# Seasonal Predictions: From Global to Regional Information for Decision Support in Water Resources Management in Semi-arid Regions

Kumulative Dissertation zur Erlangung des Doktorgrades
an der

Fakultät für Angewandte Informatik
am Institut für Geographie
der Universität Augsburg

vorgelegt von

**M.Sc. Tanja C. Portele**

November 2021

| | |
|---|---|
| Erstes Gutachten: | Prof. Dr. Harald Kunstmann |
| Zweites Gutachten: | Prof. Dr. Wolfgang Buermann |
| Drittes Gutachten: | Prof. Dr. Axel Bronstert |

Tag der mündlichen Prüfung: 16. März 2022

# Contents

# Abstract

The increasing tension within the water-energy-food nexus in face of climate change requires the sustainable use and management of available water resources. This is particularly relevant in semi-arid regions of South America, Africa and West Asia: already, these regions are facing high precipitation variability, and are additionally exposed to high population growth rates, politically induced unsustainable use of water resources, and transboundary water management conflicts. It is shown that for the climate-sensitive regions of Northeast Brazil, Iran, West Africa, and Northeast Africa, the relative frequency of droughts has increased significantly from 10 to 30 % in recent decades. This calls for proactive measures to aid climate proofing and mitigate climate risks. For water allocation during drought or for flood control, improved knowledge of the upcoming rainy season in semi-arid regions can be the basis for a more robust and sustainable water management. Here, seasonal forecasts with forecast horizons up to seven months ahead can offer great opportunities to support regional water management. State-of-the-art seasonal forecasting systems already reach resolutions that are suitable for regional applications, e.g., the latest seasonal forecasting system version 5 (SEAS5) of the European Centre for Medium-Range Weather Forecasts (ECMWF) with a horizontal resolution of 36 km.

Indeed, with skillful and reliable forecasts for the coming months, decision-makers in the water, agriculture, and energy sectors could induce a more timely, proactive and sustainable reservoir management and seed selection, thereby reducing damage and loss. Decision-makers, however, often still hesitate to use seasonal forecasts claiming their lack of reliability and the inherent uncertainty due to their probabilistic nature. In many cases, statistical performance measures for forecast quality cannot provide actual decision support. Therefore, the potential economic value (PEV) is implemented demonstrating the possible relative savings when basing decisions for preventative water management action on seasonal forecasts compared to optimal early action. This measure of forecast value therewith demonstrates the economic benefit of integrating seasonal forecasts in decisions incorporating the user's cost-loss situation and the forecast probability. The methodology used in this study allows the use of unprocessed global seasonal forecasts for regional analysis. It is shown that a proactive approach based on seasonal forecasts in the case of droughts can achieve potential economic savings of up to 70 % of those from optimal early action. For very warm months and droughts, savings of at least 20 % are achieved even with forecast horizons of up to seven months ahead. For one particular large dam in Sudan, the Upper-Atbara Dam, the savings are explicitly quantified in monetary terms if the dam was operated with early drought mitigation. In fact, avoidable losses of 16 Mio US$ are revealed in one example year.

Though shown to provide substantial economic value to users, for some applications, the seasonal forecasts SEAS5 are still not applicable. If absolute values and high spatial resolution are required, the global seasonal forecasts need to be corrected for biases

and model drifts with increasing forecast horizon (lead time), and a spatial refinement is required. In this study, the forecast quality of global seasonal forecasts is considerably improved regionally with statistical correction methods of quantile mapping for daily data. In this approach, more than 100 000 forecast days are corrected for each of the 25 (until 2016) and 51 (since 2017) ensemble members of a 1981-2019 reforecast period for each of the five analyzed variables. The forecast thus provides more accurate information about future precipitation amounts, temperatures and incoming solar radiation, allowing a concrete decision support in practice. Monthly forecast biases and lead-dependent drift effects of up to 4 mm $d^{-1}$ and 2 K are corrected and the forecasts are spatially disaggregated to the higher resolution of 0.1° of the reanalysis dataset ERA5-Land, further increasing the agreement of spatial patterns with ERA5-Land compared to the raw forecasts. The computational efficiency of this approach allows to provide region-tailored, corrected hydrometeorological seasonal forecasts for the study regions in Northeast Brazil, Sudan/Ethiopia and Iran operationally in an online decision support system (viewer).

In addition to statistical approaches, spatial refinement to higher spatial resolution and reduction of global model errors can be achieved through dynamical downscaling with regional climate models. The dynamical approach has the advantage of resolving or better representing smaller-scale processes that are not captured in the global model products, and thereby reducing possible model biases. Therefore, the potential of a physically-based refinement approach is investigated. In particular, the dependence on the physical parameters in the atmospheric model, which are known to determine the quality of the hydrometeorological information produced, is analyzed. For this purpose, the Weather Research and Forecasting (WRF) model is applied to dynamically downscale reanalysis data in the climatically sensitive study regions of Ecuador/Peru and Brazil with different parameterization setups. Downscaling is performed to a horizontal resolution of 9 km for both domains, and to an additional 3 km for the orographically highly complex domain of Ecuador/Peru. This study reveals that a high uncertainty in the simulated precipitation is introduced with the combination of physical parameterizations of four different cumulus convection (CU) physics, two microphysics (MP), two planetary boundary layer (PBL) physics and two radiation (RA) physics schemes (32 parameterization runs in total). Up to four times the monthly reference precipitation can be modeled according to different setups. A special focus of this study is on spatial patterns, which are analyzed in detail using an ensemble-tailored evaluation approach via the eSAL metric. The pooled results from ensemble-tailored distributional, temporal and the spatial pattern evaluation metrics reveal that the CU schemes, followed by the RA schemes, produce the dominant tendencies in the simulated precipitation within the WRF ensemble (too dry/wet, too large/small precipitation features etc.). No single setup is found to be superior, but the metrics allow for appropriate selection based on application needs and reference data. Future application of the setup to SEAS5 seasonal forecasts is proving to be less than straightforward. The double random perturbation of the initial conditions of SEAS5 impedes a necessary selection of

ensemble members to reduce the computational cost of the dynamical downscaling. Since the parameterization sensitivity study also shows that biases still exist after dynamical downscaling compared to the reference, whether in distributional, temporal, or spatial patterns, additional dynamical downscaling of long reforecast periods would be required for post-processing, adding an immense computational cost.

This work therewith demonstrates the opportunities and highlights the challenges in using seasonal forecasts for decision support in water resources management, and in refining and correcting these forecasts for impact studies and direct operational applications. It enables the final transfer of acquired scientific knowledge into practice and provides the tools to actively contribute to sustainable water resources management in semi-arid regions that urgently need adaptation and mitigation strategies to combat the impacts of climate change and policy-driven water use and management challenges. It turns out that global datasets of seasonal forecasts and reanalyses can provide concrete value for decision support, but depending on the application, further post-processing is needed to make these data regionally usable.

# Kurzzusammenfassung

Die zunehmende Spannung innerhalb des Wasser-, Energie-, und Ernährungssicherheitsnexus angesichts des Klimawandels erfordert die nachhaltige Nutzung und Bewirtschaftung der verfügbaren Wasserressourcen. Dies ist vor allem in den semiariden Regionen Südamerikas, Afrikas und Westasiens von Bedeutung: Diese Regionen sind bereits jetzt einer hohen Niederschlagsvariabilität ausgesetzt und haben zudem mit hohen Bevölkerungswachstumsraten, politisch veranlasster nicht-nachhaltiger Nutzung von Wasserressourcen und grenzüberschreitenden Wasserbewirtschaftungskonflikten zu kämpfen. Diese Arbeit zeigt, dass in den klimasensiblen Regionen Nordostbrasiliens, Irans, Westafrikas und Nordostafrikas die relative Häufigkeit von Dürren in den letzten Jahrzehnten von 10 auf 30 % gestiegen ist. Diese Entwicklung verlangt nach proaktiven Maßnahmen zur Unterstützung des Klimaschutzes und zur Minderung von Klimarisiken. Für die Wasserverteilung während einer Dürre bzw. für den Hochwasserschutz kann eine bessere Vorhersage über die bevorstehende Regenzeit in semiariden Regionen die Grundlage für ein robusteres und nachhaltigeres Wassermanagement bilden. Hier eröffnen saisonale Vorhersagen mit Vorhersagehorizonten von bis zu sieben Monaten im Voraus große Möglichkeiten zur Unterstützung des regionalen Wassermanagements. Moderne saisonale Vorhersagesysteme erreichen bereits Auflösungen, die für regionale Anwendungen geeignet sind, z. B. das neueste saisonale Vorhersagesystem Version 5 (SEAS5) des Europäischen Zentrums für mittelfristige Wettervorhersage (ECMWF) mit einer horizontalen Auflösung von 36 km.

Mit leistungsfähigen und zuverlässigen Vorhersagen für die kommenden Monate könnten Entscheidungsträger im Wasser-, Landwirtschafts- und Energiesektor das Management von Wasserreservoirs und die Auswahl von Saatgut zeitnaher, proaktiver und nachhaltiger gestalten und so Schäden und Verluste verringern. Entscheidungsträger zögern jedoch oft noch, saisonale Vorhersagen zu verwenden, und verweisen auf deren mangelnde Zuverlässigkeit und die ihnen innewohnende Unsicherheit aufgrund ihres wahrscheinlichkeitsbasierten Charakters. In vielen Fällen können statistische Leistungskennzahlen für die Vorhersagequalität keine wirkliche Entscheidungshilfe bieten. Daher wird der potenzielle wirtschaftliche Wert (PEV) eingeführt. Dieser zeigt die möglichen relativen Einsparungen auf, wenn Entscheidungen für vorbeugende wasserwirtschaftliche Maßnahmen auf saisonale Vorhersagen gestützt werden, im Vergleich zu optimalen frühzeitigen Maßnahmen. Dieses Maß für den Prognosewert zeigt damit den wirtschaftlichen Nutzen der Einbeziehung von saisonalen Vorhersagen in Entscheidungen unter Berücksichtigung der Kosten-Verlust-Situation des Nutzers und der Vorhersagewahrscheinlichkeit. Die in dieser Arbeit angewandte Methodik ermöglicht die Verwendung unbearbeiteter globaler saisonaler Vorhersagen für regionale Analysen. Es wird gezeigt, dass ein proaktiver Ansatz auf der Grundlage von saisonalen Vorhersagen im Falle von Dürreperioden potenzielle wirtschaftliche Einsparungen von bis zu 70 % derjenigen von optimalen frühzeitigen Maßnahmen erzielen kann. In sehr warmen Monaten und bei Dürreperioden werden sogar bei einem Vorhersagehorizont von

bis zu sieben Monaten im Voraus noch Einsparungen von mindestens 20 % erzielt. Für einen ausgewählten Großstaudamm im Sudan, den Upper-Atbara-Damm, werden die Einsparmöglichkeiten explizit in Geldwerten beziffert, wenn der Damm mit frühzeitigen Maßnahmen zur Dürrebekämpfung betrieben würde. Tatsächlich werden dabei in einem Beispieljahr vermeidbare Verluste von 16 Mio US$ festgestellt.

Obwohl sie den Nutzern nachweislich einen erheblichen wirtschaftlichen Vorteil bieten, sind die saisonalen Vorhersagen SEAS5 für einige Anwendungen noch immer nicht geeignet. Wenn absolute Werte und eine hohe räumliche Auflösung erforderlich sind, müssen die globalen saisonalen Vorhersagen um Verzerrungen und Modelldrift mit zunehmendem Vorhersagehorizont (Vorlaufzeit) korrigiert werden, und es bedarf einer räumlichen Verfeinerung. In dieser Studie wird die Vorhersagequalität globaler saisonaler Vorhersagen mit statistischen Korrekturmethoden der Quantilabbildung (Quantile Mapping) für tägliche Daten regional bedeutend verbessert. Dabei werden mehr als 100.000 Vorhersagetage für jedes der 25 (bis 2016) und 51 (seit 2017) Ensemblemitglieder des Vorhersagezeitraums 1981-2019 für jede der fünf analysierten Variablen korrigiert. Die Vorhersage liefert damit genauere Informationen über zukünftige Niederschlagsmengen, Temperaturen und die einfallende Sonnenstrahlung, und ermöglicht so eine konkrete Entscheidungshilfe in der Praxis. Monatliche Vorhersageverzerrungen und vorlaufzeitabhängige Drifteffekte von bis zu 4 mm $d^{-1}$ und 2 K werden korrigiert, und die Vorhersagen werden räumlich auf die höhere Auflösung von 0,1° des Reanalysedatensatzes ERA5-Land disaggregiert, wodurch die Übereinstimmung der räumlichen Muster mit ERA5-Land im Vergleich zu den Rohvorhersagen weiter verbessert wird. Die rechnerische Effizienz dieses Ansatzes ermöglicht es, regional angepasste, verbesserte hydrometeorologische saisonale Vorhersagen für die Untersuchungsregionen in Nordostbrasilien, Sudan/Äthiopien und im Iran in einem Online-Entscheidungsunterstützungssystem (Viewer) bereitzustellen.

Zusätzlich zu dem statistischen Ansatz kann eine räumliche Verfeinerung auf eine höhere räumliche Auflösung und die Verringerung der Fehler globaler Modelle durch dynamisches Downscaling mit regionalen Klimamodellen erreicht werden. Der dynamische Ansatz hat den Vorteil, dass er kleinskalige Prozesse, die in den globalen Modellprodukten nicht erfasst werden, auflöst bzw. besser darstellt und dadurch mögliche Modellfehler reduziert. Daher wird in dieser Arbeit auch das Potenzial eines physikalisch basierten Verfeinerungsansatzes untersucht. Insbesondere wird die Abhängigkeit von den physikalischen Parametern im Atmosphärenmodell analysiert, die bekanntlich die Qualität der erzeugten hydrometeorologischen Informationen bestimmen. Dazu wird das Weather Research and Forecasting (WRF) Modell zum dynamischen Downscaling von Reanalysedaten in den klimatisch sensiblen Untersuchungsregionen Ecuador/Peru und Brasilien mit unterschiedlichen Parametrisierungssetups eingesetzt. Das Downscaling wird für beide Gebiete mit einer horizontalen Auflösung von 9 km und für das orografisch hochkomplexe Gebiet von Ecuador/Peru mit einer zusätzlichen Auflösung von 3 km durchgeführt. Diese Studie zeigt, dass die Kombination von unterschiedlichen physikalischen Parametrisierungen mit vier Schemata

für die Kumuluskonvektion (CU), je zwei für die Mikrophysik (MP), planetare Grenzschicht (PBL) und Strahlung (RA) (insgesamt 32 Parametrisierungen) eine hohe Unsicherheit in den simulierten Niederschlag einbringt. Dabei kann bis zum Vierfachen des monatlichen Referenzniederschlags mit verschiedenen Konfigurationen modelliert werden. Ein besonderer Schwerpunkt dieser Studie liegt auf den räumlichen Mustern, die mit einem auf Ensemble zugeschnittenen Bewertungsansatz über die Metrik eSAL im Detail analysiert werden. Die zusammengefassten Ergebnisse für Verteilungs-, zeitliche und räumliche Muster zeigen, dass die CU-Schemata, gefolgt von den RA-Schemata, die dominierenden Tendenzen des simulierten Niederschlags innerhalb des WRF-Ensembles erzeugen (zu trocken/nass, zu große/kleine Niederschlagsmuster usw.). Insgesamt hat sich keine einzige Konfiguration als deutlich überlegen erwiesen, aber die angewandten Metriken ermöglichen eine angemessene Auswahl auf der Grundlage der Anwendungsanforderungen und der Referenzdaten. Eine zukünftige Übertragung des Setups auf die saisonalen Vorhersagen von SEAS5 erweist sich jedoch als nicht ganz einfach. Die doppelte zufällige Störung der Anfangsbedingungen von SEAS5 erschwert eine notwendige Auswahl von Ensemblemitgliedern, um den Rechenaufwand des dynamischen Downscaling zu reduzieren. Da die Studie zur Sensitivität der Parametrisierung auch zeigt, dass nach dem dynamischen Downscaling im Vergleich zur Referenz immer noch Abweichungen bestehen, sei es bei den Verteilungsmustern, den zeitlichen oder den räumlichen Mustern, wäre ein zusätzliches dynamisches Downscaling langer Vorhersagezeiträume in der Vergangenheit für das Post-Processing erforderlich, was einen immensen Rechenaufwand bedeutet.

Diese Arbeit zeigt damit die Möglichkeiten auf und stellt die Herausforderungen heraus, die sich bei der Verwendung von saisonalen Vorhersagen zur Entscheidungsunterstützung in der Wasserwirtschaft und bei der Verfeinerung und Korrektur dieser Vorhersagen für Impaktstudien und direkte operationelle Anwendungen ergeben. Sie ermöglicht den finalen Transfer der gewonnenen wissenschaftlichen Erkenntnisse in die Praxis und liefert die erforderlichen Instrumente, um aktiv zu einer nachhaltigen Bewirtschaftung der Wasserressourcen in semiariden Regionen beizutragen. Denn diese Regionen benötigen dringend Anpassungs- und Abmilderungsstrategien, um die Auswirkungen des Klimawandels und die politisch bedingten Herausforderungen der Wassernutzung und -bewirtschaftung zu bekämpfen. Es zeigt sich, dass globale Datensätze mit saisonalen Vorhersagen und Reanalysen einen konkreten Nutzen für die Entscheidungsfindung bieten können, aber je nach Anwendung sind weitere Nachbearbeitungen erforderlich, um diese Daten regional nutzbar zu machen.

# Acknowledgments

First of all, I would like to thank my supervisor Prof. Dr. Harald Kunstmann for all his support, professional advice, patience, and kindness. I would also like to express my sincere thanks to my two supervisors Dr. Christof Lorenz and Dr. Patrick Laux for all the interesting, inspiring, and friendly discussions, comments, hints, and scientific suggestions that made things working easier for me. Without their threefold help and support, this work would not have been possible.

I would also like to thank Prof. Dr. Wolfgang Buermann, Prof. Dr. Axel Bronstert and Prof. Dr. Peter Fiener for their kind willingness to serve in the dissertation committee.

Furthermore, I greatly appreciate the excellent camaraderie, friendliness, and helpfulness in our Regional Climate and Hydrology working group and from many other colleagues at the Institute for Meteorology and Climate Research. I would like to thank you all for the many interesting discussions and activities, the stimulating coffee breaks, and the open-minded lunches in the cooking group.

I would also like to thank the Augsburg working group, especially Dr. Jan Bliefernicht, for the valuable exchange and fruitful discussions.

My thanks also go to the great internal and external partners of the SaWaM project for interesting discussions, valuable exchanges, insights into different cultures and work practices.

Finally, I would like to thank my family and friends who always had an open ear and always supported and motivated me.

# 1 Overview and Introduction

## 1.1 Background and Motivation: The Need for Sustainable Water Resources Management

Climate change and the projected socioeconomic development make freshwater scarcity a global systemic risk. Since 1960, the proportion of people living under water scarcity has increased rapidly (Kummu et al., 2010), making it one of the major issues in the 21st century. In 2000, the water supply of around 80 % of the world's population (4.8 billion) was potentially threatened due to regionally and temporally very variable water availability (Vörösmarty et al., 2010). In 2005, 35 % (2.3 billion) of the world's population lived under high ($500 - 1000$ m$^3$ per capita and year) or even extreme ($< 500$ m$^3$ per capita and year) water shortage (Kummu et al., 2010). During at least one month of the year, 4.0 billion people, i.e. almost two thirds of the global population in 2005 according to the study of Mekonnen and Hoekstra (2016), experienced severe water scarcity. Due to population increase, the water demand is further rising (Vörösmarty et al., 2000) and about 40-50 % of the world's population will live under water scarcity by 2050 (Gosling and Arnell, 2016). Posing a severe threat on biodiversity, human welfare and economic growth, the progressing freshwater scarcity is increasingly assessed as one of the largest global risks in terms of potential impact (Mekonnen and Hoekstra, 2016; World Economic Forum, 2021).

Socioeconomic factors like technological progress, economic and population growth dominate the water stress, however, climate change will further pose a major threat to the water security and water supply of billions (Distefano and Kelly, 2017). A global warming of $+2$ °C will exacerbate the level of extreme water scarcity (percentage of world population with below 500 m$^3$ per capita and year) expected from population change alone by another 40 % (Schewe et al., 2014). Here, climate change will especially impact on countries that are already affected by water stress (Distefano and Kelly, 2017; Munia et al., 2020), such as on arid and semi-arid regions. Vörösmarty et al. (2010) identified regions of high threat to human water security and biodiversity especially in coastal Ecuador/Peru, West and East Africa, Iran, India, the western part of East Asia and northern part of Southeast Asia. This development is even more alarming due to the fact that poor countries cannot afford massive investments in water technology, which makes them even more vulnerable (Vörösmarty et al., 2010).

The projected development of global water stress outlines the severe impacts that climate change and population increase are expected to have. Not incorporating effects of current or future water resources management strategies (Arnell et al., 2011), these studies underline the need to develop and implement advanced and sustainable concepts for water resources management. Distefano and Kelly (2017) warn that business as usual is not acceptable and will result in over-exploitation of available water resources. To be able to reduce the impacts of water scarcity, a major turnaround towards sustainable water management is urgently required (AghaKouchak et al., 2015; Ashraf

et al., 2019). Man-made water stress and anthropogenic drought caused by unsustainable management of water resources must be a phenomenon of the past (AghaKouchak et al., 2015). But still far too often, decisions that depend on water resources (e.g., on land use, agriculture, energy management) are made without considering the sustainability of water use. As a result, water is increasingly becoming a source of conflict - not only in the crosshairs of the water-food-energy nexus, but also in the often insufficient transboundary river basin cooperation in terms of economic, social, and environmental consequences. The United Nations Sustainable Development Goals (SDGs) adopted at the 2015 UN Summit as part of the 2030 Agenda for Sustainable Development therefore rightly call for the sustainable management of water worldwide (United Nations, 2015).

## 1.2 Challenges of Water Resources Management in Semi-arid Regions: Five Examples

Avoiding future water stress and conflicts in the water-food-energy nexus is particularly relevant for areas that are already characterized by a lack of water, i.e., for the arid and semi-arid areas. Approximately 40 % of the land area are drylands (hyperarid, arid, semi-arid, sub-humid), inhabited by more than one third of the global population (Environment Management Group, 2011). In contrast to arid regions, semi-arid regions (15% of global land area) still have considerable water resources. However, the water systems of semi-arid regions, i.e., the river basins, wetlands, and reservoirs, are generally characterized by a pronounced seasonality of meteorological conditions and natural water availability. Consequently, water scarcity is a prevalent problem, whether in certain periods as during dry seasons or in drought years, or in sub-regions that are not connected to the water and irrigation system or to the water supply infrastructure (e.g., Loucks and van Beek, 2017). For most semi-arid regions, observations and projections indicate warming trends at rates above the global mean over land, enhanced evaporative loss, and reduced and more variable precipitation (Scholes, 2020). The high rainfall variability in semi-arid regions poses a high risk to rain-fed agriculture (Environment Management Group, 2011). Therefore, irrigated agriculture is often essential for the food supply of the population in these regions. In total, more than one third of semi-arid drylands are farmed as rainfed or irrigated land (Environment Management Group, 2011). Providing drinking water for the often large populations of these areas in times of water scarcity is another challenge that is enabled by the construction of dams for supplying reservoirs. As multi-purpose reservoirs, they often serve not only for drinking water supply and irrigation, but as hydropower plants they also form an important part of the electrical energy supply. For states comprising semi-arid regions, the importance of hydropower generation can be immense with up to two thirds of a state's total power generated at hydroelectric dams (von Sperling, 2012). Thus, in the water-food-energy nexus of semi-arid regions, the available water resources are highly used due to the high specific water demands. While there are few options for action in arid regions other than the provision of external water (e.g., through desalination

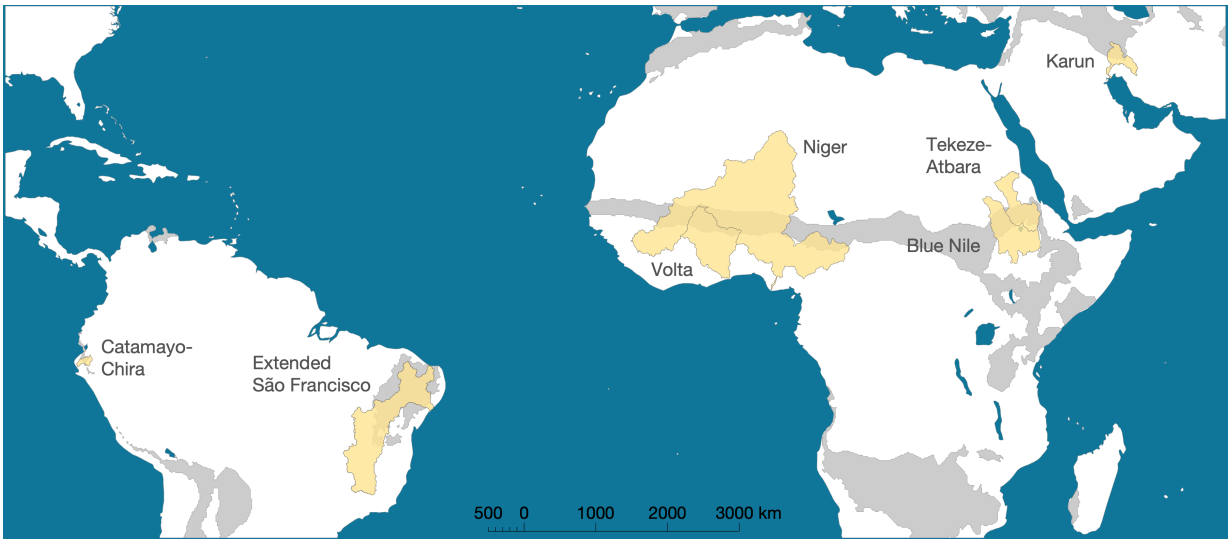


**Figure 1:** Map of the semi-arid study regions: River basins of Catamayo-Chira in Ecuador/Peru, Extended São Francisco in Northeast Brazil, Niger and Volta in West Africa, Tekeze-Atbara and Blue Nile in Ethiopia and Sudan, and Karun in Iran. Semi-arid regions according to an aridity index (ratio of precipitation to potential evapotranspiration) between 0.2 and 0.5 are shown in gray (data from Hoogeveen, 2009).

or extraction of fossil groundwater), much can be achieved in semi-arid regions with sustainable and scientifically sound water resources management.

However, in the context of water resources management, semi-arid regions worldwide (Fig. 1) face major challenges: Semi-arid Northeast Brazil with the São Francisco river is just recovering from a prolonged five-year drought between 2012 and 2016 (Marengo et al., 2018; Martins et al., 2018). The region with its wet season during austral summer and autumn experienced large negative rainfall deviations from the long-term climate mean leading to reservoirs at 5 % of its volume capacity, at inactive volume and in fact to several collapsed (empty) water reservoirs. Even an interruption of water supply occurred in many municipalities of Northeast Brazil (Martins et al., 2018). Climate change is further projected to decrease the rainfall in Northeast Brazil up to 40 % (Marengo et al., 2012). Moreover, the Caatinga biome in Northeast Brazil, the largest tropical dry forest region in South America, is expected to experience enhanced drought and desertification due to climate change, with adverse impacts especially on ecosystems and future water supply (Torres et al., 2017).

The arid to semi-arid coastal regions of Ecuador and northern Peru including the Catamayo-Chira river basin often experience intense precipitation and flooding during El Niño events - the warm phase of the El Niño-Southern Oscillation (ENSO) associated with an unusually warm band of ocean waters extending from the central equatorial Pacific to off the coast of Ecuador and Peru (Takahashi, 2004). Most recently, the 2017 coastal El Niño off Ecuador and Northern Peru caused immense flooding and land-

slides (Echevin et al., 2018). With substantial variability of rainfall from year to year, this semi-arid region not only experiences wet extreme events, but also severe droughts (Vicente-Serrano et al., 2017; Domínguez-Castro et al., 2018). With climate change, weather patterns are expected to change, increasing the blocking at the mountains (with average elevations above 4000 m. a.s.l and peaks above 6000 m) and intensifying the rainfall up to 30 % in coastal Ecuador and Northern Peru (Marengo et al., 2012).

In southwestern Iran, bordering Iraq at the Persian Gulf, the semi-arid province of Khuzestan faces serious challenges in adequately managing its available water resources. The densely populated region of Khuzestan, that includes the Karun river and a large portion of Iran's irrigated croplands, has experienced major water withdrawal and substantial population growth in the last decades (Ashraf et al., 2019). In Khuzestan, the recent years 2017 to 2019 were characterized by a rapid drought–flood transition. The impacts of such hydrometeorological extreme events are known to be particularly severe when they occur in close succession (e.g., Brunner et al., 2021). Large evacuations and inundations were caused by broken flood protectors during the big flood in 2019 (Peyravi et al., 2019). In general, water management in Iran is a critical topic. The country faces the need caused by international sanctions to become self-sufficient in the production of strategic agricultural crops. This situation together with a fatally low irrigation efficiency causes the unsustainable use of water resources (Ashraf et al., 2019). Population growth and enhanced agricultural and socio-economic activities aggravate the water withdrawal in Iran and exacerbate the water stress. The projected variability and change in precipitation will further increase the level of water stress (Ashraf et al., 2019). Cities in Khuzestan are also projected to experience the largest increases in thermal stress among Iranian cities in the future (Roshan et al., 2020). Though being expected to have a minor decline in water storage towards the end of the 21st century compared to, e.g., the Lake Urmia basin in Northwest Iran, the aggressive withdrawal of water - currently already 50 % of Karun's net outflow - will aggravate the water stress in the Karun basin (Ashraf et al., 2019).

In Sub-Saharan Africa, droughts occur more frequently than in most other regions in the world posing a major challenge to the agricultural production that continues to be mainly rain-fed (OECD/FAO, 2016; Masih et al., 2014). Drought hazards are further projected to aggravate across Africa during the 21st century, in particular in Sudan (Ahmadalipour et al., 2019). Especially the semi-arid belt of the Sahel region includes large transboundary river basins like the Niger and Volta in West Africa and the Nile with its main tributaries Blue Nile and Tekeze-Atbara in Ethiopia and Sudan. In transboundary river basins, political consensus in the operation of reservoirs would be required for fair and sustainable water resource management. This is particularly true given the increasing number of people who will depend on upstream water resources by 2050 (Munia et al., 2020; Viviroli et al., 2020). However, the current example of the construction and filling of the Grand Ethiopian Renaissance Dam (GERD) just upstream of the Ethiopian-Sudanese border in the Blue Nile basin demonstrates the difficulties of political consensus between the three involved countries Ethiopia, Sudan and Egypt. Still

agreements are missing, e.g., on how the chain of dams along the Blue Nile and Nile including the two world's largest dams - the High Aswan Dam and the new GERD - will operate in times of water scarcity (Wheeler et al., 2020).

## 1.3 Global Information for Improved Water Resources Management: Seasonal Predictions

All these examples - from increasing drought risks and precipitation variability due to climate change, to being at the mercy of the globally recurring weather and ocean phenomenon ENSO, to politically induced unsustainable use of water resources and transboundary water management conflicts - demonstrate that sustainable water management in these semi-arid regions requires transparent decisions based on evidence. Especially in the case of the targeted fair transboundary water management, publicly available hydrometeorological data are crucial to find viable solutions for all stakeholders involved. The need for high-quality precipitation and water resources information at the present time and for the future contrasts with the decline in in-situ monitoring stations observed worldwide and especially in the presented semi-arid regions (Lorenz and Kunstmann, 2012; Lorenz et al., 2014). Therefore, global gridded datasets from remote-sensing and model systems need to be used in those regions. For managing water resources in advance, such as for water allocation and prioritization, as well as for anticipated demands during drought or for the operation of reservoirs for flood control, improved knowledge of the upcoming rainy season in semi-arid regions can be the basis for sustainable water management. While climate research has been trying for many years to derive statements about climate conditions and long-term mean water availability (e.g., regional climate projections up to 2100), in practice seasonal forecasts, i.e. expected weather conditions up to 7 months in advance, can offer greater opportunities with respect to climate change adaptation and mitigation (Troccoli, 2010). Global seasonal forecasts of hydrometeorological variables can thus become a strong tool for decision support in regional water management when proofed to be skillful.

Seasonal forecasting in the form of statistical models has a long tradition since the late 1800s, but the use of complex dynamical numerical models to simulate the major Earth system components is a fairly new endeavor due to its high computational demand (Troccoli, 2010). Dynamical seasonal forecasts owe their development mainly to the great technological progress in weather forecasting (e.g., Bauer et al., 2015) with known relevance of tropical atmospheric dynamics and model initialization. Since their first operational runs in the mid-1990s, 13 Global Producing Centres (GPCs) of Long-Range Forecasts as accredited by the World Meteorological Organization (WMO) are now producing global seasonal forecasts operationally and with well-defined standards (World Meteorological Organization (WMO), 2021). Dynamical seasonal forecasts are based on the existence of interactions between the atmosphere and slowly changing land surface variables, such as sea surface temperature (SST), sea ice cover or soil moisture (i.e., drivers and feedbacks of longer duration extreme events). These slow variations due to

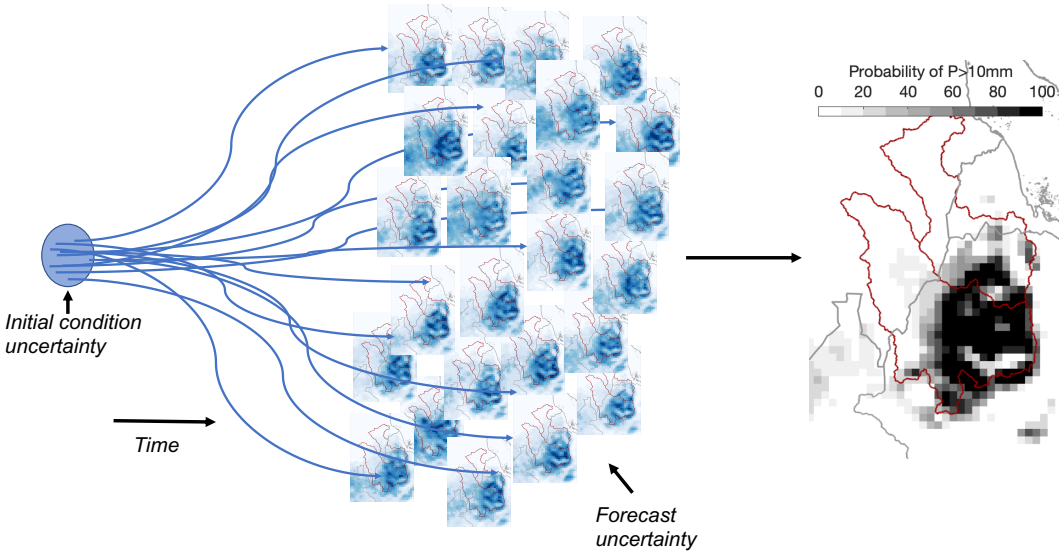


**Figure 2:** The uncertainty of the initial conditions results in ensemble predictions. Their divergence determines the uncertainty of the forecast. Predictions, e.g., of precipitation, are therefore always expressed in terms of probabilities (e.g., the probability of exceeding 10 mm $d^{-1}$ in the Tekeze-Atbara and Blue Nile region in Sudan and Ethiopia). Figure motivated from Bauer et al. (2015).

large adjustment times and heat capacities compared to the atmosphere allow the extension of the forecast horizon well beyond usual weather predictions (Troccoli, 2010). Therefore, the model atmosphere is tightly coupled to the land component, and the relationships between SST or sea ice extent and weather patterns are represented by the coupling of state-of-the-art computer models of the atmosphere, ocean and sea ice (Johnson et al., 2019; Saha et al., 2014). Especially in regions that experience strong links between SST and seasonal weather trends, like the tropics, seasonal forecasting is most promising. As the major climate signal after the seasonal cycle that impacts global climate, the most prominent atmosphere-ocean connection ENSO can be said to be the dominant driver of seasonal forecasts (Troccoli, 2010). Thus, for successful seasonal forecasting, a good representation of such large-scale and long-term phenomena in a seasonal forecasting model is key (Troccoli, 2010; Johnson et al., 2019). Although strongest links to variations in SST exist in the tropics, lower-latitude regions can likewise benefit from seasonal forecasts due to teleconnections and effects of sea ice extent and soil moisture inertia (Johnson et al., 2019). Also other signals play a role for seasonal forecasting in different regions worldwide (Johnson et al., 2019), such as the North Atlantic Oscillation (NAO, e.g., Lamb and Peppler, 1987), the Pacific-North American pattern (PNA, e.g., Frankignoul and Sennéchael, 2007) and the Indian Ocean Dipole (IOD, e.g., Luo et al., 2008). Further possible sources of seasonal predictability are dynamical processes in the stratosphere (Johnson et al., 2019). The stratosphere is not only the mediator of

teleconnections from the tropical oceans to the mid-latitudes (e.g., Bell et al., 2009), the quasi-biennal oscillation (QBO) of the tropical stratosphere may also represent one of the few purely atmospheric sources of seasonal predictability (e.g., Folland et al., 2012). To represent the model and initial condition uncertainty, seasonal forecast models are run several times for the same period of time, using slightly different atmospheric and oceanic fields (Johnson et al., 2019). Therewith, an ensemble of seasonal forecast runs predicts the future weather conditions up to seven months ahead. Seasonal forecasts are therefore not deterministic, but always stated in terms of probability (ensemble forecast with resulting probability in Fig. 2). Detailed day-to-day variations in weather are certainly not possible to predict on seasonal timescales, however, seasonal forecasts are able to give information on likely conditions averaged over a month or the next few months ahead. As the initial-value predictability decreases and the ensemble spread often increases with lead time (increasing forecast horizon), seasonal forecast skill is often higher for a few months ahead than for long forecast horizons.

## 1.4 From Global to Regional Information: Spatial Refinement and Correction of Model Errors

Progress in seasonal forecasting systems, with better representation of large-scale and long-term phenomena such as El Niño and corresponding gains in performance, particularly for affected regions (e.g., Johnson et al., 2019), is increasingly driving their application. Advances in computing weather models and substantial increases in data storage capacity have pushed us to a level where reasonably accurate forecasts and data analysis are possible for extended forecast horizons (e.g., Bauer et al., 2015; Troccoli, 2010). The spatial resolution of global seasonal forecast models, such as the latest seasonal forecast system version 5 (SEAS5) of the European Centre for Medium-Range Weather Forecasts (ECMWF) of 36 km (Johnson et al., 2019), already extends into areas formerly covered only by regional models. So it appears appropriate to use them for regional applications in semi-arid regions. In addition, the raising awareness of probability forecasts in the public helps for the dissemination of seasonal forecasting products, but suitable measures still need to be provided to translate probabilistic knowledge about the upcoming season into supportive action (Bruno Soares et al., 2018). Model biases and drifts further need to be taken into account to increase the credibility and legitimacy of the generated seasonal forecast information.

In general, weather forecasting models can exhibit large biases, i.e., large mean forecast errors, in some regions and may still fail in simulating important dynamical patterns, circulation and energy exchanges (e.g., Sillmann et al., 2017; Manzanas, 2020). Forecast model biases can arise from incomplete or imperfect representation of the current environmental state (assimilation of observations from satellites, balloons, ground stations, etc.) in the initial conditions due to misfits to the model grids or biases in the observations themselves (e.g., Privé and Errico, 2013). The resolution of the model grid can also lead to biases caused by insufficient representation of the underlying topography

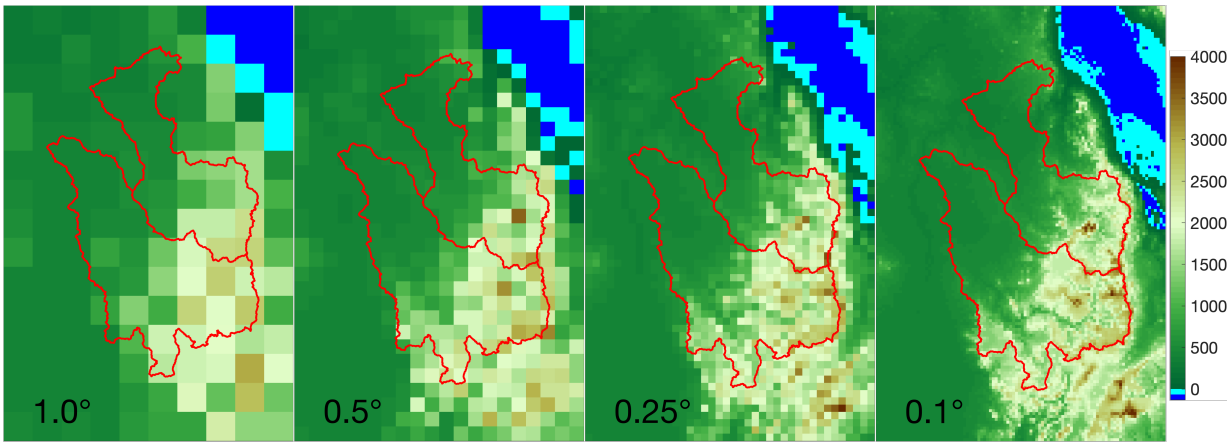


**Figure 3:** Representation of topography at different model grid resolutions for the Tekeze-Atbara and Blue Nile basins in Sudan and Ethiopia.

(see representation of topographic features at different resolutions in Fig. 3). Spatial refinement approaches can therefore help to improve the model information regionally. This so called “downscaling" can be achieved with empirical statistical methods or regional climate modeling. Statistical downscaling methods use higher-resolution observational data to improve the climatological spatial and distributional patterns of the forecast (Yuan et al., 2015). Dynamical downscaling, in turn, embeds (“nests") a regional climate model into the coarser global grid, simulating regional to local climate responses at a higher spatial (and, if needed, temporal) resolution (e.g., Yuan et al., 2015; Skamarock et al., 2008). Here, another possible reason for (global and regional) model biases comes into play that can be introduced by the employed physical parameterizations, i.e., approximations of the physical processes at the subgrid scale that incorporate certain assumptions. Parameterizations are used for processes related to, e.g., incoming solar radiation (absorption, reflection, scattering), longwave radiation emission (clouds, earth’s surface), latent and sensible heat fluxes, water phase changes, deep convection, turbulence, vegetation, and soil properties (e.g., Skamarock et al., 2008). Here, the choice of different parameterization schemes can cause large changes in the atmospheric dynamics (e.g., Klein et al., 2015) and thus large model biases. On top of the standard bias, or systematic mean error, a second-order bias occurs in seasonal forecast models that depends on lead time - the time elapsing from the model initialization to the start of the target period to be forecast (Manzanas, 2020). This is called model drift and is a consequence of initial conditions that do not match the model dynamics (e.g., Alves et al., 2004). The forecast model drifts away from the initial state well beyond the first lead month, leading to highly spatially and temporally variable biases over large regions of the world (Manzanas, 2020; Hermanson et al., 2018). From their initialization, models often do not simply drift monotonically in the direction of the free-running model climatology; for example, in the case of precipitation, asymptotic drifts frequently occur (Hermanson et al., 2018). Correction for these biases and drifts is therefore required

when absolute thresholds are to be derived from seasonal forecasts or when they are applied as input in subsequent impact modeling.

## 1.5 Research Questions

With the use of seasonal forecasts, the availability of freshwater resources and periods of drought can be predicted several months in advance. This opens up new opportunities to actively reinforce the sustainable water management in semi-arid regions. To achieve this, the understanding of seasonal forecasts needs to be improved and the advantages demonstrated in a decision-making process (e.g., Bruno Soares et al., 2018). While scientists typically focus on the technical skill (e.g., accuracy, sharpness) of a prediction, decision-makers see the value of forecasts as the more important measure, especially when the forecast skill is not clearly attributable to decision consequences, or the skill is not made sufficiently clear. Assessments of forecast performance should therefore incorporate the perspective of "users" or "decision-makers" and not just that of scientists (Bruno Soares et al., 2018; Hartmann et al., 2002; Hansen, 2002). Thus for users, it is not only the quality of the forecasts that is critical, but the potential value and benefit of incorporating seasonal forecasts. Therefore, to promote the use of seasonal forecasts in decision making, the first research question to be answered is:

1. Can seasonal forecasts support decision-making and provide economic benefit for the regional water management in semi-arid regions?

Although the spatial resolution of 36 km of ECMWF's latest global seasonal forecasting product SEAS5 is already at the scale of regional applications, further refinement of the horizontal resolution is required for grid-point based regional applications and for impact modeling of, e.g., hydrology, ecosystem, or sediment. Since this is a fairly common problem in climate science (e.g., Yuan et al., 2015), there are numerous "downscaling" methods to bridge this scale gap. For this reason, two different downscaling methods are developed and applied for the study regions to spatially refine the most important hydrometeorological variables of global fields. In this regard, corrections of model biases and drifts should also be incorporated. The second research question addressed in this thesis therefore is:

2. How can empirical-statistical regionalization and post-processing techniques improve the regional quality of global seasonal forecasts?

Apart from empirical statistical methods, the application of dynamical regional climate models (RCMs) allows the refinement of global hydrometeorological fields. The underlying idea here is that global models are capable of correctly capturing the large-scale forcing signals (e.g., ENSO), and RCMs, in turn, are able to correctly reproduce the regional-local climate response (e.g., Yuan and Liang, 2011; Yuan et al., 2015). Especially for orographically complex regions, physically-based refinement methods can improve their spatio-temporal patterns. Atmospheric physics like shallow convection or land

surface processes like frozen soil or snow can be better represented, and therewith reduce the forecast errors at daily-to-seasonal scales (e.g., Yuan and Liang, 2011). For that, however, RCMs need to be adapted and calibrated in their physical setup to the selected region prior to their use for regional weather forecasting or impact studies. With an appropriate choice of physics parameterization schemes, model biases can be reduced substantially (Klein et al., 2015). This leads to the third research question:

3. How robust are dynamically downscaled hydrometeorological fields with respect to the applied physical model setup?

This sensitivity analysis of the physical setup of an RCM involves an immense computational effort. Only if this computational battle indeed results in a substantial improvement of the spatio-temporal patterns, a dynamical downscaling of the seasonal forecasts can be justified. In the synthesis, these two provided options of regional refinement will thus be discussed in terms of operational feasibility and computational demand. Therewith the question is addressed if dynamical downscaling is suited for the regional refinement of global seasonal forecasts from both a scientific point of view and in an operational decision making environment.

The goals of this thesis therewith are i) to assess the performance in quality and value of global seasonal forecasts for decision support in regional water management of semi-arid regions and ii) to investigate the suitability and feasibility of methods for regional refinement and improvement of the global seasonal forecasts. The overall strength of this work is the development of methods and tools that can be applied in very different regions together with region-specific tailoring. The opportunities and limitations of direct application of raw global seasonal forecasts are presented together with the necessity of their regional refinement and correction. In the latter, not only the consideration of forecast operationalization in a decision support system played a role, but also the presentation of different, potentially skillful methods for regional refinement and improvement of seasonal forecasts, and the testing of horizontal resolutions achievable with those techniques.

All five semi-arid study regions (Fig. 1) were selected for the global approaches with raw seasonal forecasts:

1. Brazil - Extended Rio São Francisco catchment
2. Ecuador/Peru - Catamayo-Chira catchment
3. Iran - Karun catchment
4. Sudan/Ethiopia - Tekeze-Atbara and Blue Nile catchments
5. West Africa - Niger and Volta catchments

For four of them (1-4), the regional refinement and improvement via empirical statistical downscaling and spatial disaggregation is performed. For two of these regions (1,2), global hydrometeorological fields are refined via dynamical downscaling.

## 1.6 Structure and Framework of the Thesis: Linking the Individual Pieces

To answer the above mentioned research questions, this dissertation comprises three peer-reviewed articles:

**Article I:** *Seasonal Forecasts offer Economic Benefit for Hydrological Decision Making in Semi-Arid Regions*

Portele, T.C., Lorenz, C., Dibrani, B., Laux, P., Bliefernicht, J., and Kunstmann, H (2021). Seasonal Forecasts Offer Economic Benefit for Hydrological Decision Making in Semi-Arid Regions. *Sci. Rep.* **11**, 10581. https://doi.org/10.1038/s41598-021-89564-y

The first article addresses the potential benefits of incorporating seasonal forecasts in water management decisions to proactively prepare for drought and extreme events and to mitigate risk. For analyzing the development of past climatic extreme events and, hence, underlining the urgent need for a timely and sustainable water management, first the trend in the occurrence of drought, very warm months, and very wet months is examined in the seven basins of the five semi-arid study regions over the past decades with the latest ECMWF reanalysis ERA5. Here, seasonal climate forecasts can be a strong tool in the management of climate risks. However, they still need to overcome the perceived lack of reliability and skill, and to face the difficulties in linking their probabilistic nature to practical decision processes (Bruno Soares et al., 2018; Lopez and Haines, 2017; Rayner et al., 2005; Patt and Gwata, 2002). Classical approaches of performance evaluation fail to show the direct effect on decision makers and users of seasonal climate forecasts. Therefore, the performance measure of potential economic value ($PEV$, Richardson, 2000; Wilks, 2001; Lopez et al., 2018) is chosen to directly link ECMWF's latest global probabilistic seasonal forecasts SEAS5 to users' economic benefits when they incorporate them into their decision-making process regarding extreme events. The potential economic value includes the forecast probability and the cost-loss ratio of a user's early action. Two approaches are provided to decide on the most beneficial probability threshold at which to take an action. One approach aims to minimize expense, while the other additionally seeks to capture as many extreme events with the forecasts as possible. The employed methodology in this article allows the use of unprocessed global seasonal forecasts for the regional analysis. This approach also allows the raw forecast to be evaluated without the need for explicit bias correction. The final example case at the Upper-Atbara Dam in Sudan provides an in-depth evaluation of the economic cost-loss situation with respect to an early-action based drought reservoir operation for electrical power generation. The article *Seasonal Forecasts offer Economic Benefit for Hydrological Decision Making in Semi-Arid Regions* thus serves as a general introduction to seasonal forecasting as decision support in semi-arid regions.

