378854.pdf.
Marwala, T. (2009), *Computational Intelligence for Missing Data Imputation, Estimation, and Management: Knowledge Optimization Techniques: Knowledge Optimization Techniques*, edited, IGI Global.
Massari, C., W. Crow, and L. Brocca (2017), An assessment of the performance of global rainfall estimates without ground-based observations, *Hydrology and Earth System Sciences*, *21*(9), 4347-4361, https://doi.org/10.5194/hess-21-4347-2017.
Massari, C., L. Brocca, T. Pellarin, G. Abramowitz, P. Filippucci, L. Ciabatta, V. Maggioni, Y. Kerr, and D. Fernandez Prieto (2020), A daily 25 km short-latency rainfall product for data-scarce regions based on the integration of the Global Precipitation Measurement mission rainfall and multiple-satellite soil moisture products, *Hydrology and Earth System Sciences*, *24*(5), 2687-2710, https://doi.org/10.5194/hess-24-2687-2020.
Mathon, V., H. Laurent, and T. Lebel (2002), Mesoscale convective system rainfall in the Sahel, *J Appl Meteorol*, *41*(11), 1081-1092, https://doi.org/10.1175/1520-0450(2002)041<1081:Mcsrit>2.0.Co;2.
Mazzoleni, M., L. Brandimarte, and A. Amaranto (2019), Evaluating precipitation datasets for large-scale distributed hydrological modelling, *Journal of Hydrology*, 124076, https://doi.org/10.1016/j.jhydrol.2019.124076.
McCabe, M., E. F. Wood, R. Wójcik, M. Pan, J. Sheffield, H. Gao, and H. Su (2008), Hydrological consistency using multi-sensor remote sensing data for water and energy cycle studies, *Remote Sens Environ*, *112*(2), 430-444, https://doi.org/10.1016/j.rse.2007.03.027.
McCabe, M., A. Ershadi, C. Jimenez, D. G. Miralles, D. Michel, and E. F. Wood (2015), The GEWEX LandFlux project: evaluation of model evaporation using tower-based and globally-gridded forcing data, https://doi.org/10.5194/gmd-9-283-2016.
McCabe, M. F., et al. (2017), The Future of Earth Observation in Hydrology, *Hydrol Earth Syst Sci*, *21*(7), 3879-3914, https://doi.org/10.5194/hess-21-3879-2017.
McCabe, M. F., D. G. Miralles, T. R. Holmes, and J. B. Fisher (2019), Advances in the remote sensing of terrestrial evaporation, edited, Multidisciplinary Digital Publishing Institute.

McCartney, M., G. Forkuor, A. Sood, B. Amisigo, F. Hattermann, and L. Muthuwatta (2012), *The water resource implications of changing climate in the Volta River Basin*, edited, IWMI, https://www.iwmi.cgiar.org/Publications/IWMI_Research_Reports/PDF/PUB146/RR146.pdf.

McColl, K. A., J. Vogelzang, A. G. Konings, D. Entekhabi, M. Piles, and A. Stoffelen (2014), Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target, *Geophysical research letters*, *41*(17), 6229-6236, https://doi.org/10.1002/2014GL061322.

McDonnell, J., M. Sivapalan, K. Vaché, S. Dunn, G. Grant, R. Haggerty, C. Hinz, R. Hooper, J. Kirchner, and M. Roderick (2007), Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resources Research*, *43*(7), https://doi.org/10.1029/2006WR005467.

McMahon, T., M. Peel, L. Lowe, R. Srikanthan, and T. McVicar (2013), Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, *Hydrology and Earth System Sciences*, *17*(4), 1331-1363, https://doi.org/10.5194/hess-17-1331-2013.

McMillan, H., J. Freer, F. Pappenberger, T. Krueger, and M. Clark (2010), Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions, *Hydrological Processes: An International Journal*, *24*(10), 1270-1284, https://doi.org/10.1002/hyp.7587.

McMillan, H. K., I. K. Westerberg, and T. Krueger (2018), Hydrological data uncertainty and its implications, *Wiley Interdisciplinary Reviews: Water*, *5*(6), e1319, https://doi.org/10.1002/wat2.1319.

McVicar, T. R., M. L. Roderick, R. J. Donohue, and T. G. Van Niel (2012), Less bluster ahead? Ecohydrological implications of global trends of terrestrial near-surface wind speeds, *Ecohydrology*, *5*(4), 381-388, https://doi.org/10.1002/eco.1298.

Meerschman, E., G. Pirot, G. Mariethoz, J. Straubhaar, M. Van Meirvenne, and P. Renard (2013), A practical guide to performing multiple-point statistical simulations with the Direct Sampling algorithm, *Computers & Geosciences*, *52*, 307-324, https://doi.org/10.1016/j.cageo.2012.09.019.

Mekonnen, M. M., and A. Y. Hoekstra (2015), Global gray water footprint and water pollution levels related to anthropogenic nitrogen loads to fresh water, *Environmental science & technology*, *49*(21), 12860-12868, https://doi.org/10.1021/acs.est.5b03191.

Mekonnen, M. M., and A. Y. Hoekstra (2016), Four billion people facing severe water scarcity, *Science advances*, *2*(2), e1500323, https://doi.org/10.1126/sciadv.1500323.

Melsen, L. A., A. J. Teuling, P. J. Torfs, R. Uijlenhoet, N. Mizukami, and M. P. Clark (2016), HESS Opinions: The need for process-based evaluation of large-domain hyper-resolution models, *Hydrology and Earth System Sciences*, *20*(3), 1069-1079, https://doi.org/10.5194/hess-20-1069-2016.

Mendiguren, G., J. Koch, and S. Stisen (2017), Spatial pattern evaluation of a calibrated national hydrological model - a remote-sensing-based diagnostic approach, *Hydrology and Earth System Sciences*, *21*(12), 5987-6005, https://doi.org/10.5194/hess-21-5987-2017.

Merz, S. K. (2006), Stocktake and analysis of Australia's water accounting practice, *Victoria, Sinclair Knight Merz*, http://www.bom.gov.au/water/about/consultation/document/SKM_Stocktake_Report_2006.pdf.

Miaou, S. P. (1990), A Stepwise Time-Series Regression Procedure for Water Demand Model Identification, *Water Resources Research*, *26*(9), 1887-1897, https://doi.org/10.1029/WR026i009p01887.

Michel, D., C. Jiménez, D. G. Miralles, M. Jung, M. Hirschi, A. Ershadi, B. Martens, M. F. McCabe, J. B. Fisher, and Q. Mu (2016), The WACMOS-ET project–Part 1: Tower-scale evaluation of four remote-sensing-based evapotranspiration algorithms, *Hydrology and Earth System Sciences*, *20*(2), 803-822, https://doi.org/10.5194/hess-20-803-2016.

Michelangeli, P. A., M. Vrac, and H. Loukos (2009), Probabilistic downscaling approaches: Application to wind cumulative distribution functions, *Geophysical Research Letters*, *36*(11), https://doi.org/10.1029/2009GL038401.

Milly, P. C., K. A. Dunne, and A. V. Vecchia (2005), Global pattern of trends in streamflow and water availability in a changing climate, *Nature*, *438*(7066), 347-350, https://doi.org/10.1038/nature04312.

Minville, M., D. Cartier, C. Guay, L. A. Leclaire, C. Audet, S. Le Digabel, and J. Merleau (2014), Improving process representation in conceptual hydrological model calibration using climate simulations, *Water Resources Research*, *50*(6), 5044-5073, https://doi.org/10.1002/2013wr013857.

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

Miralles, D. G., C. Jiménez, M. Jung, D. Michel, A. Ershadi, M. McCabe, M. Hirschi, B. Martens, A. J. Dolman, and J. B. Fisher (2016), The WACMOS-ET project-Part 2: Evaluation of global terrestrial evaporation data sets, *Hydrology and Earth System Sciences*, *20*(2), 823-842, https://doi.org/10.5194/hess-20-823-2016.

Miralles, D. G., W. Brutsaert, A. Dolman, and J. H. Gash (2020), On the use of the term "Evapotranspiration", 8, https://doi.org/10.1002/essoar.10503229.1.

Mizukami, N., M. P. Clark, A. J. Newman, A. W. Wood, E. D. Gutmann, B. Nijssen, O. Rakovec, and L. Samaniego (2017), Towards seamless large-domain parameter estimation for hydrologic models, *Water Resources Research*, *53*(9), 8020-8040, https://doi.org/10.1002/2017wr020401.

Molden, D. (1997), *Accounting for water use and productivity*, edited, Iwmi, https://millenniumindicators.un.org/unsd/envAccounting/ceea/PImeetings/AC-116-10.pdf.

Molden, D., and R. Sakthivadivel (1999), Water accounting to assess use and productivity of water, *International Journal of Water Resources Development*, *15*(1-2), 55-71, https://doi.org/10.1080/07900629948934.

Moniod, F., B. Pouyaud, and P. Séchet (1977), *Le bassin du fleuve Volta*, edited, 513 pp., ORSTOM, Paris, http://horizon.documentation.ird.fr/exldoc/pleins_textes/pleins_textes_6/Mon_hydr/04351.pdf.

Montanari, A., J. Bahr, G. Bloschl, X. M. Cai, D. S. Mackay, A. M. Michalak, H. Rajaram, and G. Sander (2015), Fifty years of Water Resources Research: Legacy and perspectives for the science of hydrology, *Water Resources Research*, *51*(9), 6797-6803, https://doi.org/10.1002/2015wr017998.

Morrison, J., P. Schulte, and P. d. N. U. p. l'environnement (2010), *Corporate water accounting: An analysis of methods and tools for measuring water use and its impacts*, edited, UNEP; Pacific Institute, https://pacinst.org/wp-content/uploads/sites/21/2013/02/corporate_water_accounting_analysis3.pdf.