**Article II:** *Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-term Reference Forecast Product for the Water Sector in Semi-arid Regions*

Lorenz, C., Portele, T.C., Laux, P., and Kunstmann, H (2021). Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-term Reference Forecast Product for the Water Sector in Semi-arid Regions. *Earth Syst. Sci. Data* **13**, 2701–2722. https://doi.org/10.5194/essd-13-2701-2021

While the first article assesses the (economic) value of the forecasts, using relative data values and treating the distributions of forecast and reference data separately, the second article takes a closer look at the quality of the SEAS5 global seasonal forecasts in terms of absolute values for four study regions. The analyzed hydrometeorological variables are: Precipitation, incoming shortwave radiation, minimum, maximum and mean 2-m temperature. To minimize systematic biases and model drifts in the seasonal forecasts with respect to the reference ERA5-Land (a higher resolution replay of the land component of ECMWF's latest reanalysis ERA5) and to regionally refine the global fields to be usable in impact models, empirical statistical bias correction and spatial disaggregation (BCSD) methods (Wood, 2002) are developed and applied. Here, distribution-function based correction methods (quantile-quantile mapping) are used on a daily basis per gridcell to correct for systematic biases and model drifts with increasing lead time. The spatial disaggregation to the resolution of the reference ERA5-Land allows the regional refinement of global seasonal forecasts from 36 km to 0.1°. With the correction towards the reference ERA5-Land, the wet-day-probability, as well as topographical effects are further controlled. The quality of the SEAS5-BCSD forecasts is compared to the reference ERA5-Land and to raw SEAS5 forecasts, and is measured in terms of bias, root-mean-square error and the continuous ranked probability skill score. An entire (re-)forecast period of corrected and refined SEAS5-BCSD predictions from 1981 to 2019 is produced as driving data for hydrological, ecosystem or climate impact models. This work thus represents a natural continuation of the first article, providing corrected and regionally refined seasonal forecasts when absolute values and high-resolution gridcell information are needed.

**Article III:** *Ensemble-tailored Pattern Analysis of High-resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America*

Portele, T.C., Laux, P., Lorenz, C., Janner, A., Horna, N., Fersch, B., Iza, M., and Kunstmann, H., (2021). Ensemble-Tailored Pattern Analysis of High-Resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America. *Front. Earth Sci.* **9**, 669427. https://doi.org/10.3389/feart.2021.669427

The third article focuses on the dynamical downscaling of global hydrometeorological fields with the Weather Research and Forecast (WRF) model. For the two South American study regions, atmospheric global fields from ECMWF's former reanalysis product ERA-Interim at 79 km horizontal resolution are regionally refined to 9 km and 3 km horizontal resolution. The 3-km resolution was applied to the topographically highly complex Ecuador/Northern Peru region, including the coastal area and the Andes. In this article, the sensitivity of distributional, temporal and spatial precipitation patterns on the physical model setup is analyzed with in total 32 ensemble simulations over two full, consecutive years. The focus of this computationally demanding experiment is the investigation of ensemble effects formed by simulations that all employ one specific physics parameterization scheme but differ in other schemes (parameterization sub--ensembles). The performance of these parameterization sub-ensembles is evaluated with three different global gridded data sets that are available at the high resolution of the simulations (CHIRPS, MSWEP and ERA5-Land). For the 3-km domain, also local station data is considered. The inclusion of differently represented physical processes in the tested parameterization schemes should demonstrate which level of complexity of the model physics is required to best reproduce the reference data. The main motivation for this article is to present an alternative approach to regional refinement of global fields using a physics-based atmospheric model, and to investigate whether this is capable of substantially improving and approximating precipitation patterns relative to the reference at high resolutions.

Figure 4 provides the link between each of these research articles by illustrating the common framework of this dissertation, as well as the similarities and differences in target regions, used datasets, performance metrics and methods. Finally, the synthesis of the three pieces of this work is found in the key achievements.

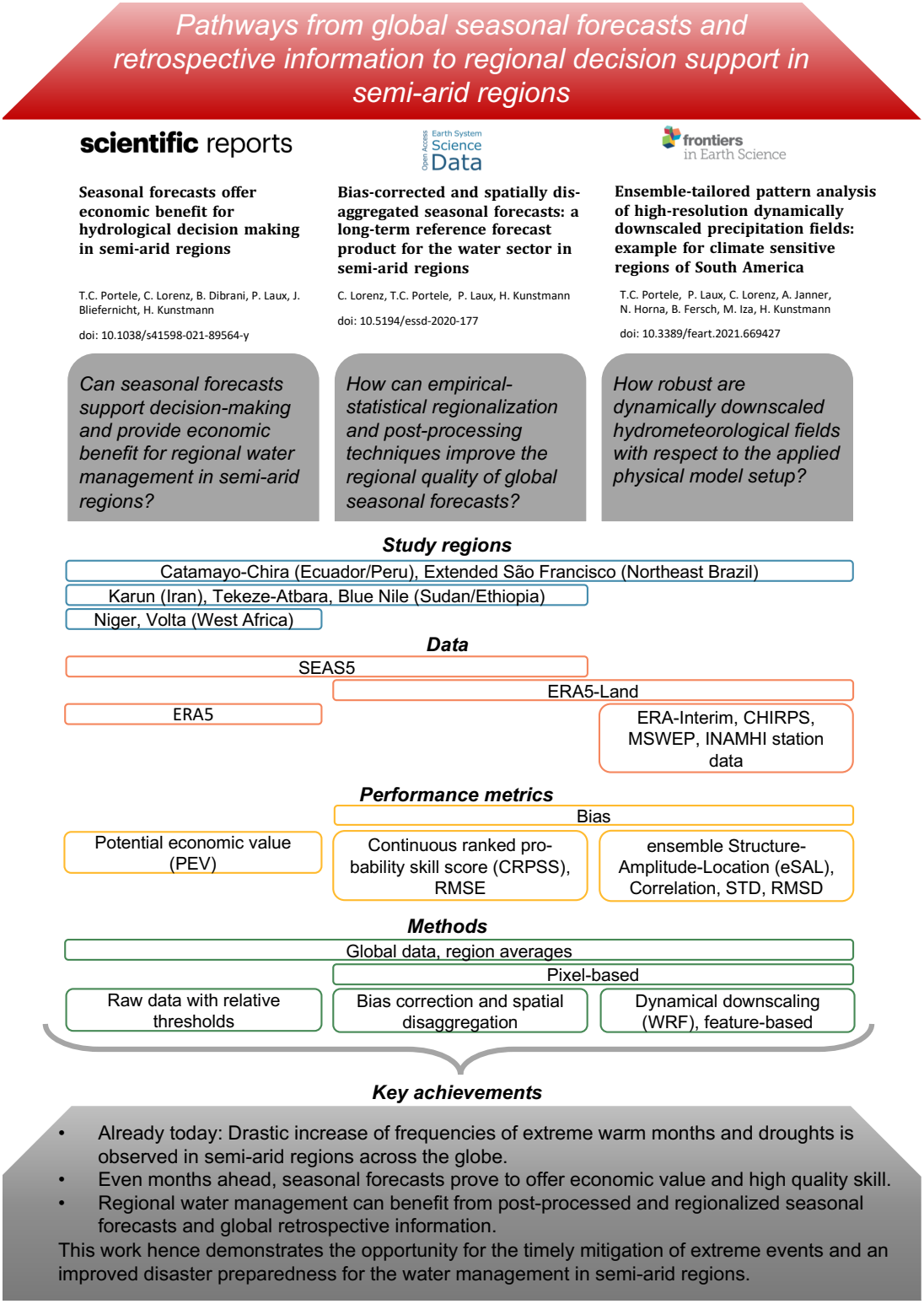


**Figure 4:** Framework of and links between the individual research articles of this thesis. Common features are represented with boxes spanning multiple columns.

## 1.7 Innovation of the Thesis

The essence of this dissertation consists of:

- Innovative, ensemble-tailored, user-oriented assessment of the performance of global and regionalized hydrometeorological fields in terms of forecast economic value and quality, as well as distributional, temporal and spatial patterns
- Development of tools and methods that can be applied in different regions worldwide, together with the use of global raw data and two very different regionalization methods
- Provision of scientifically sound, concrete decision support for regional water resources management through the estimation of economic savings from probability-triggered seasonal forecast-based action and operational implementation of a seasonal decision support system for end users
- Provision of a real-world case study at a typical, representative dam in semi-arid regions to demonstrate the final transfer into practice.

## 1.8 Key Achievements

This work demonstrates the pathways from global seasonal forecasts and retrospective information to regional decision support in the water resources management of semi-arid regions. It provides evidence that:

- Already today, the expected increase in the frequency of extreme warm periods and droughts due to climate change can be observed in semi-arid study regions.
- Especially for those extreme events, long-term seasonal forecasts offer high economic value for proactive water management decisions even for several months ahead (Fig. 5).
- Proactive reservoir management in the real-case study at a typical representative of a large dam in semi-arid regions achieves avoidable losses up to 16 Mio US$ for one example drought year.
- Empirical-statistical post-processing methods substantially increase the quality and accuracy of seasonal forecasts (Fig. 6) required for decisions based on absolute cutoff values and for impact studies.
- Lead-dependent bias correction and spatial refinement improves the consistency among different forecast horizons (Fig. 6) and increases the consistency with high-resolution reference data.
- Concrete decision support and important steps towards explicit transfer into practice are achieved by operationalizing and visualizing the post-processed seasonal forecasts in an online decision support tool.
- Independent of reference data, dynamical downscaling of hydrometeorological retrospective information can prove beneficial for very complex regions with strong climatic and topographic gradients in terms of convective or orographically-in-

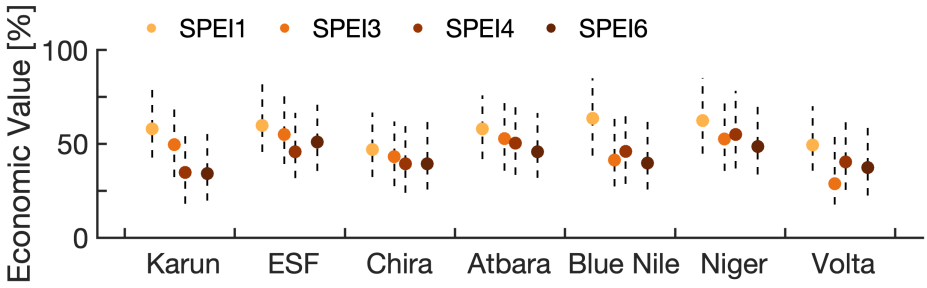


**Figure 5:** Maximum potential economic value ($PEV_{max}$) for forecast-based early action of drought (standardized potential evaporation index, SPEI $< -1$) over forecast ranges of 1 (SPEI1), 3 (SPEI3), 4 (SPEI4), and 6 (SPEI6) months. $PEV_{max}$ of SEAS5 with reference ERA5 is shown for the seven river basins for rainy seasons of the hindcast period 1981-2016. Colored dots represent the fully evaluated $PEV_{max}$ for the events of SPEI $< -1$. $PEV_{max}$ is further assessed with the confidence range between the 10% and 90% quantiles of bootstrap resampling (dashed line). (Figure adapted from Portele et al., 2021)

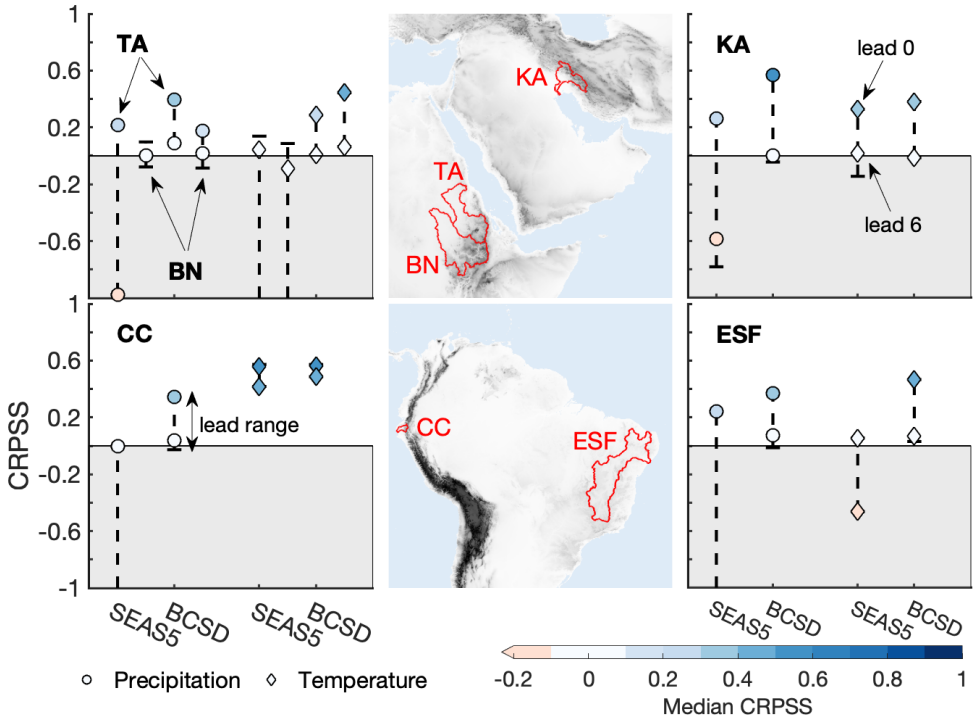


**Figure 6:** Overall performance evaluated by the Continuous Ranked Probability Skill Score (CRPSS) for precipitation and temperature forecasts of one month during the basins' rainy season of SEAS5 raw forecasts and bias-corrected and spatially disaggregated (BSCD) SEAS5. CRPSS is shown for lead 0 to lead 6 forecasts for July for the basins of Tekeze-Atbara/Blue Nile (TABN), and for February for the Karun (KA), Catamayo-Chira (CC) and Extended São Francisco (ESF) basins. A CRPSS $> 0$ defines an on median (1981-2016) better seasonal forecast than a simple climatological forecast.

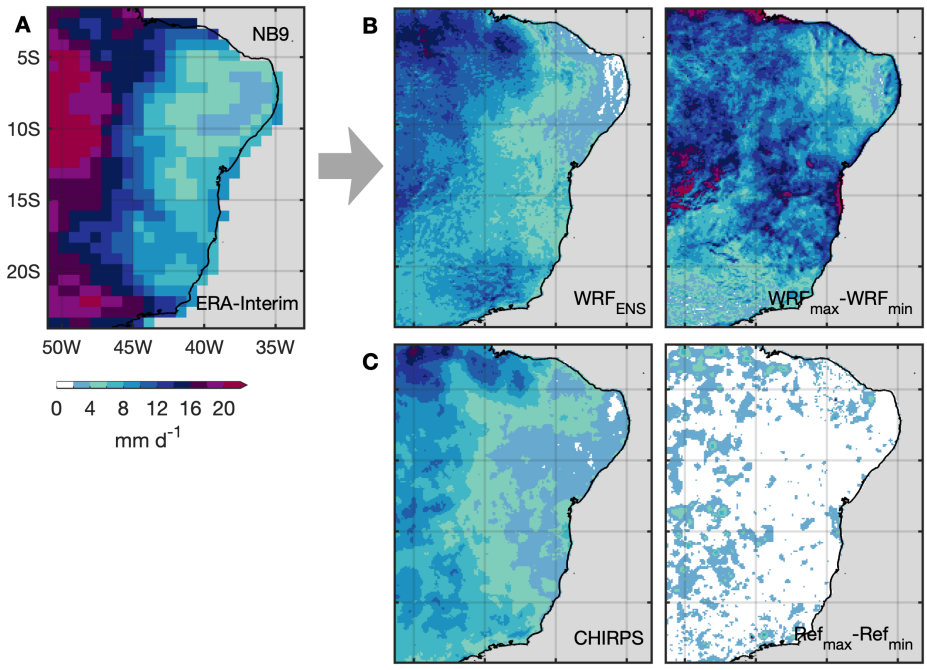


**Figure 7:** WRF dynamical downscaling of ERA-Interim (79 km) to a resolution of 9 km for the Northeast-Brazil domain (NB9). A) ERA-Interim mean precipitation for JFM 2007. B) WRF ensemble mean ($WRF_{ENS}$) precipitation (left) and WRF ensemble uncertainty expressed by the range (maximum WRF precipitation minus minimum WRF precipitation) of the individual WRF runs (right) for JFM 2007. C) CHIRPS reference precipitation (left) and reference uncertainty expressed by the range of reference precipitation from CHIRPS, MSWEP and ERA5-Land (right) for JFM 2007.

duced precipitation, where a high resolution is required to represent regional to local small-scale phenomena.

- The quality of dynamically refined information, however, highly depends on the applied physical model setup, where precipitation amounts can be up to four times the observed reference values and a high uncertainty of dynamically refined information is introduced (Fig. 7).
- As a consequence, no general best model setup could be identified, as different setups prove favorable for different applications.
- With an appropriate choice of physical setup, local meteorological services can benefit from the extensive model experiment conducted by using the setup to monitor hydrometeorological conditions in the regions.
- Overall, integrating global seasonal forecasts and retrospective information into decision making hence creates beneficial opportunities for the timely mitigation of extreme events and disaster preparedness of regional water management in semi-arid regions.

## 1.9 Contribution of the Author to the Different Articles

**Article I:** *Seasonal Forecasts offer Economic Benefit for Hydrological Decision Making in Semi-Arid Regions*

Portele, T.C., Lorenz, C., Dibrani, B., Laux, P., Bliefernicht, J., and Kunstmann, H (2021). Seasonal Forecasts Offer Economic Benefit for Hydrological Decision Making in Semi-Arid Regions. *Sci. Rep.* **11**, 10581. https://doi.org/10.1038/s41598-021-89564-y

The research concept was developed by Tanja Portele, Christof Lorenz and Harald Kunstmann. Tanja Portele and Christof Lorenz conceived and designed the analysis. The application of user-oriented forecast verification through potential economic value analysis for seasonal forecasts was designed and implemented by Tanja Portele. Christof Lorenz managed and provided the used data of seasonal forecasts SEAS5 and reanalyses ERA5. All computations regarding the seasonal forecast value, the final analysis, summary and interpretation of the results were conducted by Tanja Portele. Berhon Dibrani performed the calculations of reservoir management and energy production. All figures, except for the background map of study regions provided by Christof Lorenz, were prepared by Tanja Portele. The article was mainly written by Tanja Portele with contributions from and discussion with all coauthors.

**Article II:** *Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-term Reference Forecast Product for the Water Sector in Semi-arid Regions*

Lorenz, C., Portele, T.C., Laux, P., and Kunstmann, H (2021). Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-term Reference Forecast Product for the Water Sector in Semi-arid Regions. *Earth Syst. Sci. Data* **13**, 2701–2722. https://doi.org/10.5194/essd-13-2701-2021

The article was mainly written and designed by Christof Lorenz and Tanja Portele. Christof Lorenz collected and processed the seasonal forecasts, and developed and applied the bias-correction and spatial diaggregation algorithm. The search for a suitable reference dataset was a lengthy preparatory task carried out by Christof Lorenz and Tanja Portele. It involved testing several datasets for different variables in comparison with sparse station data and constructing a first spatially disaggregated reanalysis dataset from ERA5 before the higher resolution ERA5-Land became available. With the release of ERA5-Land with a horizontal resolution of 0.1°, the preliminary bias correction tasks were concretized using the new reference data set, and the final seasonal forecast product was computed by Christof Lorenz. The evaluation of the SEAS5 BCSD forecasts was conducted by Christof Lorenz and Tanja Portele. In particular, the performance analysis of raw and post-processed forecasts with the Continuous Ranked Probability Skill Score (CRPSS) was performed by Tanja Portele. She further provided the figure of CRPSS and computed the climatological characteristics of the study regions of ERA5-Land. Christof Lorenz initiated the data dissemination via DKRZ WDCC and the KIT Campus Alpin THREDDS Server and implemented the operationalization of

the presented approach. Christof Lorenz, Tanja Portele, Patrick Laux and Harald Kunstmann reviewed the paper and prepared the final version of the manuscript.

**Article III:** *Ensemble-tailored Pattern Analysis of High-resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America*

Portele, T.C., Laux, P., Lorenz, C., Janner, A., Horna, N., Fersch, B., Iza, M., and Kunstmann, H., (2021). Ensemble-Tailored Pattern Analysis of High-Resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America. *Front. Earth Sci.* **9**, 669427. https://doi.org/10.3389/feart.2021.669427

Tanja Portele, Patrick Laux, Christof Lorenz and Harald Kunstmann developed the research concept of this article. Tanja Portele performed the WRF model simulations and conceived the analysis of distributional, temporal and spatial patterns. In particular, the application and implementation of the ensemble structure-amplitude-location (eSAL) analysis was designed and conducted by Tanja Portele. Christof Lorenz managed and provided the used global reference data CHIRPS, MSWEP and ERA5-Land. Natalia Horna and Maylee Iza managed and provided the used Ecuadorian station reference data. Tanja Portele performed the analysis and prepared the figures. She was assisted by Annelie Janner in the temporal pattern analysis. Tanja Portele summarized and interpreted the results. The article was written primarily by Tanja Portele, with inputs from and discussion with all coauthors.

## 1.10 Further Related Article

Besides the above mentioned three articles, the author of this thesis contributed to:

Laux, P., Dieng, D., Portele, T.C., Wei, J., Shang, S., Zhang, Z., Arnault, J., Lorenz, and Kunstmann, H., (2021). A High-resolution Regional Climate Simulation Physics Parameter Ensemble for Sub-Saharan Africa. *Front. Earth Sci.* **9**, 792. https://doi.org/10.3389/feart.2021.700249

*Abstract:* "While climate information from General Circulation Models (GCMs) are usually too coarse for climate impact modelers or decision makers from various disciplines (e.g., hydrology, agriculture), Regional Climate Models (RCMs) provide feasible solutions for downscaling GCM output to finer spatiotemporal scales. However, it is well known that the model performance depends largely on the choice of the physical parameterization schemes, but optimal configurations may vary e.g., from region to region. Besides land-surface processes, the most crucial processes to be parameterized in RCMs include radiation (RA), cumulus convection (CU), cloud microphysics (MP), and planetary boundary layer (PBL), partly with complex interactions. Before conducting long-term climate simulations, it is therefore indispensable to identify a suitable combination of physics parameterization schemes for these processes. Using the *European Centre for Medium-Range Weather Forecasts* (ECMWF) reanalysis product ERA-Interim as lateral boundary conditions, we derived an ensemble of 16 physics parameterization

runs for a larger domain in Northern sub-Saharan Africa (NSSA), northwards of the equator, using two different CU-, MP-, PBL-, and RA schemes, respectively, using the *Weather Research and Forecasting* (WRF) model for the period 2006–2010 in a horizontal resolution of approximately 9 km. Based on different evaluation strategies including traditional (Taylor diagram, probability densities) and more innovative validation metrics (ensemble structure-amplitude-location (eSAL) analysis, Copula functions) and by means of different observation data for precipitation (P) and temperature (T), the impact of different physics combinations on the representation skill of P and T has been analyzed and discussed in the context of subsequent impact modeling. With the specific experimental setup, we found that the selection of the CU scheme has resulted in the highest impact with respect to the representation of P and T, followed by the RA parameterization scheme. Both, PBL and MP schemes showed much less impact. We conclude that a multi-facet evaluation can finally lead to better choices about good physics scheme combinations." (Laux et al., 2021b).

This article comprises the sensitivity analysis of precipitation and temperature fields on the physical parameterization in the WRF model for Sub-Saharan Africa. Here, a dynamical downscaling is performed to a horizontal resolution of 9 km for a five-year period and including 16 ensemble runs. This article is thus the continuation of the testing of a physics-based regional refinement approach for the West African and Sudanese-Ethiopian study regions. Tanja Portele substantially contributed to the design of the study, applied the ensemble-Structure-Amplitude-Location analysis and contributed to the writing of the article.

# 2 Article I

## Seasonal Forecasts Offer Economic Benefit for Hydrological Decision Making in Semi-arid Regions

---

Portele, T.C., Lorenz, C., Dibrani, B., Laux, P., Bliefernicht, J., and Kunstmann, H (2021). Seasonal Forecasts Offer Economic Benefit for Hydrological Decision Making in Semi-arid Regions. *Sci. Rep.* **11**, 10581. https://doi.org/10.1038/s41598-021-89564-y



---

scientific reports

Check for updates

OPEN

# Seasonal forecasts offer economic benefit for hydrological decision making in semi-arid regions

Tanja C. Portele[1,3]✉, Christof Lorenz[1], Berhon Dibrani[2], Patrick Laux[1,3,4], Jan Bliefernicht[3,4] & Harald Kunstmann[1,3,4]

**Increasing frequencies of droughts require proactive preparedness, particularly in semi-arid regions. As forecasting of such hydrometeorological extremes several months ahead allows for necessary climate proofing, we assess the potential economic value of the seasonal forecasting system SEAS5 for decision making in water management. For seven drought-prone regions analyzed in America, Africa, and Asia, the relative frequency of drought months significantly increased from 10 to 30% between 1981 and 2018. We demonstrate that seasonal forecast-based action for droughts achieves potential economic savings up to 70% of those from optimal early action. For very warm months and droughts, savings of at least 20% occur even for forecast horizons of several months. Our in-depth analysis for the Upper-Atbara dam in Sudan reveals avoidable losses of 16 Mio US$ in one example year for early-action based drought reservoir operation. These findings stress the advantage and necessity of considering seasonal forecasts in hydrological decision making.**

Climate extremes such as droughts and anomalous wet conditions greatly affect economic wealth, particularly in climate sensitive semi-arid regions[1]. 40% of the global land area are drylands, inhabited by more than one third of the global population[2] with population growth rates up to 4% per year[3]. For the majority of drylands, especially for semi-arid regions, sustainable water management can be key with respect to the food-energy-water nexus. More than one third of semi-arid drylands are cultivated rain-fed or irrigated farmland[2]. The importance of hydropower generation for states comprising semi-arid regions can be immense with up to two thirds of a state's total power generated at hydroelectric dams[4]. With increasing frequencies and intensities of droughts and hot extremes, as well as with increasing precipitation variabilities[1,5–8], conflicts of water use and impacts of climate change may become critical, increasing the importance of sustainable water management and climate proofing in semi-arid regions.

This study addresses the crucial question of whether proactive drought- and extreme event preparedness as well as risk mitigation in water management could be provided by the use of seasonal climate forecasts. Decision makers still hesitate to use seasonal forecasts[9,10]. They claim a lack of confidence and credibility due to their probabilistic nature and consequent uncertainties[10,11]. No explicit action plans could have been developed on the basis of those forecasts and the conservative decision making environment further impeded their use[10–12]. The debate about the value of seasonal forecasts already started in the 1970s when the provision of seasonal climate forecasts emerged[13], and still continues, as the optimal use and value of these forecasts is still unclear[9,10,14]. To implement and foster the usage of seasonal forecasts, the concrete economic benefits need to be demonstrated when decisions and actions are forecast-based. Uncertainties have to be accounted for and economically beneficial probability thresholds for decisions need to be developed[15]. This would allow decision makers to understand the potential benefits of using probabilistic seasonal forecasts.

The assessment of the potential economic value (*PEV*)[16–18] is a direct way to achieve action recommendations based on forecast probabilities without having detailed operational information and interaction with a decision maker. It is a superordinate base for decisions that incorporates a neutral decision environment (no risk aversion) with costs and losses that arise from taking forecast-based action, and a valuation of the decisions based on relative economic savings[16]. For water management, the estimation of costs and losses associated with decisions made at reservoirs and the consideration of a decision's economic consequences should be key approaches for dam operations, facilitating access to the economic value approach. Most studies about the economic value

[1]Karlsruhe Institute of Technology (KIT), Campus Alpin, Institute of Meteorology and Climate Research - Atmospheric Environmental Research (IMK-IFU), Garmisch-Partenkirchen, Germany. [2]Tractebel Engineering GmbH, Bad Vilbel, Germany. [3]University of Augsburg, Institute of Geography, Augsburg, Germany. [4]These authors contributed equally: Patrick Laux, Jan Bliefernicht and Harald Kunstmann. ✉email: tanja.portele@kit.edu

of forecasts only involve theoretical examples without quantification of the expected costs and losses, or used the economic value as a further measure to assess the goodness of the seasonal forecasts[18–22]. Monetary values related to forecasts are only provided by few studies, and not in the direct context of the potential economic value assessment[23,24]. Regarding the energy production at reservoirs, costs and losses of dam operations can be specified quite accurately for use in this approach. With energy production at dams playing a major role in many countries[4], a possible optimization or improvement of the dam operations by the usage of seasonal forecasts could provide a huge benefit, including proactive disaster planning.

The arising benefits when basing early action on hydrometeorological seasonal forecasts for extreme events like drought, very warm and very wet months at forecast horizons up to seven months ahead will be demonstrated for seven selected climate-sensitive, semi-arid, and in parts highly-managed river basins including mountainous and flat areas with extents from tens of thousand to millions of square kilometers. These are the river basins of São Francisco with an artificially created extension (ESF) in Northeast Brazil, of Catamayo-Chira (Chira) in South Ecuador and North Peru, of Karun in Southwest Iran, of the Nile tributaries Tekeze-Atbara (Atbara) and Blue Nile in Ethiopia and Sudan, as well as of Niger and Volta in West Africa. Hit by several severe droughts and floods during the last years[7,25–27], all these semi-arid study regions depend on sustainable multipurpose water resources management. Their severe vulnerability to climate variability can be explained by a combination of high shares of population living in the especially dry tailwaters of the basins[28], high dependencies on irrigation or rain-fed agriculture and on hydroelectric power supply by large dams[4], together with a projected aggravation of climate change signals with higher variability of seasonal precipitation[29] and prolonged, more intense droughts with intermittent heavy rainfalls[6,30,31].

For the assessment of economic value, we use the latest seasonal forecast product SEAS5 by the European Centre for Medium-Range Weather Forecasts (ECMWF) at a resolution of 35 km[32]. Providing the main input for water management, we focus on the PEV analysis of precipitation and temperature forecasts of SEAS5. The evaluation method and the selection of analyzed extreme events and evaluated quantities based on commonly used drought indices and separate probability distributions for reference and forecast data allow for the application of raw, uncorrected seasonal forecasts.

With this analysis, we reveal time horizons across which economic benefits exist, and evaluate differences between the considered regions and extreme events. For an in-depth example for the Upper Atbara Dam in Sudan, we calculate the incurred economic consequences of reservoir operations adapted to drought conditions. For water management policies, we further suggest forecast probability thresholds for early action plans based on criteria for minimizing the expenses (expense minimization) and maximizing the number of hit events (event maximization).

## Results

**Increasing frequency of extreme events in semi-arid regions.** All semi-arid study basins in South America (Chira, Extended São Francisco), Africa (Niger, Volta, Atbara and Blue Nile) and West Asia (Karun) are characterized by a distinct rainy and dry season (Fig. 1). Using the reference data of ERA5 reanalyses, peak monthly precipitation ranges between 120 and 300 mm. Monthly temperatures are rather constant, except for the Karun basin with a pronounced annual cycle of about 30 °C. For consistency in the following analysis, the rainy season is defined as the main four months of the rainy season, plus the preceding and succeeding month (Fig. 1).

For the occurrence of extreme events from 1981-2018, strong signals of climate change are observed across the study regions (Fig. 2). On average, the relative frequency of very warm $T > \mathrm{Q80}$ (cold, $T < \mathrm{Q20}$) months increases (decreases). The trend for very wet months ($P > \mathrm{Q80}$) is negative, while the frequency of six-month droughts indicated by the indices of standardized precipitation (SPI6) and of standardized precipitation and evapotranspiration (SPEI6) smaller than -1 shows a positive trend. The annual frequency range of individual river basins (shaded areas in Fig. 2) is largest for the drought indices, illustrating the large variability among the basins with respect to droughts. All trends in the basin-averaged relative frequency of extreme events shown in Fig. 2 are significant at a significance level of $\alpha = 0.05$ (non-parametric Mann-Kendall test, see Supplementary Information). Some trends of individual basins may not be significant. We excluded Chira in this trend analysis of climate extremes in Fig. 2, as not a single trend is found to be significant (also see Supplementary Information).

The occurrence of extreme events in reanalyses is shown together with the respective probability of seasonal (re-)forecasts SEAS5 from 1981–2018 in Fig. 3, for the example of drought conditions (over one month: $\mathrm{SPEI1} < -1$, and over six months: $\mathrm{SPEI6} < -1$) during the basins' rainy seasons. An increasing occurrence and accumulated SEAS5 probabilities of drought conditions can be identified in the last decade (except for Chira). Here, special emphasis can be placed on the 2011–2017 period of long-term drought conditions ($\mathrm{SPEI6} < -1$ for ERA5) in ESF, also forecast with probabilities of up to 80 and 90% ($\mathrm{SPEI6} < -1$ for SEAS5), during all included rainy seasons. However, lower forecast probabilities of around 40% are also present. For Niger and Volta, the period of 1981-2001 is mainly dominated by three drought events in 1983, 1987 and 1997; after 2001 both long-term and short-term drought events become more frequent and are frequently forecast. For all sub-Saharan African basins between 2015 and 2017, an accumulation of occurred drought periods and elevated forecast probabilities can be identified. Long-term drought conditions (SPEI6) in the study basins, e.g. for Karun in 1999–2001, 2007–2010 and 2018/2019, for the sub-Saharan African basins in 2009 and 2015, and for Chira from 2002–2005, were forecast with variable probabilities ranging from 10 to 90%. A strong variability in monthly rainy season's precipitation can be observed in the ESF (1990–1994, 2002–2007) and Karun (1985–1987, 1991–1999, 2010–2016) basins with various single months with $\mathrm{SPEI1} < -1$. For the other basins in Africa, in contrast, short-term and long-term drought conditions coincide to a higher degree, suggesting less variability within single months of the rainy season.

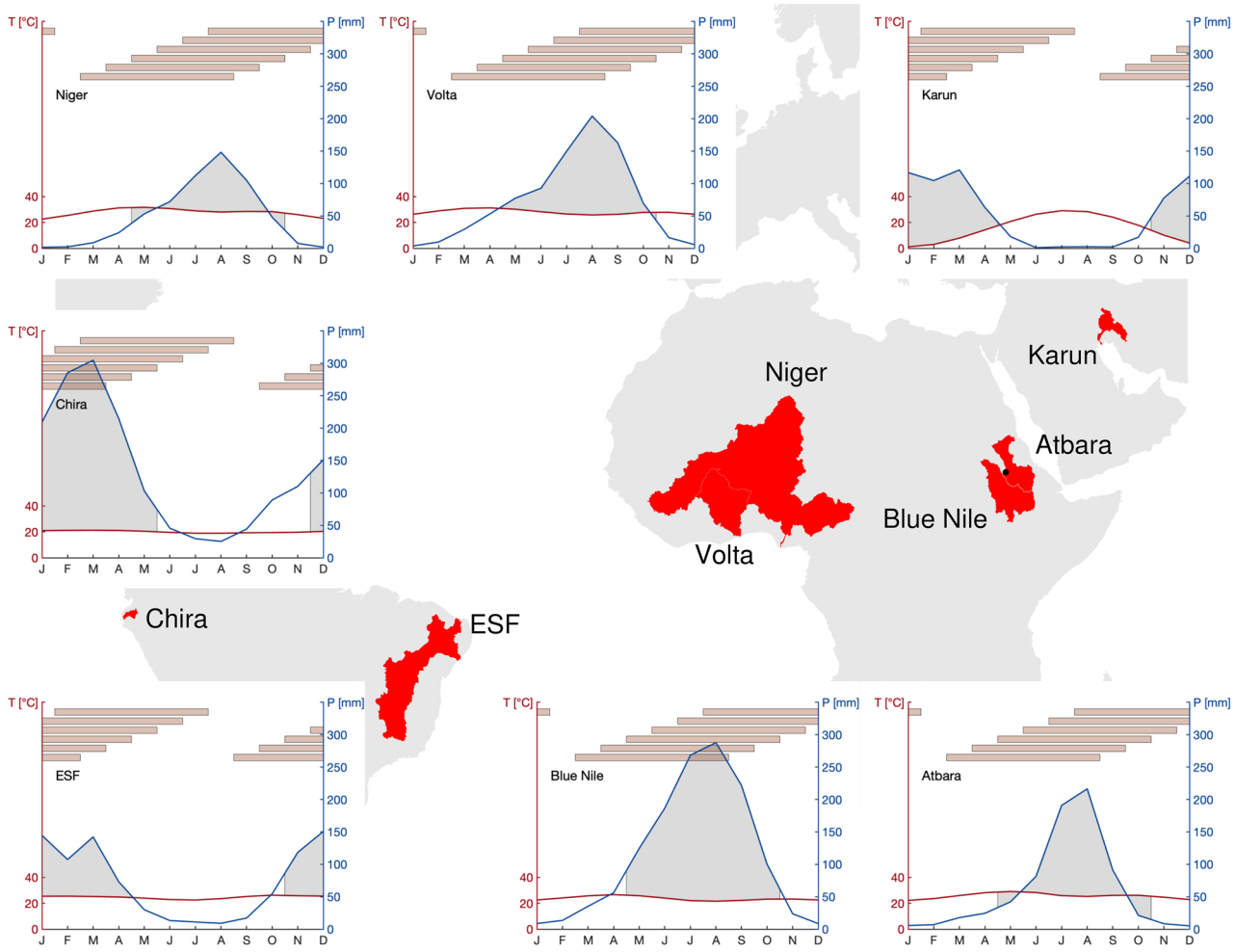


**Figure 1.** Map of basin areas and climate diagrams with temperature and precipitation. For the seven semi-arid river basins from West to East: Chira and Extended São Francisco (ESF) in South America, Niger and Volta in West Africa, Blue Nile and Atbara in Northeast Africa, and Karun in West Asia. Basin sizes range from $\mathcal{O}(10^4)\,\mathrm{km}^2$ to $\mathcal{O}(10^6)\,\mathrm{km}^2$, with basin altitudinal differences of less than 500 m to about 4000 m. The location of the Upper Atbara Dam Complex is shown with a black dot on the map. The grey shaded area in the climate diagrams defines the main four months of the rainy season, plus the preceding and succeeding month, used for the analysis as "rainy season". The red bars denote the six-month aggregation periods of a drought index for the analysis during the respective rainy seasons. Coastlines in the map originate from the Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG)[48]. Basin boundaries were produced with the HydroSHEDS dataset[49] and modified for consistency with their respective definitions by local authorities and stakeholders in the study regions.

**Economic benefit of forecast-based action for extreme events.** The comparison of forecast probabilities and drought occurrences in Fig. 3 already demonstrated the complexity of deciding whether an event was forecast or not, and the need to identify a probability threshold $p_{\mathrm{th}}$ with which this decision is made. For the seven semi-arid river basins, the potential economic value *PEV* is shown for the extreme events of very warm months ($T >$ Q80), very wet months ($P >$ Q80), and for drought months (SPEI $< -1$ and SPI $< -1$) during the rainy seasons as a function of the cost-loss ratio $C/L$ and $p_{\mathrm{th}}$ in Fig. 4. Per event, two lead times, i.e., for how many months ahead the forecast is valid (Lead 0: current month ahead, Lead 5: six months ahead), or two aggregation scales (SPI1/SPEI1: one month, SPI6/SPEI6: six months), are provided. Figure 4 includes all information needed for the choice of a beneficial $p_{\mathrm{th}}$ for each $C/L$ considering the two criteria of expense minimization and event maximization. For both criteria, the range of $p_{\mathrm{th}}$ with maximum *PEV* of each $C/L$ is relevant; however, the event maximization criterion further includes the condition that the highest possible number of extreme events should be caught by preventative early action, i.e., a maximization of the hit rate $H$. Here, we suggest $H > 0.5$ as the lowest threshold for this criterion to always ensure a higher hit rate than miss rate of an event ($H > 0.5$ is cross-hatched in Figs. 4 and 5). This therewith involves a certain degree of risk aversion of the decision maker. For the example of a warm extreme event ($T >$ Q80) with Lead 0 in the Karun basin and for expense minimization, a user with $C/L = 0.4$ would identify probability thresholds between 0.45 and 0.85 as a beneficial trigger range. However, for event maximization, the same user ($C/L = 0.4$) would need to use $0.15 < p_{th} < 0.45$ when only

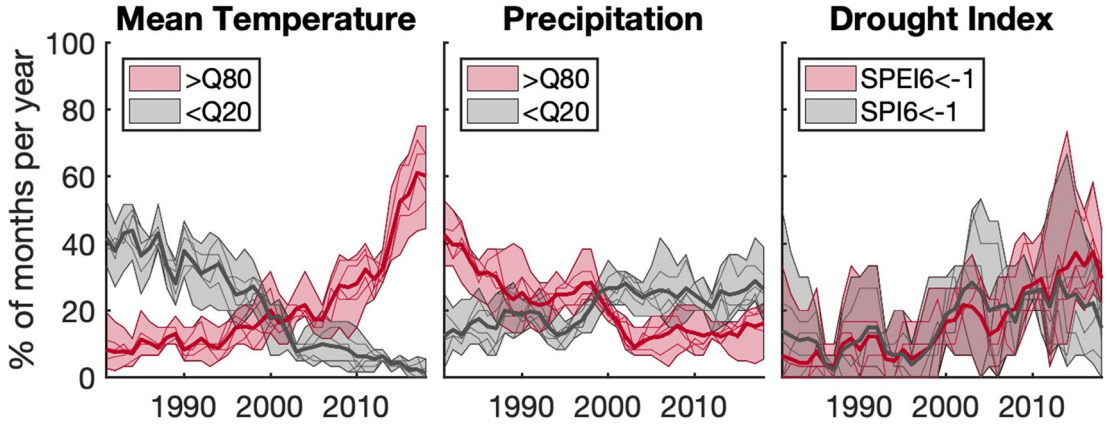


**Figure 2.** Relative frequency of climate variables from 1981–2018. From left to right the relative frequency of ERA5 mean temperature ($T$, left), total precipitation ($P$, middle) and drought index (SPEI, right) exceeding the respective quantile event thresholds (< Q20, > Q80,< −1) are shown for the basin-averages of Extended São Francisco, Niger, Volta, Atbara, Blue Nile and Karun. The thick solid lines denote the mean relative frequency over the 6 basins, and the shaded areas encompass the thin solid lines for the 5-year moving averages of relative frequencies of the individual basins.

including potential economic values with high hit rate ($H > 0.5$). When further considering the uncertainty of estimation of *PEV*, i.e., the robustness of the choice of $p_{\text{th}}$ (red line in Fig. 4, where 90% of bootstrap samples give a $PEV > 0.1$), a beneficial range of $p_{\text{th}}$ for the user with $C/L = 0.4$ could only be provided for the expense minimization criteria. For the event maximization criteria, this user would be well-advised to base decisions on the climatological strategy. In this case for $C/L = 0.4$, this would be never to act as $C/L$ is smaller than the climatological frequency of the event.

For all study regions, actions based on short forecast horizons (Lead 0, SPEI1/SPI1) possess higher potential economic values (*PEV*) and larger ranges of possible user situations and beneficial probability thesholds (ranges of $C/L$ and $p_{\text{th}}$ with $PEV > 0.1$) than actions based on the respective long-range forecasts (Lead 5 and aggregation periods over 6 months SPEI6/SPI6). Provided that action protocols are already established, the SEAS5 seasonal forecasts released on the 6th of each month can thus provide promising decision support for monthly action adaptations in addition to medium-range weather forecasts. Later release dates would diminish the usefulness of high-valued Lead-0 or SPEI1/SPI1 forecasts. The actual range of users ($C/L$) that can benefit from forecast-based action depends on the region and $p_{\text{th}}$: Among the basins, Chira provides the largest ranges of possible users and of beneficial probability thresholds for the warm ($T$ >Q80) and wet ($P$ >Q80) extreme events of both shown lead times. For drought forecasts of SPEI1 and SPI1, however, the ESF and Atbara basins show the largest areas of *PEV*. For Karun, mainly forecast horizons of one month (Lead 0 and SPEI1/SPI1) can offer economic value to users. For all regions, basing action on temperature forecasts ($T >$ Q80) of Lead 0 is beneficial for a wide range of $p_{\text{th}}$. The regions have in common that very limited *PEV* can be achieved by action based on Lead 5 forecasts of $P >$ Q80 and for forecasts of SPI6 $< -1$. Generally, the consideration of sampling uncertainty in the *PEV* estimation (red contour in Fig. 4) further reduces the ranges of users and probability thresholds with value. The comparison of the two drought indices SPEI6 and SPI6 for long-term drought events reveals that the inclusion of temperature in SPEI can increase the value of seasonal forecast-based action for most basins. Despite limited ranges of users and probability thresholds, *PEV* values above 0.4 are still possible for SPEI6 $< -1$, allowing a longer-term decision-support for drought conditions over the rainy seasons.

Considering the maximum potential economic value $PEV_{max}$ over all users (black solid line in Fig. 5), the event maximization criterion (cross-hatched) clearly cuts down the range of beneficial $p_{\text{th}}$ towards lower thresholds. Here, the definition of the maximum potential economic value as the difference between hit and false alarm rate ($PEV_{max} = H - F$) further reveals that the hit rate can still be higher than the false alarm rate when $PEV_{max} > 0$, however, the event maximization criterion of $H > 0.5$ (hit rate higher than miss rate) may no longer be valid. Exemplary for individual cost-loss ratios, we chose $C/L = 0.36$ and 0.2, that are also relevant for the real case study at the Upper Atbara Dam of the next section (also see Table 1). For $C/L = 0.36$ (blue line in Fig. 5), the effects of the different criteria on the choice of $p_{\text{th}}$ become even more pronounced: Although the maximum *PEV* for this $C/L$ would be achieved at higher threshold probabilities (expense minimization), e.g. around $p_{th} = 0.6$ for wet extreme events of Lead 0 in Chira, or around $p_{th} = 0.4$ for warm extreme events of Lead 5 in ESF, lower $p_{\text{th}}$ around 0.35 and 0.2, respectively, would be preferable in the same examples for event maximization. For other events and basins, like for the Lead 5 warm extreme events in Blue Nile, or for drought conditions SPEI1 $< -1$ in Niger, the probability threshold with maximum *PEV* for the cost-loss ratio of 0.36 does not change with the inclusion of the event maximization criterion. For some regions and events, no beneficial $p_{\text{th}}$ might even be found under the event maximization criterion for higher cost-loss ratios like $C/L = 0.36$, though showing potential economic value around 0.2 at higher $p_{\text{th}}$. This is for example the case for Chira and $P$ >Q80 for Lead 5, and for Volta and SPI1/6 $< -1$. For lower cost-loss ratios, e.g. $C/L = 0.2$ (red line in Fig. 5), this difference in the two criteria is less evident, as higher *PEV* values occur at lower probability thresholds for the extreme events, i.e., beneficial forecast-based action for those users is already provided at higher forecast uncertainty.

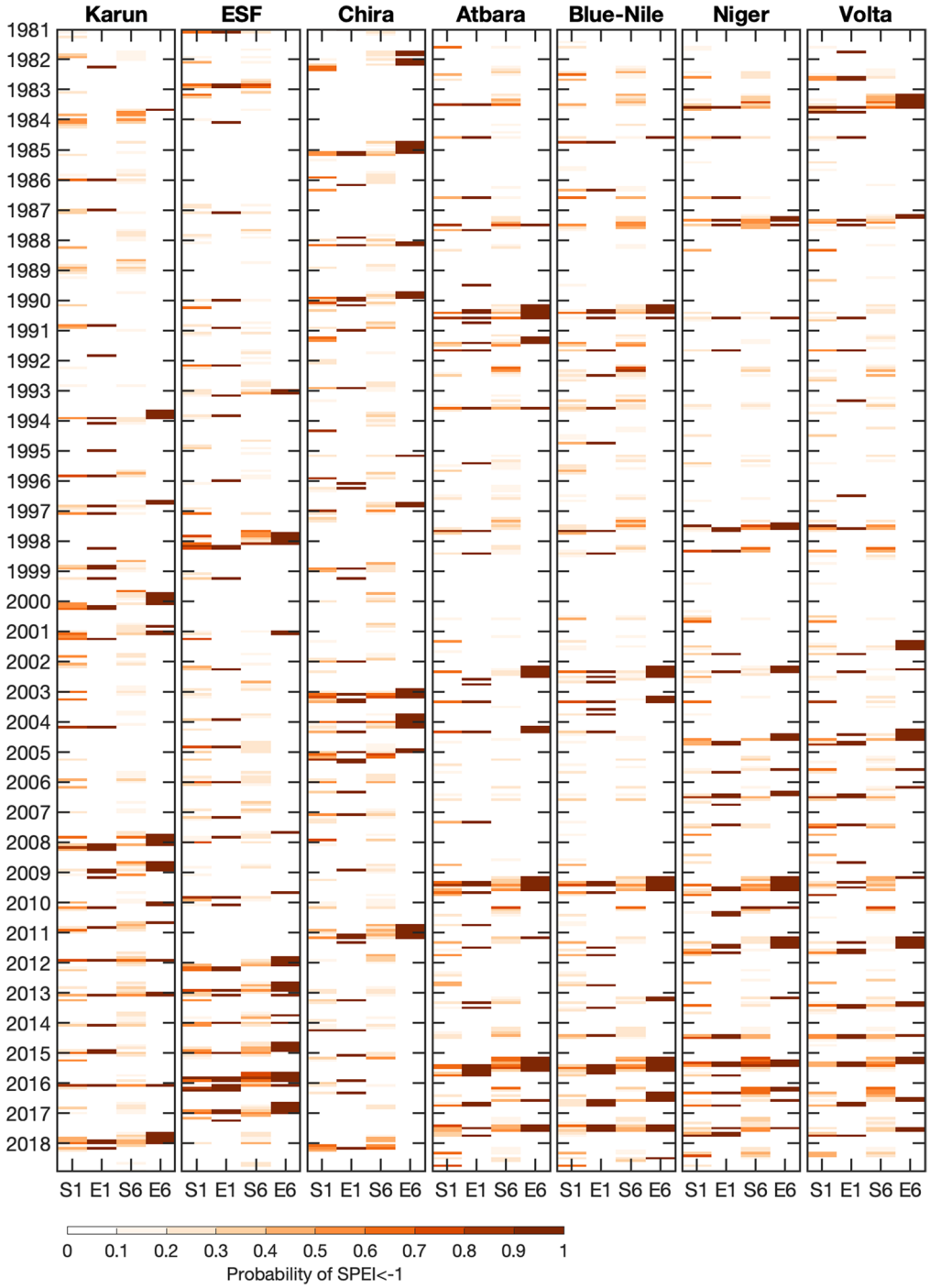


**Figure 3.** Forecast probability and occurrence of short- and long-term drought. The probability of SPEI1 $< -1$ (1) and SPEI6 $< -1$ (6) for SEAS5 (S) and ERA5 (E) are provided for the rainy seasons of the seven semi-arid regions from 1981-2018. The confinement of the analysis to the basins' rainy seasons allows a simplified overview of drought events and respective forecast probabilities when droughts are mainly driven by rainfall shortage. The used time periods for aggregation of SPEI6 are depicted with the red bars in Fig. 1.

Further summarizing $PEV_{max}$ over all probability thresholds, the effect of lead time and aggregation scale on the value of forecast-based action for the different extreme events can be studied (Fig. 6). For most regions, forecast-based action for warm extreme events ($T >$ Q80) is most valuable and maintains higher maximum

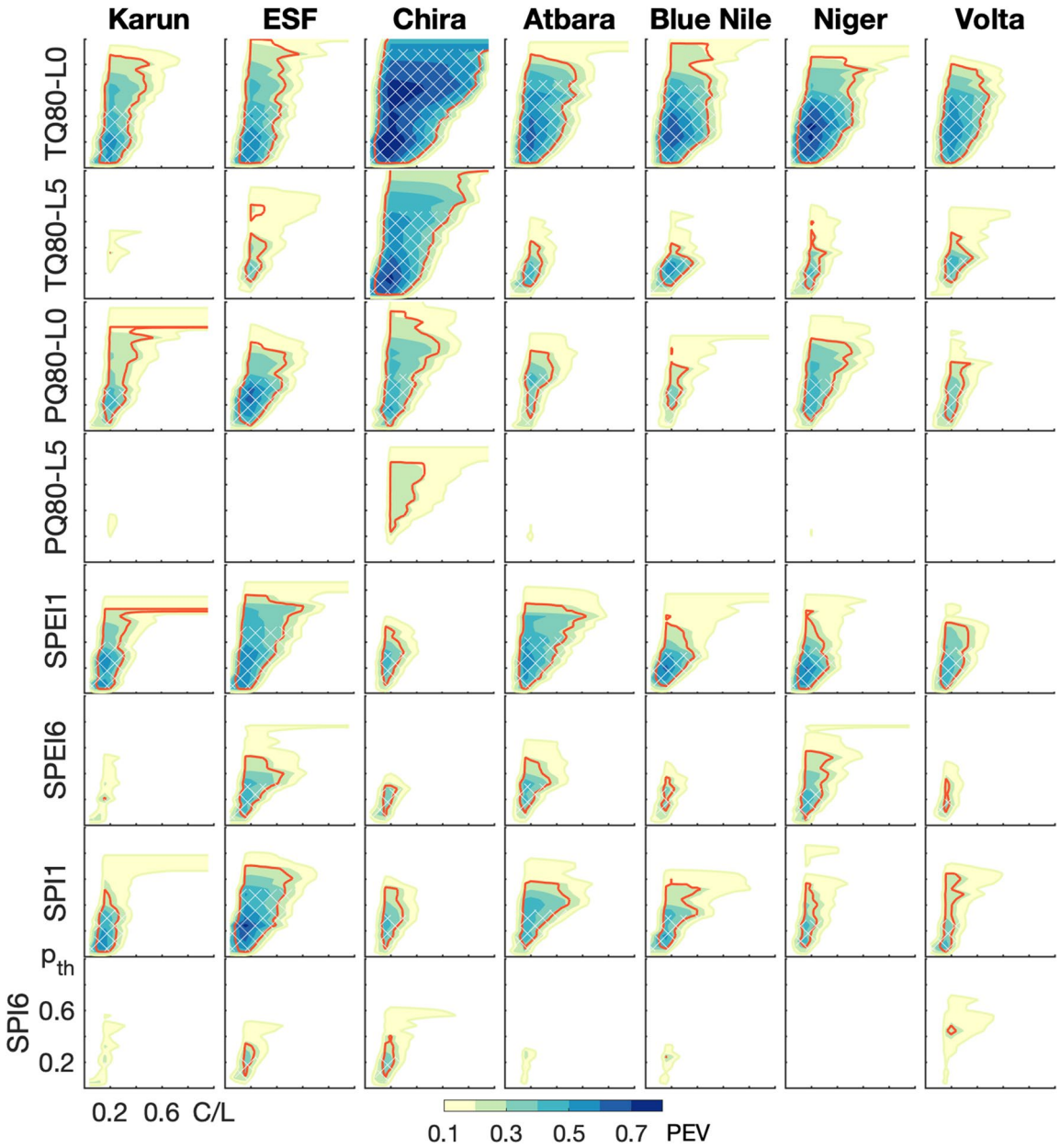


**Figure 4.** Potential economic value for forecast-based action as a function of probability threshold and user cost-loss ratio. The filled contours denote the potential economic value *PEV* of SEAS5 with reference ERA5 for extreme events of the seven river basins for rainy seasons of the hindcast period 1981-2016. The extreme events include $T > Q80$ and $P > Q80$ for leadmonths 0 (L0) and 5 (L5), as well as SPEI $< -1$ and SPI $< -1$ for aggregation scales over one (SPI1/SPEI1) and 6 months (SPI6/SPEI6). Only valuable conditions are color-coded, i.e., $PEV > 0.1$. The uncertainty of the estimation of *PEV* is assessed with bootstrap resampling, with the red contour indicating where 90% of bootstrap samples give a $PEV > 0.1$. The hatched area further represents where $PEV > 0.1$ and the hit rate $H > 0.5$ to base the choice of $p_{\text{th}}$ on the event maximization criterion.

potential economic value with increasing lead time than those for wet extreme events ($P >$ Q80). For warm extreme events, Lead 0 forecasts attain maximum potential economic values between 0.55 and 0.8, whereas for wet extreme events, $PEV_{max}$ of Lead 0 forecasts ranges between 0.38 and 0.64. The uncertainty of $PEV_{max}$ estimation is taken into account with the bootstrap confidence range (dashed lines in Fig. 6), e.g., specifying a range of $0.64 < PEV_{max} < 0.98$ for warm extreme events of Lead 0 in Chira with the 10% and 90% of bootstrap percentiles. Except for Chira with a smooth decrease of maximum potential economic value for warm extreme events with lead time, all regions show an abrupt decline of $PEV_{max}$ from Lead 0 to Lead 1. For most regions, this strong drop is larger than the decline of $PEV_{max}$ from Lead 1 to Lead 5. The most abrupt drop of $PEV_{max}$ is evident for wet extreme events in ESF from 0.75 of Lead 0 to 0.24 of Lead 1. For drought conditions, the decline with longer aggregation scales is less pronounced than the decline with single lead times for the wet extreme, as still higher-valued months of lower forecast horizon are included in SPI/SPEI. In basins like ESF, Blue Nile

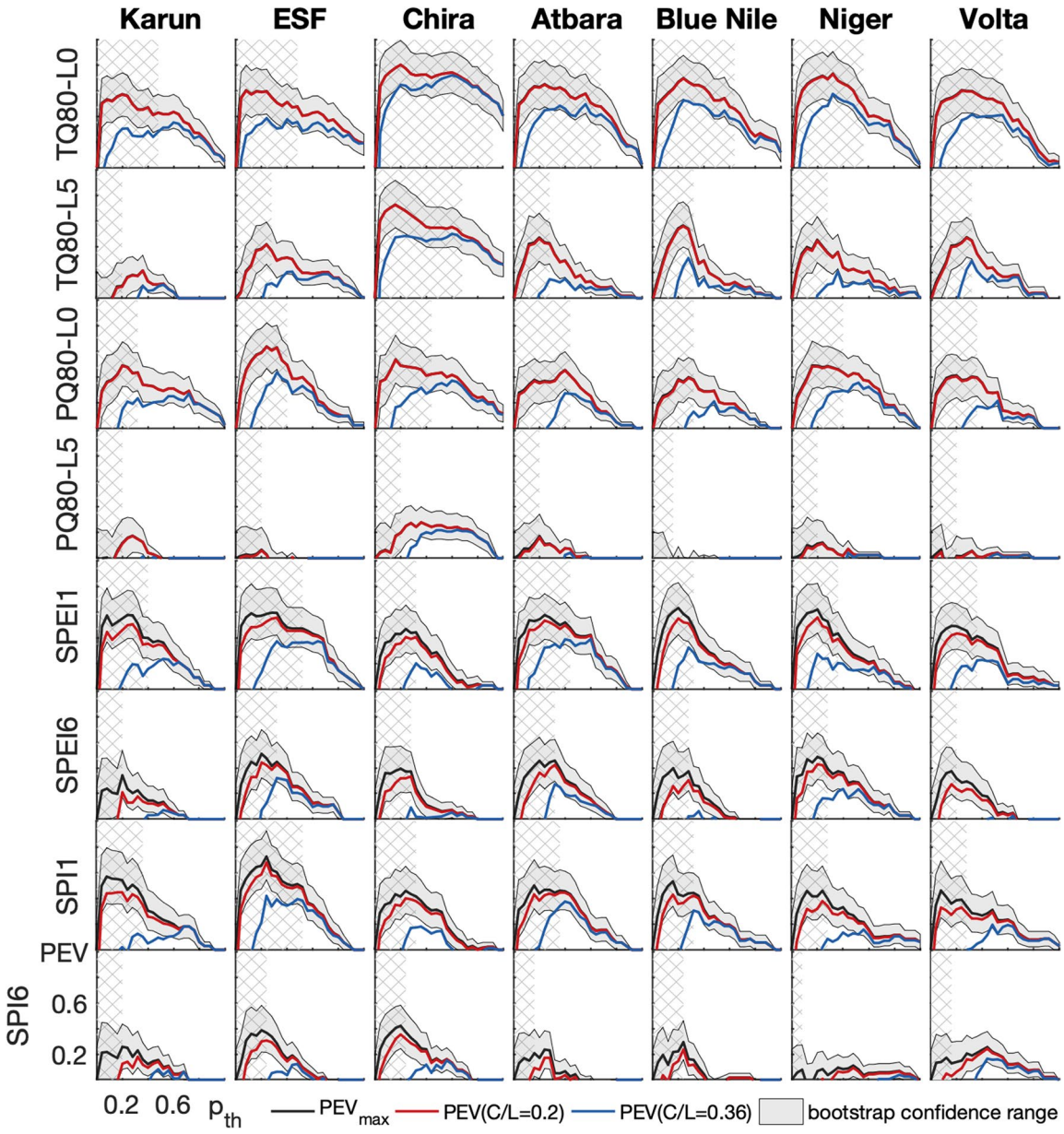


**Figure 5.** Maximum potential economic value as a function of probability threshold. $PEV_{max}$ of SEAS5 with reference ERA5 is depicted for extreme events of the seven river basins for rainy seasons of the hindcast period 1981-2016 (black line). The uncertainty of the estimation of $PEV_{max}$ is assessed with the confidence range between the 10% and 90% quantiles of bootstrap resampling (shaded area). The red and blue lines denote the $PEV$ for cost-loss ratios of 0.2 and 0.36, respectively. The extreme events include $T > Q80$ and $P > Q80$ for leadmonths 0 (L0) and 5 (L5), as well as SPEI $< -1$ and SPI $< -1$ for aggregation scales over one (SPI1/SPEI1) and 6 months (SPI6/SPEI6). The cross-hatched area further represents where the hit rate $H > 0.5$ to base the choice of $p_{\mathrm{th}}$ on the event maximization criterion. Note: From its definition, $PEV_{max}$ represents the $PEV$ when the cost-loss ratio $C/L$ is equal to the climatological frequency of the event (see Equation 2), as in this special case both options of never or always acting create the same expense. Climatology can not advise the decision maker and any forecast is of maximum value. The climatological frequency of a Q80-event and SPI/SPEI $< -1$ is 0.2 and 0.159, respectively. Therefore, $PEV\,(C/L = 0.2)$ (red line) and $PEV_{max}$ (black line) for the Q80-events overlap and accordingly cannot be seen separately.

and Volta, where $PEV_{max}$ of Lead 5 for wet extreme events is close to zero, longer aggregation scales of the dry extreme (SPI6 $< -1$) still attain values between 0.24 and 0.38. For SPEI the decrease of $PEV_{max}$ with longer aggregation scales more resembles the one of the warm extreme with increasing lead time, with relatively high maximum potential economic values between 0.28 and 0.54 still for longer aggregation scales. The inclusion of

| | No extreme event $X_{crit}$ is not exceeded | Extreme event occurs $X_{crit}$ is exceeded |
|---|---|---|
| Early **action** based on forecast $p_{\mathrm{th}}$ is exceeded | *False alarms (f)*<br>Futile action ($C$)<br>Unnecessary costs for action | *Hits (h)*<br>Valuable action<br>(Part of) losses avoided, costs for action<br>($L_{event} - L + C$) |
| **Drought reservoir operation** | $C = 6.0 \cdot 10^6$US\$ a$^{-1}$<br>$C = 0.3 \cdot 10^6$US\$ a$^{-1}$ | $L_{drought} - L + C = 20.9 \cdot 10^6$US\$a$^{-1} \rightarrow$<br>$L = 16.2 \cdot 10^6$US\$ a$^{-1}$<br>$L_{drought} - L + C = 16.4 \cdot 10^6$US\$ a$^{-1} \rightarrow$<br>$L = 1.6 \cdot 10^6$US\$a$^{-1}$ |
| No early **action** $p_{\mathrm{th}}$ is not exceeded | *Correct rejects (r)*<br>Valuable inaction<br>No costs or losses (-) | *Misses (m)*<br>Erroneous inaction<br>No losses avoided ($L_{event}$) |
| **Standard reservoir operation** | 0<br>0 | $L_{drought} = 31.1 \cdot 10^6$US\$ a$^{-1}$<br>$L_{drought} = 17.7 \cdot 10^6$US\$ a$^{-1}$ |

**Table 1.** Contingency table for forecast-based action. For the four action-event scenarios, the economic consequences are shown. For the case study of drought and standard reservoir operations at the Upper Atbara Dam in Sudan, the costs and losses for different outcomes of the decision model without operation restrictions (upper line of reservoir operations) and with sediment sluicing (lower line of reservoir operations) are provided. The resulting ratio of costs for action ($C$) and avoidable losses ($L$) for a drought operation without restrictions and with sediment sluicing are 0.37 and 0.19, respectively.

temperature attaining high values still at longer lead times in the drought index of SPEI allows for higher $PEV_{max}$ at long aggregation scales relative to the solely precipitation based drought index of SPI.

**Water level management and energy production during drought reservoir operations: the Upper Atbara Dam case.** The calculation of the cost-loss model for drought events is demonstrated for the water level management and energy production at the Upper Atbara Dam in Sudan. Though having two dam operation examples including and neglecting sediment restrictions for a reservoir drawdown, this study cannot suggest completeness of all relevant processes, decisions and measures. However, it is a demonstration of application of the cost-loss model and the above generated results.

Having no operation restrictions on the Upper Atbara Dam, water level changes are governed only by a reservoir operation to maximize electrical energy generation. The standard operation at the Upper Atbara Dam with the peak of the rainy season in July/August is a gradual lowering of the reservoir water level from the full supply level (FSL) at 521 m asl by about 10 m from January on to be able to take advantage of the stored reservoir volume and to absorb the inflow during the rainy season (red dashed line in Fig. 7a). As the expected inflow comes, the reservoir is filled again to the FSL.

In a dry year, the water level curve differs in July and August when the operator first still awaits the maximum inflow to come, but needs to fill the reservoir until September to FSL (black dashed line in Fig. 7a). The latter requires a high amount of the inflow for filling as the expected peak inflow was absent. Compared to a normal year, this results in a strong decline by about 30% in energy production from July to September (black dashed line in Fig. 7b).

With prior knowledge of drought conditions during the rainy season, i.e., applying drought operations, the reservoir is kept at a higher water level with higher heads (black solid line in Fig. 7a), lowering the energy production in the first phase by about 25%, but allowing for considerably higher energy production (+40%) in the second phase compared to the standard operations (black solid line in Fig. 7b).

If a normal year was falsely operated as a dry year, e.g., a seasonal forecast falsely predicted drought conditions for the rainy season of a normal year, the reservoir with the reduced water level by only 3 m in the first phase is not able to absorb the normal amount of inflow (red solid line in Fig. 7a). The reservoir operates at its maximum capacity and part of the inflow needs to be passed through the spillway, without the usage for energy production (red solid line in Fig. 7b).

For the decision whether to modify the reservoir operation, i.e., to take preventive action for a coming drought event or not, the decision-maker needs to evaluate the economic consequences. The losses ($L_{drought}$), when the drought occurs but no action is taken, are calculated from the yearly difference of the earned money applying standard operations during a normal year (Normal-Standard, $203.5 \cdot 10^6$US\$ a$^{-1}$) and during a dry year (Dry-Standard, $172.4 \cdot 10^6$US\$ a$^{-1}$), and amount to $31.1 \cdot 10^6$US\$ a$^{-1}$ (Table 1). The costs $C = 6.0 \cdot 10^6$US\$ a$^{-1}$ arise from the yearly difference of the earned money during a normal year with standard operations (Normal-Standard) and a normal year that is falsely treated as a drought year (Normal-Drought, $197.5 \cdot 10^6$US\$ a$^{-1}$). For correctly-taken action, costs $C$ and unavoidable losses $L_{drought} - L$ of in total $20.9 \cdot 10^6$US\$ a$^{-1}$ incur, determined from the difference of earned money in a normal year with standard operations and of a dry year with drought operations (Dry-Drought, $182.6 \cdot 10^6$US\$ a$^{-1}$). $L$ are the potential losses protected by the action, and result in $16.2 \cdot 10^6$US\$ a$^{-1}$. The characteristic cost-loss ratio $C/L$ of preventive drought operations for the Upper Atbara Reservoir, i.e., the cost of acting as a fraction of the potential loss prevented by the action, thus is 0.37.

With the closest resolvable $C/L = 0.36$, the resulting maximum $PEV$ of 0.4 (expense minimization) for drought events of SPEI1 in the Atbara basin is obtained for $p_{th} = 0.6$ (Fig. 5). Considering the event maximization criterion, a lower threshold around $p_{th} = 0.32$ is beneficial, however with lower $PEV$ of 0.35. For the larger timescales of SPEI6, both expense minimization and event maximization criteria result in most beneficial forecast-based action for $p_{th} = 0.32$ with $PEV = 0.27$.

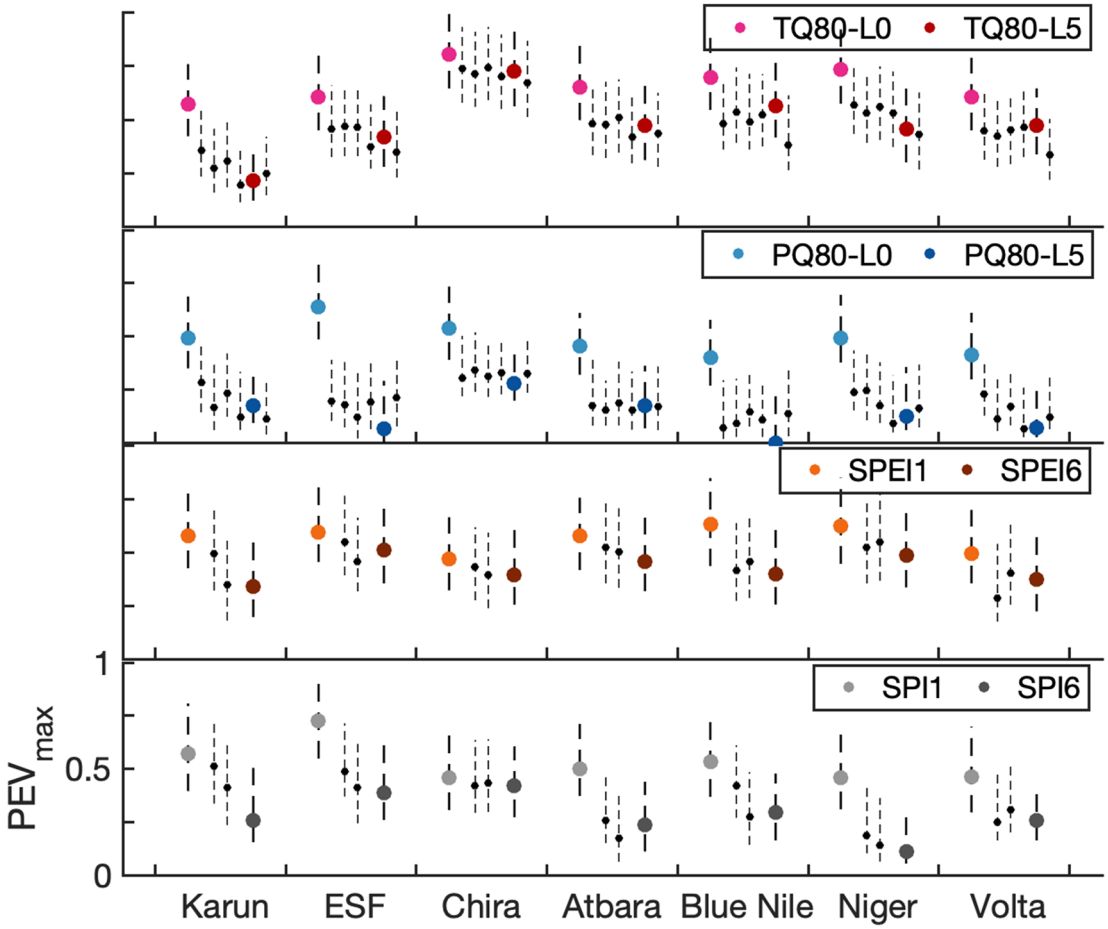


**Figure 6.** Maximum potential economic value for forecast-based action. $PEV_{max}$ of SEAS5 with reference ERA5 is shown for the seven river basins for rainy seasons of the hindcast period 1981-2016. Colored dots represent the fully evaluated $PEV_{max}$ for the events of $T > Q80$ and $P > Q80$ for leadmonths 0 (L0) and 5 (L5), and for the events SPEI $< -1$ and SPI $< -1$ for aggregation scales over one (SPI1/SPEI1) and 6 months (SPI6/SPEI6). The uncertainty of the estimation of $PEV_{max}$ is assessed with the confidence range between the 10% and 90% quantiles of bootstrap resampling (dashed line). $PEV_{max}$ is also given for temperature and precipitation of leadmonths 1-4 and 6, and for drought events of scales over 3 and 4 months (black dots) with respective bootstrap confidence ranges.

At high sediment-trapping dams in semi-arid regions often sediment sluicing operations during the peak of the rainy season are required. Here, the reduction of the reservoir water level to the minimum operation level (MOL) during high inflows increases the bottom shear stress maintaining incoming sediment in suspension and remobilizing deposited sediment with only limited or no reservoir deposition.

At the Upper Atbara Dam, the MOL at 510 m asl is usually held for six weeks during July and August and may be extended if inflows still exceed 3000 $m^3 s^{-1}$. In general, the operation restriction by sluicing reduces the energy production in a normal year by about 13% due to lower heads (dark blue dashed line in Fig. 7b).

In a dry year, the same operation restriction for sluicing with a steady drawdown to MOL is performed (light blue dashed line in Fig. 7a). As inflows are lower, about 23% less energy is produced from July until September compared to a normal year (light blue dashed line in Fig. 7b).

With the application of drought operations, i.e., having prior knowledge of drought, the water level is first more slowly lowered, also reducing the produced energy in this phase by 14% (light blue solid line in Fig. 7). Then, the water level is more rapidly lowered down to the sluicing level, thus allowing for a higher energy production (+17%) from July until September compared to standard sluicing operations due to higher heads. As the expected inflow is low and carries less sediments than high inflow, the period of sediment sluicing is reduced to four instead of six weeks.

Falsely applied drought reservoir operations in a normal year first show a lower energy production and then a higher energy production (dark blue solid line in Fig. 7b). The yearly difference in energy production between the standard (Normal-Sluicing-Standard, $177.3 \cdot 10^6$ US$ $a^{-1}$) and drought reservoir operation (Normal-Sluicing-Drought, $177.0 \cdot 10^6$ US$ $a^{-1}$) only amounts to 0.17%.

Similarly to the case without reservoir operation restrictions, the costs and losses for the drought operation including sediment sluicing can be calculated (Table 1). The costs for drought operations amount to $0.3 \cdot 10^6$ US$ $a^{-1}$. The total losses of a dry rainy season $L_{drought}$ are $17.7 \cdot 10^6$ US$ $a^{-1}$ (with Dry-Sluicing-Standard:

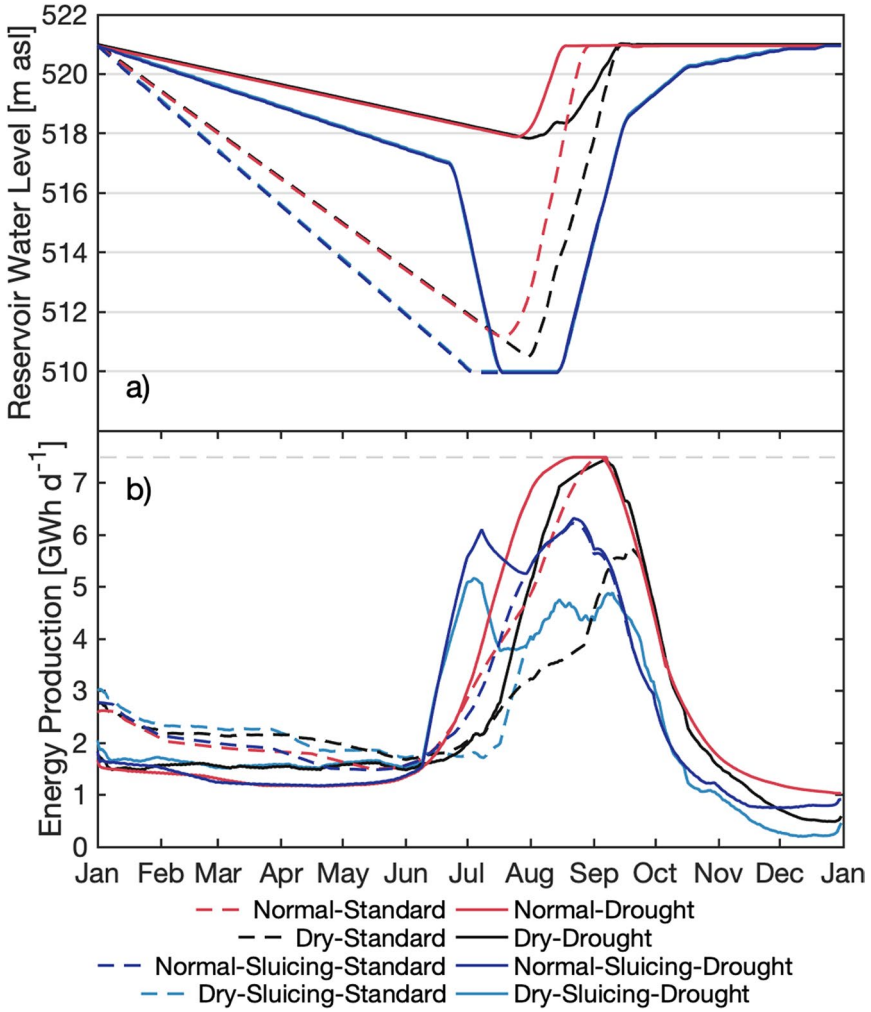


**Figure 7.** Reservoir management at the Upper Atbara Dam Complex in Sudan. **(a)** Reservoir water level (shown as 5-day moving average) and **(b)** energy production (shown as 30-day moving average) for a rainy season in a normal year without operation restrictions (red lines, "Normal"), for a rainy season under drought conditions without restrictions (black lines, "Dry"), for a rainy season in a normal year including sediment sluicing (dark blue lines, "Normal-Sluicing"), and for a rainy season under drought conditions including sediment sluicing (light blue lines, "Dry-Sluicing", in **(a)** identical to dark blue line). The dashed lines denote reservoir management with standard reservoir operations. The solid lines show the altered reservoir management for drought conditions, i.e., having prior knowledge of drought. In **(b)**, the light gray dashed line marks the maximum possible energy production (7.5 GWh $d^{-1}$), i.e., the maximum reservoir capacity. Any additional inflow would need to be passed through the spillway, without the usage for energy production.

$159.6 \cdot 10^6$ US\$ $a^{-1}$), of which $L = 1.6 \cdot 10^6$ US\$ $a^{-1}$ can be avoided by drought reservoir operations (with Dry-Sluicing-Drought: $160.9 \cdot 10^6$ US\$ $a^{-1}$). The resulting cost-loss ratio is 0.19.

With these identified costs and losses (closest resolvable $C/L = 0.2$), seasonal drought forecasts of ECWMF SEAS5 generate potential economic values between 40 and 55% for both short (SPEI1) and long (SPEI6) time horizons, and for both the event maximization and expense minimization criteria (Fig. 5). The most beneficial $p_{th}$ is therewith identified at around 30% for high hit rates above 50%.

## Discussion

This study presents how decision-makers can optimize water management by the use of seasonal forecasts and directly obtain robust beneficial probability thresholds triggering preventive action on the chosen criterion of either expense minimization or event maximization. This valuation approach is a tool for user-oriented, economic decision-related verification of probability forecasts, away from the classical approach of evaluating forecasts based on skill scores[33,34], whose statistics may not be conclusive in a decision making environment. Of course, the classical quality assessment based on skill scores and the economic value assessment should complement each other. However, the aims and advantages behind both approaches are different[35]: The skill/quality approach aims

at the improvement of correspondence of forecasting systems with observations as well as the intercomparison of different forecasting systems. It has the advantage of identifying their strengths and flaws with respect to different quality skill measures (e.g., forecast bias, distribution, sharpness). The valuation approach has so far been applied for, e.g., promoting the usage of ensemble forecasts[16], verification of seasonal forecasts[20,36], and even for decision-support of preventive humanitarian action[18]. Compared to a quality skill score analysis, the valuation approach as used in our study aims at building customer confidence in forecast products by demonstrating the economic benefit realized by decision-makers through the use of the forecasts. It has the clear advantage of providing decision-support for individual users with an individual choice of beneficial probability thresholds for a special event. A common skill score of forecast quality aggregates the forecast performance over all users (cost-loss ratios) and probability thresholds. For example, $PEV_{max} = H - F$, also known as the Kuipers Skill score[16], is the maximum economic value over all $C/L$ and $p_{th}$. In our case, skillful forecasts ($PEV_{max} > 0$) for all lead times and scales were found for all considered extreme events (except for Blue Nile $P > Q80$ of Lead 5, Fig. 6), however, the individual cost-loss situation of an action determines the individual value and action trigger. For certain users, skillful forecasts might involve non-positive economic value (Figs. 4 and 5). Considering the aggregation timescale of the drought index SPEI $< -1$, the forecast skill $PEV_{max}$ only decreases slightly with longer timescales (Fig. 6). However, the confident range of $p_{\mathrm{th}}$ and cost-loss situations with $PEV > 0.1$ may shrink drastically from SPEI1 to SPEI6 (Fig. 4), demonstrating the necessity of a detailed look at $PEV$ for each $C/L$ and $p_{\mathrm{th}}$ (Fig. 4). Therefore, skillful seasonal forecasts of hydrometeorological events *per se* are not necessarily valuable to a specific decision maker and a considered action.

In our analysis, we used uncorrected, possibly biased forecasts. This is reasonable due to the usage of relative thresholds, i.e., quantiles of the distributions of forecast and reference data separately, instead of absolute values, as well as drought indices of separate forecast and reference distributions. Other studies[37,38] merged the reference and forecasts data in their calculation of drought indices and therefore required the two datasets to have at least the same mean climate, i.e., they required linear-scaling bias correction of monthly data per lead time. With our separate treatment of distributions of forecast and reference data, mean and distributional biases that would be removed by linear scaling or quantile mapping bias correction are implicitly taken into account. Even ensemble calibration methods[39] won't have an effect on the relative quantile values of separate reference and forecast distributions, as they do not change the relative location of the quantile value within the system's distribution when only mean and standard deviation of the ensemble are changed. Forecast distributions are further used for each lead time and scale separately, diminishing the evident effect of model drift with lead time. Uncorrected raw forecasts can thus be used as long as (1) distributions are separately treated for reference and forecast data, (2) the temporal (e.g. monthly; per lead time) and spatial (basin-averages) resolutions are preserved and (3) no absolute values or event thresholds (e.g., $> 5$ mm) are used. Nevertheless, more sophisticated post-processing measures, such as multivariate bias correction[40] or even dynamical downscaling[41] may show an effect on the forecast value.

Regarding the transferability of the results, the catchment size was found to have only a minor impact on the economic value of seasonal forecasts. The results rather seem to be independent of the basin size. For differences in economic value between the basins, general circulation patterns and their predictability may play a larger role (e.g., ENSO temperature effect in Chira, influence of short predictability of westerlies in Karun). Consistently among all considered semi-arid basins that are far apart from each other in South America, Africa and West-Asia and with different dominating weather systems, higher PEV and larger beneficial ranges of probability thresholds and possible user situations were evident for lower forecast timescales than for higher ones (Fig. 4). The warm extreme events ($T > Q80$) were shown to provide the respective highest values of seasonal forecasts, even for higher lead times (Fig. 6). Similarly, drought forecasts (SPEI) allowed for high $PEV$ even at high aggregations scales (SPEI6). This consistency of a high value among all different study regions might indicate a transferability of the results to other regions. Economic benefits could likewise be expected from temperature and drought forecasts of even high forecast horizons up to six months, and from precipitation of at least one month ahead. As in our study regions, certain users may thus still benefit from large forecast horizons. However, further studies still must confirm this validity for other regions of the world.

We provided an in-depth example for water reservoir management at the Upper Atbara Dam in Sudan for energy production as a real-case practical application of the cost-loss decision model. For this example, dam operation standards are based on practical experiences of the dam management agency, however, flexibility of the reservoir manager is also possible when using other fix dates or even when considering energy demand and price. Thus, a wide range of cost-loss ratios can be obtained controlled by different underlying assumptions. Furthermore, the cost-loss situation was focusing on hydropower generation, not yet considering, e.g., costs and losses for irrigation. With this, we provided a straightforward application of the cost-loss decision model for forecast-based action. We further promote the implementation of seasonal ensemble forecasts for concrete water management operations at the Upper Atbara dam. The given example there, though including limitations and being demonstrated for only a specific year so far, showed that up to one half ($16.2 \cdot 10^6$ US$) of the losses associated with the energy production in a drought year can be avoided if altered management operations are applied.

The increase shown in the frequency of droughts and warm extremes during the last decades together with high population growth rates in semi-arid regions, raise the pressure on multipurpose reservoirs, such as the Upper Atbara Dam. Longer-term sustainable planning several months ahead prior to the next rainy season becomes even more crucial. With rising demands on the water management, switching towards beneficial seasonal-forecast-based early actions saves expenses and aids climate proofing. Consequently, we stress the advantage and necessity of considering seasonal forecasts in hydrological decision making.

## Methods

**Seasonal climate forecast and reference data.** SEAS5[32] is the latest global seasonal forecast system of the European Centre for Medium-Range Weather Forecasts (ECMWF) launched in 2017 including 51 ensemble members. For verification purposes, SEAS5 reforecasts are provided for the period 1981-2016 including 25 ensemble members. Both forecasts and reforecasts are initialized on the first day of each month and provide a forecast horizon of 7 months. Its horizontal resolution is approximately 36 km. SEAS5 is a coupled atmosphere-ocean-sea-ice model that is particularly powerful in the prediction of El Niño Southern Oscillation (ENSO)[32].

Being aware of the uncertainty of global precipitation reference products, and further requiring intrinsic dependence structure and interaction between, as well as the temporal coincidence and physical consistency of precipitation and temperature, the climate reanalysis product ERA5[42] is used as reference for both precipitation and temperature. Being system-consistent with SEAS5, ERA5 is offered by the ECMWF with a reanalysis period from 1979 to within 3 months of real time. Historical observations are combined using advanced modelling and data assimilation systems producing global, consistent estimates of climate variables at a horizontal resolution of 30 km.

For this study, monthly area-weighted basin averages of ERA5 and SEAS5 2-m mean temperature and precipitation are used. To preserve the original data and to avoid artifacts due to interpolation methods, basin averages were produced with masks for reference and forecast grids separately. The applied evaluation methods are all based on relative values, i.e., quantiles of the distribution of values per ERA5 and per SEAS5 separately. This approach of separate distributions of reference and forecast data implicitly performs a bias correction of systematic biases in mean and quantile values and therefore does not require additional prior bias correction of the forecast data. Also model biases of the seasonal forecasts, i.e., biases that drift over lead time, are implicitly taken into account by comparing only the same lead times of each month with one another. This implies different distributions and thus quantile values for individual lead times per month making use of the long period of available re-forecasts.

**Drought indices.** The standardized precipitation index (SPI) is a drought index based only on precipitation data[43]. The use of SPI assumes that droughts are controlled mainly by the temporal variability in precipitation, neglecting the variability in temperature or potential evapotranspiration (PET).

As the efficiency of drying resulting from temperature anomalies can be as high as that due to rainfall shortage, the standardized precipitation evapotranspiration index (SPEI) also includes PET[44]. Consequently, unlike SPI, SPEI is sensitive to global warming and also indicates drought conditions for increased evapotranspiration due to increased temperature despite of normal precipitation conditions[44].

Both SPI and SPEI have a multi-temporal nature and their variables of precipitation and water surplus/deficit, respectively, are aggregated at different time scales[43,44]. Unlike in historical analyses, the application of seasonal forecasts suggests the calculation of the multi-month scale of SPI/SPEI in a forecast - not in a retrospective - manner, i.e., SPI6 and SPEI6 of June specify the accumulation from June to November for both reference and forecast. For the calculation of both SPI and SPEI, probability distributions of reference and forecast data are separately treated, i.e., data are not merged, allowing again the use of uncorrected raw forecasts. With the relationship of empirical cumulative probability to the respective aggregated variable of reference and forecast data separately, the associated inverse normal of the respective cumulative probability is estimated. The then derived deviation for a standard normally distributed quantity with zero mean defines the standardized drought indices SPI and SPEI (also see Supplementary Information). In this study, we chose the threshold of SPI/SPEI $< -1$ for the drought analysis, representing all negative anomalies larger than one standard deviation. This encompasses the drought categories of moderate, severe and extreme drought and approximately 15.9% of the data[43].

**Potential economic value.** The benefit from including seasonal climate forecasts in the decision process for water management can be quantified by the potential economic value ($PEV$)[16,17]. In this decision process, actions taken in the water management would be based on seasonal climate forecasts.

For this, we evaluate the performance of the seasonal forecast by combining the outcomes of a binary forecast (hit, false alarm, miss and correct reject) with a cost-loss approach (Table 1). As seasonal forecasts are probabilistic forecasts, i. e., consist of 25 (reforecasts) or 51 (forecasts since 2017) ensemble members, the definition of "the event is forecast" depends on a probability threshold $p_{\mathrm{th}}$. Seasonal forecast-based action will be induced whenever the $p_{\mathrm{th}}$ of the ensemble forecast for an event is exceeded. No action is taken for lower probabilities. The cost-loss ratio $C/L$ defines the ratio of the cost for action over the avoidable losses by action and is a characteristic of the individual user.

$PEV$ is a measure of relative savings. The savings arise from the comparison of the expenses for the forecast-based action to expenses of a reference strategy, the climatological approach ($E_{\mathrm{climate}} - E_{\mathrm{forecast}}$). The climatological approach involves either always acting or never acting, with an optimal course defined by the relation of $C/L$ and the climatological frequency of the event $\overline{o}$: always act if $C/L < \overline{o}$, and never act otherwise. The savings $E_{\mathrm{climate}} - E_{\mathrm{forecast}}$ are set into relation with the savings of having a perfect forecast available, i.e., optimal early action due to perfect knowledge of the future weather with which the decision maker acts only when the event was going to occur. This results in:

$$PEV = \frac{E_{\mathrm{climate}} - E_{\mathrm{forecast}}}{E_{\mathrm{climate}} - E_{\mathrm{perfect}}} \tag{1}$$

The upper boundary ($PEV = 1$) is defined by the savings of a perfect forecast and the lower boundary ($PEV = 0$) by the climatological approach. The user benefits from the forecast-based action if $PEV > 0$. For $PEV \leq 0$, the

usage of the climatological approach is more beneficial. Depending on $H$ and $F$, $PEV$ can be calculated for different $C/L$ and $p_{th}$:

$$PEV\left(\frac{C}{L}, p_{th}\right) = \begin{cases} \dfrac{\frac{C}{L} - \overline{o} - \frac{C}{L}(1-\overline{o}) \cdot F(p_{th}) + \overline{o}(1-\frac{C}{L}) \cdot H(p_{th})}{\frac{C}{L}(1-\overline{o})}, & \text{if } \frac{C}{L} < \overline{o} \\ \dfrac{-\frac{C}{L}(1-\overline{o}) \cdot F(p_{th}) + \overline{o}(1-\frac{C}{L}) \cdot H(p_{th})}{\overline{o}(1-\frac{C}{L})}, & \text{else } \frac{C}{L} \geq \overline{o}. \end{cases} \quad (2)$$

To estimate the uncertainty of $H$, $F$ and $PEV$ due to sampling errors of the extreme events, bootstrap resampling is applied[20] (also see Supplementary Information for more details on the bootstrapping algorithm). The advantage of the bootstrap approach is that no assumption regarding the distribution of the values is required[45]. Here, at the equal size of the original data set, pairs of event-based forecast probability time series and binary reference time series are randomly selected with replacement. This resampling was repeated up to the bootstrap sample size of 1000. To assess the robustness of forecast thresholds[18], confidence intervals for $PEV$ are provided based on the 10 % and 90 % quantiles of the bootstrap samples.

The choice of beneficial $p_{\mathrm{th}}$ can be based on the expense minimization or event maximization criteria[18]. For expense minimization, the range of $p_{\mathrm{th}}$ with highest $PEV$ for a specific user ($C/L$) is chosen. For event maximization, this is further put under the constraint that $H$ should be high to catch as many events as possible and to avoid losses. This event maximization criterion therefore involves a certain degree of risk aversion of the decision maker.

**Water level management and energy production at the Upper Atbara Dam.** The Upper Atbara Dam Complex has a watershed area of about 100,000 km$^2$[46]. The location of the dam in the Atbara basin is shown in Fig. 1. The complex was built as a twin dam just upstream of the confluence of the Upper Atbara and Setit rivers with a power generation of up to 320 MW. The Full Supply Level (FSL) of the Upper Atbara Dam Complex is at 521 m asl, the Minimum Operation Level (MOL) or Sluicing Level is at 510 m asl.

A spreadsheet simulation model was applied based on the reservoir operation model that had been developed by the responsible dam planning company Tractebel Engineering GmbH to guide the reservoir impounding during and after the dam construction period for the Dam Implementation Unit (DIU) in Sudan, i.e. the stakeholder responsible for the erection of this scheme. The model considers the necessary components of a mass balance and power production, such as inflow, reservoir operation rules, an elevation-volume-area curve, evaporation from the reservoir surface, a tailwater rating curve, hydraulic losses, efficiencies for turbines, generator and transformer, and own electrical energy consumption.

The main purpose of the reservoir operation rule is maximizing the energy production while not leading to excessive reservoir sedimentation and in turn reducing reservoir storage. Thus, the reservoir water surface is lowered either down to sluicing level for sediment transport and removal (case of sluicing)[47], or to a reservoir level that is able to absorb the peak inflows of the rainy season (no sluicing). Subsequent to the sluicing or water surface lowering operation, reservoir filling is initiated. Energy production usually maximizes during reservoir filling due to high inflows towards the end of the rainy season.

Inflow data provided by the DIU in Sudan and used in the simulation model of Tractebel Engineering GmbH were recorded on a daily basis at gauging stations Kubur/El Sofi and Wad El Heliew at the Upper Atbara river and at the Setit river, respectively. Both gauging stations were located upstream of the dam with a negligible intermediate catchment between their locations and the twin dam. These data were also used for the construction planning of the Upper Atbara Dam. For the calculations of drought- and standard-reservoir operations, the daily inflow of the year 2015 was taken as an example for a dry year. For a normal year, an average inflow of the period 1963-2016 was chosen. The latter ensures a reservoir management designed for average conditions and average energy production. To convert the produced energy into US$, a price of 0.2 US$ per kWh was assumed.

## Data availability

Basin-averaged seasonal forecast and reanalysis data that support the findings of this study are available under https://radar.kit.edu/radar/en/dataset/XqHmFKADBYNNKqTi (doi: https://doi.org/10.35097/441). Regarding the reservoir operations at the Upper Atbara Dam, the supporting data are available from Tractebel Engineering GmbH but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of Tractebel Engineering GmbH. Data of the Multivariate ENSO Index Version 2 (MEI.v2) can be accessed via https://www.psl.noaa.gov/enso/mei/ (Supplementary Information 1).

Received: 1 September 2020; Accepted: 27 April 2021
Published online: 19 May 2021

## References

1. Du, L. *et al.* Global patterns of extreme drought-induced loss in land primary production: identifying ecological extremes from rain-use efficiency. *Sci. Total Environ.* **628–629**, 611–620. https://doi.org/10.1016/j.scitotenv.2018.02.114 (2018).
2. Environment Management Group. *Global Drylands: A UN system-wide response* (Tech. Rep, UN Environment Programme WCMC, 2011).
3. United Nations Population Fund. https://www.unfpa.org/data/world-population-dashboard (2020).