Moss, R. H., J. A. Edmonds, K. A. Hibbard, M. R. Manning, S. K. Rose, D. P. Van Vuuren, T. R. Carter, S. Emori, M. Kainuma, and T. Kram (2010), The next generation of scenarios for climate change research and assessment, *Nature*, *463*(7282), 747-756, https://doi.org/10.1038/nature08823.

Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sens Environ*, *115*(8), 1781-1800, https://doi.org/10.1016/j.rse.2011.02.019.

Mueller, B., S. I. Seneviratne, C. Jimenez, T. Corti, M. Hirschi, G. Balsamo, P. Ciais, P. Dirmeyer, J. Fisher, and Z. Guo (2011), Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations, *Geophysical Research Letters*, *38*(6), https://doi.org/10.1029/2010GL046230.

Mul, M., E. Obuobie, R. Appoh, K. Kankam-Yeboah, E. Bekoe-Obeng, B. Amisigo, F. Y. Logah, B. Ghansah, and M. McCartney (2015), Water resources assessment of the Volta River Basin*Rep. 9290908297*, 82 pp, International Water Management Institute (IWMI), http://www.iwmi.cgiar.org/Publications/Working_Papers/working/wor166.pdf.

Mulligan, M., A. van Soesbergen, and L. Sáenz (2020), GOODD, a global dataset of more than 38,000 georeferenced dams, *Scientific Data*, *7*(1), 1-8, https://doi.org/10.1038/s41597-020-0362-5.

Mwale, F., A. Adeloye, and R. Rustum (2012), Infilling of missing rainfall and streamflow data in the Shire River basin, Malawi–A self organizing map approach, *Physics and Chemistry of the Earth, Parts A/B/C*, *50*, 34-43, https://doi.org/10.1016/j.pce.2012.09.006.

MWH (1998), Water resources management study, information building block study*Rep. 2*, Ministry of Works and Housing (MWH), Nii Consulting Ltd, Accra, Ghana.

NASA (2019), Time Average Removed from Monthly Solutions, GRACE monthly mass grids-Land, Jet Propulsion Laboratory, http://grace.jpl.nasa.gov/data/get-data/monthly-mass-grids-land/, [accessed 2019.07.20].

NASEM (2019), *National Academies of Sciences Engineering and Medicine (NASEM). Thriving on our changing planet: A decadal strategy for Earth observation from space*, edited, 716 pp., The National Academies Press, Washington, DC, https://doi.org/10.17226/24938.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *Journal of hydrology*, *10*(3), 282-290, https://doi.org/10.1016/0022-1694(70)90255-6.

Nesshöver, C., T. Assmuth, K. N. Irvine, G. M. Rusch, K. A. Waylen, B. Delbaere, D. Haase, L. Jones-Walters, H. Keune, and E. Kovacs (2017), The science, policy and practice of nature-based solutions: An interdisciplinary perspective, *Science of the Total Environment*, *579*, 1215-1227, https://doi.org/10.1016/j.scitotenv.2016.11.106.

Nicholson, S. (2000), Land surface processes and Sahel climate, *Rev Geophys*, *38*(1), 117-139, https://doi.org/10.1029/1999RG900014.

Nicholson, S. E. (2013), The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability, *ISRN Meteorology*, *2013*, https://doi.org/10.1155/2013/453521.

Nicholson, S. E., A. H. Fink, and C. Funk (2018a), Assessing recovery and change in West Africa's rainfall regime from a 161-year record, *International Journal of Climatology*, *38*(10), 3770-3786, https://doi.org/10.1002/joc.5530.

Nicholson, S. E., C. Funk, and A. H. Fink (2018b), Rainfall over the African continent from the 19th through the 21st century, *Global Planet Change*, *165*, 114-127, https://doi.org/10.1016/j.gloplacha.2017.12.014.

Nijssen, B. (2004), Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites, *Journal of Geophysical Research*, *109*(D2), https://doi.org/10.1029/2003jd003497.

Nijzink, R. C., et al. (2018), Constraining Conceptual Hydrological Models With Multiple Information Sources, *Water Resources Research*, *54*(10), 8332-8362, https://doi.org/10.1029/2017wr021895.

Nikiema, P. M., M. B. Sylla, K. Ogunjobi, I. Kebe, P. Gibba, and F. Giorgi (2017), Multi-model CMIP5 and CORDEX simulations of historical summer temperature and precipitation variabilities over West Africa, *International Journal of Climatology*, *37*(5), 2438-2450, https://doi.org/10.1002/joc.4856.

Nikolopoulos, E. I., E. N. Anagnostou, F. Hossain, M. Gebremichael, and M. Borga (2010), Understanding the scale relationships of uncertainty propagation of satellite rainfall through a distributed hydrologic model, *J Hydrometeorol*, *11*(2), 520-532, https://doi.org/10.1175/2009JHM1169.1.

Nkiaka, E., N. Nawaz, and J. C. Lovett (2017), Evaluating global reanalysis datasets as input for hydrological modelling in the Sudano-Sahel region, *Hydrology*, *4*(1), 13, https://doi.org/10.3390/hydrology4010013.

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, https://doi.org/10.1175/JAMC-D-11-0238.1.

O'Neill, B. C., C. Tebaldi, D. P. Van Vuuren, V. Eyring, P. Friedlingstein, G. Hurtt, R. Knutti, E. Kriegler, J.-F. Lamarque, and J. Lowe (2016), The scenario model intercomparison project (ScenarioMIP) for CMIP6, https://doi.org/10.5194/gmd-9-3461-2016.

O'Neill, B. C., E. Kriegler, K. Riahi, K. L. Ebi, S. Hallegatte, T. R. Carter, R. Mathur, and D. P. van Vuuren (2014), A new scenario framework for climate change research: the concept of shared socioeconomic pathways, *Climatic change*, *122*(3), 387-400, https://doi.org/10.1007/s10584-013-0905-2.

O'Neill, B. C., E. Kriegler, K. L. Ebi, E. Kemp-Benedict, K. Riahi, D. S. Rothman, B. J. van Ruijven, D. P. van Vuuren, J. Birkmann, and K. Kok (2017), The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century, *Global Environmental Change*, *42*, 169-180, https://doi.org/10.1016/j.gloenvcha.2015.01.004.

Obuobie, E., and B. Barry (2012), Burkina Faso, in *Groundwater availability and use in Sub-Saharan Africa: A review of 15 countries*, edited by P. Pavelic, M. Giordano, B. Keraita, T. Rao and V. Ramesh, pp. 7-23, International Water Management Institute (IWMI), Colombo, Sri Lanka, https://www.iwmi.cgiar.org/Publications/Books/PDF/groundwater_availability_and_use_in_sub-saharan_africa_a_review_of_15_countries.pdf.

Ocio, D., T. Beskeen, and K. Smart (2019), Fully distributed hydrological modelling for catchment-wide hydrological data verification, *Hydrology Research*, https://doi.org/10.2166/nh.2019.006.

Odusanya, A. E., B. Mehdi, C. Schürz, A. O. Oke, O. S. Awokola, J. A. Awomeso, J. O. Adejuwon, and K. Schulz (2019), Multi-site calibration and validation of SWAT with satellite-based evapotranspiration in a data-sparse catchment in southwestern Nigeria, *Hydrology and Earth System Sciences*, *23*(2), 1113-1144, https://doi.org/10.5194/hess-23-1113-2019.

Okafor, G., T. Annor, S. Odai, and J. Agyekum (2019), Volta basin precipitation and temperature climatology: evaluation of CORDEX-Africa regional climate model simulations, *Theoretical and Applied Climatology*, *137*(3-4), 2803-2827, https://doi.org/10.1007/s00704-018-2746-4.

Oki, T., and S. Kanae (2006), Global hydrological cycles and world water resources, *science*, *313*(5790), 1068-1072, https://doi.org/10.1126/science.1128845

Or, D., and P. Lehmann (2019), Surface evaporative capacitance: How soil type and rainfall characteristics affect global-scale surface evaporation, *Water Resources Research*, *55*(1), 519-539, https://doi.org/10.1029/2018WR024050.

Oriani, F., J. Straubhaar, P. Renard, and G. Mariethoz (2014), Simulation of rainfall time series from different climatic regions using the direct sampling technique, *Hydrology and Earth System Sciences*, *18*(8), 3015-3031, https://doi.org/10.5194/hess-18-3015-2014.

Oriani, F., A. Borghi, J. Straubhaar, G. Mariethoz, and P. Renard (2016), Missing data simulation inside flow rate time-series using multiple-point statistics, *Environmental Modelling & Software*, *86*, 264-276, https://doi.org/10.1016/j.envsoft.2016.10.002.

Oyana, T. J., and F. Margai (2015), Chapter 3, Using Statistical Measures to Analyze Data Distributions, in *Spatial analysis: statistics, visualization, and computational methods*, edited, pp. 55-86, CRC Press, https://books.google.ch/books?id=h2FECgAAQBAJ&pg=PA55&dq=z-score+in+spatial+analysis&source=gbs_toc_r&cad=4#v=onepage&q=z-score%20in%20spatial%20analysis&f=false.

Oyerinde, G. T., D. Wisser, F. C. Hountondji, A. J. Odofin, A. E. Lawin, A. Afouda, and B. Diekkrüger (2016), Quantifying uncertainties in modeling climate change impacts on hydropower production, *Climate*, *4*(3), 34, https://doi.org/10.3390/cli4030034.

Pan, M., H. Li, and E. Wood (2010), Assessing the skill of satellite-based precipitation estimates in hydrologic applications, *Water Resources Research*, *46*(9), https://doi.org/10.1029/2009WR008290.

Pan, S., N. Pan, H. Tian, P. Friedlingstein, S. Sitch, H. Shi, V. K. Arora, V. Haverd, A. K. Jain, and E. Kato (2019), Evaluation of global terrestrial evapotranspiration by state-of-the-art approaches in remote sensing, machine learning, and land surface models, *Hydrology and Earth System Sciences Discussions*, 1-51, https://doi.org/10.5194/hess-2019-409.