4. von Sperling, E. Hydropower in Brazil: overview of positive and negative environmental aspects. *Energy Procedia* **18**, 110–118. https://doi.org/10.1016/j.egypro.2012.05.023 (2012).
5. Jia, G. *et al.* Land-climate interactions. In Shukla, P. *et al.* (eds.) *Climate Change and Land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems* (In press, 2019).
6. Touma, D., Ashfaq, M., Nayak, M. A., Kao, S.-C. & Diffenbaugh, N. S. A multi-model and multi-index evaluation of drought characteristics in the 21st century. *J. Hydrol.* **526**, 196–207. https://doi.org/10.1016/j.jhydrol.2014.12.011 (2015).
7. Masih, I., Maskey, S., Mussia, F. E. F. & Trambauer, P. A review of droughts on the African continent: a geospatial and long-term perspective. *Hydrol. Earth Syst. Sci.* **18**, 3635–3649. https://doi.org/10.5194/hess-18-3635-2014 (2014).
8. Seneviratne, S. I., Donat, M. G., Pitman, A. J., Knutti, R. & Wilby, R. L. Allowable CO2 emissions based on regional and impact-related climate targets. *Nature* **529**, 477–483. https://doi.org/10.1038/nature16542 (2016).
9. Bruno Soares, M., Daly, M. . & Dessai, S. . Assessing the value of seasonal climate forecasts for decision-making. *WIREs: Clim. Chang.* **9**, e523. https://doi.org/10.1002/wcc.523 (2018).
10. Lopez, A. & Haines, S. Exploring the usability of probabilistic weather forecasts for water resources decision-making in the United Kingdom. *Wea. Climate Soc.* **9**, 701–715. https://doi.org/10.1175/WCAS-D-16-0072.1 (2017).
11. Patt, A. & Gwata, C. Effective seasonal climate forecast applications: Examining constraints for subsistence farmers in Zimbabwe. *Global Environ. Change* **12**, 185–195. https://doi.org/10.1016/S0959-3780(02)00013-4 (2002).
12. Rayner, S., Lach, D. & Ingram, H. Weather forecasts are for wimps: why water resource managers do not use climate forecasts. *Climatic Change* **69**, 197–227. https://doi.org/10.1007/s10584-005-3148-z (2005).
13. Glantz, M. The value of a long-range weather forecast for the West African Sahel. *Bull. Amer. Meteorol. Soc.* **58**, 150–158. https://doi.org/10.1175/1520-0477(1977)058<0150:TVOALR>2.0.CO;2 (1977).
14. Merryfield, W. J. *et al.* Current and emerging developments in subseasonal to decadal prediction. *Bull. Am. Meteorol. Soc.* **19–0037**. https://doi.org/10.1175/BAMS-D-19-0037.1 (2020).
15. Fundel, V. . J. ., Fleischhut, N. ., Herzog, S. . M. ., Göber, M. . & Hagedorn, R. Promoting the use of probabilistic weather forecasts through a dialogue between scientists, developers and end-users. *Q. J. Roy. Meteorol. Soc.* **145**, 210–231. https://doi.org/10.1002/qj.3482 (2019).
16. Richardson, D. S. Skill and relative economic value of the ECMWF ensemble prediction system. *Q. J. Roy. Meteorol. Soc.* **126**, 649–667. https://doi.org/10.1002/qj.49712656313 (2000).
17. Wilks, D. S. A skill score based on economic value for probability forecasts. *Meteorol. Appl.* **8**, 209–219. https://doi.org/10.1017/S1350482701002092 (2001).
18. Lopez, A. *et al.* Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings. *Wea. Clim. Extremes,* 100167. https://doi.org/10.1016/j.wace.2018.03.006 (2018).
19. Kar, S. C., Hovsepyan, A. & Park, C. K. Economic values of the APCN multi-model ensemble categorical seasonal predictions. *Meteorol. Appl.* **13**, 267. https://doi.org/10.1017/S1350482706002271 (2006).
20. Bliefernicht, J. *et al.* Quality and Value of Seasonal Precipitation Forecasts Issued by the West African Regional Climate Outlook Forum. *J. Appl. Meteorol. Climatol.* **58**, 621–642. https://doi.org/10.1175/JAMC-D-18-0066.1 (2019).
21. Ben Bouallègue, Z., Pinson, P. & Friederichs, P. Quantile forecast discrimination ability and value. *Q. J. Roy. Meteorol. Soc.* **141**, 3415–3424. https://doi.org/10.1002/qj.2624 (2015).
22. Batté, L. & Déqué, M. Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models: Skill of the ENSEMBLES project multimodel ensemble forecasts. *Tellus A* **63**, 283–299. https://doi.org/10.1111/j.1600-0870.2010.00493.x (2011).
23. Block, P. Tailoring seasonal climate forecasts for hydropower operations. *Hydrol. Earth Sys. Sci.* **15**, 1355–1368. https://doi.org/10.5194/hess-15-1355-2011 (2011).
24. Hamlet, A. F., Huppert, D. & Lettenmaier, D. P. Economic Value of Long-Lead Streamflow Forecasts for Columbia River Hydropower. *J. Water Res. Plann. Manag.* **128**, 91–101. https://doi.org/10.1061/(ASCE)0733-9496(2002)128:2(91) (2002).
25. Elagib, N. A. & Elhag, M. M. Major climate indicators of ongoing drought in Sudan. *J. Hydrol.* **409**, 612–625. https://doi.org/10.1016/j.jhydrol.2011.08.047 (2011).
26. Martins, E. S. P. R. *et al.* A Multimethod Attribution Analysis of the Prolonged Northeast Brazil Hydrometeorological Drought (2012–16). *Bull. Amer. Meteorol. Soc.* **99**, S65–S69. https://doi.org/10.1175/BAMS-D-17-0102.1 (2018).
27. Domínguez-Castro, F., García-Herrera, R. & Vicente-Serrano, S. M. Wet and dry extremes in Quito (Ecuador) since the 17th century. *Int. J. Climatol.* **38**, 2006–2014. https://doi.org/10.1002/joc.5312 (2018).
28. Barbieri, A. F. *et al.* Climate change and population migration in Brazil's Northeast: Scenarios for 2025–2050. *Popul. Environ.* **31**, 344–370. https://doi.org/10.1007/s11111-010-0105-1 (2010).
29. Konapala, G., Mishra, A. K., Wada, Y. & Mann, M. E. Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation. *Nat. Commun.* **11**, 3044. https://doi.org/10.1038/s41467-020-16757-w (2020).
30. Vaghefi, S. A. *et al.* The future of extreme climate in Iran. *Sci. Rep.* **9**, 1464. https://doi.org/10.1038/s41598-018-38071-8 (2019).
31. Sheffield, J. & Wood, E. F. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, IPCC AR4 simulations. *Clim. Dyn.* **31**, 79–105. https://doi.org/10.1007/s00382-007-0340-z (2008).
32. Johnson, S. J. *et al.* SEAS5: The new ECMWF seasonal forecast system. *Geosci. Model Dev.* **12**, 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019 (2019).
33. Crochemore, L., Ramos, M.-H., Pappenberger, F., Andel, S. J. .v & Wood, A. . W. . An Experiment on Risk-Based Decision-Making in Water Management Using Monthly Probabilistic Forecasts. *B. Am. Meteorol. Soc.* **97**, 541–551. https://doi.org/10.1175/BAMS-D-14-00270.1 (2016).
34. Bazile, R., Boucher, M.-A., Perreault, L. & Leconte, R. Verification of ECMWF System 4 for seasonal hydrological forecasting in a northern climate. *Hydrol. Earth Syst. Sci.* **21**, 5747–5762. https://doi.org/10.5194/hess-21-5747-2017 (2017).
35. Jolliffe, I. T., & Stephenson, D. B. Introduction. In Jolliffe, I. T. & Stephenson, D. B. (eds.) *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 1–9 (Wiley, Chichester, UK, 2012).
36. Mylne, K. R. Decision-making from probability forecasts based on forecast value. *Meteorol. Appl.* **9**, 307–315. https://doi.org/10.1017/S1350482702003043 (2002).
37. Turco, M. *et al.* Summer drought predictability over Europe: empirical versus dynamical forecasts. *Environ. Res. Lett.* **12**, 084006. https://doi.org/10.1088/1748-9326/aa7859 (2017).
38. Dutra, E., Di Giuseppe, F., Wetterhall, F. & Pappenberger, F. Seasonal forecasts of droughts in African basins using the Standardized Precipitation Index. *Hydrol. Earth Syst. Sci.* **17**, 2359–2373. https://doi.org/10.5194/hess-17-2359-2013 (2013).
39. Bellprat, O., Guemas, V., Doblas-Reyes, F. & Donat, M. G. Towards reliable extreme weather and climate event attribution. *Nat. Commun.* **10**, 1732. https://doi.org/10.1038/s41467-019-09729-2 (2019).
40. Schepen, A., Everingham, Y. & Wang, Q. J. Coupling forecast calibration and data-driven downscaling for generating reliable, high-resolution, multivariate seasonal climate forecast ensembles at multiple sites. *Int. J. Climatol.* **40**, 2479–2496. https://doi.org/10.1002/joc.6346 (2020).
41. Siegmund, J., Bliefernicht, J., Laux, P. & Kunstmann, H. Toward a seasonal precipitation prediction system for West Africa: Performance of CFSv2 and high-resolution dynamical downscaling. *J. Geophys. Res. Atmos.* **120**, 7316–7339. https://doi.org/10.1002/2014JD022692 (2015).

42. Hersbach, H. *et al.* The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* qj.3803, https://doi.org/10.1002/qj.3803 (2020).
43. McKee, T. B., Doesken, N. J. & Kleist, J. The relationship of drought frequency and duration to time scales. In *Eighth Conference on Applied Climatology*, 179-184 (American Meteorological Society, Anaheim, CA., 1993).
44. Vicente-Serrano, S. M., Beguería, S. & López-Moreno, J. I. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J. Clim.* **23**, 1696–1718. https://doi.org/10.1175/2009JCLI2909.1 (2010).
45. Efron, B. & Gong, G. A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *Am. Stat.* **37**, 36–48. https://doi.org/10.1080/00031305.1983.10483087 (1983).
46. Wang, Q. Q. *et al.* Displacement monitoring of upper Atbara dam based on time series InSAR. *Surv. Rev.* **1–12**. https://doi.org/10.1080/00396265.2019.1643529 (2019).
47. Schmutz, S. & Moog, O. Dams: Ecological Impacts and Management. *Riverine Ecosystem Management* **111–127** (Springer International Publishing, Cham). https://doi.org/10.1007/978-3-319-73250-3_6 (2018).
48. Wessel, P. & Smith, W. H. F. A global, self-consistent, hierarchical, high-resolution shoreline database. *J. Geophys. Res. Solid Earth* **101**, 8741–8743. https://doi.org/10.1029/96JB00104 (1996).
49. Lehner, B. & Grill, G. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrol. Process.* **27**, 2171–2186. https://doi.org/10.1002/hyp.9740 (2013).


## Acknowledgements

This work was supported by the funding from the German Federal Ministry of Education and Research (BMBF) for the SaWaM project (*Seasonal Water Resources Management: Regionalized Global Data and Transfer to Practice*) under the financial assistance agreement No 02WGR1421. We would like to acknowledge ECMWF and the Copernicus Climate Service for providing the SEAS5 forecasts. We particularly thank our local project partners in all seven study regions for intensive discussions about regional climate-, hydrology- and water management system understanding. We further gratefully acknowledge the provision of local data for the Upper Atbara Dam Complex by the Dam Implementation Unit (DIU) in Sudan. Finally, we cordially thank Dr. Calum Brown (KIT/IMK-IFU) for his very welcome comments, corrections, and language editing.

## Author contributions

T.P., C.L. and H.K. developed the research concept. T.P. and C.L. conceived the analysis. C.L. managed and provided the used data and T.P. conducted the analysis. B.D. performed the calculations of reservoir management and energy production. T.P. summarized the results. All authors discussed and reviewed the manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interest

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-89564-y.

**Correspondence** and requests for materials should be addressed to T.C.P.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Seasonal Forecasts offer Economic Benefit for Hydrological Decision Making in Semi-Arid Regions

**Tanja C. Portele[1,*], Christof Lorenz[1], Berhon Dibrani[2], Patrick Laux[1,3,+], Jan Bliefernicht[3,+], and Harald Kunstmann[1,3,+]**

[1]Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research - Atmospheric Environmental Research (IMK-IFU), Garmisch-Partenkirchen, Germany
[2]Tractebel Engineering GmbH, Bad Vilbel, Germany
[3]University of Augsburg, Institute of Geography, Augsburg, Germany
[*]tanja.portele@kit.edu
[+]these authors contributed equally to this work

## SUPPLEMENTARY INFORMATION

### Supplementary Results

As stated in the main text, the Chira basin was excluded in the analysis of increasing occurrence of extreme events in the recent decades (Fig. 2), as it does not show significant trends there. Instead, for the Chira basin, elevated frequencies of warm and wet months are significantly correlated to El Niño (Multivariate ENSO Index, MEI, Version 2[1] $> 1$), and increased frequencies of cold months significantly to La Niña (MEI$< -1$) at $\alpha = 0.05$ (Supplementary Fig. S1; also see Supplementary Methods for MEIv2). Except for the Karun basin with significant correlation for the frequencies of $T$ <Q20 and MEI$> 1$ (El Niño), no temperature or precipitation extreme event of the other study basins was significantly correlated with either El Niño or La Niña.

### Supplementary Figure

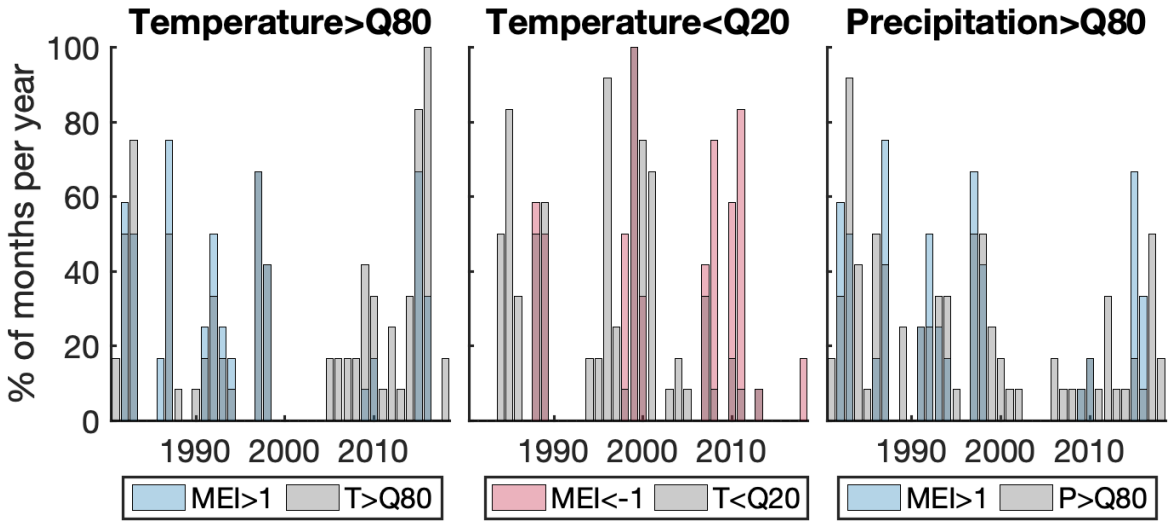


**Figure S1. Relative frequency of climate variables for the Catamayo-Chira basin from 1981-2018.** Grey bars show the relative frequency of ERA5 mean temperature $>$ Q80 (left) and $<$ Q20 (middle), and total precipitation $>$ Q80 (right). Blue and red bars depict the relative frequency of multivariate ENSO index MEIv2 $> 1$ (El Niño) and $< -1$ (La Niña), respectively.

## Supplementary Methods

### Drought indices

The Standardized Precipitation Index (SPI)[2] is calculated by the following steps:

1. Define the monthly precipitation dataset for an at least 30-year period, separately for the ensemble forecast and the reference.

2. Aggregate each dataset over the selected timescales, e.g., over $i = 1$, 3, 4 or 6 months, in a forecast - not in a retrospective - manner. This is done movingly in the sense that for each month new aggregated values are calculated from the current and the next $i-1$ months. That means for SPI6 of June, values from June to November are accumulated. For the forecasts, this implies using lead 0-5 from the forecast issued in June for each ensemble member.

3. Determine the cumulative probability distributions of the respective aggregated values for each start month including all available years for the reference, and all available years and ensemble members for the forecast. For not being dependent on the fit of parametric statistical distributions, we chose to use the empirical distribution for the calculation of SPI. Please note, the empirical cumulative probability distributions are separately calculated for reference and forecasts, instead of merging them, to allow the use of uncorrected, probably biased forecasts.

4. Estimate the inverse normal of the probability of aggregated values to determine the SPI value for the each data point, i.e., the derived deviation for a standard normally distributed probability density with zero mean.

With separate distributions for reference and forecasts, current values are only compared within the system's (reference or forecast) distribution. A chosen threshold of SPI$< -1$ thus defines a system's specific quantile value, representative of one negative standard deviation from the system's mean. Accordingly, standard bias correction approaches like linear scaling (correction of the mean) or quantile mapping (correction of absolute quantile values) at the same temporal (monthly) and spatial (basin-mean) resolutions have no effect.

For the Standardized Precipitation Evapotranspiration Index (SPEI), a simple water balance between precipitation ($P$) and potential evapotranspiration (PET) is calculated to derive the water deficit or surplus $D$. According to Vicente-Serrano et al., 2010[3], for the use of PET in the drought index, the method to calculate PET is not critical. Therefore, we followed their calculation of monthly PET (mm) with the simple approach by Thornthwaite 1948[4], requiring only data on monthly-mean temperature ($T$ in °C):

$$\mathrm{PET} = 16K\left(\frac{10T}{I}\right)^{m} \tag{S1}$$

For the temperature-dependent heat index $I$, the coefficient $m$ depending on $I$ and the correction coefficient $K$ depending on latitude and month, we refer to Vicente-Serrano et al., 2010[3]. The water deficit or surplus is then calculated for each year, month and lead time, separately for the forecasts and reference, as

$$D = P - \mathrm{PET}. \tag{S2}$$

Similarly to SPI, the derived values of $D$ are then aggregated over different timescales and above steps 2-4 are applied.

### Bootstrapping algorithm

As stated in the main text, the bootstrapping algorithm is applied to estimate the uncertainty of the hit rate $H$, false alarm rate $F$ and the potential economic value $PEV$ that may come along due to sampling errors of the considered extreme events. The applied algorithm for pairwise data is based on Bliefernicht et al., 2019[5] and Efron and Gong, 1983[6]. First, the bootstrap sample size $m$ (here 1000) is defined. For each event ($T > Q80$, $P > Q80$, SPI$< -1$ and SPEI$< -1$), each target region and each probability threshold $p_{th}$, the following steps are applied to calculate the bootstrapped $H$, $F$ and $PEV$:

1. Define the binary forecast-reference pairs of sample size $n$ each. The reference time series $\mathbf{x} = (x_1, x_2, ..., x_n)$ becomes binary according to the event threshold $x_{th}$ (e.g., $Q80$), and the event-based forecast probability $\mathbf{p} = (p_1, p_2, ..., p_n)$ is transformed to a binary array according to the current probability threshold $p_{th}$.

2. Resample the pairs of binary time series by randomly selecting with replacement (one sample pair can also be selected several times) $n$ new binary forecast-reference pairs.

3. Calculate $H$, $F$ and $PEV$ for each new binary forecast-reference pair.

4. Repeat steps 2 and 3 $m$-times to obtain $m$ bootstrapped versions of $H$, $F$ and $PEV$.

5. To assess the robustness of forecast thresholds[7], calculate the confidence intervals for $PEV$ based on the 10 % and 90 % quantiles of the $m$ bootstrap samples of $PEV$.

### Mann-Kendall trend significance test

The Mann-Kendall test is a non-parametric test for trend significance, that does not require any particular distribution of the tested timeseries[8]. The null hypothesis of the test is the absence of consistently increasing or decreasing trend in a timeseries $\mathbf{x} = (x_1, x_2, ..., x_n)$. The test analyzes differences in signs between earlier and later data points. The Mann-Kendall statistic $S$ is defined as

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \mathrm{sgn}(x_j - x_i), \tag{S3}$$

where $n$ is the length of the timeseries and sgn denotes the sign function that allows values of $+1$, 0 and $-1$. For increasing or decreasing trends, the value of the Mann-Kendall statistic $S$ should be highly positive or negative, respectively. To statistically test the trend, the probability associated with the Mann-Kendall statistic $S$ is required. A normal distribution of the Mann-Kendall statistic $S$ can be assumed for datasets with more than 10 sample points and including less equal values, i.e. ties. The normalized test statistic, the $Z$-value associated with S, is calculated as

$$Z = \begin{cases} \dfrac{(S-1)}{\sqrt{(VAR(S))}}, & \text{if S>0} \quad \text{(S4a)} \\ 0, & \text{if S=0} \quad \text{(S4b)} \\ \dfrac{S+1}{\sqrt{VAR(S)}}, & \text{if S<0,} \quad \text{(S4c)} \end{cases}$$

where the variance of the Mann-Kendall statistic $VAR(S)$ for a non-tied dataset is defined as

$$VAR(S) = \frac{1}{18} \cdot \{n(n-1)(2n+5)\}. \tag{S5}$$

Finally the probability ($p$-value) associated with the $Z$-value is calculated using the standard normal cumulative distribution function for a two-tailed test. Based on the chosen significance level $\alpha$, typically 0.05, the null hypothesis is accepted or rejected. If the $p$-value of the test is less than $\alpha$, the test rejects the null hypothesis, i.e., the test signifies the presence of trend in $\mathbf{x}$. Otherwise, if the $p$-value is greater than $\alpha$, the null hypothesis of trend absence is accepted.

### Multivariate ENSO Index Version 2 (MEIv2)

The phases of El Niño Southern Oscillation (ENSO) are described by the Multivariate El-Niño-Southern-Oscillation (ENSO) Index Version 2 (MEIv2)[1] that is based on the principal-component analysis of standardized anomalies of sea level pressure, sea surface temperature, 10-m zonal and meridional wind, and outgoing longwave radiation. For consistency with the monthly analysis, the two-month MEIv2 product (i. e., data for December-January, January-February, ... November-December) was linearly interpolated to monthly values. Data are available since 1979.

## References


1. Zhang, T. *et al.* Towards Probabilistic Multivariate ENSO Monitoring. *Geophys. Res. Lett.* **46**, 10532–10540, DOI: 10.1029/2019GL083946 (2019).

2. McKee, T. B., Doesken, N. J. & Kleist, J. The relationship of drought frequency and duration to time scales. In *Eighth Conference on Applied Climatology*, 179–184 (American Meteorological Society, Anaheim, CA., 1993).

3. Vicente-Serrano, S. M., Beguería, S. & López-Moreno, J. I. A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *J. Clim.* **23**, 1696–1718, DOI: 10.1175/2009JCLI2909.1 (2010).

4. Thornthwaite, C. W. An Approach toward a Rational Classification of Climate. *Geogr. Rev.* **38**, 55–94, DOI: 10.2307/210739 (1948).

5. Bliefernicht, J. *et al.* Quality and Value of Seasonal Precipitation Forecasts Issued by the West African Regional Climate Outlook Forum. *J. Appl. Meteorol. Clim.* **58**, 621–642, DOI: 10.1175/JAMC-D-18-0066.1 (2019).

6. Efron, B. & Gong, G. A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation. *Amer. Stat.* **37**, 36–48, DOI: 10.1080/00031305.1983.10483087 (1983).

7. Lopez, A. *et al.* Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings. *Wea. Clim. Extrem.* 100167, DOI: 10.1016/j.wace.2018.03.006 (2018).

8. Kendall, M. G. *Rank correlation methods* (Charles Griffin, London, 1975), 4th edn.

# 3 Article II

## Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-term Reference Forecast Product for the Water Sector in Semi-arid Regions

---

Lorenz, C., Portele, T.C., Laux, P., and Kunstmann, H (2021). Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-term Reference Forecast Product for the Water Sector in Semi-arid Regions. *Earth Syst. Sci. Data* **13**, 2701–2722. https://doi.org/10.5194/essd-13-2701-2021



---

Earth Syst. Sci. Data, 13, 2701–2722, 2021
https://doi.org/10.5194/essd-13-2701-2021


Open Access Earth System Science Data

# Bias-corrected and spatially disaggregated seasonal forecasts: a long-term reference forecast product for the water sector in semi-arid regions

**Christof Lorenz**[1], **Tanja C. Portele**[1], **Patrick Laux**[1,2], **and Harald Kunstmann**[1,2]

[1]Karlsruhe Institute of Technology (KIT), Campus Alpin, Institute of Meteorology and Climate Research – Atmospheric Environmental Research (IMK-IFU), Kreuzeckbahnstr. 19, 82467 Garmisch-Partenkirchen, Germany
[2]Augsburg University, Institute of Geography, Alter Postweg 118, 86159 Augsburg, Germany

**Correspondence:** Christof Lorenz (christof.lorenz@kit.edu)

Received: 2 July 2020 – Discussion started: 27 October 2020
Revised: 15 April 2021 – Accepted: 16 April 2021 – Published: 15 June 2021

**Abstract.** Seasonal forecasts have the potential to substantially improve water management particularly in water-scarce regions. However, global seasonal forecasts are usually not directly applicable as they are provided at coarse spatial resolutions of at best 36 km and suffer from model biases and drifts. In this study, we therefore apply a bias-correction and spatial-disaggregation (BCSD) approach to seasonal precipitation, temperature and radiation forecasts of the latest long-range seasonal forecasting system SEAS5 of the European Centre for Medium-Range Weather Forecasts (ECMWF). As reference we use data from the ERA5-Land offline land surface rerun of the latest ECMWF reanalysis ERA5. Thereby, we correct for model biases and drifts and improve the spatial resolution from 36 km to 0.1°. This is performed for example over four predominately semi-arid study domains across the world, which include the river basins of the Karun (Iran), the São Francisco River (Brazil), the Tekeze–Atbara river and Blue Nile (Sudan, Ethiopia and Eritrea), and the Catamayo–Chira river (Ecuador and Peru). Compared against ERA5-Land, the bias-corrected and spatially disaggregated forecasts have a higher spatial resolution and show reduced biases and better agreement of spatial patterns than the raw forecasts as well as remarkably reduced lead-dependent drift effects. But our analysis also shows that computing monthly averages from daily bias-corrected forecasts particularly during periods with strong temporal climate gradients or heteroscedasticity can lead to remaining biases especially in the lowest- and highest-lead forecasts. Our SEAS5 BCSD forecasts cover the whole (re-)forecast period from 1981 to 2019 and include bias-corrected and spatially disaggregated daily and monthly ensemble forecasts for precipitation, average, minimum, and maximum temperature as well as for shortwave radiation from the issue date to the next 215 d and 6 months, respectively. This sums up to more than 100 000 forecasted days for each of the 25 (until the year 2016) and 51 (from the year 2017) ensemble members and each of the five analyzed variables. The full repository is made freely available to the public via the World Data Centre for Climate at https://doi.org/10.26050/WDCC/SaWaM_D01_SEAS5_BCSD (Domain D01, Karun Basin (Iran), Lorenz et al., 2020b), https://doi.org/10.26050/WDCC/SaWaM_D02_SEAS5_BCSD (Domain D02: São Francisco Basin (Brazil), Lorenz et al., 2020c), https://doi.org/10.26050/WDCC/SaWaM_D03_SEAS5_BCSD (Domain D03: basins of the Tekeze–Atbara and Blue Nile (Ethiopia, Eritrea, Sudan), Lorenz et al., 2020d), and https://doi.org/10.26050/WDCC/SaWaM_D04_SEAS5_BCSD (Domain D04: Catamayo–Chira Basin (Ecuador, Peru), Lorenz et al., 2020a). It is currently the first publicly available daily high-resolution seasonal forecast product that covers multiple regions and variables for such a long period. It hence provides a unique test bed for evaluating the performance of seasonal forecasts over semi-arid regions and as driving data for hydrological, ecosystem or climate impact models. Therefore, our forecasts provide a crucial contribution for the disaster preparedness and, finally, climate proofing of the regional water management in climatically sensitive regions.

Published by Copernicus Publications.

## 1 Introduction

Since the launch of seasonal hydrometeorological forecasts, it is widely agreed that subseasonal to seasonal forecasts offer the promise of improved hydrological management strategies (Rayner et al., 2005). Various studies showed high potential when such information is used for planning the harvests from subsistence farmers (Patt et al., 2005), predicting and monitoring drought conditions in data-sparse regions (Dutra et al., 2013; Yuan et al., 2011), driving hydrological models (Thober et al., 2015), proactive drought planning (Lemos et al., 2002), predicting heavy rainfall events (Tall et al., 2012), managing irrigated agriculture (Ritchie et al., 2008), operating hydropower (Block, 2011), or predicting high and low river flow during El Niño (Emerton et al., 2019). Washington et al. (2006) even state that for the African continent, the adaptation to current (seasonal) climate anomalies through operational decision-making may reduce vulnerability to climate change. It is hence obvious that these promising perspectives led to the establishment of many global initiatives and forecast products like the Long-Range Forecast Multi-Model Ensemble from the World Meteorological Organization (https://www.wmolc.org, last access: 2 June 2021), the C3S multi-model ensemble from Copernicus (https://climate.copernicus.eu/seasonal-forecasts, last access: 2 June 2021), the North American Multi-Model Ensemble (NMME, https://www.cpc.ncep.noaa.gov/products/NMME/, last access: 2 June 2021) and the recalibrated forecasts from the International Research Institute for Climate and Society (https://iri.columbia.edu/our-expertise/climate/forecasts/seasonal-climate-forecasts/, last access: 2 June 2021). On the regional scale, the Intergovernmental Authority on Development (IGAD) Climate Prediction and Applications Centre (ICPAC) has developed operational seasonal forecasts for the IGAD region across northeast Africa (https://www.icpac.net/seasonal-forecast/, last access: 2 June 2021), while forecasts for South America were developed within the EURO–Brazilian Initiative for Improving South American seasonal forecasts (EUROBRISA, http://eurobrisa.cptec.inpe.br, last access: 2 June 2021). Nevertheless, many water managers are still unaware of most sources of seasonal climate forecasts (Bolson et al., 2013) or claim that the forecasts are not reliable enough for improving the decision-making (Rayner et al., 2005). Hence, if seasonal forecasts are to be used effectively, it is important that, along with science advances, an effort is made to develop, communicate and apply these forecasts appropriately (White et al., 2017). Patt and Gwata (2002) defined six constraints that currently limit the usefulness of seasonal forecasts particularly in developing countries: credibility, legitimacy, scale, cognitive capacity, procedural and institutional barriers, and available choices.

Some of these constraints are based on societal aspects. They hence have to be overcome through the adaptation of seasonal forecasts to accommodate for variations in the interpretive abilities of decision makers and other potential user groups (Hartmann et al., 2002). It is particularly the scale constraint (which refers to the inconsistency between the global forecast models and regional conditions) that can be addressed through post-processing techniques. Furthermore, evidence of bias, e.g., in global circulation model (GCM) and regional circulation model (RCM) precipitation data, has prompted many investigators to avoid direct use of climate model precipitation outputs for hydrological climate change impact analysis (Teutschbein and Seibert, 2013). Among other factors, one can loosely categorize these biases into systematic model errors (e.g., Xue et al., 2013; Magnusson et al., 2013) and drifts (e.g., Hermanson et al., 2018), as well as issues due to the coarse resolution of the global forecasts which prevent the models from properly representing local features in regions with complex orography (Manzanas et al., 2018a). For seasonal forecasts, particularly the model drifts are a crucial issue with their forecast horizon up to 7 months as they lead to statistical inconsistencies between forecasts from different issue dates.

Since these shortcomings of seasonal or longer-term predictions are known for a long time, there is a range of methods and techniques for correcting model biases and drifts as well as to improve the spatial resolution. For downscaling, we generally distinguish between dynamical and empirical-statistical approaches. While the dynamical methods using a RCM are computationally highly expensive (e.g., Manzanas et al., 2018b), empirical-statistical techniques usually require reliable reference data, which are often not available, particularly in data-sparse regions. Nevertheless, due to their lower computational demand and relatively simple implementation in operational systems, there have been significant developments in the empirical-statistical correction approaches in the recent years. One of the most widely used methods is the so-called *bias correction and spatial disaggregation* (BCSD, Wood et al., 2002) which was developed for downscaling climate model outputs to a higher spatial resolution. Since its introduction, there have been numerous adjustments and changes to the classic BCSD approach. One can distinguish between parametric and non-parametric techniques (e.g., Lafon et al., 2013; Crochemore et al., 2016). Abatzoglou and Brown (2012) and Ahmed et al. (2013) reversed the order in which the forecasts are bias-corrected and spatially disaggregated, which they refer to as SDBC. Thrasher et al. (2012) applied the BCSD to daily data, Voisin et al. (2010) used rank-based scaling factors between the forecasts and a random reference ensemble to allow for different daily precipitation patterns, and Hwang and Graham (2013) replaced the interpolation-based spatial disaggregation with a stochastic approach to preserve observed local rainfall character-

istics, while Vandal et al. (2019) recently combined BCSD with machine-learning methods. Besides the adjustments to the univariate BCSD technique, recent publications also aim at multivariate extensions (e.g., Cannon, 2018), which allow for the joint correction of different variables (e.g., temperature and precipitation) or values at different locations. There have also been numerous studies where different downscaling approaches have been compared (Tryhorn and DeGaetano, 2011; Chen et al., 2013; Gutmann et al., 2014).

Despite all these efforts, most studies focus on selected regions and only short periods of time. Furthermore, the corrected data are usually not made available to the public. In this study, we therefore present a comprehensive dataset which contains bias-corrected and spatially disaggregated seasonal (re-)forecasts for daily precipitation, temperature and radiation from 1981 to 2019 for four semi-arid domains in Brazil (Rio São Francisco basin), Ecuador and Peru (Catamayo–Chira Basin), Sudan and Ethiopia (basins of the Tekeze–Atbara river and Blue Nile), and Iran (Karun Basin). Our study regions are marked by a strong dependency on water, food and energy supply from water reservoirs and, hence, on a sustainable multipurpose water resources management. All regions have been hit by several severe droughts and floods particularly during the last years (e.g., Elagib and Elhag, 2011; Marengo et al., 2018; Martins et al., 2018). Moreover, the Blue Nile Basin, which will undergo tremendous changes due to the construction of the Grand Ethiopian Renaissance Dam (GERD) near the Ethiopian–Sudanese border, has been controversially debated in the public and the scientific literature (Kidus, 2019) as the filling and operation of the GERD will change downstream flow patterns significantly (e.g., Wheeler et al., 2020). This underlines the urgent need for longer-term forecasts to mitigate the impacts of climatically extreme events and improve the regions' disaster preparedness (e.g., Tall et al., 2012) as well as improve the regional water management, especially in transboundary catchments (e.g., Gerlitz et al., 2020).

Reanalysis data of the ERA5-Land (ECMWF, 2019) rerun of the land component of ERA5 climate reanalyses (Hersbach et al., 2018) are used as reference for applying the BCSD on raw forecasts from the seasonal forecasting system SEAS5 (Johnson et al., 2019) of the European Centre for Medium-Range Weather Forecasts (ECMWF). While reanalyses have clear limitations, they still provide the most comprehensive and reliable (and sometimes only) source of hydrometeorological information in such data-sparse regions.

Our final product provides a temporally and spatially consistent high-resolution dataset that can be used for assessing the skill of state-of-the-art seasonal forecasts, e.g., for drought forecasting, for driving hydrological or ecosystem models, as decision support for the regional water management, or as a comprehensive repository for teaching the use of state-of-the-art seasonal forecast products to water managers, decision makers, and other potential end users. It is made freely available to the public through the World Data Center for Climate (WDCC), which is hosted by the German Climate Computing Center (DKRZ) in Hamburg, Germany. Therefore, our approach and published dataset address several of the abovementioned constraints of seasonal forecasts and hence provide a significant contribution towards improving the usefulness of such information and practice transfer particularly in developing countries. It therefore marks a large step towards a more sustainable and timely planning of the regional water management and, hence, the adaptation to a changing climate.

## 2 Data and study areas

### 2.1 Reference data

Daily reference precipitation, average, minimum and maximum temperature at 2 m and surface solar radiation at a high spatial resolution of 0.1° is obtained from the ERA5-Land offline land surface rerun of ECMWF's latest reanalysis product ERA5.

ERA5 is currently produced within the Copernicus Climate Change Service (C3S) and is the successor of the older ERA-Interim reanalysis, which has been extensively used in numerous hydrological and hydrometeorological studies (e.g., Lorenz and Kunstmann, 2012; Lorenz et al., 2014, 2018). In contrast to ERA-Interim, ERA5 is based on the Integrated Forecasting System (IFS) cycle 41r2 which is run at a higher resolution of 31 km and is planned to cover the whole period from 1950 to 5 d before the present, which allows its usage in near-real-time and operational applications. It has been reported in numerous studies that the performance of ERA5 is superior compared to ERA-Interim (Albergel et al., 2018; Urraca et al., 2018; Mahto and Mishra, 2019). Besides improvements in the underlying model systems, this can also be attributed to the huge number of assimilated in situ, satellite and snow observations. ECMWF states that this number has increased from approximately 0.75 million per day on average in 1979 to around 24 million per day until 2018 (Hersbach et al., 2019).

ERA5-Land uses atmospheric forcing from the ERA5 reanalysis to consistently estimate hourly land surface variables at an enhanced spatial resolution of 9 km. While no observations are directly assimilated during the production of ERA5-Land, the millions of observations that are used for constraining the atmospheric forcing data from ERA5 have an indirect influence on the estimated land surface parameters. Furthermore, air temperature, air humidity and pressure are corrected to account for the altitude difference between the spatial resolution of the grids of ERA5 and ERA5-Land, respectively (ECMWF, 2019, 2020).

### 2.2 SEAS5 seasonal forecasts

The fifth generation of ECMWF's seasonal forecasting system has been operational since November 2017. The mod-

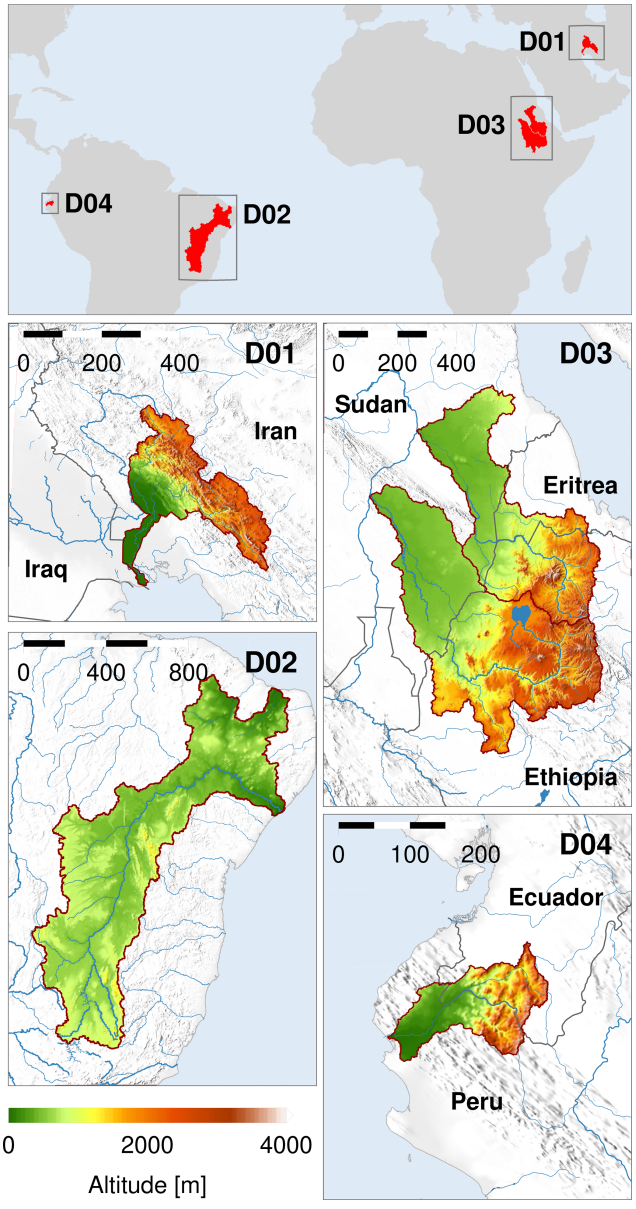


**Figure 1.** Overview of the four study domains: the Karun Basin (KA, Iran, D01, top left), the Extended São Francisco Basin (SF, Brazil, D02, bottom left), the Tekeze–Atbara (to the north) and Blue Nile (to the south) basins (TA and BN, Sudan, Ethiopia and Eritrea, D03, top right), and the Catamayo–Chira Basin (CC, Ecuador and Peru, D04, bottom right). The distance scales in the top left corner of the maps are given in units of kilometers. The basin topography is based on the high-resolution ETOPO1 Global Relief Model (Amante and Eakins, 2009), while coastlines, rivers and political borders are taken from the Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG, Wessel and Smith, 1996). The basin boundaries are based on the HydroSHEDS dataset (Lehner and Grill, 2013) with some slight modifications and adjustments for ensuring the consistency with boundary definitions from local authorities and stakeholders from the study regions.

eled variables are provided on a reduced O320 Gaussian grid, which corresponds to a spatial resolution of approximately 36 km. SEAS5 covers the period from 1981 to the present with forecasts issued on the first of each month. During the re-forecast period from 1981 until 2016, ECMWF provides 25 ensemble members, while this number increased to a total of 51 ensemble members in 2017. The seasonal forecasts are initialized with atmospheric conditions from ERA-Interim until 2016 and the ECMWF Operational Analysis since 2017. Highlights of SEAS5 compared to previous versions include a marked improvement in sea surface temperature drift, especially in the tropical Pacific, and improvements in the prediction skill of Arctic sea ice (Haiden et al., 2018).

With its release in 2017, so far only a limited number of studies exists discussing the performance of SEAS5. For a case study over Indonesia, Ratri et al. (2019) report that after bias-correcting the seasonal forecasts towards the Southeast Asia observations (SA-OBS van den Besselaar et al., 2017) gridded rainfall product, predominantly positive predictive skill during the first 2 forecast months is achieved. Recently, Gubler et al. (2019) performed a comprehensive performance analysis of bias-corrected SEAS5 forecasts against homogenized station data across South America. They found that, in general, the prediction skill of temperature forecasts is higher than the skill of precipitation forecasts and that particularly regions which are influenced by Niño 3.4 show higher skills. The highest prediction performance can be observed, amongst some other areas, over the highlands of Ecuador and the northeastern part of Brazil. This result is beneficial for our forecast product as these South American regions include two of our study domains.

### 2.3 The four semi-arid study areas

We apply the bias-correction and spatial disaggregation of the global seasonal forecasts over four domains of different size and orographic complexity which contain five semi-arid river basins: the Karun Basin (Domain D01, Iran), the Extended São Francisco Basin (D02, Brazil), the basins of the Tekeze–Atbara river and the Blue Nile (D03, Ethiopia and Sudan), and the Catamayo–Chira Basin (D04, Ecuador and Peru). Main characteristics and the location of the domains and basins are shown in Table 1 and Fig. 1, respectively. It should be noted that we label the domains with numbers from D01 to D04 while the basins are labeled with two-letter abbreviations (see Table 1). This allows us to easily add further domains and basins in the future.

All domains and basins are characterized by a semi-arid climate with an extended dry period and one rainy season. The headwaters are located in mountainous areas and exhibit relatively high seasonal precipitation amounts (e.g., due to convective effects), while the downstream conditions are mainly arid. The Karun has its source in Zard-Kuh mountain with an altitude of more than 4000 m, which is located

**Table 1.** Overview and basic characteristics of the five river basins with climate data from ERA5-Land. In addition to the yearly mean temperature, the yearly temperature range is given by the monthly minimum and maximum temperature in brackets. The main 4 months of the rainy season are provided, and the respective seasonal precipitation is given as a share of the total annual precipitation. In brackets, the percentage of precipitation of the 6-month rainy season (1 month prior to and after the main 4 months) is also provided.

| | Karun<br>KA (D01) | Ext. São Francisco<br>SF (D02) | Blue Nile<br>BN (D03) | Tekeze–Atbara<br>TA (D03) | Catamayo–Chira<br>CC (D04) |
|---|---|---|---|---|---|
| Area [km$^2$] | 67 313 | 740 820 | 308 197 | 205 097 | 17 761 |
| Annual rainfall [mm] | $640 \pm 128$ | $858 \pm 196$ | $1336 \pm 132$ | $727 \pm 95$ | $1666 \pm 399$ |
| Mean temperature [K] | $289 \pm 0.8$ | $298 \pm 0.5$ | $297 \pm 0.5$ | $298 \pm 0.5$ | $293 \pm 0.4$ |
| Min. temperature [K] | $265 \pm 0.5$ | $289 \pm 0.9$ | $287 \pm 0.8$ | $286 \pm 0.8$ | $287 \pm 0.6$ |
| Max. temperature [K] | $310 \pm 1.5$ | $307 \pm 0.5$ | $308 \pm 0.6$ | $310 \pm 0.7$ | $301 \pm 0.6$ |
| Rainy season | DJFM | DJFM | JJAS | JJAS | JFMA |
| Seasonal precipitation [%] | 73 (94) | 60 (85) | 73 (90) | 83 (92) | 65 (81) |

in the Zagros Mountains in the southwestern part of Iran. It is the main source of water for irrigation, hydropower generation and drinking water supply for the Khuzestan Province and its capital Ahvaz with a population of more than 1 million. The much longer Rio São Francisco originates in the Canastra mountain range in the state of Minas Gerais and enters, after more than 3000 km, the Atlantic Ocean. In order to transfer water to the water-scarce states of Ceará, Pernambuco, Paraíba and Rio Grande do Norte in the Brazilian northeast, it was decided in 2005 to conduct a water division project and extend the natural basin of the Rio São Francisco. This Extended São Francisco Basin has a drainage area of more than 700 000 km$^2$, and the water is heavily used for irrigation and hydropower generation. The sources for both the Blue Nile and the Tekeze–Atbara river are located in the Ethiopian Highlands, with altitudes of more than 4000 m. After they pass the Ethiopian–Sudanese border, they flow through mainly flat and dry areas. The Blue Nile joins the White Nile in the Sudanese capital Khartoum while the Tekeze–Atbara river enters the main Nile near the city of Atbara. Together, the Blue Nile and Tekeze–Atbara river deliver approx. 80 % of the mean annual discharge of the main Nile, which underlines the importance of these two tributaries for Ethiopia, but also the downstream countries of Sudan and Egypt. The Catamayo–Chira river has its source in the Andes at an altitude of more than 3000 m. After it passes the Ecuadorian–Peruvian border, it enters the Poechos Reservoir, which is mainly used for water storage, irrigation, hydropower generation and flow regulation across the Chira valley.

While the average temperatures do not change substantially from year to year across all study domains, standard deviations of up to 25 % of the total annual rainfall (e.g., in the Catamayo–Chira Basin) indicate a highly variable amount of incoming freshwater resources, which underlines the necessity for longer-term forecasts. Moreover, particularly in the Karun and Catamayo–Chira basins, there is a very strong climatic and elevation gradient from the headwater to the tailwater within only a few hundred kilometers. For obtaining realistic estimates of precipitation and temperatures for these mountainous headwaters, we hence need models and datasets with a reasonable spatial resolution capable of describing the climate dynamics in such complex terrain. Furthermore, the basins are heavily managed, including many reservoirs which are used for maintaining water security and also electrical power supply throughout the year, and all basins suffer from a lack of continuous in situ observations.

This combination of dependency from incoming water resources and lack of observations is particularly worrying in the context of climate change. Almost all of our study regions have been hit by severe extreme events during the recent years and are assumed to experience an increase in the frequency and severity of droughts and floods in the coming years (e.g., Marengo et al., 2012; Torres et al., 2017; Andrade et al., 2021). For southern Iran, Vaghefi et al. (2019) project a climate of extended dry periods interrupted by intermittent heavy rainfalls, which is a recipe for increasing the chances of floods. Accordingly, the first months of the rainy season 2017/2018 had the lowest ever recorded amounts of precipitation, which then led to water shortages and even societal unrest during the coming months. Only 1 year later, exceptionally heavy rainfall events during March and April 2019 caused severe flooding in at least 26 of Iran's 31 provinces. Similarly, the Northeast Brazil region, which also includes the Rio São Francisco Basin, suffered from a prolonged drought period from 2012 to 2016 (Martins et al., 2018) or, according to Marengo et al. (2018), even from 2010 to 2016. Elagib and Elhag (2011) report that there has been a drastic increase in temperatures over the Sudan in line with a significant decline of rainfall over the northern half of the country. Masih et al. (2014) further state that there is a clear need for increased and integrated efforts in drought mitigation to reduce the negative impacts of droughts anticipated in the future across the African continent. Finally, Domínguez-Castro et al. (2018) analyzed wet and dry extremes in Ecuador and reported that droughts have intensified in frequency and length since the middle of the 20th century.

Besides these climatically challenging conditions, the Tekeze–Atbara, Blue Nile and Catamayo–Chira basins are transboundary river basins. The headwaters of the Tekeze–Atbara and Blue Nile are located in Ethiopia and contribute, together with the Baro–Akobo–Sobat River, about 85 % of the Nile water (Yitayew and Melesse, 2011). Both rivers cross the Ethiopian–Sudanese border after several hundred kilometers. The Blue and White (which comes from the south) Nile merge in the Sudanese capital Khartoum while the Tekeze–Atbara enters the main Nile near the city of Atbara. In both Ethiopia and Sudan, reservoirs of the two rivers exist or are currently under construction. Similarly, the Catamayo–Chira river originates in Ecuador, is dammed in the Poechos reservoir right after the Ecuadorian–Peruvian border and finally flows into the Pacific Ocean. While the Catamayo–Chira Basin is jointly managed by Ecuadorian and Peruvian authorities, coordinated management of water-related infrastructure across the international borders of the Blue Nile and Tekeze–Atbara basins rarely exists (Wheeler et al., 2018). A recent study has analyzed the potential of a joint transboundary water management for hydro-economic sectors particularly through the integration of the newly built Grand Ethiopian Renaissance Dam (GERD) (Digna et al., 2018).

Hence, the four study domains not only provide an optimal test bed for the performance of seasonal forecasts in semi-arid regions but also mark regions for which a sustainable regional water management is crucial, particularly due to the increase in the frequency and severity of climatic extreme events and the transboundary challenges.

## 3 Methods

The bias-correction and spatial-disaggregation approach (BCSD, Wood et al., 2002) is a widely used two-step technique for calibrating, e.g., climate forecasts towards any kind of reference data. First, a quantile-mapping (QM) approach is used for matching the statistical distribution of the forecasts to the reference data at the coarse forecast resolution. The coarse-scale climatology of the reference data is then removed from the bias-corrected forecasts. The remaining anomalies are then interpolated to a higher-resolution grid. Finally, the high-resolution climatology is added back to the interpolated anomalies to obtain a bias-corrected and spatially disaggregated forecast. The BCSD approach has demonstrated its potential for improving particularly climate predictions and is hence still used in many recent studies (e.g., Thrasher et al., 2013; Ning et al., 2015; Briley et al., 2017; Nyaupane et al., 2018). However, we found that particularly the spatial disaggregation, which is often similar to a simple bias correction via linear scaling including interpolation to a higher-resolved grid, can lead to unrealistic values. The disaggregation of precipitation and radiation is based on a multiplicative scaling factor, which is simply the ratio between the climatology from the forecasts and the reference data. During dry months with average precipitation values close to zero, this scaling factor can become unreasonably large (especially if there are large discrepancies between the climatologies from the forecasts and the reference) and can therefore cause unreasonable, corrected values.

Consequently, to avoid the calculation of scaling factors, we also reversed the order of bias-correction and spatial disaggregation as in Abatzoglou and Brown (2012). For the spatial-disaggregation step, we apply a simple bilinear interpolation of the full precipitation, temperature and radiation forecasts. The spatially disaggregated (or interpolated) full fields are then bias-corrected using a quantile-mapping approach. However, for our final product, we still stick to the technical term BCSD, as introduced in Wood et al. (2002).