Paniconi, C., and M. Putti (2015), Physically based modeling in catchment hydrology at 50: Survey and outlook, *Water Resources Research*, *51*(9), 7090-7129, https://doi.org/10.1002/2015wr017780.
Panthou, G., T. Vischel, T. Lebel, J. Blanchet, G. Quantin, and A. Ali (2012), Extreme rainfall in West Africa: A regional modeling, *Water Resources Research*, *48*(8), https://doi.org/10.1029/2012WR012052.
Papadakis, I., J. Napiorkowski, and G. Schultz (1993), Monthly runoff generation by non-linear model using multispectral and multitemporal satellite imagery, *Advances in Space Research*, *13*(5), 181-186, https://doi.org/10.1016/0273-1177(93)90543-K.
Pappas, C., S. M. Papalexiou, and D. Koutsoyiannis (2014), A quick gap filling of missing hydrometeorological data, *Journal of Geophysical Research-Atmospheres*, *119*(15), 9290-9300, https://doi.org/10.1002/2014jd021633.
Parajka, J., V. Naeimi, G. Bloschl, and J. Komma (2009), Matching ERS scatterometer based soil moisture patterns with simulations of a conceptual dual layer hydrologic model over Austria, *Hydrology and Earth System Sciences*, *13*(2), 259-271, https://doi.org/10.5194/hess-13-259-2009.
Parker, D. J., and M. Diop-Kane (2017), *Meteorology of tropical West Africa: The forecasters' handbook*, edited, 468 pp., John Wiley & Sons, https://doi.org/10.1002/9781118391297.
Pasetto, D., et al. (2018), Integration of satellite remote sensing data in ecosystem modelling at local scales: Practices and trends, *Methods Ecol Evol*, *9*(8), 1810-1821, https://doi.org/10.1111/2041-210x.13018.
Paturel, J. E., M. Ouedraogo, G. Mahe, E. Servat, A. Dezetter, and S. Ardoin (2003), The influence of distributed input data on the hydrological modelling of monthly river flow regimes in West Africa, *Hydrolog Sci J*, *48*(6), 881-890, https://doi.org/10.1623/hysj.48.6.881.51422.
Peel, M. C., and T. A. McMahon (2020), Historical development of rainfall-runoff modeling, *Wiley Interdisciplinary Reviews: Water*, e1471, https://doi.org/10.1002/wat2.1471.
Pellarin, T., et al. (2020), The Precipitation Inferred from Soil Moisture (PrISM) near Real-Time Rainfall Product: Evaluation and Comparison, *Remote Sens-Basel*, *12*(3), 481, https://doi.org/10.3390/rs12030481
Peters-Lidard, C. D., M. Clark, L. Samaniego, N. E. C. Verhoest, T. van Emmerik, R. Uijlenhoet, K. Achieng, T. E. Franz, and R. Woods (2017), Scaling, Similarity, and the Fourth Paradigm for Hydrology, *Hydrol Earth Syst Sci*, *21*(7), 3701-3713, https://doi.org/10.5194/hess-2016-695.
Peters-Lidard, C. D., F. Hossain, L. R. Leung, N. McDowell, M. Rodell, F. J. Tapiador, F. J. Turk, and A. Wood (2019), 100 years of progress in hydrology, *Meteorological Monographs*, https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0019.1.
Peters, G. P., R. M. Andrew, T. Boden, J. G. Canadell, P. Ciais, C. Le Quéré, G. Marland, M. R. Raupach, and C. Wilson (2013), The challenge to keep global warming below 2 C, *Nature Climate Change*, *3*(1), 4-6, https://doi.org/10.1038/nclimate1783.
Peterson, T. J., and A. W. Western (2018), Statistical Interpolation of Groundwater Hydrographs, *Water Resources Research*, *54*(7), 4663-4680, https://doi.org/10.1029/2017wr021838.
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*, *55*(1), 23-39, https://doi.org/10.1175/Jamc-D-15-0065.1.
Pfister, L., and J. W. Kirchner (2017), Debates—Hypothesis testing in hydrology: Theory and practice, *Water Resources Research*, *53*(3), 1792-1798, https://doi.org/10.1002/2016WR020116.
Philippon, N., F. Doblas-Reyes, and P. Ruti (2010), Skill, reproducibility and potential predictability of the West African monsoon in coupled GCMs, *Clim Dynam*, *35*(1), 53-74, https://doi.org/10.1007/s00382-010-0856-5.
Piazza, A. D., F. L. Conti, F. Viola, E. Eccel, and L. V. Noto (2015), Comparative analysis of spatial interpolation methods in the Mediterranean area: application to temperature in Sicily, *Water*, *7*(5), 1866-1888, https://doi.org/10.3390/w7051866
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), https://doi.org/10.1029/2007wr006615.
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, https://doi.org/10.1016/j.jhydrol.2017.01.055.
Poméon, T., B. Diekkrüger, and R. Kumar (2018), Computationally Efficient Multivariate Calibration and Validation of a Grid-Based Hydrologic Model in Sparsely Gauged West African River Basins, *Water*, *10*(10), 1418, https://doi.org/10.3390/w10101418.
Potter, G. L., L. Carriere, J. Hertz, M. Bosilovich, D. Duffy, T. Lee, and D. N. Williams (2018), Enabling reanalysis research using the collaborative reanalysis technical environment (CREATE), *Bulletin of the American Meteorological Society*, *99*(4), 677-687, https://doi.org/10.1175/Bams-D-17-0174.1.
Priestley, C. H. B., and R. Taylor (1972), On the assessment of surface heat flux and evaporation using large-scale parameters, *Monthly weather review*, *100*(2), 81-92, https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2.

Pushpalatha, R., C. Perrin, N. Le Moine, and V. Andreassian (2012), A review of efficiency criteria suitable for evaluating low-flow simulations, *Journal of Hydrology*, *420*, 171-182, https://doi.org/10.1016/j.jhydrol.2011.11.055.

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, https://doi.org/10.5194/hess-20-903-2016.

Raimonet, M., L. Oudin, V. Thieu, M. Silvestre, R. Vautard, C. Rabouille, and P. Le Moigne (2017), Evaluation of gridded meteorological datasets for hydrological modeling, *J Hydrometeorol*, *18*(11), 3027-3041, https://doi.org/10.1175/JHM-D-17-0018.1.

Rajib, A., G. R. Evenson, H. E. Golden, and C. R. Lane (2018a), Hydrologic model predictability improves with spatially explicit calibration using remotely sensed evapotranspiration and biophysical parameters, *Journal of hydrology*, *567*, 668-683, https://doi.org/10.1016/j.jhydrol.2018.10.024.

Rajib, A., V. Merwade, and Z. Yu (2018b), Rationale and efficacy of assimilating remotely sensed potential evapotranspiration for reduced uncertainty of hydrologic models, *Water Resources Research*, *54*(7), 4615-4637, https://doi.org/10.1029/2017WR021147.

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, https://doi.org/10.1002/2016wr019430.

Rakovec, O., R. Kumar, J. Mai, M. Cuntz, S. Thober, M. Zink, S. Attinger, D. Schafer, M. Schron, and L. Samaniego (2016b), Multiscale and Multivariate Evaluation of Water Fluxes and States over European River Basins, *J Hydrometeorol*, *17*(1), 287-307, https://doi.org/10.1175/Jhm-D-15-0054.1.

Ramoelo, A., N. Majozi, R. Mathieu, N. Jovanovic, A. Nickless, and S. Dzikiti (2014), Validation of global evapotranspiration product (MOD16) using flux tower data in the African savanna, South Africa, *Remote Sens-Basel*, *6*(8), 7406-7423, https://doi.org/10.3390/rs6087406.

Rees, G. (2008), Hydrological data, in *Manual on Low-flow Estimation and Prediction. Operational Hydrology Report No. 50*, edited by A. Gustard and S. Demuth, World Meteorological Organization, Geneva, Switzerland.

Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. Mahanama, and G. S. Partyka (2017), Land surface precipitation in MERRA-2, *Journal of Climate*, *30*(5), 1643-1664, https://doi.org/10.1175/JCLI-D-16-0570.1.

Renard, B., D. Kavetski, G. Kuczera, M. Thyer, and S. W. Franks (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resources Research*, *46*(5), https://doi.org/10.1029/2009WR008328.

Revilla-Romero, B., F. A. Hirpa, J. Thielen-del Pozo, P. Salamon, R. Brakenridge, F. Pappenberger, and T. De Groeve (2015), On the Use of Global Flood Forecasts and Satellite-Derived Inundation Maps for Flood Monitoring in Data-Sparse Regions, *Remote Sens-Basel*, *7*(11), 15702-15728, https://doi.org/10.3390/rs71115702.

Riahi, K., D. P. Van Vuuren, E. Kriegler, J. Edmonds, B. C. O'neill, S. Fujimori, N. Bauer, K. Calvin, R. Dellink, and O. Fricko (2017), The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview, *Global Environmental Change*, *42*, 153-168, https://doi.org/10.1016/j.gloenvcha.2016.05.009.

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

Ritchie, J., and H. Dowlatabadi (2017), Why do climate change scenarios return to coal?, *Energy*, *140*, 1276-1291, https://doi.org/10.1016/j.energy.2017.08.083.

Robin, Y., M. Vrac, P. Naveau, and P. Yiou (2019), Multivariate stochastic bias corrections with optimal transport, *Hydrology and Earth System Sciences*, *23*(2), 773-786, https://doi.org/10.5194/hess-23-773-2019.

Roca, R., L. V. Alexander, G. Potter, M. Bador, R. Jucá, S. Contractor, M. G. Bosilovich, and S. Cloché (2019), FROGS: a daily 1°× 1° gridded precipitation database of rain gauge, satellite and reanalysis products, *Earth System Science Data*, *11*(3), 1017-1035, https://doi.org/10.5194/essd-11-1017-2019.