The different steps are depicted in Appendix A. Here, we only summarize the key characteristics of our BCSD implementation.

- Spatial disaggregation from the coarse SEAS5 to the higher-resolved ERA5-Land grid is achieved by applying a bilinear interpolation technique to the full precipitation, temperature and radiation forecasts and not, as in most other studies, to the anomalies.
- The cumulative distribution functions (CDFs) for the (re-)forecasts and reference data are based on daily data from the period 1981 to 2016.
- To estimate the CDF of the reference data and (re-)forecasts, we apply a $\pm 15$ d window around the forecasted day.
- The seasonality is taken into account by estimating the forecast distributions with forecasts from the same issue date only.
- To avoid inconsistencies in the temperature data, we correct the deviations from the mean daily temperature instead of the full maximum and minimum temperatures. After bias correction, maximum and minimum temperature are restored by adding and removing the corrected deviations from the corrected mean temperature, respectively.
- Forecast values above or below the maximum and minimum reference quantile are corrected using the *constant correction technique* from Boé et al. (2007).
- Precipitation intermittency is corrected using the method by Voisin et al. (2010) to ensure the agreement of the wet- and dry-day frequencies from ERA5-Land and SEAS5 BCSD.

While several studies report that using parametric distributions can lead to more stable results (e.g., Lafon et al., 2013), we prefer to use an empirical distribution as we (a) have a fairly large number of samples for both the reference and

forecast CDFs and (b) do not want to force precipitation, temperature and radiation to follow a certain fixed parametric distribution. A drawback of empirical distributions, however, is the need for extrapolation when a forecasted value is beyond the maximum or minimum values from the reference period. While this is crucial for climate projections where we are interested in the occurrence of, e.g., temperatures beyond the current maximum values, we do not assume such extremely high or low values in current seasonal forecasts. If, however, a forecast contains a value beyond the maximum or minimum reference quantile, we apply the constant correction method from Boé et al. (2007).

The usage of moving windows for estimating the distributions requires special attention. This step is necessary to obtain a reasonable sample size for the reference data and to allow for some climatic variability during the calibration period. However, such a moving window can lead to blurred distributions, particularly during pronounced transition phases, e.g., from dry to wet, wet to dry, cold to warm, or warm to cold seasons. If the onset and end of the rainy season is well known and less variable throughout the years, it might be more appropriate to adapt the moving window to such climatic conditions. However, in this study, for our approach to be as general as possible, we are using a window with constant length.

## 4 Results

To assess the performance of the bias-corrected and spatially disaggregated SEAS5, we compare the seasonal forecasts before and after applying the BCSD against the reference data from ERA5-Land. For a better understanding of the impact of the correction, we separate the results according to model biases, lead-dependent effects, topographic and resolution-dependent effects, and overall performance. That being said, it is difficult to separate these effects completely. As an example, the low spatial resolution of the global data can result in different amounts of convective and large-scale precipitation compared to higher-resolved reference data. The result is a bias between the forecasts and the reference data due to the different spatial resolutions and the resulting description of precipitation. It is therefore not within the scope of this study to discuss all details of differences between seasonal forecasts and reference data. This holds also true for the detailed discussion of results across all five variables and four domains. We focus on selected results which should show the example differences between SEAS5 and ERA5-Land and how the BCSD is able to correct for these.

### 4.1 Model biases

The most obvious effect of the bias correction is the correction of systematic model biases in the raw forecasts. Figure 2 shows the bias of area-averaged SEAS5 forecasts before and after applying the BCSD for the five study basins. First, no simple over- or underestimation of SEAS5 with respect to ERA5-Land can be observed. During certain months, the precipitation, temperature and radiation biases can reach values of up to $\pm 5$ mm/d (CC), $\pm 2$ K (KA) and $\pm 30$ W/m$^2$ (TA and BN), respectively. The biases show strong annual cycles for most basins and variables. Precipitation biases reach peak values during the main months of the rainy season (KA, SF, TA, BN, CC) and also during the transition from the dry to the rainy season (SF).

The temperature biases show much more complex patterns. While the SF and CC basins show an annual cycle in the temperature forecasts, particularly the forecasts for the TA and BN basins also reveal highly lead dependent effects. As an example, over the TA Basin, the lead-0 and lead-2 forecasts for July (i.e., the forecasts that have been issued in July and May, respectively) show biases of about $-1.3$ and $-0.4$ K, respectively. Such large gaps between the temperature forecasts from different issue dates can also be observed over the KA and BN basins.

For the KA Basin, the biases of average and minimum temperature from SEAS5 reach peak values during the main months of the rainy season (around January), while the biases of maximum temperature show an opposite behavior with maximum biases during the dry season. Over the SF Basin, the biases of mean and maximum temperature both show a tendency towards positive biases during the dry season. Minimum temperature, however, is generally underestimated in the raw SEAS5 forecasts of the SF Basin, which leads to a negative bias throughout the year.

The biases of mean and minimum temperature forecasts over the CC Basin also show an annual cycle with the maximum deviations from ERA5-Land around April, which also marks the end of the rainy season. But in contrast to the SF Basin, the biases of mean and minimum temperatures show a very similar cycle, while there is a positive bias of maximum temperature almost throughout the year in the CC Basin.

The bias from the radiation forecasts also shows an annual cycle with peak values either at the transition from the dry to the wet season (SF) or during the main months of the rainy season (KA, SF, TA, BN, CC).

All these effects are almost completely removed after applying the BCSD. The bias-corrected forecasts do not show any systematic positive or negative biases when compared against ERA5-Land. However, while the biases of the BCSD temperature forecasts for lead 1 to lead 6 for the KA Basin are reduced compared to the raw forecasts, there are remaining biases of up to 0.7 K for the lead-0 and lead-7 forecasts.

The effect of the BCSD on the root mean squared errors, which are shown in Fig. 3, is much more diverse. In general, the RMSE of SEAS5 BCSD is lower compared to the raw forecasts. As an example, the RMSEs of the precipitation and radiation forecasts for February over the CC Basin are reduced from around 6 mm/d and 30 W/m$^2$ to less than 4 mm/d and 20 W/m$^2$, respectively. The minimum tempera-

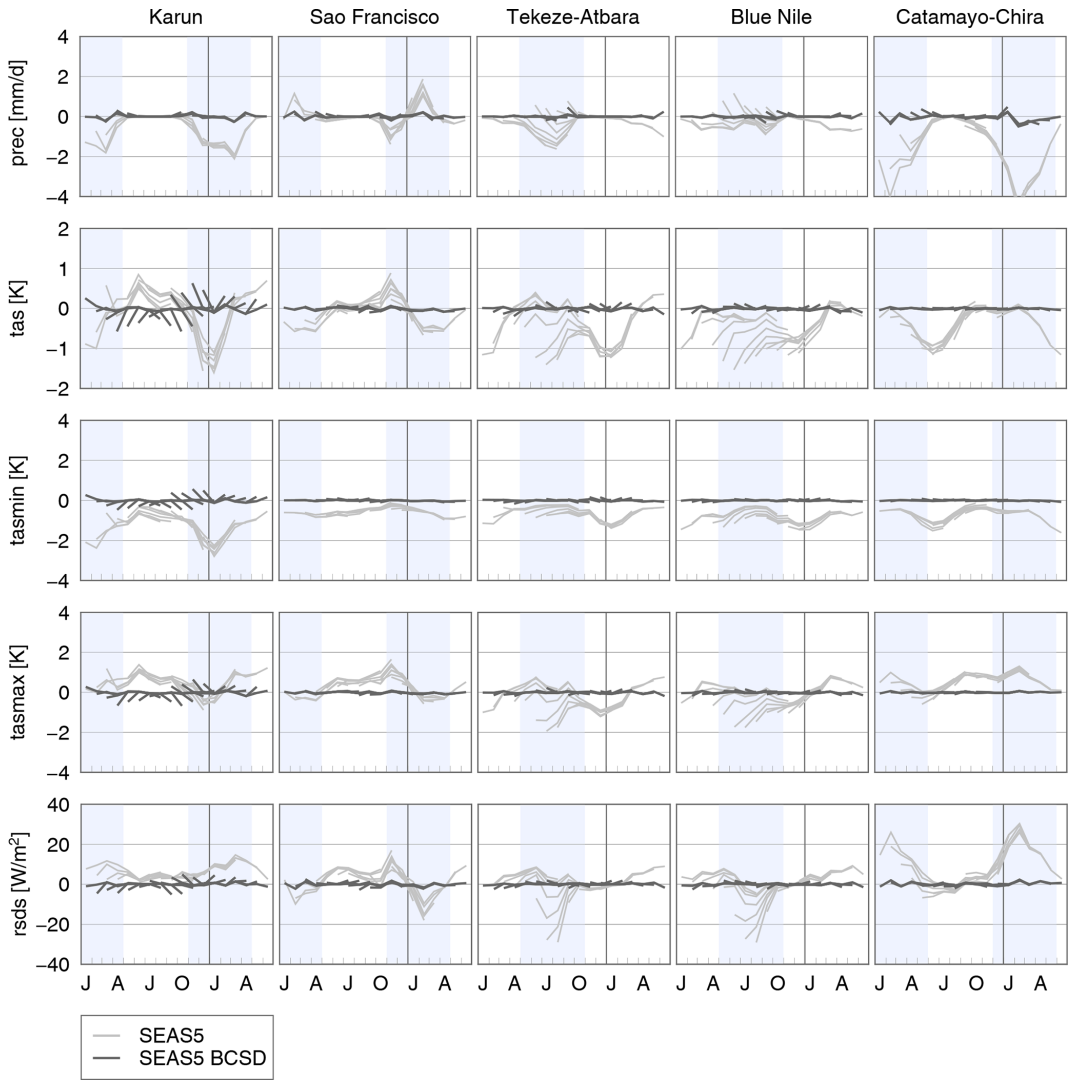


**Figure 2.** Bias of monthly (from top to bottom) precipitation (pr), average temperature (tas), minimum temperature (tasmin), maximum temperature (tasmax) and radiation (rsds) forecasts from SEAS5 (dashed lines) and SEAS5 BCSD (solid lines) with respect to ERA5-Land during the period 1981 to 2016 for all five study basins. The colors of the lines indicate the 12 different issue months from January to December. The vertical line between December and January indicates the end of the year; the values right of this line belong to the higher lead forecasts issued in September until December. The blue-shaded area depicts the 6 months of the rainy season.

ture RMSE during the rainy season of the KA Basin could be reduced from 3 K to less than 2 K.

However, there are other cases where the bias correction shows almost no improvement. While the biases of the raw precipitation, temperature or radiation forecasts for the SF Basin are much lower after applying the BCSD, the RMSE remains almost unchanged. The same holds true for the precipitation forecasts for the BN Basin. Moreover, there are still some lead-dependent effects of the bias-corrected forecasts, which can be seen from the gaps and jumps in the RMSEs of the temperature and radiation forecasts over the SF Basin.

### 4.2 Topographic and resolution-dependent effects

To evaluate the impact of the improved spatial resolution, Fig. 4 shows the total accumulated precipitation and its standard deviation during the 4 months of the rainy season from the raw and bias-corrected lead-0 forecasts, respectively, and ERA5-Land. The precipitation sums of the raw forecasts are generally lower than those from ERA5-Land. This is particularly visible for the mountainous headwaters of the KA (D01) and CC (D04) basins, where the raw forecasts predict seasonal precipitation sums of less than 500 mm (D01) and 2000 mm (D04), respectively, while ERA5-Land show values of more than 750 mm (D01) and 2400 mm (D04), respectively. Furthermore, there is a single connected precipi-

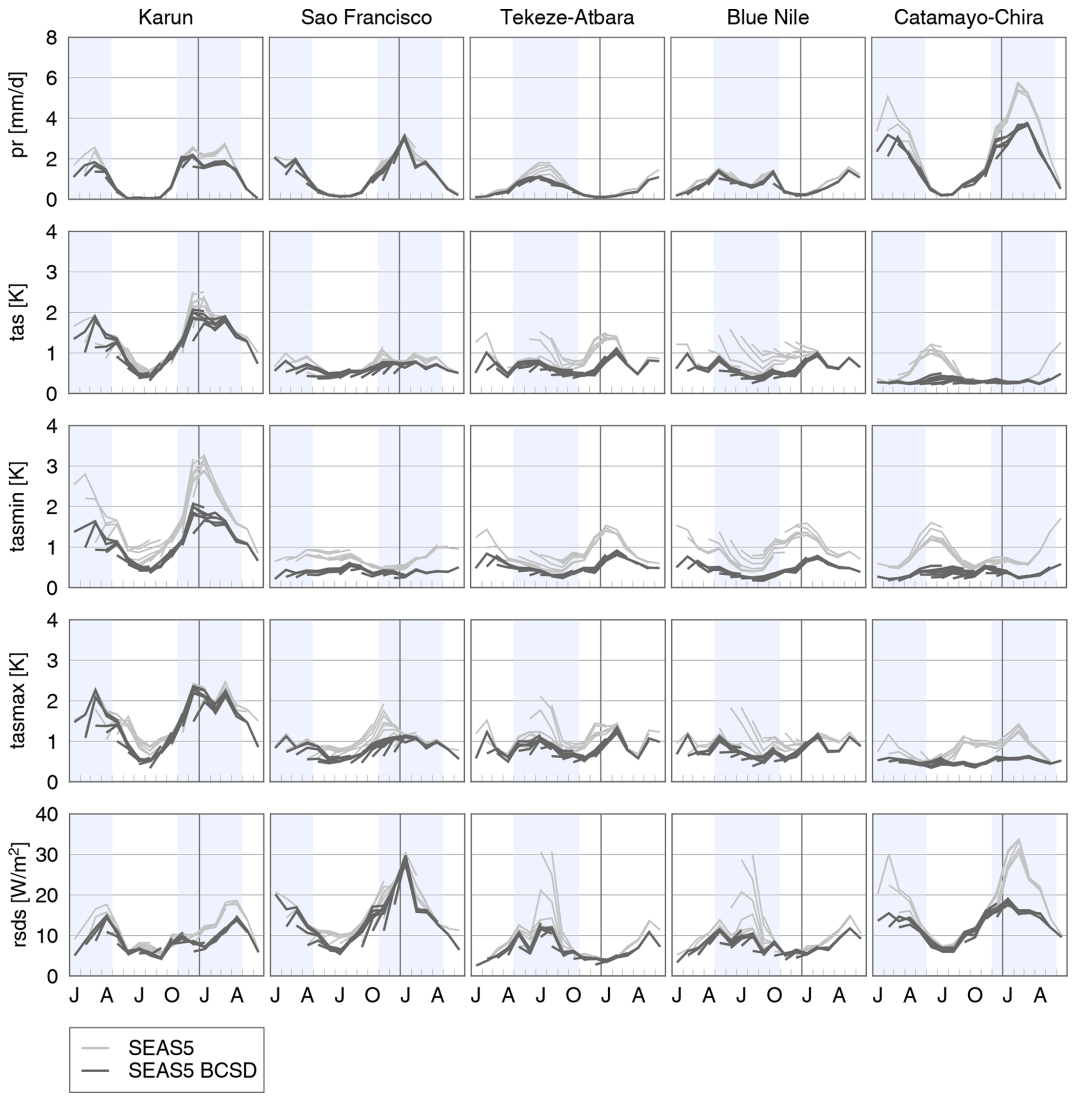


**Figure 3.** Same as Fig. 2 but for the RMSE of SEAS5 and SEAS5 BCSD compared to ERA5-Land.

tation pattern along the Andes in the D04 domain and multiple precipitation bands in the D03 domain over the Ethiopian Highlands. ERA5-Land, however, shows much more distinct spatial precipitation peaks, which agree very well with the complex topography in the mountainous headwaters. After applying the BCSD, the patterns and values of the seasonal forecasts match almost perfectly with the reference data and, hence, also the higher-resolved topography. The same holds true for the standard deviation of seasonal precipitation. The raw forecasts tend to underestimate the precipitation variability across all four domains. Especially in the mountainous areas of D01 and D04, ERA5-Land and, hence, SEAS5 BCSD show maximum standard deviations of ±600 mm and more, while the raw forecasts only reach values of less than ±450 mm. On the other hand, the raw forecasts show a higher precipitation variability particularly across the southwestern corner of D03, which is reduced in SEAS5 BCSD and therefore agrees better with ERA5-Land.

### 4.3 Lead-time-dependent effects

The BCSD approach further corrects for lead-dependent effects. The magnitude of these effects strongly depends on the lead time, which can be seen when comparing, e.g., the climatologies of the raw forecasts from different issue months. As an example, Fig. 5 shows the difference between the July forecasts for D03 from different issue months. These differences obviously increase with increasing lead time. While precipitation amounts are decreasing for higher lead times, temperatures and radiation are increasing.

One reason for these drifts is a shift of higher temperatures and higher radiations with increasing lead times towards south. This is visible in Fig. 6, which shows the July

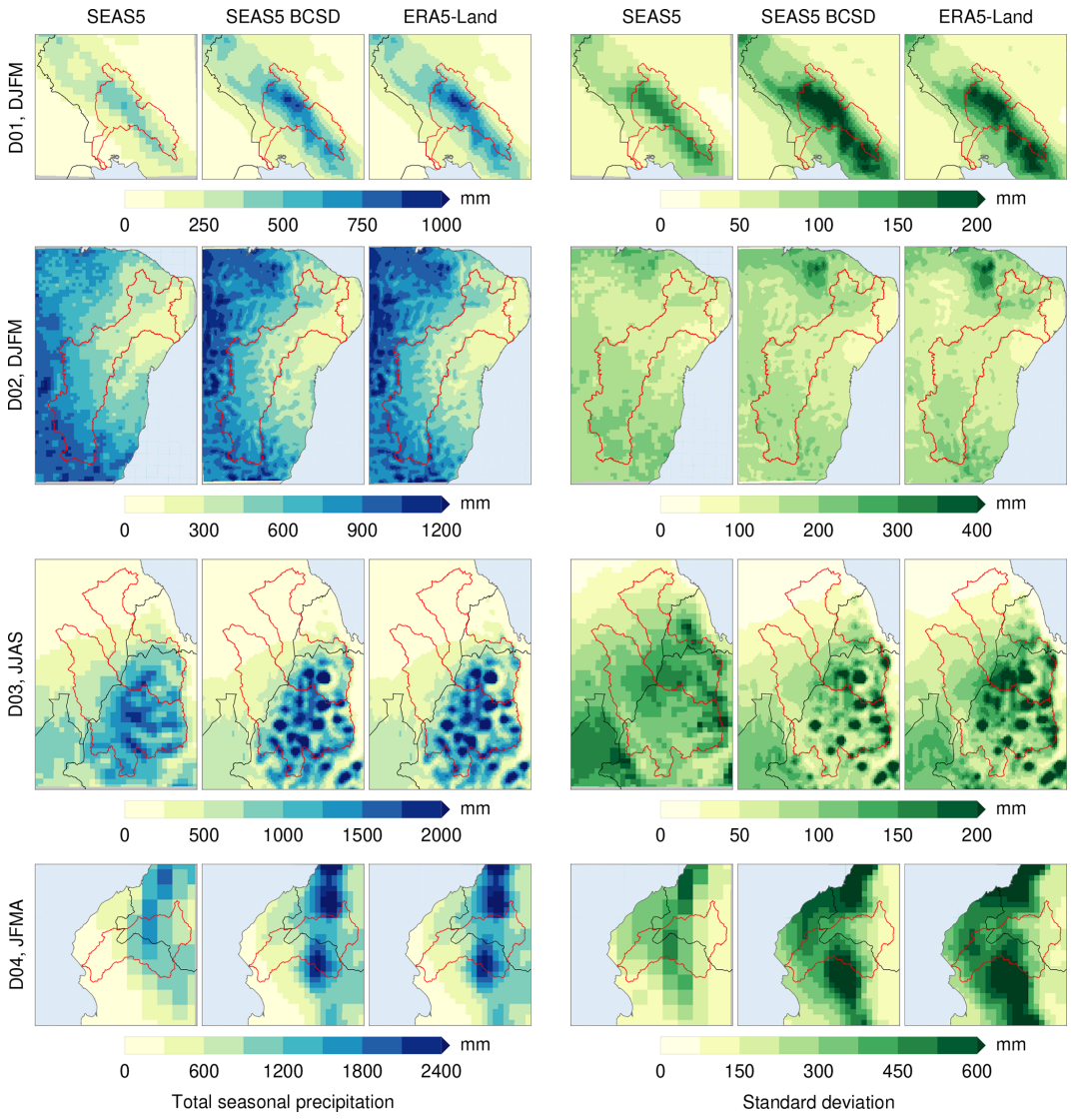


**Figure 4.** Total precipitation (left three columns) and the corresponding standard deviation (right three columns) during the 4 main months of the rainy season for the four study domains from raw SEAS5 and SEAS5 BCSD lead-0 forecasts and ERA5-Land, averaged over the period 1981 to 2016.

forecasts from SEAS5 and SEAS5 BCSD from different lead times, compared to ERA5-Land. Despite the biases in absolute magnitudes, the climatology of the higher-lead SEAS5 temperature and radiation forecasts match better with the climatology from ERA5-Land. The differences between the lead-0 and lead-5 forecasts in Fig. 5 also show the largest deviations in the northern part of D03.

In most other cases, however, the climatologies from lower lead times show a better agreement with the ERA5-Land climatology.

The SEAS5 BCSD forecasts show only minor lead-dependent effects (Fig. 5). The remaining differences for precipitation, temperature and radiation between the low- and higher-lead forecasts are below 0.5 mm/d, 0.5 K and 10 W/m$^2$, respectively. Similarly, as depicted in Fig. 5, the lead-0 and lead-5 forecasts for precipitation, temperature and radiation forecasts in July as well as the ERA5-Land based estimates agree almost perfectly in magnitude and spatial patterns, indicating that the model drift of SEAS5 is almost completely removed after applying the BCSD approach. This is also true for the other three study domains (not shown).

### 4.4 Wet- and dry-day frequencies

Besides biases in the absolute values from raw forecasts, we usually also have to take into account biases in the frequencies of wet and dry days. Figure 7 shows the wet-day probability from the lead-0 and lead-5 forecasts, respectively, for a single month for all four domains. Similar to the drifts in

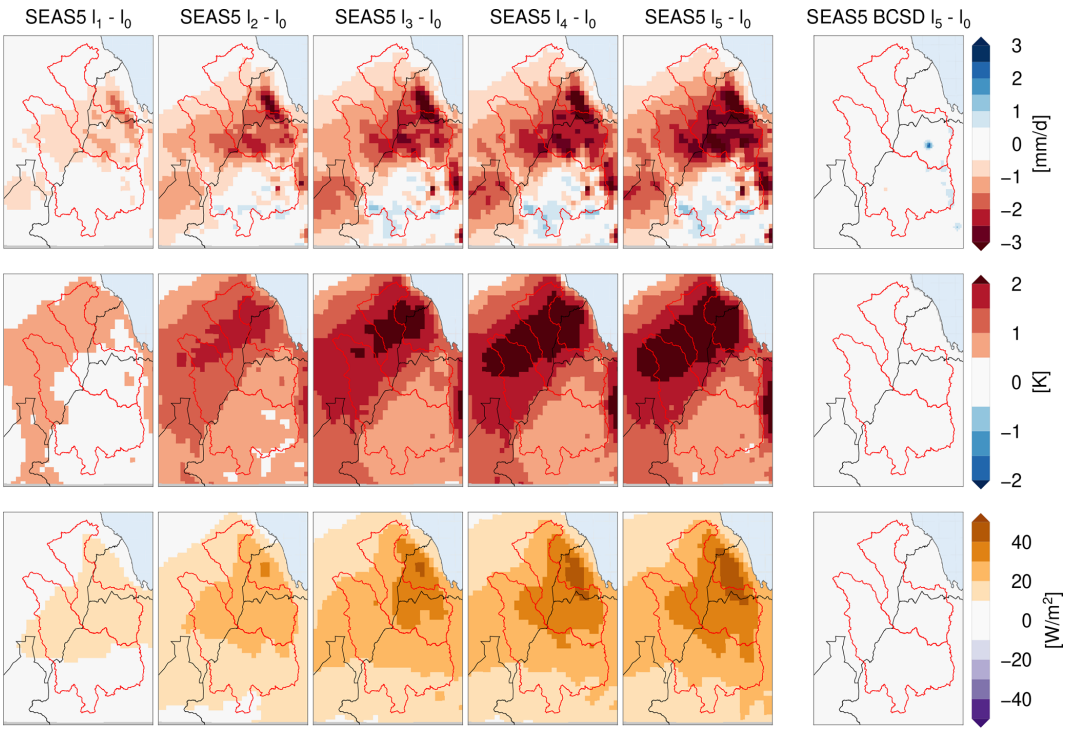


**Figure 5.** Differences between the July forecasts from different issue months for precipitation (top row), temperature (middle row) and radiation (bottom row) for the African Domain D03.

the absolute values and patterns, there is also a clear difference between the wet-day probabilities from different lead times. For example, the lead-0 forecasts for June over the D03 domain predict a wet-day probability during June of about 100 % across large parts of the Ethiopian Highlands. This means that there is at least 1 mm of precipitation on every single day in June. In the lead-5 forecasts (which are issued in January), this probability is reduced to 80 % and less. In other words, only 80 % of the forecasted days in June receive precipitation amounts of at least 1 mm per day. Similarly, the higher-lead forecasts over the D01 domain also predict lower wet-day probabilities across the Zagros Mountains. After correcting for this lead-dependent wet-day frequency, the BCSD forecasts show spatial patterns very similar to the reference data and more consistent frequencies across the different lead times.

### 4.5 Overall performance

The change in overall performance of the seasonal forecasts due to the bias-correction and spatial-disaggregation approach with respect to ERA5-Land is evaluated with the continuous ranked probability skill score (CRPSS, Sect. B). In general, the overall performance of SEAS5 BCSD improves compared to raw SEAS5; i.e., the cumulative distribution functions (CDFs) of SEAS5 BCSD better correspond to the reference ERA5-Land than the CDFs of raw SEAS5 (Fig. 8). Largest improvements for all basins are found for the minimum temperature, with frequent CRPSS values $> 0.4$ indicating an improvement in the distributional distances by 40 % compared to raw SEAS5. Among the basins, largest improvements by BCSD are produced for the CC Basin, especially for precipitation and maximum temperature. For the TA and BN basins, the abovementioned lead-dependent effect is evident with larger improvements for the lower lead times of the temperature and radiation forecasts. For the KA Basin, precipitation forecasts for November and April may be worsened in their performance by BCSD. In contrast, the main 4 months of the rainy season of KA show improvements mainly above 30 %. Also for the SF Basin, the December precipitation forecasts may be worsened, whereas other months only show little improvement of SEAS5 BCSD compared to raw SEAS5 for precipitation forecasts. Similarly, there is only slight improvement for the maximum temperature forecasts for the KA Basin or the higher-lead temperature forecasts from December to March for the CC Basin.

## 5 Discussion

The BCSD forecasts show a much better agreement with ERA5-Land as the raw SEAS5 product across most variables, domains and forecast months. However, to understand the performance of the BCSD across the different study regions, the regional climatic conditions are important.

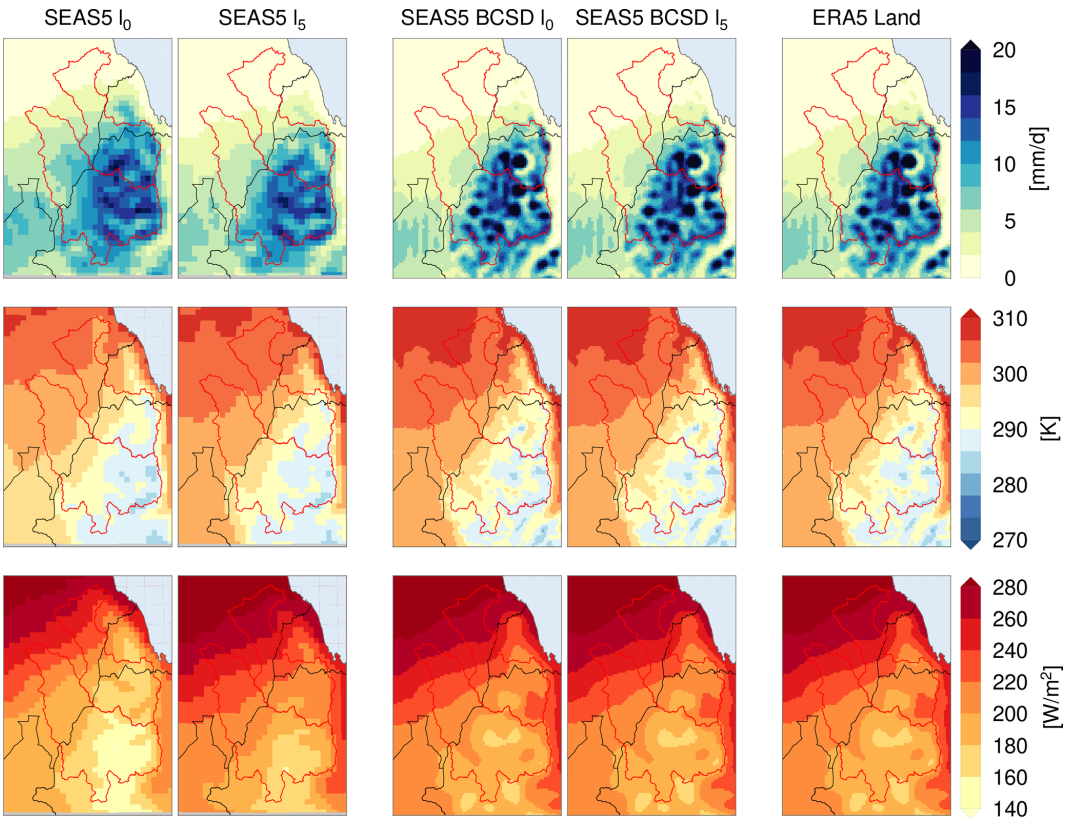


**Figure 6.** Average July predictions from SEAS5 lead 0 and 5 (first two columns), SEAS5 BCSD lead 0 and 5 (center columns), and reference values from ERA5-Land (right column) for precipitation (top row), average temperature (middle row) and radiation (bottom row) for the African Domain D03.

Over regions like the East African D03 domain, the rainy season is dominated by the East African monsoon. This is usually associated with daily convective precipitation and, hence, many continuous wet days during the rainy season. According to Figs. 6 and 7, the African monsoon is predicted with higher rainfall intensities, higher wet-day frequencies, and especially towards the northern parts of the domain with lower temperature and radiation at shorter forecast horizons than at longer lead times. Hence, it is assumed that in contrast to the model climatologies from SEAS5 and ERA5-Land, the initial conditions (which strongly influence the low-lead forecasts) cause the low-lead forecast to show higher intensities as well as a more northern extension of the summer monsoon. The comparison with the reference ERA5-Land reveals that the spatial extent of the monsoon is predicted too far towards the north at low lead times. However, the rainfall intensity and wet-day frequency is more realistic than at long forecast horizons.

Such spatial and temporal inconsistencies in the forecasted spatial extent and intensity of the monsoon from different issue dates impede the direct application of raw forecasts for the regional water management. Therefore, as we correct the raw forecasts to the same ERA5-Land reference data across all lead times, this lead dependency is eliminated during the bias correction. This simplifies the use as well as the interpretation of our BCSD forecasts compared to the raw SEAS5 products.

For some regions and variables, e.g., for the maximum temperature forecasts across the KA Basin, the precipitation or radiation forecasts across the SF Basin or the precipitation forecasts across the BN Basin, almost no or little improvements for most forecast months can be identified, indicated by CRPSS values of around 0 in Fig. 8 and almost identical RMSE values of SEAS5 and SEAS5 BCSD in Fig. 3. In such cases, there is only a limited effect of the bias correction which can be explained by, e.g., an already good correspondence between the raw forecasts and ERA5-Land (indicated by low biases in Fig. 2) and/or rather random biases. According to, e.g., Thrasher et al. (2012), variables that are systematically biased usually benefit more from a quantile-mapping-based bias correction than randomly biased variables.

Hence, large improvements of overall performance, indicated by high CRPSS values (Fig. 8), usually point to large systematic discrepancies between the raw SEAS5 and the reference ERA5-Land. This is obvious for, e.g., the minimum temperature forecasts which show a negative bias of the raw forecasts (Fig. 2), high CRPSS values (Fig. 8) and reduced

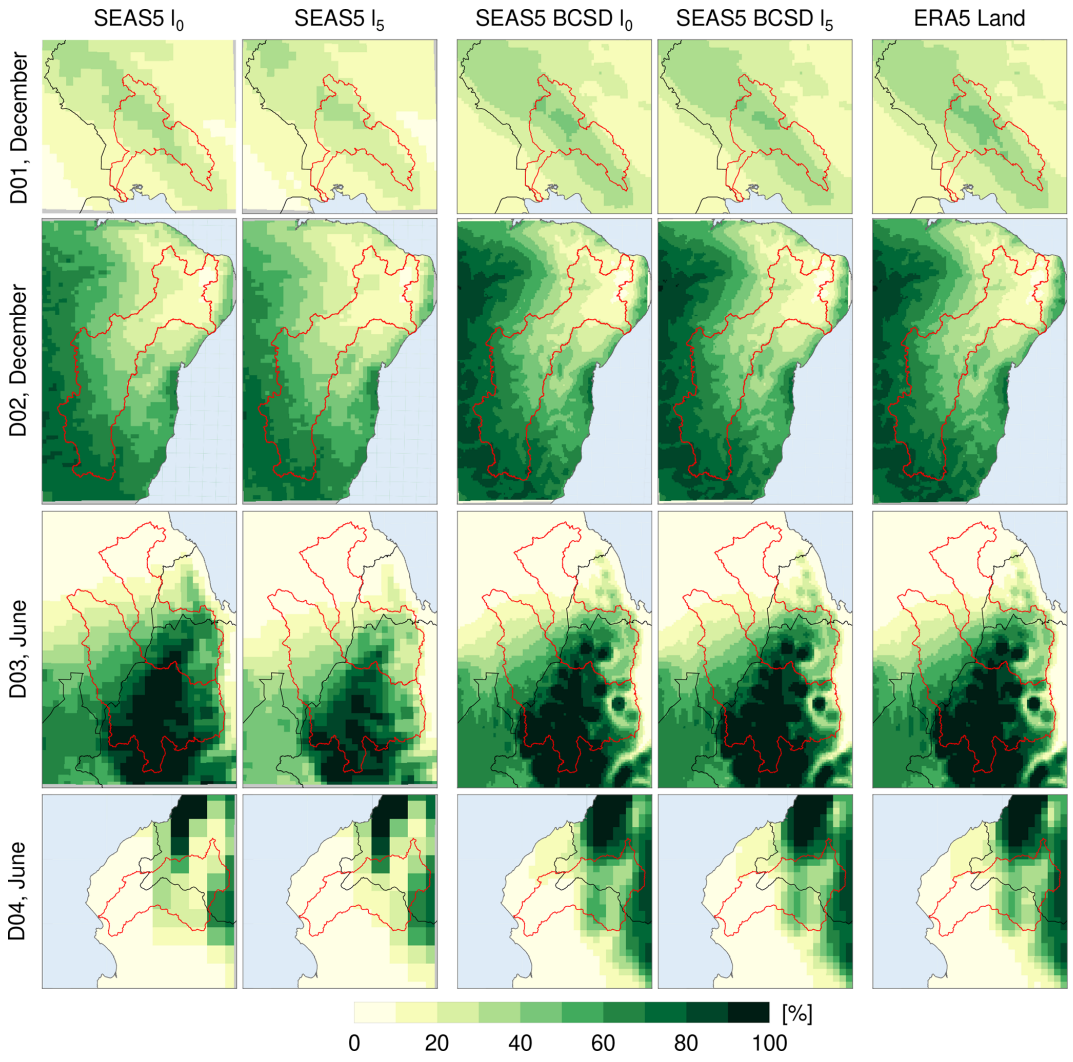


**Figure 7.** Wet-day (> 1 mm/d) probability for the four study domains for a single month (D01: December, D02: December, D03: June, D04: May) from SEAS5 and SEAS5 BCSD for both lead 0 and lead 5, respectively, and the reference ERA5-Land.

RMSEs after the bias correction (Fig. 3) across almost all basins and forecast months.

Besides these mostly positive results for BCSD, the mixed impact of the BCSD approach across the D01 domain, indicated by negative CRPSS values for the KA Basin in Fig. 8, requires a closer look. Iran's climate during the rainy season is dominated by migrating low-pressure systems mainly from the west and the Mediterranean Sea (Khalili and Rahimi, 2014). The precipitation over the D01 domain hence occurs intermittently and spatially variable, which is usually difficult to predict especially with higher lead times and over the mountainous headwaters of the Karun. For such regions, it is necessary to correct for the amount and spatial location of precipitation as well as the wet- and dry-day frequency. While, according to Fig. 7, the wet-day frequency even for high lead times could be improved, the CRPSS values in Fig. 8 show worse agreement with ERA5-Land after the bias correction during the transition from the dry to the wet season (November) or the transition from the wet to the dry season (April). We assume that this is caused by the application of a 31 d window for estimating the distribution functions, which might not be adequate in such strongly varying climate conditions during the transition months between the wet and dry seasons. This can also be seen in Fig. 2, where particularly the lead-0 and lead-7 temperature forecasts show remaining biases with values of up to 0.7 °K. The temperature bias in the first and last forecast months appears strongest in the KA Basin due to the large annual temperature variations with an annual temperature range of up to 45° (Table 1). Nevertheless, mostly positive CRPSS values at these lead times (Fig. 8) result from still reduced biases compared to the raw forecasts (Fig. 2).

In general, during the first and last days of a forecast, we cannot fill the complete 31 d window for estimating the fore-

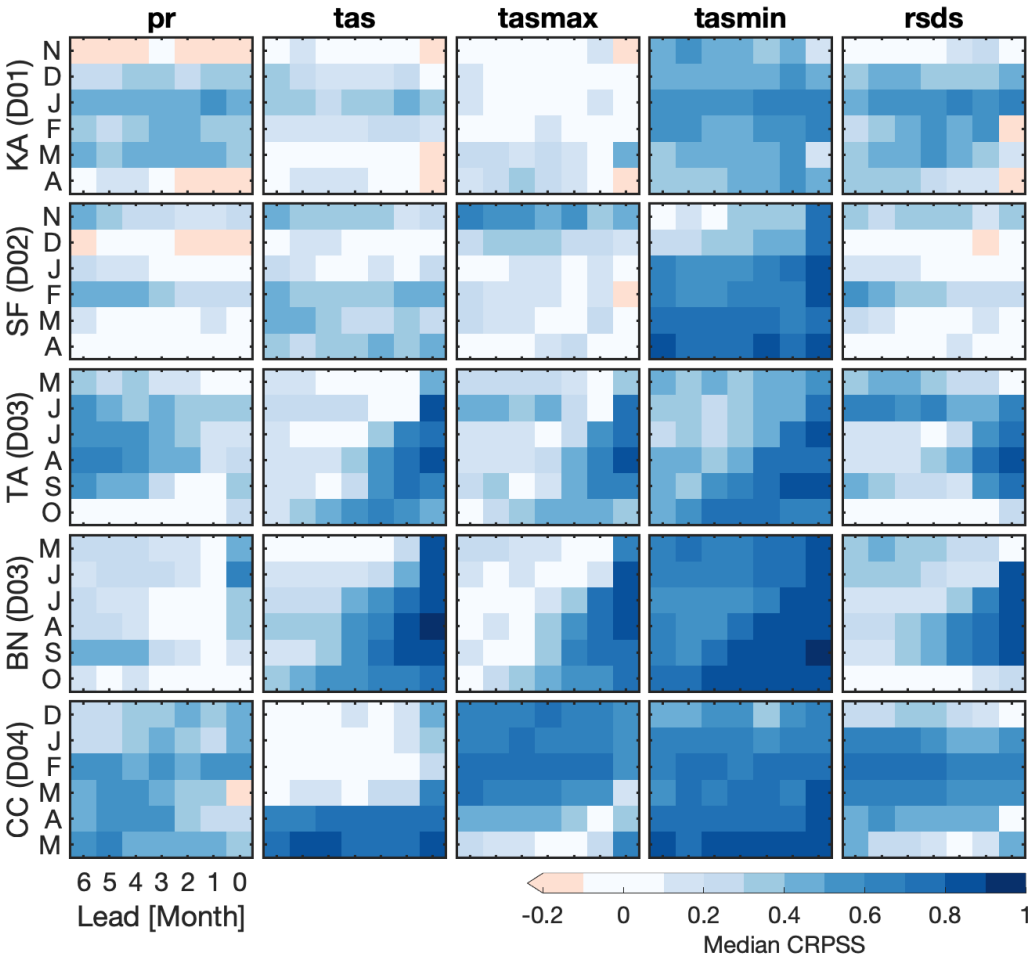


**Figure 8.** Median continuous ranked probability skill score (CRPSS) of area averages over the five basins (from top to bottom) of SEAS5 BCSD against raw SEAS5 forecasts with respect to the reference ERA5-Land. The CRPSS values are derived for precipitation (pr); mean (tas), maximum (tasmax), and minimum (tasmin) temperature; and shortwave radiation (rsds) as monthly medians for each of the 6 months during the wet season ($x$ axis) of the period 1981 to 2016 for each lead time ($y$ axis) separately. Blueish (reddish) colors indicate better (worse) correspondence with ERA5-Land after applying the BCSD to the SEAS5 forecasts.

cast CDF. As an example, the reference CDF for 1 January is based on the values from 17 December until 16 January, while the CDF of the January forecast for 1 January only uses the values from 1 to 16 January. If there are strong temporal climate gradients or heteroscedasticity, e.g., during the transition from a cold to a warm period, a bias correction using moving windows can lead to remaining biases and, hence, to statistical inconsistencies particularly on non-daily timescales. An approach to account for such gradients would be to use a dynamic moving window, where the length of the window is based on, e.g., the gradient of the daily climatology. It will be subject of future studies if such an approach is able to improve the statistical consistency particularly during the first and last days of the forecast.

The representation of small-scale features in SEAS5 BCSD, particularly in complex and mountainous terrain, benefits from the explicit altitude correction in ERA5-Land, which was necessary due to the higher spatial resolution compared to ERA5: when correcting the SEAS5 forecasts towards such a reference, we automatically include an indirect correction for altitude. For the small basins of KA and CC with large altitude differences, the added value of spatial disaggregation (better representation of small-scale features) and bias correction (indirect altitude correction) is therefore most evident. Particularly at high elevations of the Zagros Mountains in KA and the Andes in CC, the higher resolution and subsequent bias correction allows for locally distinct precipitation intensities (Fig. 4). Also in the Ethiopian Highlands of D03 higher resolution produces more complex (at this resolution circular shaped) structures around the Ethiopian mountains. Independent of the accuracy of the seasonal forecasts, we strongly assume that the higher spatial resolution and, hence, better representation of small-scale precipitation patterns make the BCSD SEAS5 forecasts

more suitable for the regional water management. As already shown in, e.g., Westrick and Mass (2001), a higher spatial resolution of the atmospheric forcing (i.e., precipitation) usually leads to more accurate streamflow modeling.

We would also like to discuss the choice of the bias-correction method used. As reported by, e.g., Anghileri et al. (2019), bias correction is crucial to improve both forecast quality and value. The quantile-mapping method that is used in this study serves this purpose and is a widely used, well understood and robust technique that is not computationally demanding and can be easily implemented (Siegmund et al., 2015). During the recent years, there have been numerous studies in which new approaches were presented. While other techniques can lead to more skillful, reliable and accurate forecasts (e.g., Schepen et al., 2018; Manzanas et al., 2019; Khajehei et al., 2018), or lower biases (e.g., Alidoost et al., 2019) as quantile mapping for example tends to produce negatively skillful forecasts when the raw forecasts are not significantly positively correlated with observations (Zhao et al., 2017), it should be considered that quantile mapping still serves as the reference method in most of the recent bias-correction studies. In other words, there is currently no other bias-correction method that is similarly widespread. This not only improves the comparability of our data with similar studies, but also marks our SEAS5 BCSD forecasts as a reference product for exploring new forecast products and developing and evaluating new bias-correction techniques.

## 6 Data availability

The bias-corrected and spatially disaggregated seasonal forecasts are published via the World Data Center for Climate (WDCC), which is hosted by the German Climate Computing Center (DKRZ), within the project Seasonal Water Resources Management for Semiarid Areas: Regionalized Global Data and Transfer to Practise (SaWaM, https://cera-www.dkrz.de/WDCC/ui/cerasearch/project?acronym=SaWaM, last access: 2 June 2021). In this project, we have created the four experiments SaWaM D01, SaWaM D02, SaWaM D03 and SaWaM D04 (i.e., one experiment for each study domain), which contain all products for the respective region. Our SEAS5 BCSD forecasts are available via the dataset group SaWaM SEAS5 BCSD, which contains all daily and monthly forecasts:

- SaWaM D01 SEAS5 BCSD (https://doi.org/10.26050/WDCC/SaWaM_D01_SEAS5_BCSD): Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Karun Basin (Iran) (Lorenz et al., 2020b).

- SaWaM D02 SEAS5 BCSD (https://doi.org/10.26050/WDCC/SaWaM_D02_SEAS5_BCSD): Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Rio São Francisco Basin (Brazil) (Lorenz et al., 2020c).

- SaWaM D03 SEAS5 BCSD (https://doi.org/10.26050/WDCC/SaWaM_D03_SEAS5_BCSD): Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Tekeze–Atbara and Blue Nile Basins (Ethiopia/Eritrea/Sudan) (Lorenz et al., 2020d).

- SaWaM D04 SEAS5 BCSD (https://doi.org/10.26050/WDCC/SaWaM_D04_SEAS5_BCSD): Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Catamayo–Chira Basin (Ecuador/Peru) (Lorenz et al., 2020a).

Each of these four groups contains six datasets: BCSD_daily_pr and BCSD_monthly_pr (daily and monthly precipitation forecasts), BCSD_daily_tas and BCSD_monthly_tas (daily and monthly average, minimum and maximum temperature forecasts), and BCSD_daily_rsds and BCSD_monthly_rsds (daily and monthly surface solar radiation forecasts). All datasets contain forecasts from the issue date (i.e., the first of each month) for the next 215 d (daily) and 6 months (monthly), respectively.

Users interested in a teaser product are advised to use the monthly averaged forecasts. They have a maximum download size of around 2 GB for precipitation, 2.4 GB for radiation and 6 GB for the three temperature variables over our largest domain across Brazil for the whole period from 1981 to 2019 and all ensemble members. The data size for the other domains is of course much smaller. Some of the products as well as derived forecast measures like categorical precipitation and temperature forecasts are visualized through an online decision support system for the regional water management at https://sawam.gaf.de/ (last access: 2 June 2021). This system is currently under joint development with the company GAF AG (https://www.gaf.de, last access: 2 June 2021, Munich, Germany). As of now, forecasts for the Brazilian and Sudanese and Ethiopian domain are included. The data for Iran will be implemented in the near future.

We also publish the BCSD forecasts through the Karlsruhe Institute of Technology (KIT) – Campus Alpin THREDDS Data Server. In contrast to the products available via the WDCC, the operational forecasts are only available via the THREDDS Data Server. These are published with a delay of about 1 d after the release of the official seasonal forecasts from ECMWF on the fifth of each month. For getting access to the operational products, contact christof.lorenz@kit.edu.

## 7 Conclusions

In this study we present a comprehensive dataset of bias-corrected and spatially disaggregated seasonal forecasts for four different semi-arid domains across three continents. The forecasts are based on the most recent version of ECMWF's seasonal forecast system SEAS5, which are corrected towards the ERA5-Land land surface rerun of the ERA5 reanalysis with an enhanced spatial resolution of 9 km (here: 0.1°). The final SEAS5 BCSD repository contains seasonal forecasts at daily and monthly resolution for precipitation, mean, minimum and maximum temperature as well as for radiation for the period from 1981 until 2019 with a spatial resolution of 0.1°. For each of the 468 issue dates, we provide ensemble forecasts with 25 or 51 members (since 2017) for the coming 214 d. Currently, the data cover domains in Iran (D01), Brazil (D02), Ethiopia/Sudan (D03) and Ecuador/Peru (D04), but it is planned to extend this list to further domains.

The comparison of our SEAS5 BCSD product with the raw SEAS5 forecasts against reference data from ERA5-Land clearly indicated a reduction of biases and root mean squared errors across most study regions and variables. Further, the spatial resolution of the forecasts is improved from 36 km to 0.1°, and the patterns of precipitation, temperature and radiation show much better agreement with the reference data. Finally, model drifts are reduced, which leads to temporally more consistent forecasts.

Besides these improvements, we could also observe remaining biases after bias correction particularly during the low- and high-lead temperature forecasts for the KA Basin. This is explained with a highly variable climate with strong gradients and heteroscedasticity, where a moving window can introduce statistical inconsistencies when, e.g., monthly averages are derived from daily data. The impact of a highly dynamic climate on the statistical consistency of bias-corrected forecasts obviously has also huge implications for, e.g., global approaches. Therefore, future works have to further examine methods and approaches to account for such strong gradients in the reference climatologies.

Our bias-corrected and spatially disaggregated seasonal forecasts are freely available on both daily and monthly temporal resolution via the World Data Center for Climate. To our knowledge, this is the first multi-variable and multi-domain high-resolution seasonal forecast for a period of almost 40 years. It provides a unique product for a wide variety of researchers, stakeholders and other experts from the water sector who are interested in a consistent dataset of post-processed seasonal forecasts. This product gives local experts from the four study domains, who often do not have the computational framework conditions or access to the operational products from ECMWF, the opportunity to investigate the potential of high-resolution seasonal forecasts for, e.g., the regional water management, drought forecasting or irrigation planning. Derived products like categorical forecasts, based on our SEAS5 BCSD product, have therefore already been adopted by several weather services and other higher-level authorities in the study domains. As SEAS5 BCSD also covers all months during the re-forecasting period since 1981, it can be used to review and refine currently existing decision calendars and dates when actions and management strategies for the coming rainy season are defined.

That being said, the published SEAS5 BCSD dataset currently contains only the daily and monthly ensemble forecasts and no derived information like, e.g., probabilistic forecasts or drought indicators. Hence, for transferring this product into practice, we have to (a) identify and compute regionally suitable forecasting measures and indicators, (b) make this information accessible, e.g., via a user-friendly online platform and (c) ensure a proper communication of the potential but also the limitations of such current seasonal forecasting products to all end-user sectors like authorities, water managers or farmers.

For these subsequent steps, SEAS5 BCSD can be used as consistent data resource and, hence, serves as a further element for ensuring a sustainable and timely regional water management in our semi-arid study regions. Ultimately, in the context of global climate change with increasing risks of climatic extreme events, we have to ensure that longer-term forecasts like SEAS5 BCSD are adopted by authorities and included in the water resources planning for improving future disaster preparedness.

## Appendix A: Empirical quantile mapping

### A1 Empirical quantile mapping

We follow the classical empirical quantile-mapping approach as depicted in, e.g., Wood et al. (2002); Voisin et al. (2010). The bias correction is performed separately for each single pixel. For each forecasted day, we select a window of ±15 d around the forecasted day from all ensemble members and all re-forecasts of the period 1981 to 2016 which have been initialized in the same month. The CDF for the forecasts is then computed from this large sample. As an example, the CDF for 1 January from the January forecast is based on 16 (days in January) ×25 (ensemble members) ×36 (years) = 14 400 values. The CDF for the 31 January from the same issue month is based on 27 900 values (as we use the 15 d before and after the 31 January). Obviously, during the first and last 15 d of a forecast, we have to use fewer samples as we cannot extend the window before or after the forecasted period. While this might introduce inconsistencies particularly during the first and last days of a forecast, we think that this approach of a daily moving window is still more climatically reasonable than estimating a single CDF for each month. This approach, while computationally less expensive, can lead to large jumps between consecutive days at the end of the previous and beginning of the next month. The CDF for the reference data is estimated in the same way by using 31 (days) × 36 (years) = 1116 samples for each forecasted day. Using the CDFs for the forecasts $F_{i,\theta,\mathrm{mod}}$ and the reference $F_{i,\theta,\mathrm{ref}}$ for the forecasted day $i$ and pixel $\theta$, the bias correction of each forecasted value $X_{i,\theta}$ is performed through

$$\widehat{X}_{i,m,\theta} = F^{-1}_{\mathrm{doy,ref}}\left(F_{\mathrm{doy,mod}}\left(X_{i,m,\theta}\right)\right). \tag{A1}$$

This is shown for example in Fig. A1a.

### A2 Bias correction of extremes

As we are using ensemble-based forecasts, we usually have more samples from which we can estimate the forecast CDF compared to the reference data. If further an empirical quantile mapping is applied, we can get forecasted probabilities outside the range of the empirical reference quantiles (i.e., below $1/(n+1)$ or above $n/(n+1)$, where $n$ is the number of samples from which the reference CDF is derived; upper dashed line in Fig. A1). To apply the bias correction to such extreme values, some extrapolation is required. Here, we use the constant correction method from Boé et al. (2007). We first calculate the correction that corresponds to the lowest and highest reference quantile, respectively. These constant corrections are then applied to all forecast values below or above the lowest and highest reference quantile. Due to its simple application and robustness, this approach is a good choice particularly in cases where no climate change signal is expected or when the parametric distributions for the variables to be corrected are unknown or difficult to estimate.

### A3 Correction of the precipitation intermittency

Besides the correction of the absolute values, there might also be a bias between the wet- and dry-day frequencies. For correcting these frequencies, we apply the same approach as in Voisin et al. (2010). We first compute the dry-day probability of the forecasts and the reference data. If a forecast falls below the reference dry-day probability, it is set to zero. This is demonstrated for example in Fig. A1c. If, however, the dry-day probability of the forecasts is higher than the reference data and we obtain a zero-precipitation forecast, we draw a uniform random sample between 0 and the dry-day probability of the forecasts (depicted by the arrow in Fig. A1d). If this value is below the dry-day probability of the reference (the lower dashed line in Fig. A1d), the forecast is again set to zero. If not, it is corrected using the inverse CDF of the reference.

### A4 Consistent correction of minimum, maximum and average temperature

Due to the very nature of QM-based bias correction, the corrected temperatures can become physically unrealistic as

1. the corrected minimum (maximum) temperatures are higher (lower) than the corrected maximum (minimum) temperatures or
2. the corrected average temperature is higher (lower) than the corrected maximum (minimum) temperature, respectively.

The first two cases are comprehensively discussed in Thrasher et al. (2012). For bias-correcting daily temperature data, they propose to first calculate the diurnal temperature range from the maximum and minimum temperature. The bias correction is then applied to the maximum temperature and diurnal temperature range (instead of the minimum temperature). Afterwards, the corrected minimum temperature is derived by subtracting the bias-corrected diurnal temperature range from the bias-corrected maximum temperature. As we are interested in correcting the average temperature as well, we apply a slightly modified version of this approach. Instead of computing the diurnal temperature range, we compute the difference between the maximum (minimum) and average temperature, respectively. Then, we apply the bias correction to the two ranges as well as the average temperature. Corrected maximum and minimum temperatures are then computed by adding and subtracting the ranges to and from the corrected average temperature, respectively.

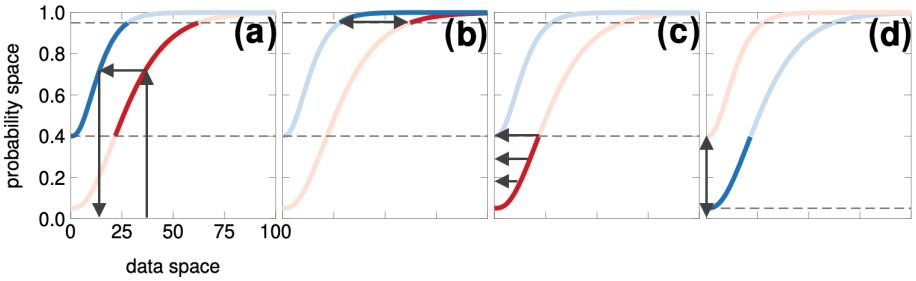


**Figure A1.** **(a)** Empirical quantile mapping between model-based (red) and reference (blue) data; **(b)** delta approach for correcting extreme values above the maximum Weibull quantiles; **(c)** correction of precipitation intermittency when the dry-day probability of the reference (lower dashed line) is higher; **(d)** correction of precipitation intermittency when the dry-day probability of the reference is lower.

## Appendix B: Verification metrics

To compare the bias-corrected and spatially disaggregated forecasts against the reference data, we use three different verification metrics. First, for evaluating the overall level of agreement of the corrected data, we compare the BCSD and raw SEAS5 forecasts against the ERA5-Land dataset using the bias and root mean squared error (RMSE). For a pixel or basin average, these are

$$\text{Bias} = \frac{1}{N}\left(X_\text{p} - X_\text{o}\right) \tag{B1}$$

and

$$\text{RMSE} = \sqrt{\frac{1}{N}\left(X_\text{p} - X_\text{o}\right)^2}, \tag{B2}$$

respectively, where $X_\text{p}$ and $X_\text{o}$ are the predictions and reference values, respectively and $N$ the number of samples. In our case, for monthly forecasts, $N$ is 36 (years). Obviously, the bias and RMSE can be computed for different lead times, which then gives information about the dependency of the error from the lead time.

While the bias and RMSE are suitable for comparing the overall agreement of the ensemble mean against a reference, there is a wide range of metrics particularly for ensemble forecasts. A comprehensive overview and discussion of ensemble forecast verification measures can be found in, e.g., Casati et al. (2008). In this study, we use the continuous ranked probability skill score (CRPSS, e.g Hersbach, 2000) for the evaluation of overall performance. The CRPSS is the continuous extension of the widely used Brier skill score (BSS Brier, 1950) and compares relative distributional distances of forecast and reference data. While the BSS aims at the verification of specific events like the probability of precipitation amounts $> 10$ mm/d, the CRPSS extends this to all possible events. As a skill score, comparing the prediction skill of different forecasts, the CRPSS is defined as

$$\text{CRPSS} = 1 - \frac{\text{CRPS}_\text{forecast}}{\text{CRPS}_\text{reference forecast}}. \tag{B3}$$

Here, the continuous ranked probability score (CRPS) for a forecasted quantity $x$ is defined as

$$\text{CRPS}(F_\text{p}, x) = \int_{-\infty}^{+\infty} \left(F_\text{p}(x) - F_\text{o}(x)\right)^2 \text{d}x, \tag{B4}$$

where $F_\text{p}\left(x_\text{p}\right)$ and $F_\text{o}(x_\text{o})$ are the cumulative distribution functions from the ensemble seasonal forecasts and the reference data, respectively. In this study, we use the empirical CDF for approximating the ensemble CDF of the forecasts, and the CDF of the reference data is defined as

$$F_\text{o}(x) = H(x - x_\text{o}), \tag{B5}$$

where $H(x)$ is the Heaviside function. Obviously, for each observation, we compute a single CRPSS. For evaluating the CRPSS across multiple points in time and/or space, we calculate the average from the single CRPSS values.

In most cases, the comparison with the reference forecast in the denominator of the CRPSS simply uses climatology. However, in this study, we are interested in the increase in the level of agreement of the SEAS5 forecasts with the ERA5-Land reference data after applying the BCSD. We therefore compute the skill score from the CRPS between the SEAS5 BCSD and the raw SEAS5 forecasts against the ERA5-Land reference, respectively.

**Author contributions.** CL collected and processed the seasonal forecasts from ECMWF, developed and applied the bias-correction and spatial-disaggregation algorithm, computed the final seasonal forecast products, initiated the data dissemination via DKRZ WDCC and the KIT Campus Alpin THREDDS Data Server, and implemented the operationalization of the presented approach; CL and TCP conducted the evaluation of the BCSD forecasts and wrote the first draft of the paper; CL, TCP, PL and HK reviewed the paper and prepared the final version of the manuscript.

**Competing interests.** The authors declare that they have no conflict of interest.

**Acknowledgements.** The authors would like to thank the German Federal Ministry of Education and Research (BMBF) for funding this research within the Seasonal Water Resources Management: Regionalized Global Data and Transfer to Practise (SaWaM, http://grow-sawam.org, last access: 2 June 2021) project. We would also like to thank the ECMWF and the Copernicus Climate Service for providing the SEAS5 seasonal forecasts as well as the IT team of the Karlsruhe Institute of Technology Institute of Meteorology and Climate Research for providing access to the HPC environment. Finally, we would like to thank the DKRZ and WDCC teams and, particularly, Heinrich Widmann and Andrea Lammert for uploading, publishing and supporting our product.

**Review statement.** This paper was edited by David Carlson and reviewed by two anonymous referees.

## References

Abatzoglou, J. T. and Brown, T. J.: A comparison of statistical downscaling methods suited for wildfire applications, Int. J. Climatol., 32, 772–780, https://doi.org/10.1002/joc.2312, 2012.

Ahmed, K. F., Wang, G., Silander, J., Wilson, A. M., Allen, J. M., Horton, R., and Anyah, R.: Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U.S. northeast, Global Planet. Change, 100, 320–332, https://doi.org/10.1016/j.gloplacha.2012.11.003, 2013.

Albergel, C., Dutra, E., Munier, S., Calvet, J.-C., Munoz-Sabater, J., de Rosnay, P., and Balsamo, G.: ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better?, Hydrol. Earth Syst. Sci., 22, 3515–3532, https://doi.org/10.5194/hess-22-3515-2018, 2018.

Alidoost, F., Stein, A., Su, Z., and Sharifi, A.: Multivariate copula quantile mapping for bias correction of reanalysis air temperature data, J. Spatial Sci., 66, 299–315, https://doi.org/10.1080/14498596.2019.1601138, 2019.

Amante, C. and Eakins, B. W.: ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, NOAA Technical Memorandum NESDIS NGDC-24. National Geophysical Data Center, NOAA, https://doi.org/10.7289/V5C8276M, 2009.

Andrade, C. W. L., Montenegro, S. M. G. L., Montenegro, A. A. A., Lima, J. R. D. S., Srinivasan, R., and Jones, C. A.: Climate change impact assessment on water resources under RCP scenarios: A case study in Mundaú River Basin, Northeastern Brazil, Int. J. Climatol., 41, E1045–E1061, https://doi.org/10.1002/joc.6751, 2021.

Anghileri, D., Monhart, S., Zhou, C., Bogner, K., Castelletti, A., Burlando, P., and Zappa, M.: The Value of Subseasonal Hydrometeorological Forecasts to Hydropower Operations: How Much Does Preprocessing Matter?, Water Resour. Res., 55, 10159–10178, https://doi.org/10.1029/2019WR025280, 2019.

Block, P.: Tailoring seasonal climate forecasts for hydropower operations, Hydrol. Earth Syst. Sci., 15, 1355–1368, https://doi.org/10.5194/hess-15-1355-2011, 2011.

Boé, J., Terray, L., Habets, F., and Martin, E.: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies, Int. J. Climatol., 27, 1643–1655, https://doi.org/10.1002/joc.1602, 2007.

Bolson, J., Martinez, C., Breuer, N., Srivastava, P., and Knox, P.: Climate information use among southeast US water managers: beyond barriers and toward opportunities, Reg. Environ. Change, 13, 141–151, https://doi.org/10.1007/s10113-013-0463-1, 2013.

Brier, G. W.: Verification of forecasts expressed in terms of probability, Mon. Weather Rev., 78, 1–3, https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2, 1950.

Briley, L. J., Ashley, W. S., Rood, R. B., and Krmenec, A.: The role of meteorological processes in the description of uncertainty for climate change decision-making, Theor. Appl. Climatol., 127, 643–654, https://doi.org/10.1007/s00704-015-1652-2, 2017.

Cannon, A. J.: Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables, Clim. Dynam., 50, 31–49, https://doi.org/10.1007/s00382-017-3580-6, 2018.

Casati, B., Wilson, L. J., Stephenson, D. B., Nurmi, P., Ghelli, A., Pocernich, M., Damrath, U., Ebert, E. E., Brown, B. G., and Mason, S.: Forecast verification: current status and future directions, Meteorol. Appl., 15, 3–18, https://doi.org/10.1002/met.52, 2008.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, Water Resour. Res., 49, 4187–4205, https://doi.org/10.1002/wrcr.20331, 2013.

Crochemore, L., Ramos, M.-H., and Pappenberger, F.: Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts, Hydrol. Earth Syst. Sci., 20, 3601–3618, https://doi.org/10.5194/hess-20-3601-2016, 2016.

Digna, R., Castro-Gama, M., van der Zaag, P., Mohamed, Y., Corzo, G., and Uhlenbrook, S.: Optimal Operation of the Eastern Nile System Using Genetic Algorithm, and Benefits Distribution of Water Resources Development, Water, 10, 921, https://doi.org/10.3390/w10070921, 2018.

Domínguez-Castro, F., García-Herrera, R., and Vicente-Serrano, S. M.: Wet and dry extremes in Quito (Ecuador) since the 17th century, Int. J. Climatol., 38, 2006–2014, https://doi.org/10.1002/joc.5312, 2018.

Dutra, E., Di Giuseppe, F., Wetterhall, F., and Pappenberger, F.: Seasonal forecasts of droughts in African basins using the Standardized Precipitation Index, Hydrol. Earth Syst. Sci., 17, 2359–2373, https://doi.org/10.5194/hess-17-2359-2013, 2013.

ECMWF: ERA5-Land hourly data from 1981 to present, Tech. rep., ECMWF, https://doi.org/10.24381/cds.e2161bac, 2019.

ECMWF: ERA5-Land: data documentation, Tech. rep., 2020.

Elagib, N. A. and Elhag, M. M.: Major climate indicators of ongoing drought in Sudan, J. Hydrol., 409, 612–625, https://doi.org/10.1016/j.jhydrol.2011.08.047, 2011.

Emerton, R. E., Stephens, E. M., and Cloke, H. L.: What is the most useful approach for forecasting hydrological extremes during El Niño?, Environ. Res. Commun., 1, 031002, https://doi.org/10.1088/2515-7620/ab114e, 2019.

Gerlitz, L., Vorogushyn, S., and Gafurov, A.: Climate informed seasonal forecast of water availability in Central Asia: State-of-the-art and decision making context, Water Security, 10, 100061, https://doi.org/10.1016/j.wasec.2020.100061, 2020.

Gubler, S., Sedlmeier, K., Bhend, J., Avalos, G., Coelho, C. A. S., Escajadillo, Y., Jacques-Coper, M., Martinez, R., Schwierz, C., de Skansi, M., and Spirig, C.: Assessment of ECMWF SEAS5 Seasonal Forecast Performance over South America, Weather Forecast., 35, 561–584, https://doi.org/10.1175/WAF-D-19-0106.1, 2019.

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., and Rasmussen, R. M.: An intercomparison of statistical downscaling methods used for water resource assessments in the United States, Water Resour. Res., 50, 7167–7186, https://doi.org/10.1002/2014WR015559, 2014.

Haiden, T., Janousek, M., Bidlot, J., Buizza, R., Ferranti, L., Prates, F., and Vitart, F.: Evaluation of ECMWF forecasts, including the 2018 upgrade, Tech. rep., ECMWF Tech. Memo. 831, 2018.

Hartmann, H. C., Pagano, T. C., Sorooshian, S., and Bales, R.: Confidence Builders: Evaluating Seasonal Climate Forecasts from User Perspectives, B. Am. Meteorol. Soc., 83, 683–698, https://doi.org/10.1175/1520-0477(2002)083<0683:CBESCF>2.3.CO;2, 2002.

Hermanson, L., Ren, H.-L., Vellinga, M., Dunstone, N. D., Hyder, P., Ineson, S., Scaife, A. A., Smith, D. M., Thompson, V., Tian, B., and Williams, K. D.: Different types of drifts in two seasonal forecast systems and their dependence on ENSO, Clim. Dynam., 51, 1411–1426, https://doi.org/10.1007/s00382-017-3962-9, 2018.

Hersbach, H.: Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, Weather Forecast., 15, 559–570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2, 2000.

Hersbach, H., De Rosnay, P., Bell, B., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Balmaseda, A., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J., De Boisséson, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P., Dee, D., Dragani, R., Diamantakis, M., Flemming, J., Forbes, R., Geer, A., Haiden, T., Hólm, E., Haimberger, L., Hogan, R., Horányi, A., Janisková, M., Laloyaux, P., Lopez, P., Muñoz-Sabater, J., Peubey, C., Radu, R., Richardson, D., Thépaut, J.-N., Vitart, F., Yang, X., Zsótér, E., and Zuo, H.: Operational global reanalysis: progress, future directions and synergies with NWP including updates on the ERA5 production status, Tech. rep., ECMWF, ERA Report Series 27, https://doi.org/10.21957/tkic6g3wm, 2018.

Hersbach, H., Bell, B., Berrisford, P., Horányi, A., Munoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., and Dee, D.: Global reanalysis: goodbye ERA-Interim, hello ERA5, ECMWF Newsletter, 159, Spring 2019.

Hwang, S. and Graham, W. D.: Development and comparative evaluation of a stochastic analog method to downscale daily GCM precipitation, Hydrol. Earth Syst. Sci., 17, 4481–4502, https://doi.org/10.5194/hess-17-4481-2013, 2013.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S. P. E., Mogensen, K., Zuo, H., and Monge-Sanz, B. M.: SEAS5: the new ECMWF seasonal forecast system, Geosci. Model Dev., 12, 1087–1117, https://doi.org/10.5194/gmd-12-1087-2019, 2019.

Khajehei, S., Ahmadalipour, A., and Moradkhani, H.: An effective post-processing of the North American multi-model ensemble (NMME) precipitation forecasts over the continental US, Clim. Dynam., 51, 457–472, https://doi.org/10.1007/s00382-017-3934-0, 2018.

Khalili, A. and Rahimi, J.: High-resolution spatiotemporal distribution of precipitation in Iran: a comparative study with three global-precipitation datasets, Theor. Appl. Climatol., 118, 211–221, https://doi.org/10.1007/s00704-013-1055-1, 2014.

Kidus, A. E.: Long-term potential impact of Great Ethiopian Renaissance Dam (GERD) on the downstream eastern Nile High Aswan Dam (HAD), Sustainable Water Resources Management, 5, 1973–1980, https://doi.org/10.1007/s40899-019-00351-0, 2019.

Lafon, T., Dadson, S., Buys, G., and Prudhomme, C.: Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods, Int. J. Climatol., 33, 1367–1381, https://doi.org/10.1002/joc.3518, 2013.

Lehner, B. and Grill, G.: Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems, Hydrol. Process., 27, 2171–2186, https://doi.org/10.1002/hyp.9740, 2013.

Lemos, M. C., Finan, T. J., Fox, R. W., Nelson, D. R., and Tucker, J.: The Use of Seasonal Climate Forecasting in Policymaking: Lessons from Northeast Brazil, Clim. Change, 55, 479–507, https://doi.org/10.1023/A:1020785826029, 2002.

Lorenz, C. and Kunstmann, H.: The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis, J. Hydrometeorol., 13, 1397–1420, https://doi.org/10.1175/JHM-D-11-088.1, 2012.

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N., Riegger, J., and Kunstmann, H.: Large-scale runoff from landmasses: a global assessment of the closure of the hydrological and atmospheric water balances, J. Hydrometeorol., 15, 2111–2139, https://doi.org/10.1175/JHM-D-13-0157.1, 2014.

Lorenz, C., Montzka, C., Jagdhuber, T., Laux, P., and Kunstmann, H.: Long-Term and High-Resolution Global Time Series of Brightness Temperature from Copula-Based Fusion of SMAP Enhanced and SMOS Data, Remote Sensing, 10, 1842, https://doi.org/10.3390/rs10111842, 2018.

Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H.: Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Catamayo-Chira Basin (Ecuador/Peru) [dataset], World Data Center for Climate (WDCC) at DKRZ, https://doi.org/10.26050/WDCC/SaWaM_D04_SEAS5_BCSD, 2020a.

Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H.: Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Karun Basin (Iran) [dataset],

World Data Center for Climate (WDCC) at DKRZ, https://doi.org/10.26050/WDCC/SaWaM_D01_SEAS5_BCSD, 2020b.

Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H.: Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Rio São Francisco Basin (Brazil) [dataset], World Data Center for Climate (WDCC) at DKRZ, https://doi.org/10.26050/WDCC/SaWaM_D02_SEAS5_BCSD, 2020c.

Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H.: Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Tekeze-Atbara and Blue Nile Basins (Ethiopia/Eritrea/Sudan) [dataset], World Data Center for Climate (WDCC) at DKRZ, https://doi.org/10.26050/WDCC/SaWaM_D03_SEAS5_BCSD, 2020d.

Magnusson, L., Alonso-Balmaseda, M., Corti, S., Molteni, F., and Stockdale, T.: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors, Clim. Dynam., 41, 2393–2409, https://doi.org/10.1007/s00382-012-1599-2, 2013.

Mahto, S. S. and Mishra, V.: Does ERA-5 Outperform Other Reanalysis Products for Hydrologic Applications in India?, J. Geophys. Res.-Atmos., 124, 9423–9441, https://doi.org/10.1029/2019JD031155, 2019.

Manzanas, R., Gutiérrez, J., Fernández, J., van Meijgaard, E., Calmanti, S., Magariño, M., Cofiño, A., and Herrera, S.: Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: Added value for user applications, Climate Services, 9, 44–56, https://doi.org/10.1016/j.cliser.2017.06.004, 2018a.

Manzanas, R., Lucero, A., Weisheimer, A., and Gutiérrez, J. M.: Can bias correction and statistical downscaling methods improve the skill of seasonal precipitation forecasts?, Clim. Dynam., 50, 1161–1176, https://doi.org/10.1007/s00382-017-3668-z, 2018b.

Manzanas, R., Gutiérrez, J. M., Bhend, J., Hemri, S., Doblas-Reyes, F. J., Torralba, V., Penabad, E., and Brookshaw, A.: Bias adjustment and ensemble recalibration methods for seasonal forecasting: a comprehensive intercomparison using the C3S dataset, Clim. Dynam., 53, 1287–1305, https://doi.org/10.1007/s00382-019-04640-4, 2019.

Marengo, J. A., Chou, S. C., Kay, G., Alves, L. M., Pesquero, J. F., Soares, W. R., Santos, D. C., Lyra, A. A., Sueiro, G., Betts, R., Chagas, D. J., Gomes, J. L., Bustamante, J. F., and Tavares, P.: Development of regional future climate change scenarios in South America using the Eta CPTEC/HadCM3 climate change projections: climatology and regional analyses for the Amazon, São Francisco and the Paraná River basins, Clim. Dynam., 38, 1829–1848, https://doi.org/10.1007/s00382-011-1155-5, 2012.

Marengo, J. A., Alves, L. M., Alvala, R. C., Cunha, A. P., Brito, S., and Moraes, O. L.: Climatic characteristics of the 2010-2016 drought in the semiarid Northeast Brazil region, Anais da Academia Brasileira de Ciências, 90, 1973–1985, https://doi.org/10.1590/0001-3765201720170206, 2018.

Martins, E. S. P. R., Coelho, C. A. S., Haarsma, R., Otto, F. E. L., King, A. D., Jan van Oldenborgh, G., Kew, S., Philip, S., Vasconcelos Júnior, F. C., and Cullen, H.: A Multimethod Attribution Analysis of the Prolonged Northeast Brazil Hydrometeorological Drought (2012–16), B. Am. Meteorol. Soc., 99, S65–S69, https://doi.org/10.1175/BAMS-D-17-0102.1, 2018.

Masih, I., Maskey, S., Mussá, F. E. F., and Trambauer, P.: A review of droughts on the African continent: a geospatial and long-term perspective, Hydrol. Earth Syst. Sci., 18, 3635–3649, https://doi.org/10.5194/hess-18-3635-2014, 2014.

Ning, L., Riddle, E. E., and Bradley, R. S.: Projected Changes in Climate Extremes over the Northeastern United States, J. Climate, 28, 3289–3310, https://doi.org/10.1175/JCLI-D-14-00150.1, 2015.

Nyaupane, N., Thakur, B., Kalra, A., and Ahmad, S.: Evaluating Future Flood Scenarios Using CMIP5 Climate Projections, Water, 10, 1866, https://doi.org/10.3390/w10121866, 2018.

Patt, A. and Gwata, C.: Effective seasonal climate forecast applications: examining constraints for subsistence farmers in Zimbabwe, Global Environ. Change, 12, 185–195, https://doi.org/10.1016/S0959-3780(02)00013-4, 2002.

Patt, A., Suarez, P., and Gwata, C.: Effects of seasonal climate forecasts and participatory workshops among subsistence farmers in Zimbabwe, P. Natl. Acad. Sci., 102, 12623–12628, https://doi.org/10.1073/pnas.0506125102, 2005.

Ratri, D. N., Whan, K., and Schmeits, M.: A Comparative Verification of Raw and Bias-Corrected ECMWF Seasonal Ensemble Precipitation Reforecasts in Java (Indonesia), J. Appl. Meteorol. Climatol., 58, 1709–1723, https://doi.org/10.1175/JAMC-D-18-0210.1, 2019.

Rayner, S., Lach, D., and Ingram, H.: Weather Forecasts are for Wimps: Why Water Resource Managers Do Not Use Climate Forecasts, Clim. Change, 69, 197–227, https://doi.org/10.1007/s10584-005-3148-z, 2005.

Ritchie, J. W., Abawi, G. Y., Dutta, S. C., Harris, T. R., and Bange, M.: Risk management strategies using seasonal climate forecasting in irrigated cotton production: a tale of stochastic dominance, Aust. J. Agr. Resour. Ec., 48, 65–93, https://doi.org/10.1111/j.1467-8489.2004.00236.x, 2008.

Schepen, A., Zhao, T., Wang, Q. J., and Robertson, D. E.: A Bayesian modelling method for post-processing daily subseasonal to seasonal rainfall forecasts from global climate models and evaluation for 12 Australian catchments, Hydrol. Earth Syst. Sci., 22, 1615–1628, https://doi.org/10.5194/hess-22-1615-2018, 2018.

Siegmund, J., Bliefernicht, J., Laux, P., and Kunstmann, H.: Toward a seasonal precipitation prediction system for West Africa: Performance of CFSv2 and high-resolution dynamical downscaling, J. Geophys. Res.-Atmos., 120, 7316–7339, https://doi.org/10.1002/2014JD022692, 2015.

Tall, A., Mason, S. J., van Aalst, M., Suarez, P., Ait-Chellouche, Y., Diallo, A. A., and Braman, L.: Using Seasonal Climate Forecasts to Guide Disaster Management: The Red Cross Experience during the 2008 West Africa Floods, Int. J. Geophys., 2012, 1–12, https://doi.org/10.1155/2012/986016, 2012.

Teutschbein, C. and Seibert, J.: Is bias correction of regional climate model (RCM) simulations possible for non-stationary conditions?, Hydrol. Earth Syst. Sci., 17, 5061–5077, https://doi.org/10.5194/hess-17-5061-2013, 2013.

Thober, S., Kumar, R., Sheffield, J., Mai, J., Schäfer, D., and Samaniego, L.: Seasonal Soil Moisture Drought Prediction over Europe Using the North American Multi-

Model Ensemble (NMME), J. Hydrometeorol., 16, 2329–2344, https://doi.org/10.1175/JHM-D-15-0053.1, 2015.

Thrasher, B., Maurer, E. P., McKellar, C., and Duffy, P. B.: Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping, Hydrol. Earth Syst. Sci., 16, 3309–3314, https://doi.org/10.5194/hess-16-3309-2012, 2012.

Thrasher, B., Xiong, J., Wang, W., Melton, F., Michaelis, A., and Nemani, R.: Downscaled Climate Projections Suitable for Resource Management, Eos, Transactions American Geophysical Union, 94, 321–323, https://doi.org/10.1002/2013EO370002, 2013.

Torres, R. R., Lapola, D. M., and Gamarra, N. L. R.: Future Climate Change in the Caatinga, in: Caatinga, Springer International Publishing, Cham, 383–410, https://doi.org/10.1007/978-3-319-68339-3_15, 2017.

Tryhorn, L. and DeGaetano, A.: A comparison of techniques for downscaling extreme precipitation over the Northeastern United States, Int. J. Climatol., 31, 1975–1989, https://doi.org/10.1002/joc.2208, 2011.

Urraca, R., Huld, T., Gracia-Amillo, A., Martinez-de Pison, F. J., Kaspar, F., and Sanz-Garcia, A.: Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data, Solar Energ., 164, 339–354, https://doi.org/10.1016/j.solener.2018.02.059, 2018.

Vaghefi, S. A., Keykhai, M., Jahanbakhshi, F., Sheikholeslami, J., Ahmadi, A., Yang, H., and Abbaspour, K. C.: The future of extreme climate in Iran, Sci. Rep.-UK, 9, 1464, https://doi.org/10.1038/s41598-018-38071-8, 2019.

van den Besselaar, E. J. M., van der Schrier, G., Cornes, R. C., Iqbal, A. S., and Klein Tank, A. M. G.: SA-OBS: A Daily Gridded Surface Temperature and Precipitation Dataset for Southeast Asia, J. Climate, 30, 5151–5165, https://doi.org/10.1175/JCLI-D-16-0575.1, 2017.

Vandal, T., Kodra, E., and Ganguly, A. R.: Intercomparison of machine learning methods for statistical downscaling: the case of daily and extreme precipitation, Theor. Appl. Climatol., 137, 557–570, https://doi.org/10.1007/s00704-018-2613-3, 2019.

Voisin, N., Schaake, J. C., and Lettenmaier, D. P.: Calibration and Downscaling Methods for Quantitative Ensemble Precipitation Forecasts, Weather Forecast., 25, 1603–1627, https://doi.org/10.1175/2010WAF2222367.1, 2010.

Washington, R., Harrison, M., Conway, D., Black, E., Challinor, A., Grimes, D., Jones, R., Morse, A., Kay, G., and Todd, M.: African Climate Change: Taking the Shorter Route, B. Am. Meteorol. Soc., 87, 1355–1366, https://doi.org/10.1175/BAMS-87-10-1355, 2006.

Wessel, P. and Smith, W. H. F.: A global, self-consistent, hierarchical, high-resolution shoreline database, J. Geophys. Res.-Sol. Ea., 101, 8741–8743, https://doi.org/10.1029/96JB00104, 1996.

Westrick, K. J. and Mass, C. F.: An Evaluation of a High-Resolution Hydrometeorological Modeling System for Prediction of a Cool-Season Flood Event in a Coastal Mountainous Watershed, J. Hydrometeorol., 2, 161–180, https://doi.org/10.1175/1525-7541(2001)002<0161:AEOAHR>2.0.CO;2, 2001.

Wheeler, K. G., Hall, J. W., Abdo, G. M., Dadson, S. J., Kasprzyk, J. R., Smith, R., and Zagona, E. A.: Exploring Cooperative Transboundary River Management Strategies for the Eastern Nile Basin, Water Resour. Res., 54, 9224–9254, https://doi.org/10.1029/2017WR022149, 2018.

Wheeler, K. G., Jeuland, M., Hall, J. W., Zagona, E., and Whittington, D.: Understanding and managing new risks on the Nile with the Grand Ethiopian Renaissance Dam, Nat. Commun., 11, 5222, https://doi.org/10.1038/s41467-020-19089-x, 2020.

White, C. J., Carlsen, H., Robertson, A. W., Klein, R. J., Lazo, J. K., Kumar, A., Vitart, F., Coughlan de Perez, E., Ray, A. J., Murray, V., Bharwani, S., MacLeod, D., James, R., Fleming, L., Morse, A. P., Eggen, B., Graham, R., Kjellström, E., Becker, E., Pegion, K. V., Holbrook, N. J., McEvoy, D., Depledge, M., Perkins-Kirkpatrick, S., Brown, T. J., Street, R., Jones, L., Remenyi, T. A., Hodgson-Johnston, I., Buontempo, C., Lamb, R., Meinke, H., Arheimer, B., and Zebiak, S. E.: Potential applications of subseasonal-to-seasonal (S2S) predictions, Meteorol. Appl., 24, 315–325, https://doi.org/10.1002/met.1654, 2017.

Wood, A., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-range experimental hydrologic forecasting for the eastern United States, J. Geophys. Res., 107, 4429, https://doi.org/10.1029/2001JD000659, 2002.

Xue, Y., Chen, M., Kumar, A., Hu, Z.-Z., and Wang, W.: Prediction Skill and Bias of Tropical Pacific Sea Surface Temperatures in the NCEP Climate Forecast System Version 2, J. Climate, 26, 5358–5378, https://doi.org/10.1175/JCLI-D-12-00600.1, 2013.

Yitayew, M. and Melesse, A. M.: Critical Water Resources Issues in the Nile River Basin, in: Nile River Basin, pp. 401–416, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-0689-7_20, 2011.

Yuan, X., Wood, E. F., Luo, L., and Pan, M.: A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction, Geophys. Res. Lett., 38, L13402, https://doi.org/10.1029/2011GL047792, 2011.

Zhao, T., Bennett, J. C., Wang, Q. J., Schepen, A., Wood, A. W., Robertson, D. E., and Ramos, M.-H.: How Suitable is Quantile Mapping For Postprocessing GCM Precipitation Forecasts?, J. Climate, 30, 3185–3196, https://doi.org/10.1175/JCLI-D-16-0652.1, 2017.

# 4 Article III

## Ensemble-Tailored Pattern Analysis of High-Resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America

---

Portele, T.C., Laux, P., Lorenz, C., Janner, A., Horna, N., Fersch, B., Iza, M., and Kunstmann, H., (2021). Ensemble-Tailored Pattern Analysis of High-Resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America. *Front. Earth Sci.* **9**, 669427. https://doi.org/10.3389/feart.2021.669427



---

frontiers in Earth Science

ORIGINAL RESEARCH
published: 13 May 2021
doi: 10.3389/feart.2021.669427

# Ensemble-Tailored Pattern Analysis of High-Resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America

*Tanja C. Portele[1,2]*, Patrick Laux[1,2], Christof Lorenz[1], Annelie Janner[3], Natalia Horna[4], Benjamin Fersch[1], Maylee Iza[4] and Harald Kunstmann[1,2]*

[1]Karlsruhe Institute of Technology (KIT), Campus Alpin, Institute of Meteorology and Climate Research - Atmospheric Environmental Research (IMK-IFU), , Garmisch-Partenkirchen, Germany, [2]Institute of Geography, University of Augsburg, Augsburg, Germany, [3]University of Erlangen Nuremberg, Erlangen, Germany, [4]Instituto Nacional de Meteorología e Hidrología, Dirección de Estudios, Investigación y Desarrollo Hidrometeorológico, Quito, Ecuador

OPEN ACCESS

***Edited by:***
*Tomas Halenka,*
*Charles University, Czechia*

***Reviewed by:***
*Changhai Liu,*
*National Center for Atmospheric Research (UCAR), United States*
*Xander Wang,*
*University of Prince Edward Island, Canada*

****Correspondence:***
*Tanja C. Portele*
*tanja.portele@kit.edu*

***Specialty section:***
*This article was submitted to Interdisciplinary Climate Studies, a section of the journal Frontiers in Earth Science*

***Received:*** *18 February 2021*
***Accepted:*** *29 April 2021*
***Published:*** *13 May 2021*

***Citation:***
*Portele TC, Laux P, Lorenz C, Janner A, Horna N, Fersch B, Iza M and Kunstmann H (2021) Ensemble-Tailored Pattern Analysis of High-Resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America. Front. Earth Sci. 9:669427. doi: 10.3389/feart.2021.669427*

For climate adaptation and risk mitigation, decision makers in water management or agriculture increasingly demand for regionalized weather and climate information. To provide these, regional atmospheric models, such as the Weather Research and Forecasting (WRF) model, need to be optimized in their physical setup to the region of interest. The objective of this study is to evaluate four cumulus physics (CU), two microphysics (MP), two planetary boundary layer physics (PBL), and two radiation physics (RA) schemes in WRF according to their performance in dynamically downscaling the precipitation over two typical South American regions: one orographically complex area in Ecuador/Peru (horizontal resolution up to 9 and 3 km), and one area of rolling hills in Northeast Brazil (up to 9 km). For this, an extensive ensemble of 32 simulations over two continuous years was conducted. Including the reference uncertainty of three high-resolution global datasets (CHIRPS, MSWEP, ERA5-Land), we show that different parameterization setups can produce up to four times the monthly reference precipitation. This underscores the urgent need to conduct parameterization sensitivity studies before weather forecasts or input for impact modeling can be produced. Contrarily to usual studies, we focus on distributional, temporal and spatial precipitation patterns and evaluate these in an ensemble-tailored approach. These ensemble characteristics such as ensemble Structure-, Amplitude-, and Location-error, allow us to generalize the impacts of combining one parameterization scheme with others. We find that varying the CU and RA schemes stronger affects the WRF performance than varying the MP or PBL schemes. This effect is even present in the convection-resolving 3-km-domain over Ecuador/Peru where CU schemes are only used in the parent domain of the one-way nesting approach. The G3D CU physics ensemble best represents the CHIRPS probability distribution in the 9-km-domains. However, spatial and temporal patterns of CHIRPS are best captured by Tiedtke or BMJ CU schemes. Ecuadorian station data in the 3-km-domain is best simulated by the ensemble whose parent domains use the KF CU

scheme. Accounting for all evaluation metrics, no general-purpose setup could be identified, but suited parameterizations can be narrowed down according to final application needs.

**Keywords: weather research and forecasting (WRF) model, sensitivity, CHIRPS 2.0, northeast Brazil, Ecuador, Peru, eSAL, parameterizations**

# 1 INTRODUCTION

Increasing demands for water until 2050 will aggravate the impacts of water scarcity on agricultural production and livelihood activities (Food and Agriculture Organization of the United Nations, 2015). To successfully deal with this projected development, sustainable management of available water resources is mandatory. For improved decision support in regional water management or agriculture, high-quality regionalized weather and climate forecasting increasingly matters (Soares et al., 2012; Müller et al., 2016) and is a key requirement for successful hydrological and crop impact modeling (Decharme and Douville, 2006; Barbosa and Lakshmi Kumar, 2016; Parkes et al., 2019). Especially in semi-arid regions and areas affected by El-Niño-Southern-Oscillation (ENSO) more profound knowledge of variations in seasonal rainfall, estimated streamflow and crop yield from regionalized model output, can help to better prepare against climate extremes. Here, our focus is on the downscaling of global hydrometeorological fields over two typical regions of South America: one area of rolling hills in Northeast Brazil and one orographically complex area in Ecuador/Peru. Hit by a multiyear drought in the last decade (Marengo et al., 2018; Martins et al., 2018) and strongly influenced by El Niño (Domínguez-Castro et al., 2018), respectively, those regions are facing high climate variability especially during their rainy seasons. Water availability in those regions not only determines the production of rainfed and irrigated agriculture, it also limits the hydroelectric energy production in their highly managed river basins (von Sperling, 2012).

One way to achieve regionalized hydrometeorological information is the application of regional atmospheric models, such as the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008), to dynamically downscale coarse-grained, large-scale fields over selected areas. The performance of regional simulations highly depends on the used physical parameterization model setup, and it may vary from region to region. Therefore, these models need to be adapted in their physical setting to the region of interest. Sub-grid scale processes such as cloud microphysics (MP), cumulus convection (CU), planetary boundary layer physics (PBL), and radiation physics (RA) are parameterized in these models, and their choice and combination highly influences the modeled hydrometeorological variables (Flaounas et al., 2011; Crétat et al., 2012; Yang et al., 2012; Zhang et al., 2012; Efstathiou et al., 2013; Klein et al., 2015; Que et al., 2016; Gbode et al., 2019; Yang et al., 2021). As precipitation amounts result from the complex interaction of the applied physical schemes, they combine all uncertainties due to the model setup and are a major source of impact model uncertainty (e.g., Decharme and Douville, 2006; Barbosa and Lakshmi Kumar, 2016; Parkes et al., 2019). This study therefore focuses on the performance of the WRF setup in terms of reproduction of precipitation patterns.

Comprehensive parameterization studies are common practice to decide on the best-suited regional model setup before conducting operational weather forecasts, long-term climate projections or impact studies. Previous WRF parameterization studies, relevant for our focus regions, did not reach high resolutions below 10 km for regional applications (e.g., 30 km in the Amazon basin; Fersch and Kunstmann, 2014), chose a very limited domain (e.g., 15 km resolution over a 225 × 225 km domain in Southern Ecuador; Ochoa et al., 2016), included only parts of our study domains (e.g., 40 km resolution over South America between −10° and −50°S; Ruiz et al., 2010), did not pursue a full factorial combination of applied physics schemes (e.g., high-resolution runs at 18–6–3 km over the Peruvian Andes; Moya-Álvarez et al., 2018) or studied only short-term events (e.g., high-resolution runs at 25–5–1.66 km in Northeast Brazil for a 10-day period; Comin et al., 2020). Ochoa et al. (2016) and Ruiz et al. (2010) further lacked a profound analysis with respect to the effects of different physics schemes on precipitation. Fersch and Kunstmann (2014) already showed difficulties of the different WRF setups in reproducing the observed precipitation. Our study therefore aims at a detailed, multi-target evaluation tailored for the use of ensembles (ensemble-tailored) of the conducted high-resolution WRF physics ensemble runs including four CU, two MP, two PBL, and two RA physics in a full-factorial combination of physics schemes over a continuous 2-years period. With high-resolution, daily, global gridded precipitation reference data available, the regionalized WRF precipitation at target resolutions of 9 and 3 km can be extensively validated: gridpoint-based, domain-based and feature-based, allowing the consideration of different aspects needed when used as input for spatially distributed impact models or for running regional climate projections or weather forecasts. We first demonstrate the necessity of model parameterization studies in these regions of South America, and then differentiate the impact of different physical parameterization schemes on the distributional, temporal and spatial patterns of the high-resolution dynamically downscaled precipitation fields.

# 2 METHODS

## 2.1 WRF Model Configuration

The Weather Research and Forecasting (WRF) model is a numerical weather prediction (NWP) model, used for both

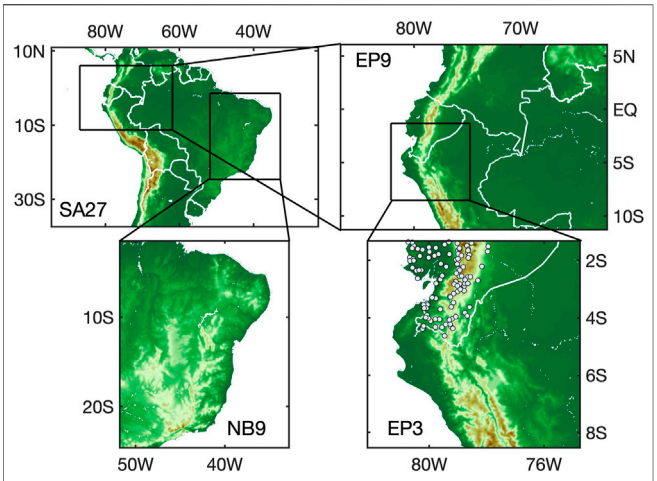


**FIGURE 1 |** Regional dynamical downscaling over the Ecuadorian-Peruvian and Brazilian domain. In a nested approach, the input data of the ERA-Interim reanalyses is downscaled to 27 km (SA27), 9 km (EP9) and 3 km (EP3) over Ecuador-Peru and to 9 km (NB9) over Northeast Brazil. The dots in EP3 represent 106 selected precipitation gauges of the Ecuadorian Weather Service INAMHI that have less than 5% of missing values during Jul 2006–Jun 2008.

research and operational applications (Skamarock et al., 2008). In this project, WRF Version 3.9.1 with the Advanced Research WRF (ARW) solver was used for dynamical downscaling. Initial and lateral boundary conditions for our simulations were used from interpolated data of the European Center for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA) Interim (Dee et al., 2011) with a spatial resolution of 79 km, 60 vertical levels and a model top at 0.1 hPa. The initial and lateral boundary conditions used in WRF are wind components, temperature, water vapor, surface pressure, sea surface temperature (SST) and soil moisture, all of them in different atmospheric and soil height levels, respectively. The boundary conditions were updated at every WRF model time step from the linearly interpolated 6-hourly global fields.

The spatial setup of WRF consists of four domains, with the first domain (SA27) representing the 27-km-resolution parent domain with 288 × 206 grid points. In a 1-way nested approach (no feedback), the horizontal grid resolution was stepwise increased from 27 km (SA27) to 9 km (EP9 and NB9), and finally to 3 km (EP3, **Figure 1**). This last step to 3 km was only performed over the highly mountainous Ecuadorian-Peruvian region to be able to represent topographically induced processes. Over the topographically more homogeneous Brazilian region, the final horizontal grid resolution was 9 km. The simulated time step was 90 s for SA27 with a parent time step ratio of 1:3. For some simulations, the time step of SA27 needed to be reduced to 30 s for numerically stable simulations. In the vertical, the domains consisted of 100 terrain-following hydrostatic pressure levels with a model top at 10 hPa.

Simulations were continuously run over a period of 2.5 years (Jan 2006 to Jun 2008) to capture different wet and dry conditions during the rainy and dry seasons of the regions and including a 6-months spin-up for soil moisture equilibrium adjustment. In their sensitivity study on the effect of spin-up length to atmospheric variables like precipitation and 2-m temperature, Jerez et al. (2020) recommend a six-months spin-up period as the best compromise between required computational expense and any remaining imbalance of the soil subsystem. Analyzing the behavior of relative soil moisture content in different parameterization experiments for the 2.5 years simulation period suggests that a 6-months period is sufficient for obtaining the dynamical soil moisture equilibrium in all domains (not shown). The time period from Jan 2006 to Jun 2008 was chosen to obtain continuous simulations of medium years in the domains with respect to precipitation (**Figure 2A**). To allow for a better representation of regionalized climatological conditions, these years were further selected to be within the period of maximum assimilated observations in the atmospheric analysis component of ERA-Interim from 2002 till 2010 (Dee et al., 2011; Simmons et al., 2014). Ochoa et al. (2016) further recommend to include ENSO neutral years for physics parameterization studies, as high uncertainties can be introduced by ENSO especially in the western coastal regions of Ecuador and Peru. The inclusion of El-Niño-years would be an additional research question and would require the simulation of at least 10 years to average out the strong effects of El Niño. This study focuses on years with typical average climatological conditions. By excluding the spin-up period, we focus our analysis on the period Jul 2006–Jun 2008. For the validation against gridded reference data, the WRF grid was bilinearly interpolated from 9 km to 0.1° for EP9 and NB9, and from 3 km to 0.05° for EP3. For the comparison with station observation data, the nearest grid cell to the station location of the 3 km EP3 grid was used. The relaxation zones of the nested domains were omitted for the analysis.