Rodgers, C., N. van de Giesen, W. Laube, P. L. Vlek, and E. Youkhana (2006), The GLOWA Volta Project: A framework for water resources decision-making and scientific capacity building in a transnational West African basin, in *Integrated assessment of water resources and global change*, edited, pp. 295-313, Springer, https://doi.org/10.1007/978-1-4020-5591-1_18.

Rodrigues, L. R. L., J. García-Serrano, and F. Doblas-Reyes (2014), Seasonal forecast quality of the West African monsoon rainfall regimes by multiple forecast systems, *Journal of Geophysical Research: Atmospheres*, *119*(13), 7908-7930, https://doi.org/10.1002/2013JD021316.

Roebeling, R., E. Wolters, J. Meirink, and H. Leijnse (2012), Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data, *J Hydrometeorol*, *13*(5), 1552-1566, https://doi.org/10.1175/JHM-D-11-089.1.
Romilly, T. G., and M. Gebremichael (2011), Evaluation of satellite rainfall estimates over Ethiopian river basins, *Hydrology and Earth System Sciences*, *15*(5), 1505-1514, https://doi.org/10.5194/hess-15-1505-2011.
Roudier, P., A. Ducharne, and L. Feyen (2014a), Climate change impacts on runoff in West Africa: a review, *Hydrology and Earth System Sciences*, *18*(7), 2789-2801, https://doi.org/10.5194/hess-18-2789-2014.
Roudier, P., A. Ducharne, and L. Feyen (2014b), Climate change impacts on runoff in West Africa: a review, https://doi.org/10.5194/hess-18-2789-2014.
Sakumura, C., S. Bettadpur, and S. Bruinsma (2014), Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models, *Geophysical Research Letters*, *41*(5), 1389-1397, https://doi.org/10.1002/2013GL058632.
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), https://doi.org/10.1029/2008wr007327.
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, https://doi.org/10.2166/nh.2011.156.
Samaniego, L., et al. (2017), Toward seamless hydrologic predictions across spatial scales, *Hydrology and Earth System Sciences*, *21*(9), 4323-4346, https://doi.org/10.5194/hess-21-4323-2017.
Sandwidi, W. J. P. (2007), Groundwater potential to supply population demand within the Kompienga dam basin in Burkina Faso, Universitäts-und Landesbibliothek Bonn.
Santos, L., G. Thirel, and C. Perrin (2018), Technical note: Pitfalls in using log-transformed flows within the KGE criterion, *Hydrology and Earth System Sciences*, *22*(8), 4583-4591, https://doi.org/10.5194/hess-22-4583-2018.
Satgé, F., D. Ruelland, M.-P. Bonnet, J. Molina, and R. Pillco (2019), Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region, *Hydrology and Earth System Sciences*, *23*(1), 595-619, https://doi.org/10.5194/hess-23-595-2019.
Satgé, F., D. Defrance, B. Sultan, M.-P. Bonnet, F. Seyler, N. Rouché, F. Pierron, and J.-E. Paturel (2020), Evaluation of 23 gridded precipitation datasets across West Africa, *Journal of Hydrology*, *581*, 124412, https://doi.org/10.1016/j.jhydrol.2019.124412.
Savenije, H. H. (2001), Equifinality, a blessing in disguise?, *Hydrological processes*, *15*(14), 2835-2838, https://doi.org/10.1002/hyp.494.
Savenije, H. H. (2004), The importance of interception and why we should delete the term evapotranspiration from our vocabulary, *Hydrological processes*, *18*(8), 1507-1511, https://doi.org/10.1002/hyp.5563.
Savenije, H. H. (2009), HESS Opinions" The art of hydrology", *Hydrology and Earth System Sciences*, *13*(2), 157-161, https://doi.org/10.5194/hess-13-157-2009.
Schaefli, B., and H. V. Gupta (2007), Do Nash values have value?, *Hydrological Processes*, *21*(15), 2075-2080, https://doi.org/10.1002/hyp.6825.
Schaefli, B., C. Harman, M. Sivapalan, and S. Schymanski (2010), Hydrologic predictions in a changing environment: behavioral modeling, *Hydrology and Earth System Sciences Discussions*, *7*, 7779-7808, https://doi.org/10.5194/hess-15-635-2011.
Schneider, T., T. Bischoff, and G. H. Haug (2014), Migrations and dynamics of the intertropical convergence zone, *Nature*, *513*(7516), 45, https://doi.org/10.1038/nature13636.
Schoups, G., J. W. Hopmans, C. Young, J. Vrugt, and W. W. Wallender (2005), Multi-criteria optimization of a regional spatially-distributed subsurface water flow model, *Journal of Hydrology*, *311*(1-4), 20-48, https://doi.org/10.1016/j.jhydrol.2005.01.001.
Schröder, M., M. Lockhoff, F. Fell, J. Forsythe, T. Trent, R. Bennartz, E. Borbas, M. G. Bosilovich, E. Castelli, and H. Hersbach (2018), The GEWEX Water Vapor Assessment archive of water vapour products from satellite observations and reanalyses, *Earth system science data*, *10*(2), 1093-1117, https://doi.org/10.5194/essd-10-1093-2018.
Schultz, G. A., and E. T. Engman (2012), *Remote sensing in hydrology and water management*, edited, 483 pp., Springer Science & Business Media.
Semenova, O., and K. Beven (2015), Barriers to progress in distributed hydrological modelling, *Hydrological Processes*, *29*(8), 2074-2078, https://doi.org/10.1002/hyp.10434.
Senay, G. B., M. Budde, J. P. Verdin, and A. M. Melesse (2007), A coupled remote sensing and simplified surface energy balance approach to estimate actual evapotranspiration from irrigated fields, *Sensors*, *7*(6), 979-1000, https://doi.org/10.3390/s7060979.
Senay, G. B., S. Bohms, R. K. Singh, P. H. Gowda, N. M. Velpuri, H. Alemu, and J. P. Verdin (2013), Operational evapotranspiration mapping using remote sensing and weather datasets: A new parameterization for the

SSEB approach, *JAWRA Journal of the American Water Resources Association*, *49*(3), 577-591, https://doi.org/10.1111/jawr.12057.
Senay, G. B., N. M. Velpuri, S. Bohms, M. Budde, C. Young, J. Rowland, and J. Verdin (2015), Drought monitoring and assessment: remote sensing and modeling approaches for the famine early warning systems network, in *Hydro-Meteorological Hazards, Risks and Disasters*, edited, pp. 233-262, Elsevier.
Senkondo, W., S. E. Munishi, M. Tumbo, J. Nobert, and S. W. Lyon (2019), Comparing remotely-sensed surface energy balance evapotranspiration estimates in heterogeneous and data-limited regions: A case study of Tanzania's Kilombero Valley, *Remote Sens-Basel*, *11*(11), 1289, https://doi.org/10.3390/rs11111289.
Serdeczny, O., S. Adams, F. Baarsch, D. Coumou, A. Robinson, W. Hare, M. Schaeffer, M. Perrette, and J. Reinhardt (2017), Climate change impacts in Sub-Saharan Africa: from physical changes to their social repercussions, *Reg Environ Change*, *17*(6), 1585-1600, https://doi.org/10.1007/s10113-015-0910-2.
Serrat-Capdevila, A., D. A. García Ramírez, and N. Tayebi (2016), Key Data Needs for Good Water Management, in *Earth Observation for Water Resources Management: Current Use and Future Opportunities for the Water Sector*, edited by L. E. García, D. J. Rodríguez, M. Wijnen and I. Pakulski, p. 267, The World Bank, Washington, D.C., https://doi.org/10.1596/978-1-4648-0475-5.
Serrat-Capdevila, A., J. B. Valdes, and E. Z. Stakhiv (2014), Water management applications for satellite precipitation products: Synthesis and recommendations, *JAWRA Journal of the American Water Resources Association*, *50*(2), 509-525, https://doi.org/10.1111/jawr.12140.
Seyyedi, H., E. N. Anagnostou, E. Beighley, and J. McCollum (2015), Hydrologic evaluation of satellite and reanalysis precipitation datasets over a mid-latitude basin, *Atmos Res*, *164*, 37-48, https://doi.org/10.1016/j.atmosres.2015.03.019.
Shafii, M., and B. A. Tolson (2015), Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives, *Water Resources Research*, *51*(5), 3796-3814, https://doi.org/10.1002/2014wr016520.
Shawul, A. A., and S. Chakma (2020), Suitability of global precipitation estimates for hydrologic prediction in the main watersheds of Upper Awash basin, *Environmental Earth Sciences*, *79*(2), 53, https://doi.org/10.1007%2Fs12665-019-8801-3.
Shayeghi, A., A. Azizian, and L. Brocca (2020), Reliability of reanalysis and remotely sensed precipitation products for hydrological simulation over the Sefidrood River Basin, Iran, *Hydrological Sciences Journal*, *65*(2), 296-310, https://doi.org/10.1080/02626667.2019.1691217.
Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *Journal of climate*, *19*(13), 3088-3111, https://doi.org/10.1175/JCLI3790.1.
Sheffield, J., E. F. Wood, M. Pan, H. Beck, G. Coccia, A. Serrat-Capdevila, and K. Verbist (2018), Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions, *Water Resources Research*, *54*(12), 9724-9758, https://doi.org/10.1029/2017wr022437.
Shilpakar, R. L., W. G. Bastiaanssen, and D. J. Molden (2011), A remote sensing-based approach for water accounting in the East Rapti River Basin, Nepal, *Himalayan Journal of Sciences*, *7*(9), 15-30, https://doi.org/10.3126/hjs.v7i9.5785
Shuttleworth, W. (1993), Evaporation, in *Handbook of hydrology*, edited by M. DR, pp. 4.1–4.53, McGraw-Hill New York, New York.
Sidibe, M., B. Dieppois, G. Mahé, J.-E. Paturel, E. Amoussou, B. Anifowose, and D. Lawler (2018), Trend and variability in a new, reconstructed streamflow dataset for West and Central Africa, and climatic interactions, 1950–2005, *Journal of Hydrology*, *561*, 478-493, https://doi.org/10.1016/j.jhydrol.2018.04.024.
Sidibe, M., B. Dieppois, J. Eden, G. Mahé, J.-E. Paturel, E. Amoussou, B. Anifowose, M. Van De Wiel, and D. Lawler (2020), Near-term impacts of climate variability and change on hydrological systems in West and Central Africa, *Clim Dynam*, *54*(3-4), 2041-2070, https://doi.org/10.1007/s00382-019-05102-7.
Silvestro, F., S. Gabellani, R. Rudari, F. Delogu, P. Laiolo, and G. Boni (2015), Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data, *Hydrology and Earth System Sciences*, *19*(4), 1727-1751, https://doi.org/10.5194/hess-19-1727-2015.
Singh, V. P. (2018), Hydrologic modeling: progress and future directions, *Geosci Lett*, *5*(1), 15, https://doi.org/10.1186/s40562-018-0113-z.
Sivapalan, M. (2003), Prediction in ungauged basins: a grand challenge for theoretical hydrology, *Hydrological Processes*, *17*(15), 3163-3170, https://doi.org/10.1002/hyp.5155.
Smakhtin, V., C. Revenga, and P. Döll (2004), A pilot global assessment of environmental water requirements and scarcity, *Water Int*, *29*(3), 307-317, https://doi.org/10.1080/02508060408691785.

Sood, A., L. Muthuwatta, and M. McCartney (2013), A SWAT evaluation of the effect of climate change on the hydrology of the Volta River basin, *Water Int*, *38*(3), 297-311, https://doi.org/10.1080/02508060.2013.792404.

Sörensson, A. A., and R. C. Ruscica (2018), Intercomparison and uncertainty assessment of nine evapotranspiration estimates over South America, *Water Resources Research*, *54*(4), 2891-2908, https://doi.org/10.1002/2017WR021682.

Sorooshian, S., A. AghaKouchak, P. Arkin, J. Eylander, E. Foufoula-Georgiou, R. Harmon, J. M. Hendrickx, B. Imam, R. Kuligowski, and B. Skahill (2011), Advanced concepts on remote sensing of precipitation at multiple scales, *Bulletin of the American Meteorological Society*, *92*(10), 1353-1357, https://doi.org/10.1175/2011bams3158.1.

Spaaks, J. H., and W. Bouten (2013), Resolving structural errors in a spatially distributed hydrologic model using ensemble Kalman filter state updates, *Hydrology and Earth System Sciences*, *17*(9), 3455-3472, https://doi.org/10.5194/hess-17-3455-2013.

Sposito, G. (2017), Understanding the Budyko equation, *Water*, *9*(4), 236, https://doi.org/10.3390/w9040236

Stanzel, P., H. Kling, and H. Bauer (2018), Climate change impact on West African rivers under an ensemble of CORDEX climate projections, *Climate Services*, *11*, 36-48, https://doi.org/10.1016/j.cliser.2018.05.003.

Steduto, P., J.-M. Faurès, J. Hoogeveen, J. Winpenny, and J. Burke (2012), Coping with water scarcity: an action framework for agriculture and food security, *FAO water reports*, *16*, 78, http://www.fao.org/3/i3015e/i3015e.pdf.

Stephens, G. L., and C. D. Kummerow (2007), The remote sensing of clouds and precipitation from space: A review, *J Atmos Sci*, *64*(11), 3742-3765, https://doi.org/10.1175/2006jas2375.1.

Stisen, S., M. F. McCabe, J. C. Refsgaard, S. Lerer, and M. B. Butts (2011), Model parameter analysis using remotely sensed pattern information in a multi-constraint framework, *Journal of Hydrology*, *409*(1-2), 337-349, https://doi.org/10.1016/j.jhydrol.2011.08.030.

Stisen, S., A. Højberg, L. Troldborg, J. Refsgaard, B. Christensen, M. Olsen, and H. Henriksen (2012), On the importance of appropriate precipitation gauge catch correction for hydrological modelling at mid to high latitudes, *Hydrology and Earth System Sciences*, *16*(11), https://doi.org/10.5194/hess-16-4157-2012.

Stisen, S., J. Koch, T. O. Sonnenborg, J. C. Refsgaard, S. Bircher, R. Ringgaard, and K. H. Jensen (2018), Moving beyond run-off calibration—Multivariable optimization of a surface–subsurface–atmosphere model, *Hydrological Processes*, *32*(17), 2654-2668, https://doi.org/10.1002/hyp.13177.

Straubhaar, J. (2017), *DeeSse User's Guide*, edited, 36 pp., CHYN - Stochastic Hydrogeology and Geostatistics, University of Neuchâtel, Switzerland, www.unine.ch/chyn.

Su, F., Y. Hong, and D. P. Lettenmaier (2008), Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin, *J Hydrometeorol*, *9*(4), 622-640, https://doi.org/10.1175/2007jhm944.1.

Su, Z. (2002), The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes, *Hydrology and Earth System Sciences*, *6*(1), 85-100, https://doi.org/10.5194/hess-6-85-2002.

Sultan, B., and M. Gaetani (2016), Agriculture in West Africa in the twenty-first century: climate change and impacts scenarios, and potential for adaptation, *Frontiers in Plant Science*, *7*, 1262, https://doi.org/10.3389/fpls.2016.01262.

Sun, Q., C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K. L. Hsu (2018), A review of global precipitation data sets: Data sources, estimation, and intercomparisons, *Rev Geophys*, *56*(1), 79-107, https://doi.org/10.1002/2017rg000574.

Sutanudjaja, E., L. Van Beek, S. De Jong, F. Van Geer, and M. Bierkens (2014), Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data, *Water Resources Research*, *50*(1), 687-705, https://doi.org/10.1002/2013WR013807.

Swenson, S. C. (2012), GRACE monthly land water mass grids NETCDF release 5.0. Ver. 5.0. PO.DAAC, CA, USA, https://doi.org/10.5067/TELND-NC005, [Dataset accessed 2018.11.01].

Sylla, M., F. Giorgi, E. Coppola, and L. Mariotti (2013), Uncertainties in daily rainfall over Africa: assessment of gridded observation products and evaluation of a regional climate model simulation, *International Journal of Climatology*, *33*(7), 1805-1817, https://doi.org/10.1002/joc.3551.

Sylla, M. B., P. M. Nikiema, P. Gibba, I. Kebe, and N. A. B. Klutse (2016), Climate change over West Africa: Recent trends and future projections, in *Adaptation to climate change and variability in rural West Africa*, edited, pp. 25-40, Springer, https://doi.org/10.1007/978-3-319-31499-0_3.

Sylla, M. B., A. Faye, N. A. B. Klutse, and K. Dimobe (2018a), Projected increased risk of water deficit over major West African river basins under future climates, *Climatic Change*, *151*(2), 247-258, https://doi.org/10.1007/s10584-018-2308-x.

Sylla, M. B., J. S. Pal, A. Faye, K. Dimobe, and H. Kunstmann (2018b), Climate change to severely impact West African basin scale irrigation in 2 C and 1.5 C global warming scenarios, *Scientific reports*, *8*(1), 1-9, https://doi.org/10.1038/s41598-018-32736-0.

Talsma, C. J., S. P. Good, C. Jimenez, B. Martens, J. B. Fisher, D. G. Miralles, M. F. McCabe, and A. J. Purdy (2018), Partitioning of evapotranspiration in remote sensing-based models, *Agricultural and Forest Meteorology*, *260*, 131-143, https://doi.org/10.1016/j.agrformet.2018.05.010.

Tang, Q. H., H. L. Gao, H. Lu, and D. P. Lettenmaier (2009), Remote sensing: hydrology, *Prog Phys Geog*, *33*(4), 490-509, https://doi.org/10.1177/0309133309346650.

Tang, X., J. Zhang, C. Gao, G. B. Ruben, and G. Wang (2019), Assessing the Uncertainties of Four Precipitation Products for Swat Modeling in Mekong River Basin, *Remote Sens-Basel*, *11*(3), 304, https://doi.org/10.3390/rs11030304.

Tangdamrongsub, N., S. C. Steele-Dunne, B. C. Gunter, P. G. Ditmar, E. H. Sutanudjaja, Y. Sun, T. Xia, and Z. J. Wang (2017), Improving estimates of water resources in a semi-arid region by assimilating GRACE data into the PCR-GLOBWB hydrological model, *Hydrology and Earth System Sciences*, *21*(4), 2053-2074, https://doi.org/10.5194/hess-21-2053-2017.

Tapiador, F., F. J. Turk, W. Petersen, A. Y. Hou, E. García-Ortega, L. A. Machado, C. F. Angelis, P. Salio, C. Kidd, and G. J. Huffman (2012), Global precipitation measurement: Methods, datasets and applications, *Atmos Res*, *104*, 70-97, https://doi.org/10.1016/j.atmosres.2011.10.021.

Tapiador, F., A. Navarro, V. Levizzani, E. García-Ortega, G. Huffman, C. Kidd, P. Kucera, C. Kummerow, H. Masunaga, and W. Petersen (2017), Global precipitation measurements for validating climate models, *Atmos Res*, *197*, 1-20, https://doi.org/10.1016/j.atmosres.2017.06.021.

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), https://doi.org/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, https://doi.org/10.1175/JAMC-D-14-0016.1.

Tauro, F., J. Selker, N. Van De Giesen, T. Abrate, R. Uijlenhoet, M. Porfiri, S. Manfreda, K. Caylor, T. Moramarco, and J. Benveniste (2018), Measurements and Observations in the XXI century (MOXXI): innovation and multi-disciplinarity to sense the hydrological cycle, *Hydrological sciences journal*, *63*(2), 169-196, https://doi.org/10.1080/02626667.2017.1420191.