## 2.2 Tested WRF Physics Parameterization Schemes

Testing all possible combinations of all available physics schemes for several years is computationally too expensive. Therefore, we chose a representative sample of WRF physics schemes, including different levels of complexity and formulation of particular processes. For precipitation generation, the cumulus and microphysics schemes are mainly responsible, representing convective and non-convective precipitation, respectively. Testing four different CU schemes, our main focus is on the performance of CU schemes that are already reported to have major impact on the simulated precipitation among the different physics scheme groups (Crétat et al., 2012; Sikder and Hossain, 2016; Ochoa et al., 2016; Yang et al., 2021), particularly in regions producing predominantly convective precipitation like in the tropics. However, also different MP, PBL and RA schemes largely contribute to precipitation uncertainties (Klein et al., 2015; Ulate et al., 2014; Efstathiou et al., 2013; Flaounas et al., 2011), especially for high spatial resolution <5 km for which convective processes are expected to be captured and CU schemes should not be used (Hsiao et al., 2013; Skamarock et al., 2008). For

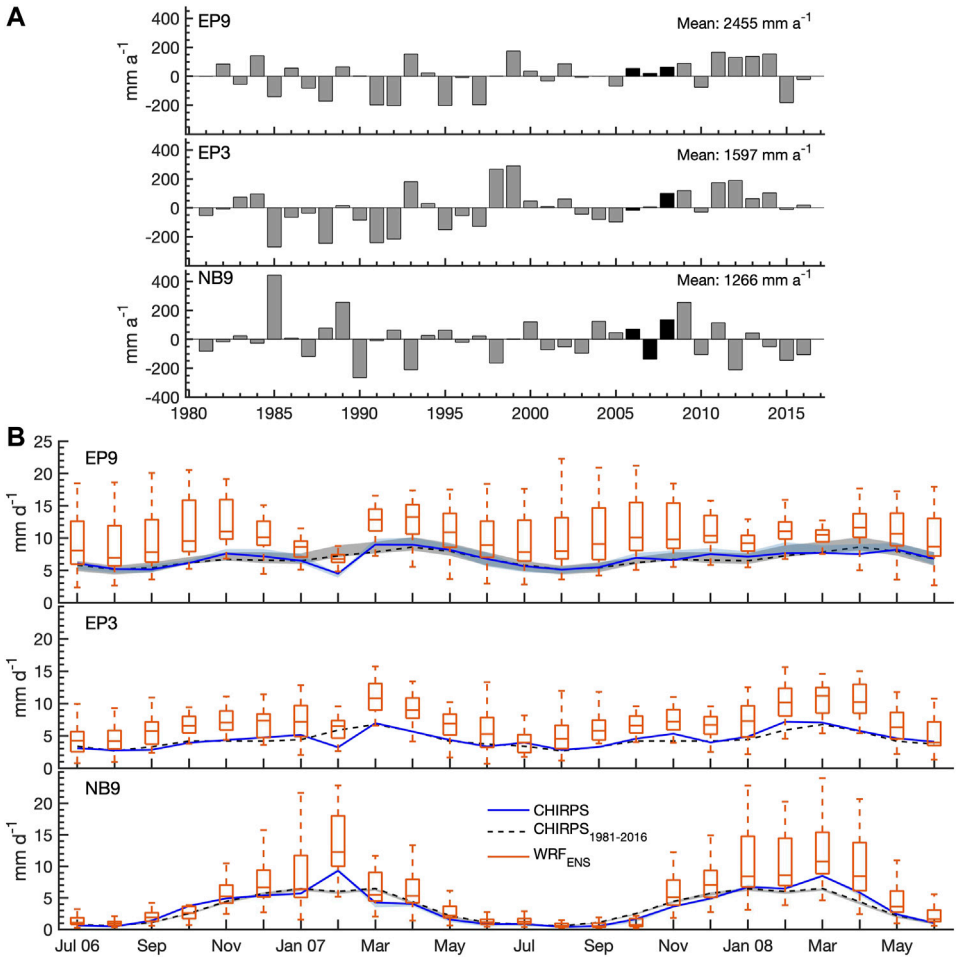


**FIGURE 2 | (A)** Annual domain-averaged precipitation anomalies (mm $a^{-1}$) of CHIRPS for EP9 **(upper)**, EP3 **(middle)** and NB9 **(lower)** with respect to the climatological mean from 1981 to 2016. Simulated years of 2006–2008 are marked with black. **(B)** CHIRPS timeseries and 1981–2016 climatology, together with WRF ensemble uncertainty of monthly domain-averaged precipitation (mm $d^{-1}$) for Jul 2006 till Jun 2008 for EP9 **(upper)**, EP3 **(middle)** and NB9 **(lower)**. The uncertainty in the reference data by including MSWEP and ERA5-Land for EP9 and NB9 is depicted by shaded areas.

this reason, we also tested two different schemes each for MP, PBL and RA. In total, 32 simulations from the full factorial combination of the selected schemes were run (**Table 1**). Regarding land surface physics parameterizations, we fixed the Noah Land surface model (LSM; Chen and Dudhia, 2001), with soil temperature and moisture in four levels, for all runs, as previous studies (e.g., Crétat et al., 2012; Tariku and Gan, 2017) found smaller effects of different LSMs over the amount of precipitation compared to the other mentioned physics parameterization groups. Using the nested approach, all study domains (**Figure 1**) were run in one simulation at the same time using the same physics setup. For the 3-km-domain, CU schemes were not used, but explicitly resolved convection was assumed.

The cumulus physics schemes manage subgrid-scale convection processes and shallow clouds. The timing and location of cumulus convection is controlled by a trigger function, and the adjustment method controls how profiles of temperature and moisture are modified according to the triggering in the convection scheme. The used schemes of Grell-3 (improved Grell-Devenyi scheme; G3D, Grell and Dévényi, 2002; Skamarock et al., 2008), New Tiedtke (Tiedtke, Zhang and Wang, 2017) and the modified version of Kain-Fritsch (KF, Kain and Fritsch, 1990; Kain, 2004) are mass-flux schemes. Mass-flux schemes have explicit updrafts to transport air from the updraft source layer upwards while reducing the convective available potential energy (CAPE). They also include compensating environmental subsidence, i.e., downdrafts, around clouds. Among the mass-flux schemes, G3D produces an ensemble of triggers and closures (determines cloud strength) and feeds back the ensemble mean to the model. G3D is especially suitable for small grid sizes as it allows the spreading of subsidence effects to neighboring grid columns (Skamarock et al., 2008). However, this subsidence spreading is only recommended for grid sizes ≤5 km (Skamarock et al., 2008; Grell and Freitas, 2014) and is therefore not used for the 9 km domain. For G3D, also the shallow convection ("ishallow"), i.e., an additional scheme for non-precipitating shallow clouds by enhanced mass-flux, is used. In contrast to the other schemes,

**TABLE 1 |** Tested 32 WRF setups of parameterization combinations of 4 cumulus convection (CU), 2 microphysics (MP), 2 planetary boundary layer (PBL), and 2 radiation (RA) physics schemes.

| Run # | CU | MP | PBL | RA |
|---|---|---|---|---|
| 1 | G3D + shallow | WSM3 | YSU | RRTMG |
| 2 | G3D + shallow | WSM3 | YSU | RRTM + Dudhia |
| 3 | G3D + shallow | WSM3 | ACM2 | RRTMG |
| 4 | G3D + shallow | WSM3 | ACM2 | RRTM + Dudhia |
| 5 | G3D + shallow | WSM6 | YSU | RRTMG |
| 6 | G3D + shallow | WSM6 | YSU | RRTM + Dudhia |
| 7 | G3D + shallow | WSM6 | ACM2 | RRTMG |
| 8 | G3D + shallow | WSM6 | ACM2 | RRTM + Dudhia |
| 9 | Tiedtke | WSM3 | YSU | RRTMG |
| 10 | Tiedtke | WSM3 | YSU | RRTM + Dudhia |
| 11 | Tiedtke | WSM3 | ACM2 | RRTMG |
| 12 | Tiedtke | WSM3 | ACM2 | RRTM + Dudhia |
| 13 | Tiedtke | WSM6 | YSU | RRTMG |
| 14 | Tiedtke | WSM6 | YSU | RRTM + Dudhia |
| 15 | Tiedtke | WSM6 | ACM2 | RRTMG |
| 16 | Tiedtke | WSM6 | ACM2 | RRTM + Dudhia |
| 17 | BMJ | WSM3 | YSU | RRTMG |
| 18 | BMJ | WSM3 | YSU | RRTM + Dudhia |
| 19 | BMJ | WSM3 | ACM2 | RRTMG |
| 20 | BMJ | WSM3 | ACM2 | RRTM + Dudhia |
| 21 | BMJ | WSM6 | YSU | RRTMG |
| 22 | BMJ | WSM6 | YSU | RRTM + Dudhia |
| 23 | BMJ | WSM6 | ACM2 | RRTMG |
| 24 | BMJ | WSM6 | ACM2 | RRTM + Dudhia |
| 25 | KF | WSM3 | YSU | RRTMG |
| 26 | KF | WSM3 | YSU | RRTM + Dudhia |
| 27 | KF | WSM3 | ACM2 | RRTMG |
| 28 | KF | WSM3 | ACM2 | RRTM + Dudhia |
| 29 | KF | WSM6 | YSU | RRTMG |
| 30 | KF | WSM6 | YSU | RRTM + Dudhia |
| 31 | KF | WSM6 | ACM2 | RRTMG |
| 32 | KF | WSM6 | ACM2 | RRTM + Dudhia |

Tiedtke is able to include momentum transport. The fourth used scheme of Betts–Miller–Janjic (BMJ, Janjic, 1994; Janjic, 2000) is based on convective adjustment of atmospheric profiles to post-convective (mixed) soundings and does not have explicit updrafts or downdrafts. In contrast to the others, BMJ also does not include cloud and ice detrainment at the cloud top.

Microphysics schemes produce clouds by resolved-scale radiative, dynamical or convective processes. They include latent heat release, cloud and precipitation processes and different particle types associated with the phase changes of water between the vapor, liquid and solid phases (Skamarock et al., 2008). The microphysics therewith provide atmospheric heat and moisture tendencies, as well as resolved-scale surface precipitation. The conditions at grid cells, e.g., supersaturation or temperature below 0°C, −40°C, determine the distribution of water among the represented species. For our study, we chose the WRF single moment 3-class simple ice scheme (WSM3) and the more complex WRF single moment 6-class (WSM6) scheme. WSM3 (Hong et al., 2004) includes simulated mixing ratios for water vapor (Qv), cloud ice crystals (Qi), and snow (Qs) and their interactions below 0°C, as well as water vapor (Qv), cloud liquid water (Qc), and rain liquid water (Qr) and their interactions above 0°C. Compared to WSM3, WSM6 (Hong and Lim, 2006) also represents graupel (Qg) and more interactions between the classes. Both WSM3 and WSM6 do not include hail (frozen drops, Qh) and do not predict number concentrations of species like double moment schemes.

Planetary boundary layer physics schemes control sub-grid-scale vertical fluxes of heat, momentum and other quantities like moisture via eddy transports in the whole atmospheric column. To obtain turbulent fluxes, a closure scheme is needed to relate the unknown turbulent flux terms in the equations to known variables like mean state or gradients. Whether fluxes depend on local or nonlocal values and gradients defines if mixing with nearby (local) or distant (nonlocal) grid cells is allowed. The two selected schemes of Yonsei University (YSU, Hong et al., 2006) and Asymmetrical Convective Model version 2 (ACM2, Pleim, 2007) are both first-order closures. YSU is a nonlocal closure scheme and explicitly treats the entrainment layer at the PBL top based on results from large-eddy simulations. As a hybrid local-nonlocal scheme, ACM2 combines explicit nonlocal exchange of the surface layer to all layers above with local eddy diffusion between the other layers and the immediately next layer above, together better representing the effect of diurnal heating (Pleim, 2007).

Radiation physics schemes are used to describe longwave and shortwave radiation processes, including absorption, reflection, and scattering, and they determine the heating of the ground and the atmosphere. We tested the combination of the rapid radiative transfer model (RRTM, Mlawer et al., 1997) for longwave radiation and the Dudhia shortwave scheme (Dudhia, 1989), and the rapid radiative transfer model for general circulation models (RRTMG, Iacono et al., 2008) for both longwave and shortwave radiation. Whereas shortwave Dudhia only performs downward integration of the solar flux per column, shortwave RRTMG considers both downward and upward (reflected) fluxes. In contrast to shortwave RRTMG, Dudhia has no ozone effect that would maintain a warm stratosphere. Both shortwave Dudhia and longwave RRTM work with a binary cloud fraction (0/1), whereas RRTMG includes subgrid-scale cloud variability with a statistical method of maximum-random cloud overlap in different layers. Longwave RRTMG uses the same basic physics and absorption coefficients as longwave RRTM, but includes some modifications to improve computational efficiency.

## 2.3 Validation Reference Datasets

For the validation of simulated WRF precipitation, the steady decrease in the number of precipitation gauges and the lack of a consistently best-performing global dataset (Lorenz et al., 2014) has to be kept in mind. Especially in Northeast Brazil and coastal Ecuador-Peru, the number of rain gauges dramatically decreased between 1980 and 2009 (Lorenz et al., 2014). Hence, for the validation of WRF precipitation, we focus on global publicly available daily datasets of different observational origin, that are natively provided at roughly the horizontal resolution of the WRF domains of 9 and 3 km. We use 1) blended gauge-satellite precipitation estimates of the Climate Hazards Group Infrared Precipitation with Stations version 2.0 (CHIRPS, Funk et al., 2015), 2) merged precipitation estimates from satellite and reanalysis data of the Multi-Source Weighted-Ensemble

Precipitation version 2.1 (MSWEP, Beck et al., 2019), and 3) a pure reanalysis product of a land component replay of the ECMWF ERA5 (Hersbach et al., 2020) climate reanalysis (ERA5-Land, ECMWF, 2019).

The CHIRPS dataset is quasi-global (50°S–50°N) and has a spatial resolution of 0.05°. It is available from 1981 to near present, as a daily, pentadal and monthly precipitation dataset (Funk et al., 2015). In their final product, best available precipitation gauge data are blended with high resolution cold-cloud-duration-based rainfall estimates producing precipitation fields that are, as a monthly product, similar to gridded monthly station data of those like Global Precipitation Climatology Center (GPCC, Becker et al., 2013) or University of Eeast Anglia's Climate Research Unit (CRU, Harris et al., 2014).

The MSWEP dataset wants to take full advantage of the complementary strengths of gauge, satellite and reanalysis data (Beck et al., 2017, 2019). Applying weighted averages of precipitation anomalies of gauge observations, satellite remote sensing and atmospheric model reanalyses, MSWEP provides reliable global 3-hourly precipitation estimates at a horizontal resolution of 0.1° from 1979 to 2017 (Beck et al., 2019).

ERA5-Land (ECMWF, 2019) provides hourly data at a horizontal resolution of 0.1° from 1981 to 2–3 months before the present. Despite no direct usage of observations in the production of ERA5-Land, the assimilation of observations in the ERA5 atmospheric forcing as well as the lapse rate correction of input air temperature, air humidity and pressure in the interpolation step ensure a high quality and high resolution information of surface variables.

For the comparison with the high resolution domain EP3, precipitation station data was provided by the national weather service of Ecuador (INAMHI) for the coastal and Andes region of Southern Ecuador. For our analysis, we included up to 132 stations having less than 5% of missing values during the WRF simulation period. Precipitation was measured from 07 to 07 am local time (12–12 UTC). The usage of local station data, though only for a limited part of the domain, should provide an indication for the robustness of obtained results.

## 2.4 Analysis Methods

For the performance analysis of the WRF ensemble, the WRF runs are grouped according to their applied physics parameterization schemes (similar to Klein et al., 2015). This ensemble approach generalizes the impact of and allows to identify the uncertainties introduced by individual physics schemes. For the four cumulus physics groups, each ensemble consists of eight members using the respective scheme. The two microphysics, two planetary boundary layer and two radiation physics groups have 16 members each.

For the validation of timeseries of domain average precipitation, we use the Taylor diagram (Taylor, 2001) which allows to compare the standard deviation, Pearson correlation coefficient and root mean square difference (RMSD) of the simulations and the reference in a single plot.

For the spatial verification of the simulated precipitation fields, we apply the ensemble Structure-Amplitude-Location (eSAL) analysis (Radanovics et al., 2018; Wernli et al., 2008, 2009) to daily precipitation fields of each domain. This feature-based method compares the simulated ensemble with a reference field in terms of amplitude A, i.e., the total precipitation of the domain, location L, i.e., the location of the center of mass of the total domain and the location of the centers of mass of individual precipitation features, and structure S, i.e., the size and shape, or volume, of the precipitation features. Any contiguous gridpoints of precipitation above a given threshold are defined as a precipitation feature or object. This threshold is independently calculated for the simulation ensemble and the reference field as $R^{95} \times f$, with $R^{95}$ being the 95th percentile of all nonzero gridpoint values in the domain for the current timestep (Radanovics et al., 2018). For the simulations, $R^{95}_{\mathrm{sim}}$ is defined by all nonzero gridpoint values of all relevant ensemble members together. $f$ is a threshold factor, for daily values determined as

$$f = \max\left(\frac{1}{15}, \frac{0.01\ \mathrm{mm}}{R^{95}_{\mathrm{sim}}}, \frac{0.01\ \mathrm{mm}}{R^{95}_{\mathrm{ref}}}\right). \tag{1}$$

$f$ is thus always set to 1/15, except if this resulted in a threshold below the minimum reasonable nonzero value of 0.01 mm in the data (Radanovics et al., 2018). Both S and L depend on this threshold. In general, eSAL is only defined for nonzero precipitation fields in both the reference field and in at least one ensemble member of the simulations.

The ensemble amplitude error is defined as the relative difference of the ensemble mean ($\langle\rangle$) of domain average precipitation ($\overline{rr}$) in the simulation ensemble and the reference field:

$$eA = \frac{\langle \overline{rr_{\mathrm{sim}}} \rangle - \langle \overline{rr_{\mathrm{ref}}} \rangle}{0.5\left( \langle \overline{rr_{\mathrm{sim}}} \rangle + \overline{rr_{\mathrm{ref}}} \right)}. \tag{2}$$

$eA$ ranges from −2 to 2, with perfect agreement for $eA = 0$, too little simulated precipitation for $eA < 0$ and too high simulated precipitation for $eA > 0$. The ensemble structure error ($eS$) determines the relative difference of the ensemble mean ($\langle\rangle$) of weighted averaged scaled precipitation volumes ($V$) in the simulation ensemble and the reference field:

$$eS = \frac{\langle V_{\mathrm{sim}} \rangle - V_{\mathrm{ref}}}{0.5\left( \langle V_{\mathrm{sim}} \rangle + V_{\mathrm{ref}} \right)}, \tag{3}$$

with

$$V = \frac{\sum_i \left( rr_i \frac{rr_i}{rr_i^{\max}} \right)}{\sum_i rr_i}. \tag{4}$$

Here, $rr_i$ is the precipitation sum of all connected gridpoints in feature $i$ and $rr_i^{\max}$ the maximum gridpoint precipitation of this feature. $eS$ thus indicates if the ensemble as a whole is able to simulate, on average, the right scaled precipitation volumes. Similar to $eA$, $eS$ ranges from −2 to 2, with equally average scaled volumes for $eS = 0$, too small or too peaked simulated

features for $eS < 0$ and too large or too flat simulated features for $eS > 0$.

The ensemble location error ($eL$) consists of two parts, relating to both the entire domain ($eL_1$) and individual features ($eL_2$):

$$eL = eL_1 + eL_2. \tag{5}$$

$eL_1$ is defined as the relative distance of the ensemble mean ($\langle\rangle$) centers of mass in the simulation ensemble and the reference field:

$$eL_1 = \frac{\left|\langle \boldsymbol{x}(rr_{\text{sim}})\rangle - \boldsymbol{x}\left(rr_{\text{ref}}\right)\right|}{d}, \tag{6}$$

where $d$ is the largest distance between two domain borders and $\boldsymbol{x}(rr)$ is the coordinate vector of the center of mass of all precipitation in the domain. $eL_2$ is finally specified as twice the squared distance between the cumulative distribution functions P, i.e., the continuous ranked probability score (CRPS, Hersbach, 2000), of the relative weighted average distances between the centers of mass of individual features and the total center of mass in the simulation ensemble and in the reference field:

$$eL_2 = 2 \times \text{CRPS}\left[P\left(\frac{r_{\text{sim}}}{d}\right), P\left(\frac{r_{\text{ref}}}{d}\right)\right], \tag{7}$$

with

$$r = \frac{\sum_i rr_i|\boldsymbol{x}_i - \boldsymbol{x}|}{\sum_i rr_i}, \tag{8}$$

and

$$\text{CRPS}\left(P_{\text{sim}}, P_{\text{ref}}\right) = \int_{-\infty}^{\infty}\left[P_{\text{sim}}(x) - P_{\text{ref}}(x)\right]^2 \mathrm{d}x. \tag{9}$$

Here, $\boldsymbol{x}_i$ is the coordinate vector of the center of mass of precipitation in the feature $i$. For the reference field, the cumulative distribution function $P_{\text{ref}}$ is a step-function. Both $L_1$ and $L_2$ range between 0 and 1. A value of $L = 0$ defines a perfect ensemble in terms of location. $L = 2$ would be found for total centers of mass located at the opposite domain border ($eL_1$ close to 1) and for contrarily organized features, e.g., far from each other in one field and close to each other in the other field ($eL_2$ close to 1).

## 2.5 Climatological Domain Characteristics

According to CHIRPS, the 9-km-Ecuador-Peru (EP9) domain that includes parts of Brazil, Colombia and Venezuela, has a mean annual precipitation of 2,455 mm. The smaller 3-km-Ecuador-Peru (EP3) domain shows a long-term mean of 1597 mm $a^{-1}$. With 1266 mm $a^{-1}$, the 9-km-Northeast-Brazil (NB9) domain exhibits the lowest value. Yearly precipitation anomalies for EP9 and EP3 reach up to ±200 and ±300 mm $a^{-1}$, respectively, and for NB9 even up to ±400 mm $a^{-1}$ (**Figure 2A**). In most years, annual domain-averaged precipitation anomalies of EP9 and EP3 have the same sign, but with different magnitudes. The years 2006–2008 are characterized by small positive precipitation anomalies around 50 mm $a^{-1}$ for EP9. In the smaller domain EP3, the conditions in 2006 and 2007 roughly correspond to the long-term mean value, and the year 2008 is wetter by 100 mm $a^{-1}$. For NB9 a wet year in 2006 alternates with a dry year in 2007 and another wet year in 2008 with anomalies up to ±140 mm $a^{-1}$. In sum, the selected years 2006–2008 are a continuous period of mean precipitation conditions in the domains well suited for the WRF physics ensemble simulations.

Most of South America is dominated by a monsoon-like seasonal cycle of precipitation, with strongly contrasting conditions during austral winter and summer (Grimm, 2003; Gan et al., 2004). Correspondingly, NB9 has a distinct wet and dry season in austral summer and winter, respectively, with mean-monthly precipitation ranging between 0 and 6 mm $d^{-1}$ for CHIRPS (**Figure 2B**). EP9 and EP3 show a less pronounced seasonal cycle with a range between 5 and 9 mm $d^{-1}$, and 3 and 7 mm $d^{-1}$, respectively, which seems to be related to the spatial average over different precipitation regimes (Grimm, 2003; Celleri et al., 2007). Maximum domain-averaged monthly mean precipitation in EP9 is found between March and May, in EP3 between February and April, and in NB9 between January and March (**Figure 2B**). The observational uncertainty when further including MSWEP and ERA5-Land as reference data is up to 2 mm $d^{-1}$ for EP9 and up to 0.5 mm $d^{-1}$ for NB9. For the analysis period of Jul 2006–Jun 2008, major deviations from the climatological mean are evident for February 2007 for all domains, with drier conditions in EP9 and EP3, and wetter conditions in NB9. For March 2007, anomalies reverse for EP9 and NB9 with wetter and drier conditions, respectively. Distinct positive anomalies are further found for NB9 in March and April 2008.

# 3 RESULTS

## 3.1 Necessity of Model Parameterization Studies

For the selected 2-years analysis period, the simulated WRF ensemble uncertainty in domain-averaged monthly mean precipitation is shown in the boxplots of **Figure 2B**. Among differently parameterized simulations, the simulated precipitation can deviate by up to 20 mm $d^{-1}$ when the respective reference precipitation is around 6 mm $d^{-1}$. Both too dry and too wet WRF simulations are possible, though with an overall tendency of overestimation in all three domains. By demonstrating this large ensemble spread of simulated precipitation over EP9, EP3 and NB9, the necessity for the present model parameterization study becomes evident. The ensemble spread is by far higher than the uncertainty introduced by the different reference datasets that range between 1 and 2 mm $d^{-1}$ for EP9 and around 0.5 mm $d^{-1}$ for NB9. A remarkably lower WRF ensemble spread (<5 mm $d^{-1}$) is only found for the distinct dry season of the Northeast Brazil domain (NB9) and for the anomalously dry February 2007 in the 9-km-Ecuador-Peru- domain (EP9).

To further demonstrate the necessity of model parameterization studies based on a spatial pattern evaluation, the WRF-simulated precipitation is compared against gridded observations from CHIRPS. This is shown exemplary for the

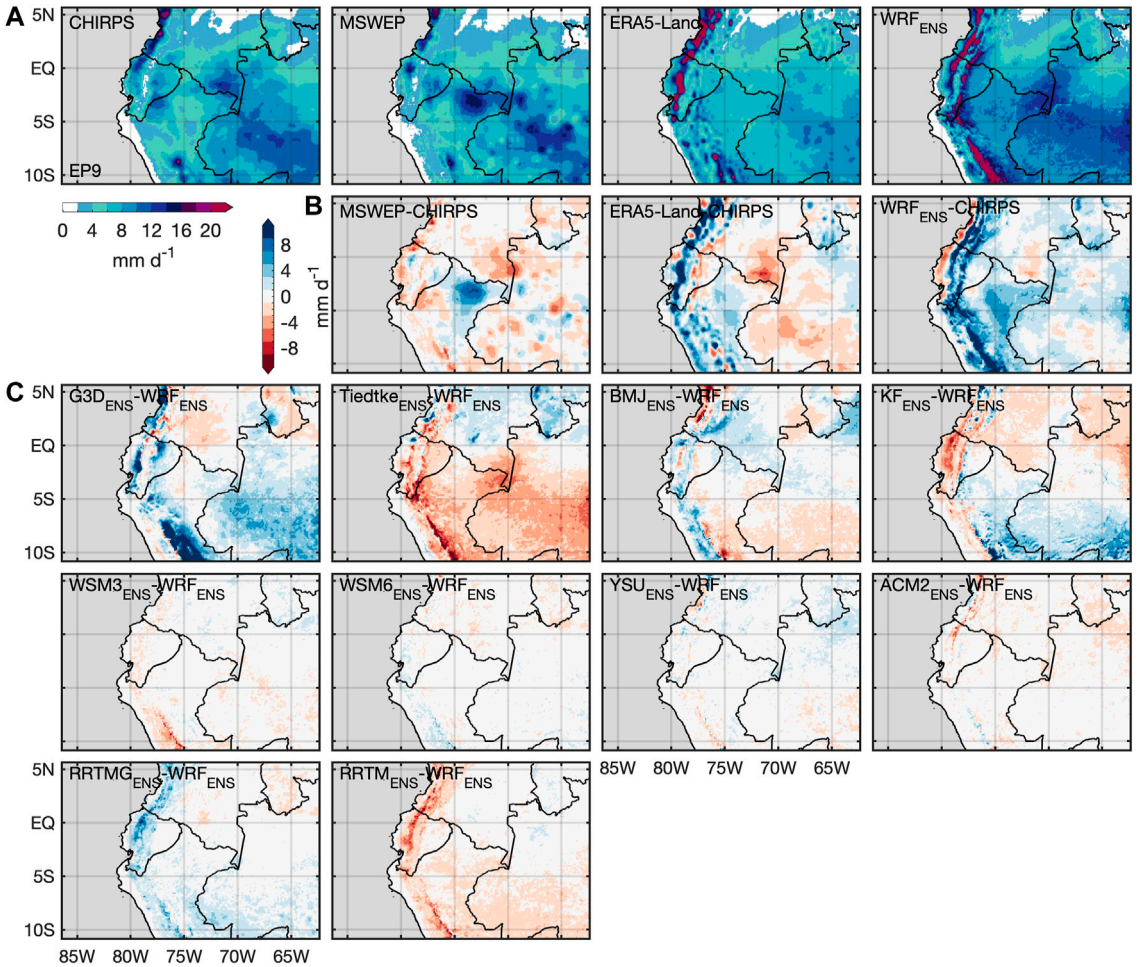


**FIGURE 3 | (A)** Observed and simulated mean precipitation in JFM 2007 of CHIRPS, MSWEP, ERA5-Land and WRF ensemble median ($WRF_{ENS}$) over EP9. **(B)** Bias of MSWEP, ERA5-Land and $WRF_{ENS}$ against CHIRPS. **(C)** Deviation of the parameterization sub-ensemble median from $WRF_{ENS}$.

three-months season of JFM within the rainy season in 2007 (**Figures 3–5**) in which largest interannual variability was evident for all three domains (**Figure 2B**). For CHIRPS, the JFM season is characterized by mean precipitation mainly above 4 mm $d^{-1}$ in all three domains (**Figures 3–5** CHIRPS). In EP9 (**Figure 3A** CHIRPS), high precipitation values >10 mm $d^{-1}$ are evident especially in the Amazon region of West Brazil, Southeast Bolivia, at the Border of Ecuador–Peru (also for EP3, **Figure 5A** CHIRPS) and in the Andes region of North Ecuador, Bolivia and Peru. Values below 4 mm $d^{-1}$ are found in the northern domain borders of Bolivia and Venezuela, as well as at the Peruvian coast and in the lee of the Ecuadorian Andes (also for EP3, **Figure 5A** CHIRPS). In NB9 (**Figure 4A** CHIRPS), values above 10 mm $d^{-1}$ concentrate at the northwestern domain border, whereas low precipitation amounts of <4 mm d are found in the easternmost region and in the core of Northeast Brazil around 10°S and 40°W.

To also provide a spatial overview of reference uncertainty, fields of precipitation (**Figures 3A**, **4A**) and biases against CHIRPS (**Figures 3B**, **4B**) of MSWEP and ERA5-Land are presented for EP9 and NB9. For EP3, precipitation of INAMHI stations is shown along with CHIRPS (**Figure 5A**). Whereas only small and little structured deviations around ±2 mm $d^{-1}$ against CHIRPS can be observed for MSWEP and ERA5-Land in NB9 (**Figure 4B**), larger bias of MSWEP or ERA5-Land against CHIRPS, i.e., larger reference uncertainty, is found for EP9 (**Figure 3B**). Here, ERA5-Land shows wetter conditions especially for the mountainous Andes region, in parts by more than 10 mm $d^{-1}$. Drier conditions by around 4 mm $d^{-1}$ are found for the Amazon region in West Brazil and Southeast Colombia. MSWEP is rather drier than CHIRPS by around 4 mm $d^{-1}$, only the Amazon region in Northeast Peru shows wetter conditions. INAMHI station data and CHIRPS generally coincide for Southwest Ecuador in EP3, only some stations deviate from CHIRPS where both higher and lower values of mean precipitation are possible.

In all three domains, the WRF ensemble median ($WRF_{ENS}$) shows an overall positive bias against CHIRPS, especially for the mountainous regions with more than 10 mm $d^{-1}$ (**Figures 3–5**

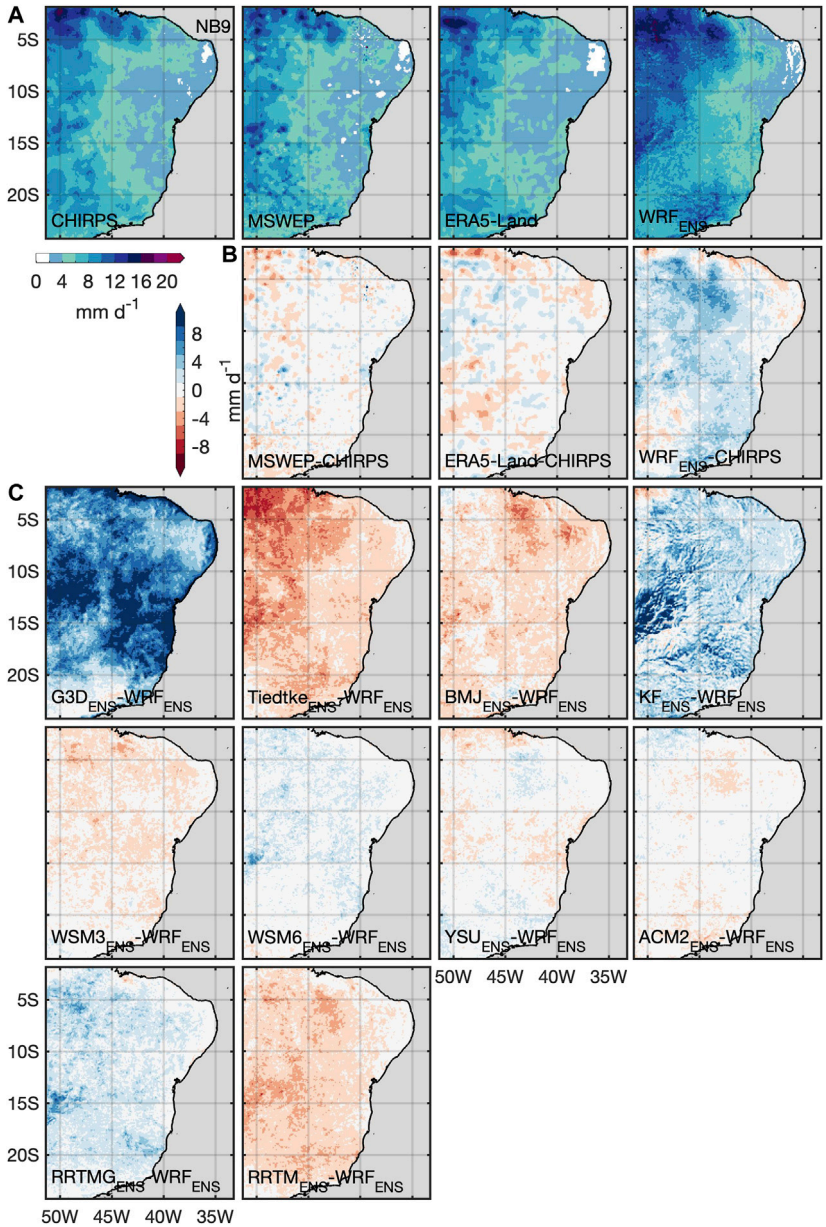


**FIGURE 4 |** As **Figure 3**, but for NB9.

$WRF_{ENS}$). Only some coastal and inland areas are drier than CHIRPS. For EP3 (**Figure 5B** $WRF_{ENS}$), $WRF_{ENS}$ bias against INAMHI station data generally follows the patterns of bias against CHIRPS, with a too dry simulated Ecuadorian coast and too wet simulated Ecuadorian mountain regions. Especially for EP3, higher resolved topography in WRF seems to play a role as more distinct features of higher precipitation amounts are evident in the Andes region at the Ecuador-Peru border and in the southern part of the domain. CHIRPS does not show these distinct features, resulting in a large positive bias of $WRF_{ENS}$. For the Ecuadorian Andes, the lack of station data within the area of largest positive bias of $WRF_{ENS}$ against CHIRPS can also not bring further clarity to this high simulated precipitation by WRF.

To illustrate the divergence within the WRF ensemble for different applied parameterization schemes, the WRF ensemble is grouped into several sub-ensembles that all use one fixed parameterization scheme. For example, the $G3D_{ENS}$ or $RRTMG_{ENS}$ encompass all eight or 16 runs, respectively, that use the G3D cumulus or RRTMG radiation physics scheme. The deviation of the median of these sub-ensembles from the total $WRF_{ENS}$ is shown in **Figures 3–5**. In all three domains, the variation of the deviations from $WRF_{ENS}$ is largest within the CU and RA sub-ensembles. The MP and PBL sub-ensemble medians

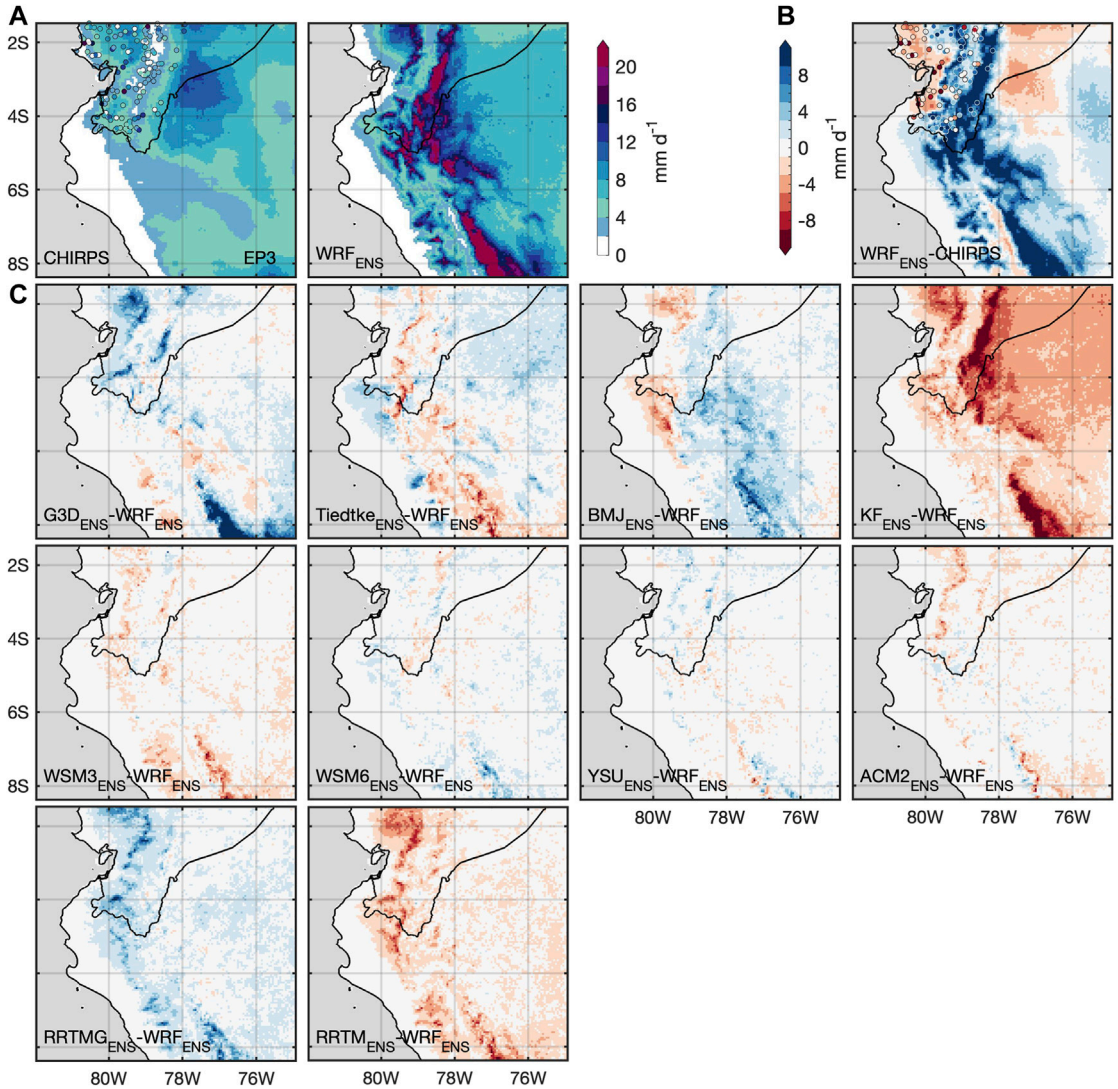


**FIGURE 5 |** As **Figure 3**, but for EP3. Apart from CHIRPS, in **(A)** JFM mean precipitation values of and **(B)** $WRF_{ENS}$ bias against 132 INAMHI stations having less than 5% of missing values during JFM 2007 are shown.

rather represent $WRF_{ENS}$. Only slight tendencies of wetter runs for $WSM6_{ENS}$ and drier runs of $WSM3_{ENS}$ are found. For the cumulus sub-ensembles in EP9 (**Figure 3C**), $G3D_{ENS}$ and $Tiedtke_{ENS}$, as well as $BMJ_{ENS}$ and $KF_{ENS}$ seem to behave in opposite ways: the $G3D_{ENS}$ simulates more precipitation especially for the mountainous (more than +10 mm $d^{-1}$) and Brazilian Amazon (up to +6 mm $d^{-1}$) region whereas $Tiedtke_{ENS}$ gives less precipitation in these regions with similar magnitudes. Opposite behavior for $G3D_{ENS}$ and $Tiedtke_{ENS}$ is also found north of the equator with a drier tendency for $G3D_{ENS}$ and a wetter tendency for $Tiedtke_{ENS}$. A similar north-south division of positive and negative deviations from $WRF_{ENS}$ is evident for $BMJ_{ENS}$ and $KF_{ENS}$ that also behave in opposite ways, but with smaller magnitudes than $G3D_{ENS}$ and $Tiedtke_{ENS}$. For NB9 (**Figure 4C**), the picture is more clear: $G3D_{ENS}$ and $KF_{ENS}$ compose the wetter runs of $WRF_{ENS}$, whereas $Tiedtke_{ENS}$ and $BMJ_{ENS}$ yield drier conditions. Similar to EP9, highest magnitudes of deviations from $WRF_{ENS}$ are found for $G3D_{ENS}$ with large areas above +10 mm $d^{-1}$ and for $Tiedtke_{ENS}$ with regions especially in the Northwest and West up to −8 mm $d^{-1}$. For EP3 (**Figure 5C**), the deviations from $WRF_{ENS}$ among the CU sub-ensembles are more diverse. The pattern of higher (less) precipitation in the Andes region of $G3D_{ENS}$ and $BMJ_{ENS}$ ($Tiedtke_{ENS}$ and $KF_{ENS}$) of EP9 passes through to EP3. However, the negative deviations from $WRF_{ENS}$ in the Andes region seem to be even amplified for EP3 $KF_{ENS}$, whereas the negative deviations of EP3 $Tiedtke_{ENS}$ are dampened compared to EP9. Here, it needs to be recalled that EP3 was simulated with explicit convection, i.e., the CU parameterization was switched off. However, for the parent domain EP9 the respective CU

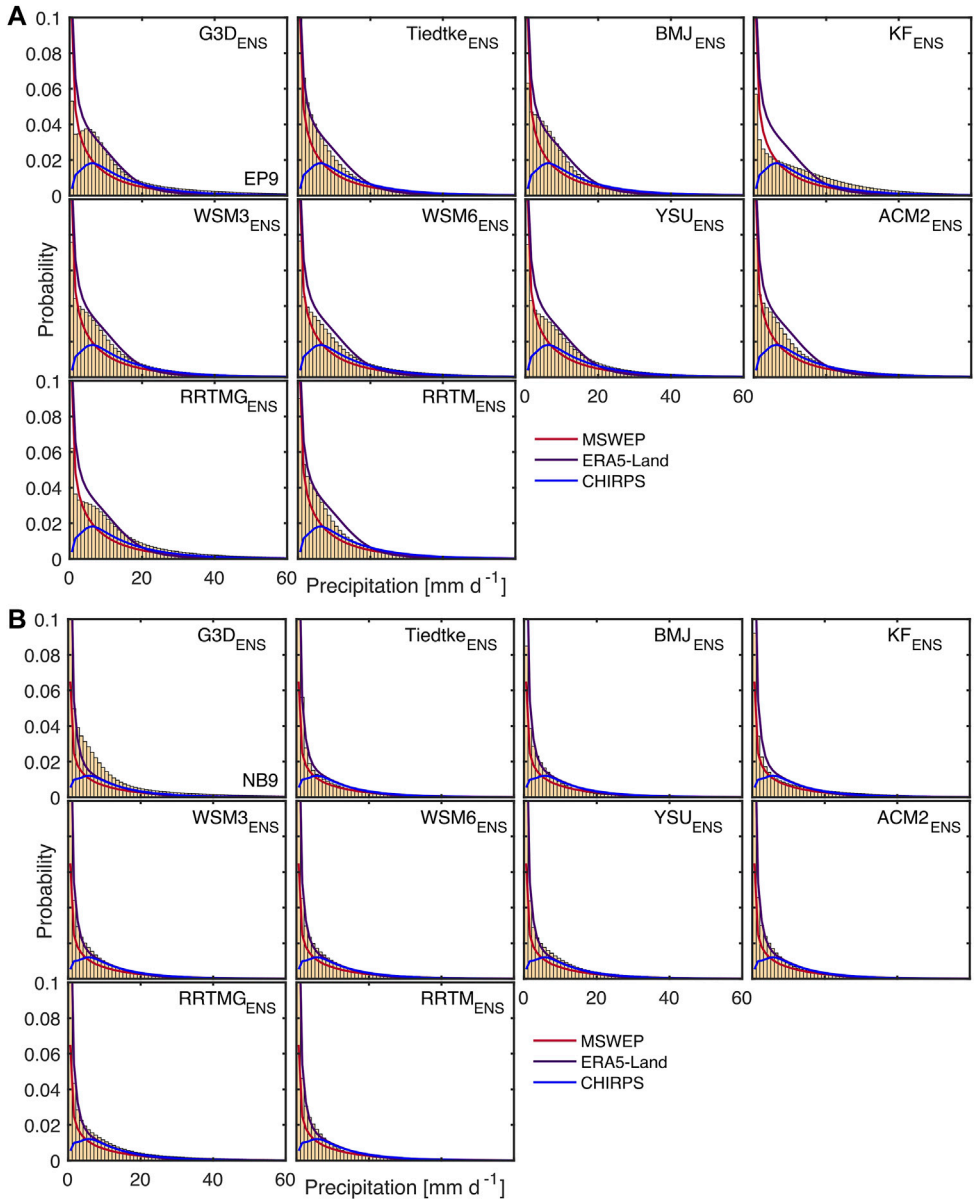


**FIGURE 6 |** Probability distribution of the parameterization sub-ensembles compared to CHIRPS, MSWEP and ERA5-Land for daily gridpoint-based precipitation of **(A)** EP9 and **(B)** NB9 from Jul 2006 till Jun 2008.

parameterizations were active. For the radiation sub-ensembles, in all three domains, $RRTMG_{ENS}$ composes the wetter part of $WRF_{ENS}$, whereas $RRTM_{ENS}$ builds the drier part of $WRF_{ENS}$ with similar magnitudes up to around ±6 mm d$^{-1}$.

## 3.2 Gridpoint-Based Validation

**Figures 6**, **7** show the discrete probability distribution for daily gridpoint-based precipitation of the parameterization sub-ensembles compared to the different reference datasets for EP9, NB9 and EP3. Regardless of the domain, probability distributions of CHIRPS, MSWEP and ERA5-Land differ strongly and seem to be characterized by different shape and scale parameters when compared to theoretical probability distributions like the Gamma distribution. Whereas highest probabilities of MSWEP and ERA5-Land are found for lowest daily precipitation amounts below 1 mm d$^{-1}$, the CHIRPS probability distribution has its maximum around 7 mm d$^{-1}$ and shows a strong decline of probabilities toward lower precipitation amounts. For both EP9 and NB9, MSWEP has an exponential decrease of probabilities and shows lower probabilities than CHIRPS for precipitation values between 8 and 20 mm d$^{-1}$. For EP9 (**Figure 6A**), greatly enhanced probabilities of ERA5-Land precipitation between 5 and 18 mm d$^{-1}$ are found compared to both CHIRPS and

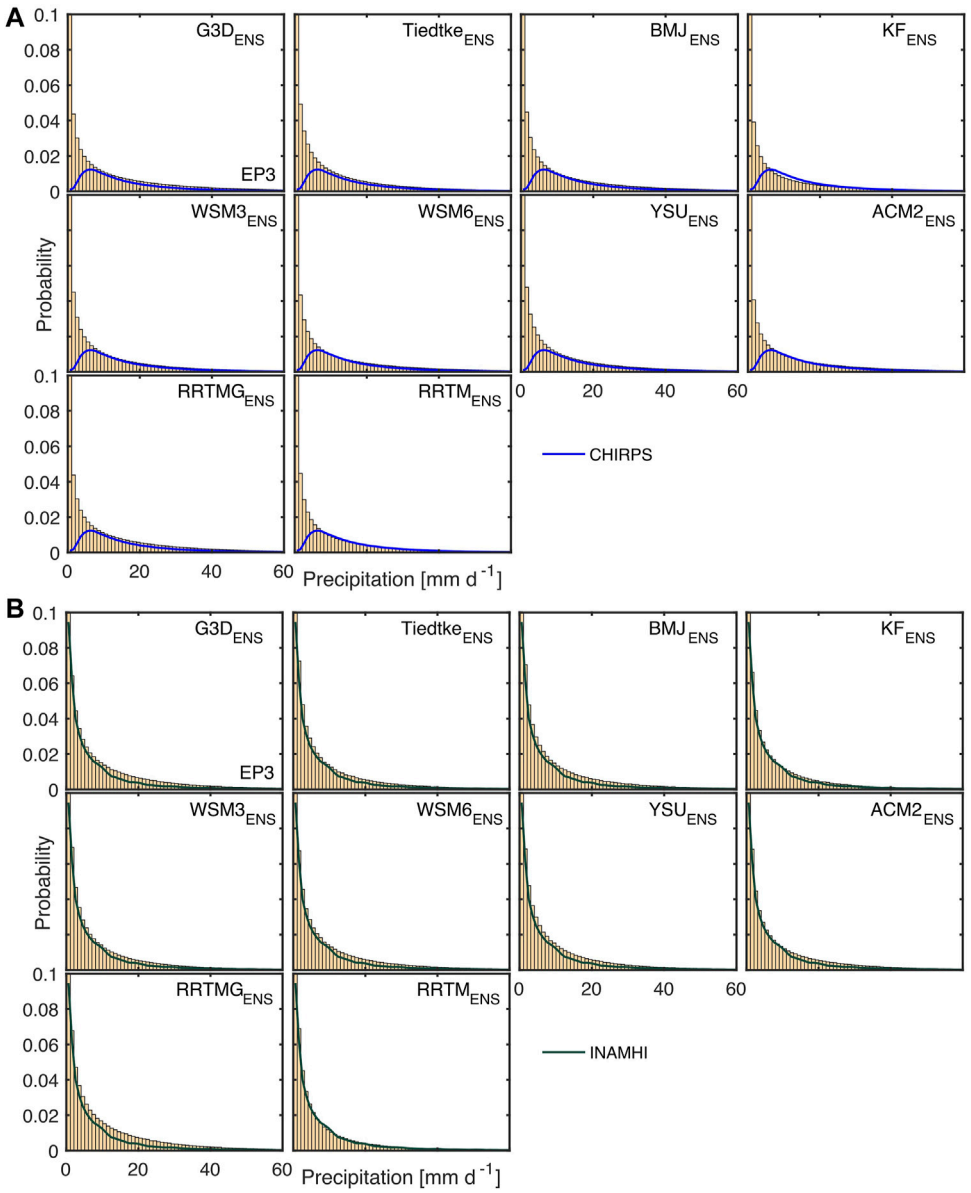


**FIGURE 7 |** As **Figure 6**, but for EP3, with **(A)** CHIRPS reference data and **(B)** INAMHI station data. For **(B)**, WRF data interpolated to the location of the 106 selected stations was used.