Taylor, C. M., D. Belušić, F. Guichard, D. J. Parker, T. Vischel, O. Bock, P. P. Harris, S. Janicot, C. Klein, and G. Panthou (2017), Frequency of extreme Sahelian storms tripled since 1982 in satellite observations, *Nature*, *544*(7651), 475, https://doi.org/10.1038/nature22069.

Taylor, J. C., N. van de Giesen, and T. S. Steenhuis (2006), West Africa: Volta discharge data quality assessment and use, *Journal of the American Water Resources Association*, *42*(4), 1113-1126, https://doi.org/10.1111/j.1752-1688.2006.tb04517.x.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bulletin of the American Meteorological Society*, *93*(4), 485-498, https://doi.org/10.1175/BAMS-D-11-00094.1.

Teegavarapu, R. S. V. (2014), Missing precipitation data estimation using optimal proximity metric-based imputation, nearest-neighbour classification and cluster-based interpolation methods, *Hydrolog Sci J*, *59*(11), 2009-2026, https://doi.org/10.1080/02626667.2013.862334.

Tencaliec, P., A. C. Favre, C. Prieur, and T. Mathevet (2015), Reconstruction of missing daily streamflow data using dynamic regression models, *Water Resources Research*, *51*(12), 9447-9463, https://doi.org/10.1002/2015wr017399.

Teng, J., A. J. Jakeman, J. Vaze, B. F. W. Croke, D. Dutta, and S. Kim (2017), Flood inundation modelling: A review of methods, recent advances and uncertainty analysis, *Environmental Modelling & Software*, *90*, 201-216, https://doi.org/10.1016/j.envsoft.2017.01.006.

Tfwala, S. S., Y. M. Wang, and Y. C. Lin (2013), Prediction of Missing Flow Records Using Multilayer Perceptron and Coactive Neurofuzzy Inference System, *Scientific World Journal*, *2013*, https://doi.org/10.1155/2013/584516.

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, *J Hydrometeorol*, *13*(6), 1760-1783, https://doi.org/10.1175/Jhm-D-12-032.1.

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, https://doi.org/10.1016/j.jhydrol.2013.07.012.

Thober, S., M. Cuntz, M. Kelbling, R. Kumar, J. Mai, and L. J. G. M. D. D. Samaniego (2019), The multiscale Routing Model mRM v1. 0: simple river routing at resolutions from 1 to 50 km, *2019*, 1-26, https://doi.org/10.5194/gmd-12-2501-2019.

Thorncroft, C. D., H. Nguyen, C. Zhang, and P. Peyrillé (2011), Annual cycle of the West African monsoon: regional circulations and associated water vapour transport, *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 129-147, https://doi.org/10.1002/qj.728.

Tian, S. Y., P. Tregoning, L. J. Renzullo, A. I. J. M. van Dijk, J. P. Walker, V. R. N. Pauwels, and S. Allgeyer (2017), Improved water balance component estimates through joint assimilation of GRACE water storage and SMOS soil moisture retrievals, *Water Resources Research*, *53*(3), 1820-1840, https://doi.org/10.1002/2016wr019641.

Tobin, K. J., and M. E. Bennett (2014), Satellite precipitation products and hydrologic applications, *Water Int*, *39*(3), 360-380, https://doi.org/10.1080/02508060.2013.870423.

Tobin, K. J., and M. E. Bennett (2017), Constraining SWAT calibration with remotely sensed evapotranspiration data, *JAWRA Journal of the American Water Resources Association*, *53*(3), 593-604, https://doi.org/10.1111/1752-1688.12516.

Todzo, S., A. Bichet, and A. Diedhiou (2020), Intensification of the hydrological cycle expected in West Africa over the 21st century, *Earth System Dynamics*, *11*(1), https://doi.org/10.5194/esd-11-319-2020.

Tolson, B. A., and C. A. Shoemaker (2007), Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, *Water Resources Research*, *43*(1), https://doi.org/10.1029/2005wr004723.

Trabucco, A., and R. Zomer (2018), Global Aridity Index and Potential Evapo-Transpiration (ET0) Climate Database v2. CGIAR Consortium for Spatial Information (CGIAR-CSI). Published online, available from the CGIAR-CSI GeoPortal at https://cgiarcsi.community, https://figshare.com/articles/Global_Aridity_Index_and_Potential_Evapotranspiration_ET0_Climate_Database_v2/7504448/1accessed 11.07.2019].

Trambauer, P., E. Dutra, S. Maskey, M. Werner, F. Pappenberger, L. Van Beek, and S. Uhlenbrook (2014), Comparison of different evaporation estimates over the African continent, *Hydrology and Earth System Sciences*, *18*(1), 193-212, https://doi.org/10.5194/hess-18-193-2014.

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, https://doi.org/10.1080/02626667.2016.1154149.

Tucker, C. J., J. E. Pinzon, M. E. Brown, D. A. Slayback, E. W. Pak, R. Mahoney, E. F. Vermote, and N. El Saleous (2005), An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, *International Journal of Remote Sensing*, *26*(20), 4485-4498, https://doi.org/10.1080/01431160500168686.

Turner, A., K. R. SPERBER, J. Slingo, G. Meehl, C. R. MECHOSO, M. Kimoto, and A. Giannini (2011), Modelling monsoons: Understanding and predicting current and future behaviour, in *The Global Monsoon System: Research and Forecast*, edited, pp. 421-454, World Scientific, https://doi.org/10.1142/9789814343411_0025.

UN-Water (2009), Water in a Changing World, United Nations World Water Development Report 3, Paris, France, https://www.unwater.org/publications/water-changing-world/.

UN-Water (2015), Water and Sustainable Development: From Vision to Action, Annual International Zaragoza Conference, Zaragoza, Spain, https://www.un.org/waterforlifedecade/pdf/WaterandSD_Vision_to_Action-2.pdf.

UN-Water (2020), The United Nations world water development report 2020: water and climate change, UNESCO, Paris, France, https://unesdoc.unesco.org/ark:/48223/pf0000372985.locale=en.

UNEP-GEF (2013), Volta Basin Transboundary Diagnostic Analysis, 154 pp, UNEP-GEF Volta Project, Ghana, http://gefvolta.iwlearn.org/project-resources/studies-reports/tda-final/regional-tda/volta-basin-tda-english.

Upton, G., and I. Cook (2014), A dictionary of statistics, edited, Oxford university press.

Ushio, T., K. Sasashige, T. Kubota, S. Shige, K. i. Okamoto, K. Aonashi, T. Inoue, N. Takahashi, T. Iguchi, and M. Kachi (2009), A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data, *Journal of the Meteorological Society of Japan. Ser. II*, *87*, 137-151, https://doi.org/10.2151/jmsj.87A.137.

Ushio, T., T. Mega, and T. Kubota (2019), Multi-satellite Global Satellite Mapping of Precipitation (GSMaP)-Design and Products, paper presented at 2019 URSI Asia-Pacific Radio Science Conference (AP-RASC), IEEE, https://doi.org/10.23919/URSIAP-RASC.2019.8738594.

Valente, F., J. S. David, and J. H. C. Gash (1997), Modelling interception loss for two sparse eucalypt and pine forests in central Portugal using reformulated Rutter and Gash analytical models, *Journal of Hydrology*, *190*(1-2), 141-162, https://doi.org/10.1016/S0022-1694(96)03066-1.

van de Giesen, N., M. Andreini, A. van Edig, and P. Vlek (2001), Competition for water resources of the Volta basin, *IAHS publication*, 199-206, https://pdfs.semanticscholar.org/c674/8a5d648bbeffb02cd4ba92dd655f97d5c1b3.pdf.

Van der Ent, R. J., H. H. Savenije, B. Schaefli, and S. C. Steele-Dunne (2010), Origin and fate of atmospheric moisture over continents, *Water Resources Research*, *46*(9), https://doi.org/10.1029/2010WR009127.
Van Stan, J. T., E. Gutmann, and J. Friesen (2020), *Precipitation Partitioning by Vegetation: A Global Synthesis*, edited by J. T. Van Stan, E. Gutmann and J. Friesen, 295 pp., Springer Nature, Switzerland, https://doi.org/10.1007/978-3-030-29702-2.
Van Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. C. Hurtt, T. Kram, V. Krey, and J.-F. Lamarque (2011), The representative concentration pathways: an overview, *Climatic change*, *109*(1-2), 5, https://doi.org/10.1007/s10584-011-0148-z.
Velpuri, N. M., G. B. Senay, R. K. Singh, S. Bohms, and J. P. Verdin (2013), A comprehensive evaluation of two MODIS evapotranspiration products over the conterminous United States: Using point and gridded FLUXNET and water balance ET, *Remote Sens Environ*, *139*, 35-49, https://doi.org/10.1016/j.rse.2013.07.013.
Vereecken, H., J. A. Huisman, H. Bogena, J. Vanderborght, J. A. Vrugt, and J. W. Hopmans (2008), On the value of soil moisture measurements in vadose zone hydrology: A review, *Water Resources Research*, *44*(4), https://doi.org/10.1029/2008wr006829.
Vervoort, R. W., S. F. Miechels, F. F. van Ogtrop, and J. H. Guillaume (2014), Remotely sensed evapotranspiration to calibrate a lumped conceptual model: Pitfalls and opportunities, *Journal of hydrology*, *519*, 3223-3236, https://doi.org/10.1016/j.jhydrol.2014.10.034.
Vinukollu, R. K., R. Meynadier, J. Sheffield, and E. F. Wood (2011a), Multi-model, multi-sensor estimates of global evapotranspiration: Climatology, uncertainties and trends, *Hydrological Processes*, *25*(26), 3993-4010, https://doi.org/10.1002/hyp.8393.
Vinukollu, R. K., E. F. Wood, C. R. Ferguson, and J. B. Fisher (2011b), Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches, *Remote Sens Environ*, *115*(3), 801-823, https://doi.org/10.1016/j.rse.2010.11.006.
Vizy, E. K., and K. H. Cook (2018), Mesoscale convective systems and nocturnal rainfall over the West African Sahel: role of the Inter-tropical front, *Clim Dynam*, *50*(1-2), 587-614, https://doi.org/10.1007/s00382-017-3628-7.
Voisin, N., A. W. Wood, and D. P. Lettenmaier (2008), Evaluation of precipitation products for global hydrological prediction, *J Hydrometeorol*, *9*(3), 388-407, https://doi.org/10.1175/2007jhm938.1.
Vörösmarty, C. J., P. B. McIntyre, M. O. Gessner, D. Dudgeon, A. Prusevich, P. Green, S. Glidden, S. E. Bunn, C. A. Sullivan, and C. R. Liermann (2010), Global threats to human water security and river biodiversity, *nature*, *467*(7315), 555-561, https://doi.org/10.1038/nature09440.
Vrac, M., P. Drobinski, A. Merlo, M. Herrmann, C. Lavaysse, L. Li, and S. Somot (2012), Dynamical and statistical downscaling of the French Mediterranean climate: uncertainty assessment, https://doi.org/10.5194/nhess-12-2769-2012.
Vrac, M. (2018), Multivariate bias adjustment of high-dimensional climate simulations: the Rank Resampling for Distributions and Dependences (R2D2) bias correction, *Hydrol. Earth Syst. Sci*, *22*, 3175-3196, https://doi.org/10.5194/hess-22-3175-2018.
Vrac, M., and S. Thao (2020), R2D2: Accounting for temporal dependences in multivariate bias correction via analogue ranks resampling, *Geoscientific Model Development Discussions*, 1-29, https://doi.org/10.5194/gmd-2020-132.
Wagener, T., H. Wheater, and H. V. Gupta (2004), *Rainfall-runoff modelling in gauged and ungauged catchments*, edited, World Scientific, https://doi.org/10.1142/p335.
Wagener, T., M. Sivapalan, P. Troch, and R. Woods (2007), Catchment classification and hydrologic similarity, *Geography compass*, *1*(4), 901-931, https://doi.org/10.1111/j.1749-8198.2007.00039.x.
Wagner, P. D., P. Fiener, F. Wilken, S. Kumar, and K. Schneider (2012a), Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions, *Journal of Hydrology*, *464*, 388-400, https://doi.org/10.1016/j.jhydrol.2012.07.026.
Wagner, W., W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, and M. Ertl (2012b), Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture, paper presented at ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Annals), https://doi.org/10.5194/isprsannals-I-7-315-2012.
Wambura, F. J., O. Dietrich, and G. Lischeid (2018), Improving a distributed hydrological model using evapotranspiration-related boundary conditions as additional constraints in a data-scarce river basin, *Hydrological processes*, *32*(6), 759-775, https://doi.org/10.1002/hyp.11453.
Wan, Z., S. Hook, and G. Hulley (2015), MYD11A2 MODIS/Aqua Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V00610.5067/MODIS/MYD11A2.006, [Dataset accessed 2019.02.13].
Wanders, N., M. F. Bierkens, S. M. de Jong, A. de Roo, and D. Karssenberg (2014), The benefits of using remotely sensed soil moisture in parameter identification of large-scale hydrological models, *Water resources research*, *50*(8), 6874-6891, https://doi.org/10.1002/2013WR014639.

Wang, C., S. Wang, B. Fu, and L. Zhang (2016), Advances in hydrological modelling with the Budyko framework: A review, *Prog Phys Geog*, *40*(3), 409-430, https://doi.org/10.1177/0309133315620997.
Wang, G., J. Pan, C. Shen, S. Li, J. Lu, D. Lou, and D. F. Hagan (2018), Evaluation of evapotranspiration estimates in the Yellow River Basin against the water balance method, *Water*, *10*(12), 1884, https://doi.org/10.3390/w10121884
Wang, K., and R. E. Dickinson (2012), A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, *Rev Geophys*, *50*(2), https://doi.org/10.1029/2011RG000373.
Wang, L., J. Qu, S. Zhang, X. Hao, and S. Dasgupta (2007), Soil moisture estimation using MODIS and ground measurements in eastern China, *International Journal of Remote Sensing*, *28*(6), 1413-1418, https://doi.org/10.1080/01431160601075525.
WDPA (2016), World Database on Protected Areashttps://www.protectedplanet.net/c/world-database-on-protected-areas, [Dataset accessed 01.01.2017].
Wealands, S. R., R. B. Grayson, and J. P. Walker (2005), Quantitative comparison of spatial fields for hydrological model assessment—some promising approaches, *Advances in Water Resources*, *28*(1), 15-32, https://doi.org/10.1016/j.advwatres.2004.10.001.
Weedon, G. P., G. Balsamo, N. Bellouin, S. Gomes, M. J. Best, and P. Viterbo (2014), The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resources Research*, *50*(9), 7505-7514, https://doi.org/10.1002/2014WR015638.
Weerasinghe, I., A. v. Griensven, W. Bastiaanssen, M. Mul, and L. Jia (2019), Can we trust remote sensing ET products over Africa?, *Hydrology and Earth System Sciences Discussions*, 1-27, https://doi.org/10.5194/hess-2019-233.
Werth, S., A. Guntner, S. Petrovic, and R. Schmidt (2009), Integration of GRACE mass variations into a global hydrological model, *Earth Planet Sc Lett*, *277*(1-2), 166-173, https://doi.org/10.1016/j.epsl.2008.10.021.
West, H., N. Quinn, and M. Horswell (2019), Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities, *Remote Sens Environ*, *232*, 111291, https://doi.org/10.1016/j.rse.2019.111291.
Westerberg, I. K., and C. Birkel (2015), Observational uncertainties in hypothesis testing: investigating the hydrological functioning of a tropical catchment, *Hydrological Processes*, *29*(23), 4863-4879, https://doi.org/10.1002/hyp.10533.
Wilby, R. L. (2019), A global hydrology research agenda fit for the 2030s, *Hydrology Research*, https://doi.org/10.2166/nh.2019.100.
Wilkinson, M. D., et al. (2016), The FAIR Guiding Principles for scientific data management and stewardship, *Sci Data*, *3*, 160018, https://doi.org/10.1038/sdata.2016.18.
Williams, T. O., M. L. Mul, C. A. Biney, and V. Smakhtin (2016), *The Volta River Basin: Water for food, economic growth and environment*, edited, 282 pp., Routledge.
Winsemius, H., H. Savenije, and W. Bastiaanssen (2008), Constraining model parameters on remotely sensed evaporation: justification for distribution in ungauged basins?, *Hydrology and Earth System Sciences*, *12*(6), 1403-1413, https://doi.org/10.5194/hess-12-1403-2008.
WMO (2008), *Guide to Hydrological Practices, Hydrology – From Measurement to Hydrological Information*, edited, 6 ed., 296 pp., World Meteorological Organization, WMO-No. 168, Geneva, Switzerland, http://www.wmo.int/pages/prog/hwrp/publications/guide/english/168_Vol_I_en.pdf.
Woodhouse, C. A., S. T. Gray, and D. M. Meko (2006), Updated streamflow reconstructions for the Upper Colorado River Basin, *Water Resources Research*, *42*(5), https://doi.org/10.1029/2005wr004455.
Woodhouse, P., G. J. Veldwisch, J.-P. Venot, D. Brockington, H. Komakech, and Â. Manjichi (2017), African farmer-led irrigation development: re-framing agricultural policy and investment?, *The Journal of Peasant Studies*, *44*(1), 213-233, https://doi.org/10.1080/03066150.2016.1219719.
World Bank Group (2016), *High and dry: Climate change, water, and the economy*, edited, World Bank, http://hdl.handle.net/10986/23665.
Wu, H., R. F. Adler, Y. Tian, G. J. Huffman, H. Li, and J. Wang (2014), Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model, *Water Resources Research*, *50*(3), 2693-2717, https://doi.org/10.1002/2013WR014710.
Wulder, M. A., and N. C. Coops (2014), Satellites: Make Earth observations open access, *Nature News, 513*(7516), 30, https://doi.org/10.1038/513030a.
Xie, P., and P. A. Arkin (1996), Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions, *Journal of climate*, *9*(4), 840-858, https://doi.org/10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO;2.
Xie, P., R. Joyce, S. Wu, S.-H. Yoo, Y. Yarosh, F. Sun, and R. Lin (2017), Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998, *J Hydrometeorol*, *18*(6), 1617-1641, https://doi.org/10.1175/JHM-D-16-0168.1.

Xu, X., J. Li, and B. A. Tolson (2014), Progress in integrating remote sensing data and hydrologic modeling, *Prog Phys Geog*, *38*(4), 464-498, https://doi.org/10.1177/0309133314536583.
Xue, Y., F. De Sales, W.-M. Lau, A. Boone, J. Feng, P. Dirmeyer, Z. Guo, K.-M. Kim, A. Kitoh, and V. Kumar (2010), Intercomparison and analyses of the climatology of the West African Monsoon in the West African Monsoon Modeling and Evaluation project (WAMME) first model intercomparison experiment, *Clim Dynam*, *35*(1), 3-27, https://doi.org/10.1007/s00382-010-0778-2.
Yang, Z., Q. Zhang, Y. Yang, X. Hao, and H. Zhang (2016), Evaluation of evapotranspiration models over semi-arid and semi-humid areas of China, *Hydrological Processes*, *30*(23), 4292-4313, https://doi.org/10.1002/hyp.10824.
Yassin, F., S. Razavi, H. Wheater, G. Sapriza-Azuri, B. Davison, and A. Pietroniro (2017), Enhanced identification of a hydrologic model using streamflow and satellite water storage data: A multicriteria sensitivity analysis and optimization approach, *Hydrological Processes*, *31*(19), 3320-3333, https://doi.org/10.1002/hyp.11267.
Yilmaz, M. T., M. C. Anderson, B. Zaitchik, C. R. Hain, W. T. Crow, M. Ozdogan, J. A. Chun, and J. Evans (2014), Comparison of prognostic and diagnostic surface flux modeling approaches over the Nile River basin, *Water Resources Research*, *50*(1), 386-408, https://doi.org/10.1002/2013WR014194.
Yira, Y., B. Diekkrüger, G. Steup, and A. Y. 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, https://doi.org/10.5194/hess-21-2143-2017.
Zambrano-Bigiarini, M., A. Nauditt, C. Birkel, K. Verbist, and L. Ribbe (2017), Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile, *Hydrology and Earth System Sciences*, *21*(2), 1295-1320, https://doi.org/10.5194/hess-21-1295-2017.
Zandler, H., I. Haag, and C. Samimi (2019), Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions, *Scientific reports*, *9*(1), 1-15, https://doi.org/10.1038/s41598-019-51666-z.
Zeitoun, M., B. Lankford, T. Krueger, T. Forsyth, R. Carter, A. Y. Hoekstra, R. Taylor, O. Varis, F. Cleaver, and R. Boelens (2016), Reductionist and integrative research approaches to complex water security policy challenges, *Global Environmental Change*, *39*, 143-154, https://doi.org/10.1016/j.gloenvcha.2016.04.010.
Zeng, R. J., and X. M. Cai (2016), Climatic and terrestrial storage control on evapotranspiration temporal variability: Analysis of river basins around the world, *Geophysical Research Letters*, *43*(1), 185-195, https://doi.org/10.1002/2015gl066470.
Zhang, B., Y. Xia, B. Long, M. Hobbins, X. Zhao, C. Hain, Y. Li, and M. C. Anderson (2020), Evaluation and comparison of multiple evapotranspiration data models over the contiguous United States: Implications for the next phase of NLDAS (NLDAS-Testbed) development, *Agricultural and Forest Meteorology*, *280*, 107810, https://doi.org/10.1016/j.agrformet.2019.107810.
Zhang, D., X. Liu, P. Bai, and X.-H. Li (2019), Suitability of satellite-based precipitation products for water balance simulations using multiple observations in a humid catchment, *Remote Sens-Basel*, *11*(2), 151, https://doi.org/10.3390/rs11020151.
Zhang, K., J. S. Kimball, and S. W. Running (2016), A review of remote sensing based actual evapotranspiration estimation, *Wiley Interdisciplinary Reviews: Water*, *3*(6), 834-853, https://doi.org/10.1002/wat2.1168.
Zheng, H., Z. L. Yang, P. Lin, J. Wei, W. Y. Wu, L. Li, L. Zhao, and S. Wang (2019), On the sensitivity of the precipitation partitioning into evapotranspiration and runoff in land surface parameterizations, *Water Resources Research*, *55*(1), 95-111, https://doi.org/10.1029/2017WR022236.
Zhu, Z. C., J. Bi, Y. Z. Pan, S. Ganguly, A. Anav, L. Xu, A. Samanta, S. L. Piao, R. R. Nemani, and R. B. Myneni (2013), Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011, *Remote Sens-Basel*, *5*(2), 927-948, https://doi.org/10.3390/rs5020927.
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, https://doi.org/10.1002/2017wr021346.

List of publications related to this PhD thesis

As the lead author of the publications listed below, my contribution was the methodological development, data collection and analysis, programming code writing, model setup, manuscript writing and funding acquisition. The final versions of the publications were possible with valuable contributions of the co-authors, mainly on the supervision of the work, conceptualization, manuscript review and editing, and funding acquisition.

Published in peer-reviewed scientific journals

Dembélé, M., Ceperley, N., Zwart, S. J., Mariéthoz, G., & Schaefli, B. (2020). *Potential of Satellite and Reanalysis Evaporation Datasets for Hydrological Modelling under Various Model Calibration Strategies*. Advances in Water Resources. https://doi.org/10.1016/j.advwatres.2020.103667

Dembélé, M., Hrachowitz, M., Savenije, H., Mariéthoz, G., & Schaefli, B. (2020). *Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite datasets*. Water Resources Research. https://doi.org/10.1029/2019WR026085

Dembélé, M., Oriani, F., Tumbulto, J., Mariéthoz, G., & Schaefli, B. (2019). *Gap-filling of daily streamflow time series using Direct Sampling in various hydroclimatic settings*. Journal of Hydrology, 569, 573-586. https://doi.org/10.1016/j.jhydrol.2018.11.076

Submitted in peer-reviewed scientific journals

Dembélé, M., Schaefli, B., van de Giesen, N., & Mariéthoz, G. (in review). *Suitability of 17 rainfall and temperature gridded datasets for large-scale hydrological modelling in West Africa*. Hydrology and Earth System Sciences Discussions (HESS). https://www.hydrol-earth-syst-sci-discuss.net/hess-2020-68/

In preparation for submission in peer-reviewed scientific journals

Dembélé, M., et al. (in preparation). *Impacts of Climate Change on Hydrological Processes in the Volta River Basin*. (Chapter 7 of this thesis)

Dembélé, M., et al. (in preparation). *Water Accounting for Sustainable Water Management under Climate Change in the Transboundary Volta River Basin*. (Chapter 8 of this thesis)

Contributions to conferences & seminars

Oral presentations

Dembélé, M., Zwart, S. J., Ceperley, N., Mariéthoz, G., & Schaefli, B. (2020). Multivariate and spatially calibrated hydrological model for assessing climate change impacts on hydrological processes in West Africa. EGU General Assembly (online). https://presentations.copernicus.org/EGU2020/EGU2020-9143_presentation.pdf

Dembélé, M. (2020). *Improving process representation in hydrological models: from satellite and reanalysis data accuracy assessment to multivariate model calibration*. *Invited talk* at the Hydro-JULES seminar, CEH, Wallingford, UK, 24 Feb., 2020.

Dembélé, M., Mariéthoz, G., & Schaefli, B. (2019). *How skillful are satellite and reanalysis precipitation products at reproducing hydrological processes in semi-arid catchments?* AGU Fall Meeting, San Francisco, USA. https://agu.confex.com/agu/fm19/meetingapp.cgi/Paper/547064

Dembélé, M., Ceperley, N., Zwart, S. J., Mariéthoz, G., & Schaefli, B. (2019). *What is the impact of calibration procedure on improving hydrological simulations with reanalysis and satellite-based evaporation products?* AGU Fall Meeting, San Francisco, USA. https://agu.confex.com/agu/fm19/meetingapp.cgi/Paper/545356

Dembélé, M., Ceperley, N., Zwart, S. J., Hrachowitz, M., Mariéthoz, G., & Schaefli, B. (2019). *Improving Spatial Patterns of Hydrological Processes for Water Accounting in the Volta River Basin using Earth Observation*. Water Earth Systems PhD conference, Lausanne, Switzerland. https://www.unine.ch/files/live/sites/phdschool-wes/files/PDF/PHDconference_PROGRAM_2019.pdf

Dembélé, M., Hrachowitz, M., Mariéthoz, G., & Schaefli, B. (2019). *Improving the representation of spatial patterns with a distributed hydrologic model through multivariate parameter estimation using multiple satellite data sources*. EGU General Assembly, Vienna, Austria. https://meetingorganizer.copernicus.org/EGU2019/EGU2019-8590.pdf

Dembélé, M. (2018). *Data availability and accessibility to young researchers: Challenges & Opportunities*. *Invited talk* at the ECS-Short course on "Crowd-solving problems in earth science research" by the European Young Geomorphologists at the EGU 2018, Vienna, Austria. http://meetingorganizer.copernicus.org/EGU2018/session/29006

Dembélé, M., Schaefli, B., & Mariéthoz, G. (2017). *Water Accounting Plus in the Volta river basin: quantification of water flows, fluxes and stocks from space*. *Invited talk* at the SGHL/SSHL, CHy, SCNAT, and OFEV congress on remote sensing: applications for hydrology and limnology, Bienne, Switzerland, 13 September 2017.

Posters

Dembélé, M., Schaefli, B., Mariéthoz, G., Ceperley, N., & Zwart, S. (2018). *Supporting water management in the Volta river basin with Water Accounting Plus*. Remote Sensing and Hydrology Symposium (ICRS-IAHS): Applications of remote sensing data in water resources management. RSHS18-31. Cordoba, Spain. https://meetingorganizer.copernicus.org/RSHS18/RSHS18-31.pdf

Dembélé, M., Mariéthoz, G., & Schaefli, B. (2018). *Filling gaps in streamflow data with the Direct Sampling technique*. EGU General Assembly, Vienna, Austria. https://meetingorganizer.copernicus.org/EGU2018/EGU2018-1324.pdf

Dembélé, M., Mariéthoz, G., & Schaefli, B. (2017). *Gap filling of streamflow time series using Direct Sampling in data scarce regions.* 15th Swiss Geoscience Meeting, Davos, Switzerland
https://geoscience-meeting.ch/sgm2017/moctar-dembele-has-submitted-an-abstract/

Dembélé, M., Schaefli, B., Mariéthoz, G., Ceperley, N., & Zwart, S. J. (2017). *Water Accounting Plus for sustainable water management in the Volta river basin, West Africa.* EGU General Assembly, Vienna. http://meetingorganizer.copernicus.org/EGU2017/EGU2017-10220-1.pdf

Co-convenor of scientific sessions

As a co-convener in the scientific events listed below, I actively contributed to designing the content of the sessions, preparing the working materials, reviewing the abstracts and moderating the sessions.

Gravey, M., Dembélé, M., Oriani, F. (2019). *A hands-on introduction to Multiple Point Statistics.* Short course at the EGU General Assembly, Vienna, Austria, 12 Apr. 2019. https://meetingorganizer.copernicus.org/EGU2019/session/30924

Ceperley, N., Mandé, T., & Dembélé, M. (2017). *Citizen-Driven Water Management for Healthy Lives and Sustainable Agriculture.* 4th International Conference on Research for Development (ICRD), Bern, Switzerland, 5-8 September 2017. https://www.icrd.ch/sessions/citizen-driven-water-management-for-healthy-lives-and-sustainable-agriculture/