MSWEP, with a rather constant than exponential decrease. For NB9 (**Figure 6B**), the probability distribution of ERA5-Land has a more exponential decrease and resembles the one of CHIRPS after CHIRPS's maximum. For EP3, CHIRPS shows the lowest probabilities of precipitation below 2 mm d$^{-1}$. The INAMHI station data in EP3, however, show high probabilities of low precipitation values and low probabilities of precipitation above 20 mm d$^{-1}$.

Regarding the probability distributions of the WRF sub-ensembles, largest differences in their shape and scale are found among the sub-ensembles for EP9 (**Figure 6A**). With respect to CHIRPS, all probability distributions of the sub-ensembles of EP9 show too high probabilities especially for precipitation amounts <19 mm d$^{-1}$. The only exception is KF$_{ENS}$, that matches the probabilities for CHIRPS's maximum, but overestimates those for both lower and higher precipitation amounts. Apart from the in general overestimated probabilities, the shape of CHIRPS's probability distribution is best represented by the one of G3D$_{ENS}$ for values >2 mm d$^{-1}$. No probability distribution of any sub-ensemble resembles the one of MSWEP. With respect to ERA5-Land, Tiedtke$_{ENS}$ and RRTM$_{ENS}$ best reproduce the shape, only with slightly underestimated probabilities for precipitation amounts between 8 and 19 mm d$^{-1}$. Like those of MSWEP and ERA5-Land, the probability distributions of the WRF sub-ensembles show the general problem of overestimated probabilities for small

precipitation values <1 mm $d^{-1}$, i.e., model drizzle. Lowest probabilities of drizzle rain is found for the $G3D_{ENS}$ among the cumulus groups and for $RRTMG_{ENS}$ among the radiation physics groups. For drizzle rain, about equal performance is found for the microphysics and planetary boundary layer physics groups. In sum for EP9, the parameterization sub-ensembles can be grouped in mainly two classes (classification also in comparison with the other used parameterization scheme for the same physics topic): 1) monotonous to exponential decrease of probabilities with highest probabilities for drizzle rain ($Tiedtke_{ENS}$, $KF_{ENS}$, $WSM6_{ENS}$, $ACM2_{ENS}$, $RRTM_{ENS}$), and 2) drizzle rain with highest probabilities, but enhanced probabilities also for medium precipitation amounts around 5 mm $d^{-1}$ and exponential decrease thereafter ($G3D_{ENS}$, $BMJ_{ENS}$, $WSM3_{ENS}$, $YSU_{ENS}$, $RRTMG_{ENS}$).

For NB9 (**Figure 6B**) and EP3 (**Figure 7**), such a clear distinction between the shape of the probability distribution of the different sub-ensembles is not valid. Rather, their shape is characterized by an exponential decrease with highest probabilities for drizzle rain. For NB9 (**Figure 6B**), only $G3D_{ENS}$ does not show an explicit exponential decrease for values between 4 and 15 mm $d^{-1}$. $YSU_{ENS}$ further has slightly enhanced probabilities around 8 mm $d^{-1}$. For EP3, the WRF gridpoints interpolated to INAMHI station locations in general show higher probabilities of precipitation than for the entire EP3 domain. $KF_{ENS}$ and $RRTM_{ENS}$ best represent the distribution of the INAMHI station data.

## 3.3 Domain-Based Validation

In the Taylor diagrams (**Figures 8**, **9A**), domain averaged timeseries of daily precipitation of all individual WRF runs are evaluated with respect to their Pearson correlation coefficient (CORR), standard-deviation (STD), and root-mean square difference (RMSD) as compared to CHIRPS. For EP9 and NB9 (**Figure 8**), the values of CORR, RMSD and STD of MSWEP and ERA5-Land compared to CHIRPS demonstrate a lower uncertainty within the reference than within the WRF ensemble. Both MSWEP and ERA5-Land have high CORR (both around 0.93 for NB9; around 0.78 and 0.87 for ERA5-Land and MSWEP for EP9), and low RMSD (both around 2.2 mm $d^{-1}$ for NB9; around 2 and 1.6 mm $d^{-1}$ for ERA5-Land and MSWEP for EP9) as compared to CHIRPS. The respective STD of CHIRPS is around 3.2 and 3.3 mm $d^{-1}$ for EP9 and NB9. The STDs of MSWEP and ERA5-Land are around 2.4 and 2.8 mm $d^{-1}$ for EP9, and around 2.9 and 3.4 mm $d^{-1}$ for NB9, respectively. The comparison of $WRF_{ENS}$ against CHIRPS for EP9 and NB9 (**Figure 8**) reveals in general higher CORR with smaller spread for NB9 (0.74–0.85) than for EP9 (0.28–0.64), but also higher spread of RMSD and STD for NB9 (RMSD: 1.7–5.6 mm $d^{-1}$, and STD: 2–7.6 mm $d^{-1}$) than for EP9 (RMSD: 2.6–5.2 mm $d^{-1}$, and STD: 2.4–5.1 mm $d^{-1}$).

For EP3, in addition to CHIRPS, further comparison is made against timeseries of INAMHI station data (**Figure 9B**). Here, higher CORR is found against INAMHI station data (0.48–0.75) than against CHIRPS (0.25–0.49). RMSD ranges between 4.6 and 7.3 mm $d^{-1}$ against CHIRPS and between 2.7 and 6.5 mm $d^{-1}$ against INAMHI station data. CHIRPS has a STD of around 4.6 mm $d^{-1}$, whereas INAMHI station data has its STD around 3.7 mm $d^{-1}$. Due to interpolation to station locations and considered gaps in the timeseries of INAMHI stations, the STD of $WRF_{ENS}$ differs in the comparison against CHIRPS and INAMHI station data. It ranges between 3.2 and 7.2 mm $d^{-1}$ for the entire domain and entire analysis period (CHIRPS), and between 2.8 and 7.6 mm $d^{-1}$ for INAMHI station locations and periods.

Regarding different parameterization sub-ensembles, largest differences are found for the CU schemes for all three domains. The two sub-ensembles of both the MP and PBL schemes cover the entire possible range of values without clear tendencies. Among the CU sub-ensembles, $G3D_{ENS}$ has the largest variance with respect to CORR, RMSD and STD with a tendency toward the ensemble's lowest CORR, highest RMSD and highest STD. $Tiedtke_{ENS}$ encompasses runs with lowest RMSD and STD and medium to high CORR. $BMJ_{ENS}$ is characterized by medium to high CORR, medium RMSD and STD, and $KF_{ENS}$ shows medium CORR, high RMSD and STD. Only for EP3 (**Figure 9**), the rating especially of $KF_{ENS}$ within the entire WRF ensemble differs, with low to medium RMSD and STD when compared both against CHIRPS and INAMHI station reference data. Among the RA sub-ensembles, $RRTMG_{ENS}$ gives lower CORR, higher RMSD and STD than $RRTM_{ENS}$ for all three domains.

## 3.4 Feature-based Validation

The overall domain precipitation as well as the feature-based location and volume of individual precipitation patterns in the WRF sub-ensembles are validated in the ensemble Structure-Amplitude-Location analysis (eSAL) of **Figures 10–12**. In general, for all three domains, reference domain average precipitation <5 mm $d^{-1}$ of CHIRPS is mostly overestimated by $WRF_{ENS}$ ($eA > 0$) and associated with too large or too flat modeled precipitation features ($eS > 0$) and large location error $eL$. The overestimation of small reference domain average precipitation is more pronounced for EP9 and EP3 than for NB9. Larger reference domain average precipitation is correctly estimated ($eA \approx 0$), or slightly over- or underestimated, and shows smaller eL and too small or too peaked modeled precipitation features ($eS < 0$) in comparison to CHIRPS.

As for previous analysis measures in **Sections 3.2** and **3.3**, largest impact on the performance of WRF runs in eSAL, especially for EP9 and NB9, is given by the choice of different CU schemes, followed by different RA schemes (**Figures 10**, **11A**,**B**,**D**). eSAL characteristics for different evaluated PBL (**Figures 10**, **11C**), or MP (see **Supplementary Material**) sub-ensembles are almost identical. For EP9 (**Figure 10**), best eSAL characteristics among the CU schemes are found for $Tiedtke_{ENS}$. $Tiedtke_{ENS}$ shows highest density for correlated $eA \approx 0$ and $eS \approx 0$, together with a high density of small $eL$. The overestimation of reference domain average precipitation <5 mm $d^{-1}$ is smallest. For NB9 (**Figure 11**), even both over- and underestimation for $Tiedtke_{ENS}$ is possible, but highest

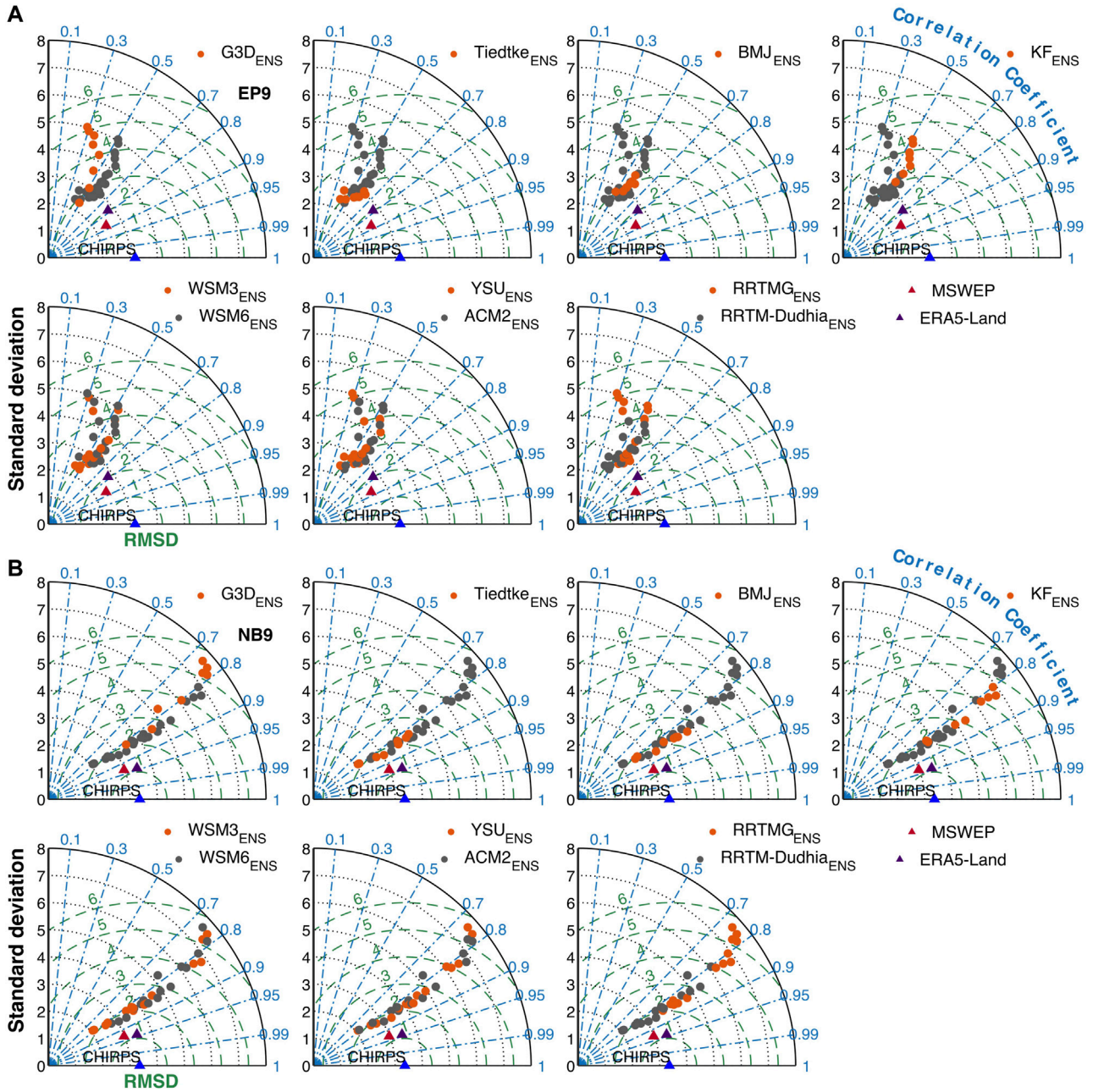


**FIGURE 8 |** Taylor diagram for **(A)** EP9 and **(B)** NB9 including the Pearson correlation coefficient (dashed-dotted lines), the standard deviation (dotted lines, in mm $d^{-1}$) and the root-mean-square difference (RMSD) of the anomalies (dashed lines, in mm $d^{-1}$) between the simulated daily precipitation timeseries of the 32 WRF physics ensemble members and CHIRPS reference data for the period Jan 2006–Jun 2008. Each panel shows the differences in the physics parameterization sub-ensembles, i.e., cumulus convection (G3D, Tiedtke, BMJ, KF), microphysics (WSM3, WSM6), planetary boundary layer (YSU, ACM2), and radiation physics (RRTMG, RRTM-Dudhia).

density is found for $eA \approx 0$. $BMJ_{ENS}$ and $KF_{ENS}$ for both EP9 and NB9 have the disadvantage of high density of either $eA > 0$ or $eS < 0$. For NB9, however, both $BMJ_{ENS}$ and $KF_{ENS}$ show even better $eL$ than $Tiedtke_{ENS}$. For both domains, $G3D_{ENS}$ generally overestimates domain average precipitation. Its structure and location error is similar to $KF_{ENS}$ for EP9, whereas for NB9 both largest location errors and positive structure errors of $G3D_{ENS}$ especially for low reference precipitation are evident. Among the RA schemes, $RRTM_{ENS}$ seems to be preferable with a generally smaller amplitude error, as well as smaller structure and location errors in particular for EP9. Whereas no distinction can be made for EP9 regarding the PBL sub-ensembles, $ACM2_{ENS}$ tends toward smaller amplitude, structure and location error than $YSU_{ENS}$ for NB9.

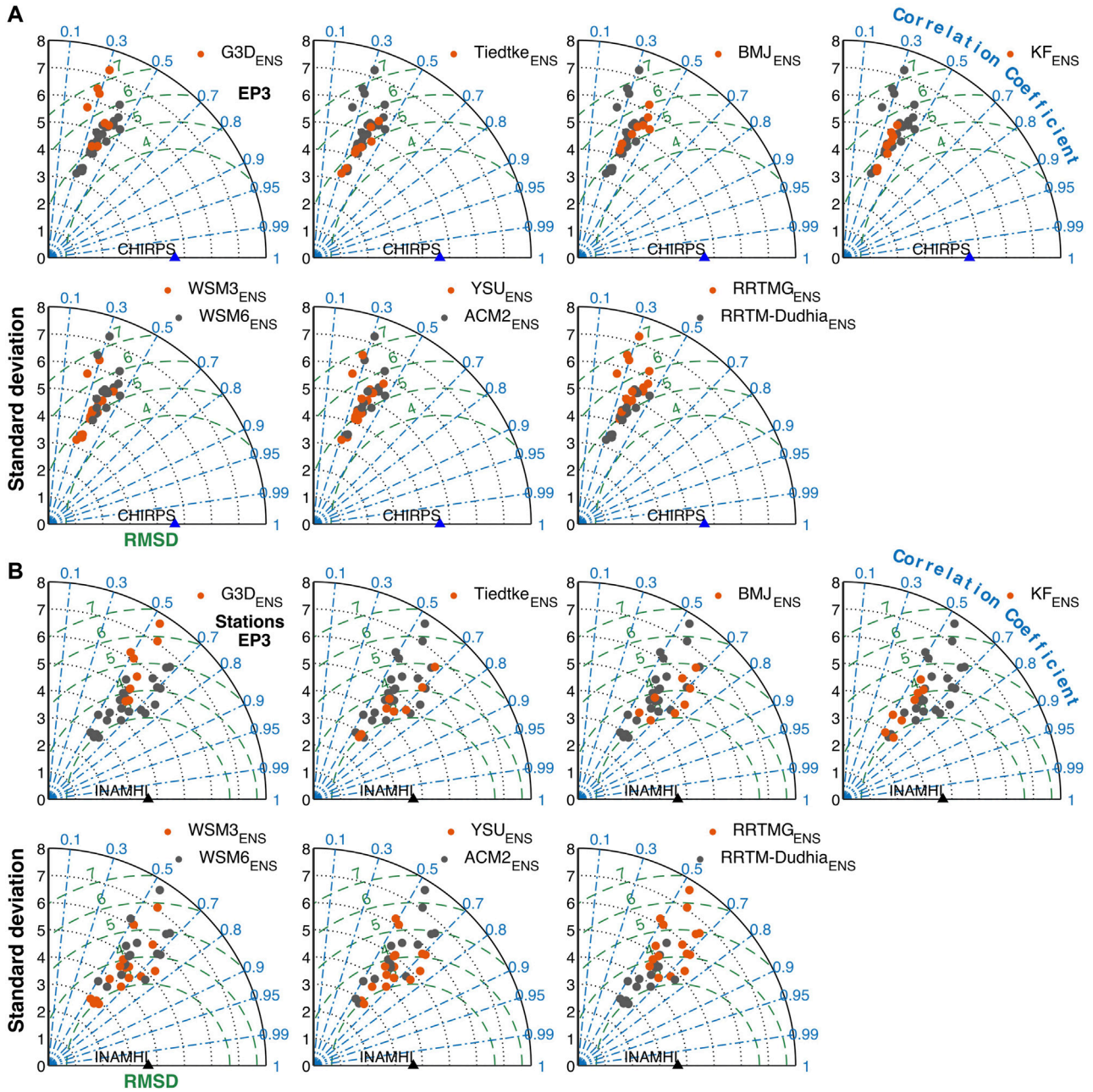


**FIGURE 9 |** As **Figure 8**, but for EP3, with **(A)** CHIRPS reference data and **(B)** INAMHI station data. A direct comparison of CHIRPS and INAMHI station data on a daily scale in one plot as in **Figure 8** for MSWEP and ERA5-Land is not possible due to different accumulation times of 00 UTC–00 UTC (CHIRPS) and 12 UTC-12 UTC (07–07 am local time, INAMHI stations), respectively. For the comparison with INAMHI stations, the mean of the 106 selected stations or respective modeled nearest grid-cell data was used.

For EP3 (**Figure 12**), the performance of $WRF_{ENS}$ against CHIRPS with respect to *eA* and *eS* is poorer and shows a higher variance than for EP9 or NB9. High positive *eA* is common, and both positive and negative structure errors occur frequently. Location errors are similar to NB9, with lower values than for EP9. A clear distinction between different sub-ensembles turns out to be harder than for EP9 or NB9. Only the $KF_{ENS}$ has clearly different behavior with a tendency of underestimation of domain average precipitation ($eA < 0$) and too small or too peaked modeled precipitation features ($eS < 0$). As in EP9, $RRTM_{ENS}$ tends toward smaller amplitude error than $RRTMG_{ENS}$, and rather too small or too peaked modeled precipitation ($eS < 0$) compared to $RRTMG_{ENS}$ with rather too large or too flat modeled precipitation ($eS < 0$).

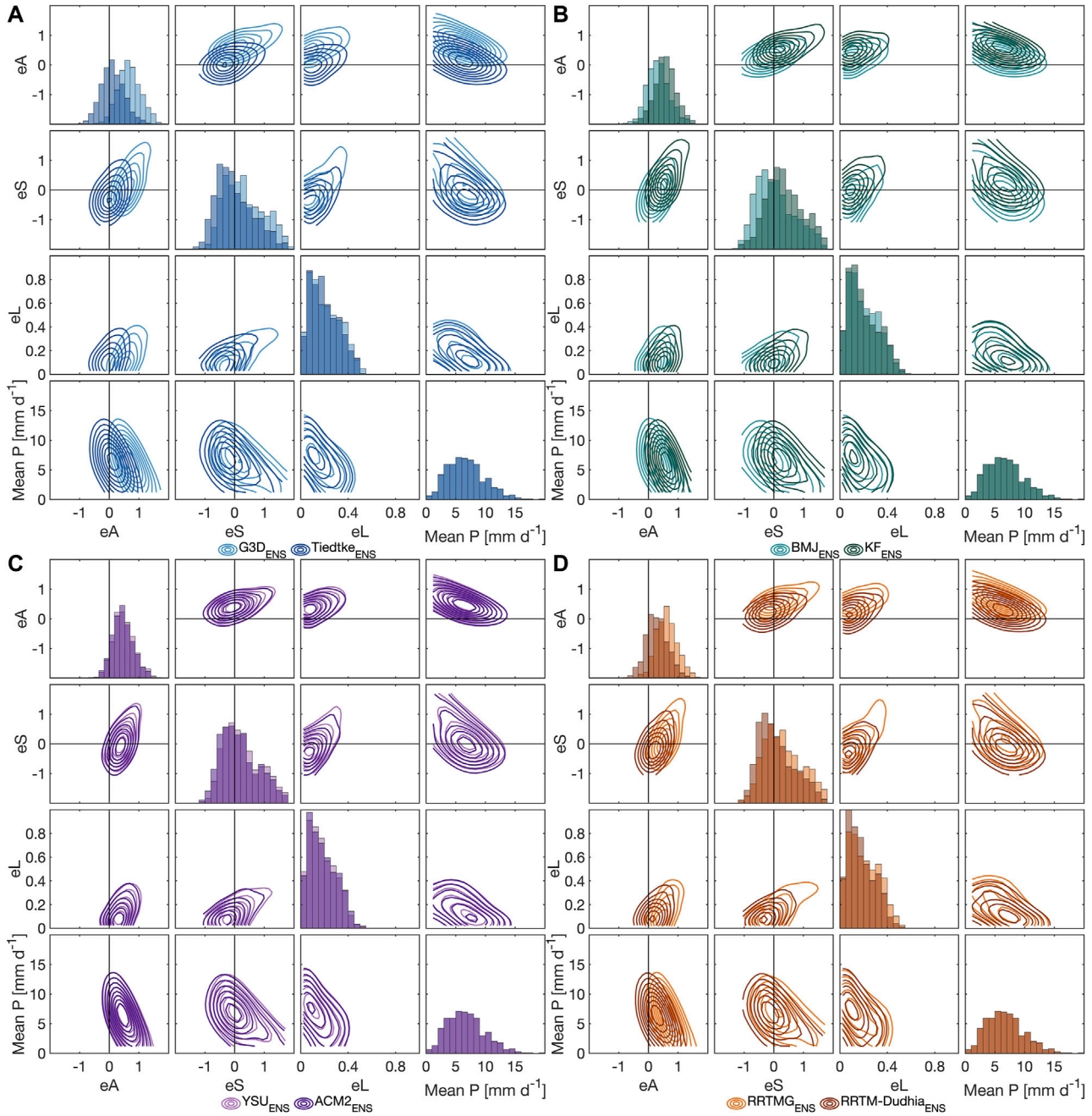


**FIGURE 10 |** Ensemble Structure-Amplitude-Location (eSAL) analysis for EP9 with respect to CHIRPS for the physical parameterization sub-ensembles of **(A)** G3D and Tiedtke **(B)** BMJ and KF **(C)** YSU and ACM2, and **(D)** RRTMG and RRTM-Dudhia. For each physics sub-ensemble, *eA*, *eS*, *eL* and the reference domain average precipitation (Mean P) are juxtaposed in density contour plots of daily values from Jul 2006 till Jun 2008. The diagonal shows the probability histograms of *eA*, *eS*, *eL* of the respective physics ensemble and of the reference domain average precipitation (Mean P). When the transparently colored probability histograms of *eA*, *eS*, or *eL* overlap, darker colors are produced. As all different parameterization groups use CHIRPS as reference, histograms of reference mean P always overlap.

## 4 DISCUSSION AND CONCLUSION

In all evaluated categories for the physics sub-ensembles in terms of deviation from the total WRF physics ensemble (gridpoint-based probability distributions, domain-averaged characteristics of standard deviation, correlation, root-mean square difference and bias, and spatial feature-based structure and location), switching between different cumulus physics (CU) schemes resulted in the largest performance variability, followed by the radiation physics (RA) schemes. The large spread among the CU sub-ensembles confirms the strong impact of different CU schemes on the simulated precipitation in the study regions.

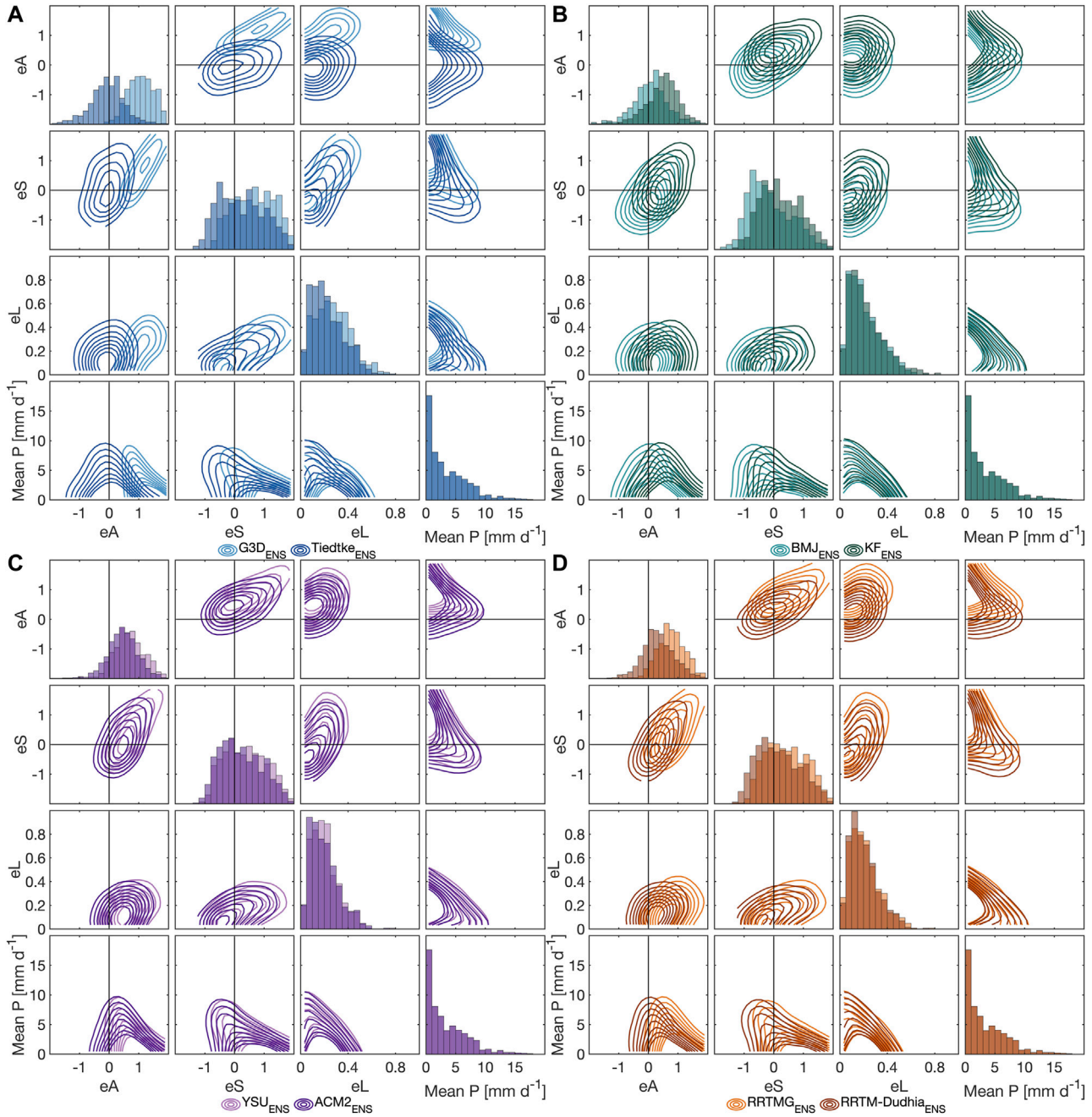


**FIGURE 11 |** As **Figure 10**, but for NB9.

Even for the explicitly resolved convection in EP3, the effects of CU schemes used in the parent domains is still larger than those of different microphysics (MP) or planetary boundary layer physics (PBL) schemes, but in the same order of magnitude as between different RA schemes. Regarding the deviations from the ensemble median for the CU sub-ensembles, many areas reached more than ±6 mm $d^{-1}$, whereas different MP and PBL sub-ensembles deviated only by around ±2 mm $d^{-1}$ (**Figures 3–5**). In Klein et al. (2015), such large variations between different CU sub-ensembles were not observed for West Africa. Instead, these deviations were mainly in the order of ±3 mm $d^{-1}$ for almost all tested CU, and in the same range as those of the MP and PBL sub-ensembles. Klein et al. (2015) further showed an overall dry bias for WSM3 compared to more sophisticated schemes, and for ACM2 compared to YSU in the Sahel and Sudano-Sahel regions. Similar, but not as pronounced tendencies within the MP and PBL schemes as in Klein et al. (2015) could be observed for our South American regions. An overestimation of precipitation by

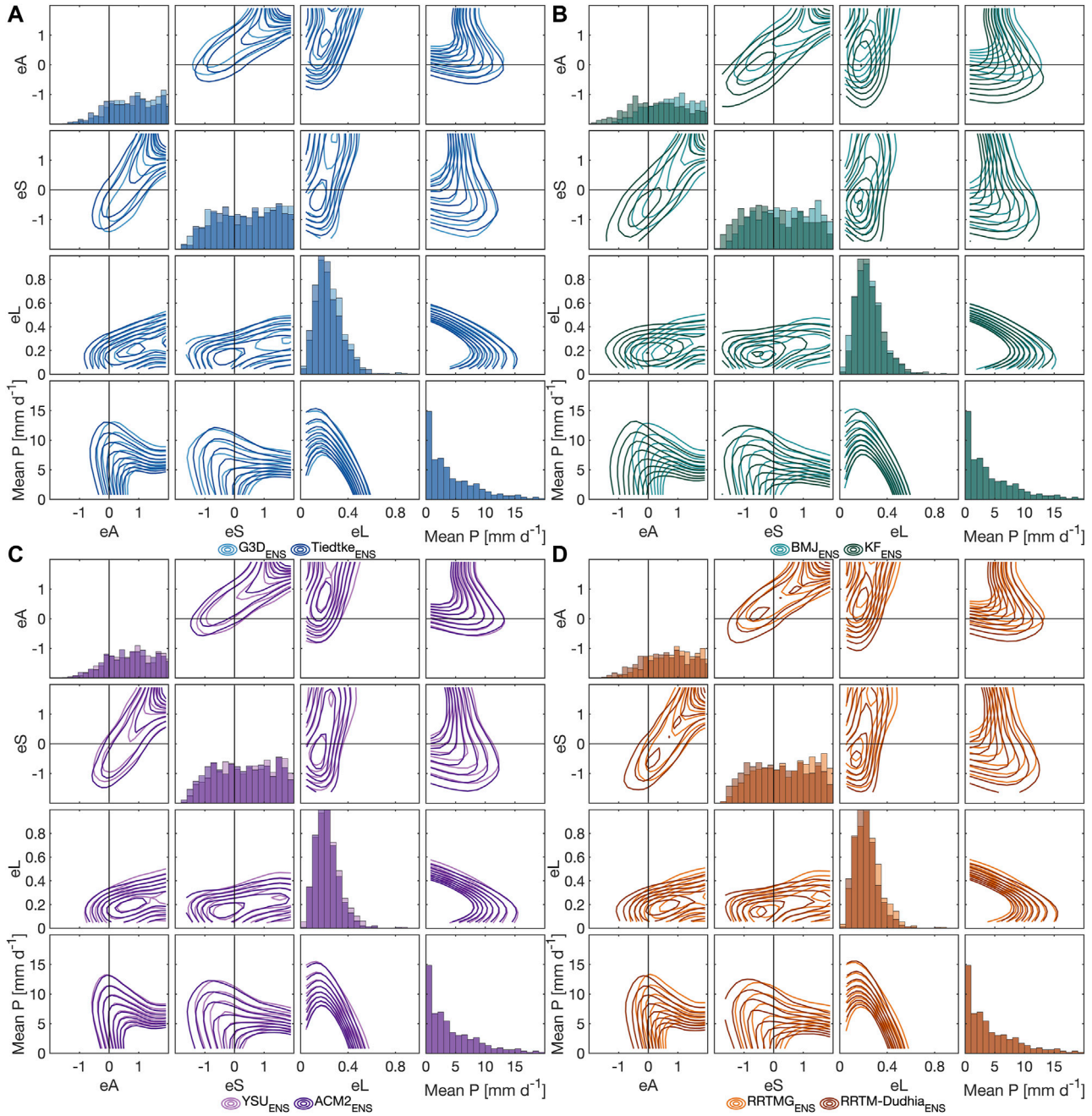


**FIGURE 12 |** As **Figure 10**, but for EP3.

KF and an underestimation by BMJ for Northeastern Africa found by Pohl et al. (2011) and Tariku and Gan (2017), is of course relative to their respective reference, however, opposing behavior of BMJ and KF sub-ensembles with respect to the deviation from the WRF ensemble median was also evident for our focus regions. An observed wetter climate for RRTMG than for RRTM reported by Tariku and Gan (2017) can also be confirmed for all our domains. Especially for EP9, the generally small differences between WSM3 and WSM6 sub-ensembles could be related to the limited necessity of using a graupel scheme (e.g., WSM6) for horizontal resolutions above 10 km (Skamarock et al., 2008). Obviously, the 9-km-domain is still in a "gray zone" of only partly resolved updrafts producing graupel and the simpler WSM3 scheme gives similar results. However, also the convection-resolving 3-km-grid of EP3 does not produce large differences between WSM3 and WSM6 sub-ensembles for all evaluated metrics, which should be expected from Hong and Lim (2006). Thus for our study, independent of the grid

resolution, the tested MP schemes introduced less uncertainties than the CU or RA schemes but similar uncertainties as the PBL schemes with respect to analyzed precipitation distributional, temporal and spatial patterns. The generally weak dependency of the modeled precipitation on the MP treatment in our study could also have its origin in the choice of tested MP schemes. Other studies (e.g., Liu et al., 2011; Klein et al., 2015) that found larger differences between MP schemes not only used single moment schemes like WSM3 and WSM6, but also used more complex double moment schemes predicting number concentration of the water and ice species that are not part of the WRF-MP series (WRF single or double moment schemes). Here, further studies are needed to investigate if the weak dependency on MP schemes still holds for more complex MP schemes.

Regarding our choice of CU schemes, it should be noticed that the 9-km grid size is at the upper end of the convective "gray zone" and we only used conventional parameterizations of convection such as BMJ, Tiedtke, and KF. These are constructed so that the parameterization acts self-contained within one grid column, assuming that the proportion of the grid column covered by active convection is small. With small grid sizes, however, this assumption is no longer valid (Grell and Freitas, 2014). G3D allows subsidence spreading to neighboring grid cells, however, this is recommended only for even smaller grid sizes (Skamarock et al., 2008; Grell and Freitas, 2014). For the 9-km-domain, it is thus assumed that the subsidence still takes place in the same grid column and G3D therefore acts conventionally. Scale-aware cumulus schemes like Grell–Freitas (Grell and Freitas, 2014) or the updated multi-scale KF (Zheng et al., 2016) attempt to smooth the transition to cloud-resolving scales by increasing deactivation of the parameterization with increasing grid resolution. However, the parameterization is still quite active at 9 km and Jeworrek et al. (2019) found a comparable performance of the scale-aware msKF and GF schemes to the conventional cumulus schemes such as BMJ, Tiedtke and KF at 9 km. For the scale-aware msKF and GF, the contribution of the resolved-scale precipitation from the microphysics scheme was comparable to BMJ at 9 km; Tiedtke even produced substantially more resolved-scale precipitation at 9 km (Jeworrek et al., 2019). This example shows that the application of conventional cumulus schemes is still justifiable at 9 km grid resolution. However, especially with respect to the conventional KF, the msKF may be preferred: In their comparison of conventional KF and msKF at 9 km, Zheng et al. (2016) found superior performance of the msKF in terms of precipitation bias, location and intensity. Future studies should therefore consider to additionally test the scale-aware GF and msKF schemes at 9 km resolution. The inclusion of the simple conventional KF in our study should also be a test in terms of required computational costs, i.e., if the simple and computationally cheap KF scheme is also able to produce reasonable results. Especially as parent domain for the 3 km EP3, $KF_{ENS}$ proved to be closest to the reference data.

The comparison of the WRF performance in the 3-km-domain EP3 with explicit convection against the performance in the 9-km-domain EP9 showed that the bias of the WRF ensemble against CHIRPS in this region was not reduced. Also Sikder and Hossain (2016) could not find substantial improvements at higher resolution (3 km) for their mountainous study regions around the Himalaya by disabling cumulus parameterization. Noteworthy, many performance measures are substantially worse for EP3 than for EP9 with respect to CHIRPS: Correlation coefficients below 0.5, RMSD above 4.5 mm $d^{-1}$ and frequent large amplitude and structure errors for EP3. However, considering the main patterns of precipitation in $WRF_{ENS}$ and in CHIRPS for the JFM season of 2007, CHIRPS seems to contain only little topographic precipitation effects from the Andes. An underestimation of precipitation in some mountainous regions by CHIRPS and other thermal-infrared based satellite products is well known (Bai et al., 2018; Dinku et al., 2018). The better performance of $WRF_{ENS}$ in terms of correlation and root-mean square difference with respect to INAMHI station data as compared to CHIRPS may thus suggest possible limitations for the use of CHIRPS as reference data for high-resolution simulations in mountainous domains. But still, the areas with largest positive bias above 9 mm $d^{-1}$ of WRF vs. CHIRPS are also not covered by INAMHI and the general trend of overestimated precipitation around those regions by $WRF_{ENS}$ is similarly represented if compared against INAMHI. For the WRF physics sub-ensembles, only $KF_{ENS}$ (KF only used in parent domains) is able to diminish these peaks of precipitation overestimation, however leading to a drier Amazon basin in the Northeast of the EP3 domain.

Another aspect revealed for the 3-km-domain EP3 is that the lack of cumulus parameterizations obviously produces very similarly shaped probability distribution functions of precipitation for all parameterization sub-ensembles, including the high probabilities for precipitation below 1 mm $d^{-1}$, i.e., drizzle, that were partly compensated by G3D, BMJ and KF schemes in the larger 9-km-domain. Interestingly, the INAMHI stations confirm this shape of the WRF histograms with an exponential decrease and with high probabilities for drizzle. CHIRPS contrasts this for EP3 by showing remarkably low probabilities of gridpoint-based daily precipitation values not only up to 1 mm $d^{-1}$ (as in EP9), but also up to 3 mm $d^{-1}$, possibly further indicating a limited use of CHIRPS in this region or at the 0.05° resolution. With the harmonization of the probability distributions among the sub-ensembles for EP3, also the strong overestimation of probabilities for heavy precipitation above 20 mm $d^{-1}$ within the $KF_{ENS}$ in EP9 vanished. For EP3, $KF_{ENS}$ even represented best the probability distribution of INAMHI station data.

For the two 9-km-domains EP9 and NB9, the strong similarity of probability distributions of $Tiedtke_{ENS}$ and ERA5-Land is striking. ERA5-Land gets its atmospheric forcing from ERA5 whose cumulus parameterization is based on Tiedtke (1989) (Hersbach et al., 2020). Precipitation here is only spatially disaggregated from 31 to 9 km and a "lapse rate correction" for altitude is performed in ERA5-Land (ECMWF, 2019). Also bias/deviation patterns for $Tiedtke_{ENS}$ and ERA5-Land highly resemble each other, e.g., with a drier Amazon basin and drier Peruvian Andes in EP9 (**Figure 3**), that seem to be associated with the same underlying cumulus physics.

Finally, we discuss several examples of how the choice of a suitable WRF setup could be made based on the provided performance metrics. Firstly, if one is interested in how precipitation probability distributions might change with future regional climate scenarios, the observed current probability distribution needs to be first and foremost met. The shape of CHIRPS's statistical distribution of precipitation values for EP9 would be best represented by $G3D_{ENS}$ among the CU schemes, and by $RTTMG_{ENS}$ among the RA schemes. However, these WRF sub-ensembles generally overestimate the wet-day probability as compared to CHIRPS, thus requiring a frequency correction (see e.g., Lorenz et al., 2020). For other physics sub-ensembles, the more their distributions and the one of the reference differ from each other, the more severe corrections would be required and the deeper one intervenes in precipitation processes. This could probably produce inconsistencies in the precipitation fields especially when not using correction methods including neighboring gridpoints. Such large differences in the distributions would be the case, e.g., when choosing physics schemes like Tiedtke for CU or RRTM for RA when using CHIRPS as a reference. A choice of G3D and RRTMG for best-fit probability distributions for EP9, however, is connected with the highest overestimation of precipitation, largest RMSD and lowest correlation of domain-averaged precipitation of the WRF ensemble. On the other hand, $Tiedtke_{ENS}$ or $BMJ_{ENS}$ and $RRTM_{ENS}$ have low structural errors of precipitation features as compared to CHIRPS, as well as high correlation, low RMSD and relatively low amplitude errors for domain averages of EP9. Moreover, the spread of evaluation measures for individual ensemble members within $Tiedtke_{ENS}$ or $BMJ_{ENS}$ is smaller than for other sub-ensembles for EP9, indicating a high robustness of the performance of runs when combining Tiedtke or BMJ with different MP, PBL and RA physics schemes in this domain. So secondly, the schemes of Tiedtke and BMJ would be best suited if the WRF simulations are used for distributed impact models that require spatially and temporally differentiated precipitation input. In particular, these spatial criteria, i.e., where we can expect how much precipitation, that means the location and volume of precipitation patterns, are of interest and have major impact on the model results of, e.g., simulated local streamflow or ecosystem functionality. For NB9, similar choices would be made 1) with G3D for precipitation distributional studies, and 2) with Tiedtke or BMJ when precipitation amplitude and structure, as well as RMSD and correlation of domain-averages should be best comparable to CHIRPS. In the case of NB9, the disadvantages of G3D with respect to precipitation overestimation and too large or too flat structured precipitation patterns is even more pronounced than in EP9, suggesting an even more limited suitability. We further faced the difficulty to have suitable reference data to select a model configuration for the 3-km-resolution EP3 domain, as CHIRPS seemed to fail to provide reasonable precipitation for the mountainous regions. All these examples show the challenges when attempting to identify an overall best suited WRF setup. Similarly to previous studies (e.g., Klein et al., 2015; Yang et al., 2021), also in our two South American example regions, we can not clearly conclude on a single "best" setup for all cases and all evaluation metrics. Synthesizing all applied performance measures in one metric (like in, e.g., Yang et al., 2021; Gbode et al., 2019) looses the contained information on individual performance criteria that are critical for one application but not for another. Whatever choice is finally met, not for all areas and not for all performance metrics this choice is optimal. It continues to depend on the specific requirements and on the application of the WRF output, as well as on the observational reference. Nonetheless, with our applied ensemble-tailored analysis methods for distributional, temporal and spatial patterns, it is possible to narrow down the number of suited parameterizations significantly, allowing this choice according to application needs and reference data.

The use of explicitly resolved convection for the highest-resolved 3-km-domain (while still using CU schemes in the parent domains) revealed that even in this domain without CU schemes still large differences were present between runs whose parent domains used different CU schemes. This effect was even larger than the effect of differently applied MP or PBL physics. This has far-reaching implications: tuning a WRF setup with one input data for initial and boundary conditions, e.g. with the latest ERA5 reanalysis that uses a modified Tiedtke CU scheme (Hersbach et al., 2020), does not allow a direct transfer of the setup to other input data sets that may have different underlying physics parameterizations. Sensitivity tests therefore always should be performed with the final input data set.

Concluding, this study clearly demonstrated the necessity of having a very detailed look at the effects of different physics parameterization schemes on simulated precipitation. In our case, up to four times as high monthly precipitation can be generated compared to three different reference data. This shows that a tremendous gain in model performance can be achieved when the computationally expensive effort of conducting a 32-member parameterization ensemble over two consecutive years is undertaken. Especially for distributed impact modeling, it is further essential that such sensitivity studies are not only based on region-averages, but also comprise spatial pattern analysis tools as used in this study. But only the detailed look at all distributional, temporal and spatial patterns allows a comprehensive overview on how different parameterization schemes effect the simulated precipitation. The presented ensemble-tailored strategy allowed to generalize the impacts of and uncertainties introduced by the chosen parameterization schemes. Though failing in identifying an overall best suited setup for each study region in South America, still a final choice of WRF setup according to the application and reference data is possible that can satisfy the need for the required reliable regional information on precipitation to be employed, e.g., for regional climate adaptation and risk mitigation.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. Reference products of CHIRPS, MSWEP and ERA5-Land used in this study are available at the Climate Hazards Center https://

data.chc.ucsb.edu/products/CHIRPS-2.0/, at http://www.gloh2o.org/, and at the Copernicus Climate Data Store https://doi.org/10.24381/cds.e2161bac, respectively.

## AUTHOR CONTRIBUTIONS

TP, PL, CL, and HK developed the research concept. TP performed the model simulations and conceived the analysis. CL managed and provided the used global reference data. NH and MI managed and provided the used station reference data. TP and AJ conducted the analysis. TP summarized the results and wrote the first version of the manuscript. All authors discussed and reviewed the manuscript.

## FUNDING

This work was supported by the funding from the German Federal Ministry of Education and Research (BMBF) for the SaWaM project (*Seasonal Water Resources Management: Regionalized Global Data and Transfer to Practice*) under the financial assistance agreement No 02WGR1421.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the Gauss Center for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SUPERMUC-NG at Leibniz Supercomputing Center (www.lrz.de) under the project ID pr27ko.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feart.2021.669427/full#supplementary-material

## REFERENCES

Bai, L., Shi, C., Li, L., Yang, Y., and Wu, J. (2018). Accuracy of CHIRPS Satellite-Rainfall Products over Mainland China. *Remote Sensing* 10, 362. doi:10.3390/rs10030362

Barbosa, H. A., and Lakshmi Kumar, T. V. (2016). Influence of Rainfall Variability on the Vegetation Dynamics over Northeastern Brazil. *J. Arid Environ.* 124, 377–387. doi:10.1016/j.jaridenv.2015.08.015

Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., et al. (2017). MSWEP: 3-hourly 0.25 Global Gridded Precipitation (1979–2015) by Merging Gauge, Satellite, and Reanalysis Data. *Hydrol. Earth. Syst. Sc.* 21, 589–615. doi:10.5194/hess-21-589-2017

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., et al. (2019). MSWEP V2 Global 3-Hourly 0.1 Precipitation: Methodology and Quantitative Assessment. *Bull. Amer. Meteorol. Soc.* 100, 473–500. doi:10.1175/BAMS-D-17-0138.1

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., et al. (2013). A Description of the Global Land-Surface Precipitation Data Products of the Global Precipitation Climatology Centre with Sample Applications Including Centennial (Trend) Analysis from 1901–present. *Earth Syst. Sci. Data* 5, 71–99. doi:10.5194/essd-5-71-2013

Celleri, R., Willems, P., Buytaert, W., and Feyen, J. (2007). Space–time Rainfall Variability in the Paute Basin, Ecuadorian Andes. *Hydrol. Process.* 21, 3316–3327. doi:10.1002/hyp.6575

Chen, F., and Dudhia, J. (2001). Coupling an Advanced Land Surface-Hydrology Model with the Penn State-NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Wea. Rev.* 129, 569. doi:10.1175/1520-0493(2001)129<0569:caalsh>2.0.co;2

Comin, A. N., Justino, F., Pezzi, L., de Sousa Gurjão, C. D., Shumacher, V., Fernández, A., et al. (2020). Extreme Rainfall Event in the Northeast Coast of Brazil: a Numerical Sensitivity Study. *Meteorol. Atmos. Phys.* 112, 39. doi:10.1007/s00703-020-00747-0

Crétat, J., Pohl, B., Richard, Y., and Drobinski, P. (2012). Uncertainties in Simulating Regional Climate of Southern Africa: Sensitivity to Physical Parameterizations Using WRF. *Clim. Dyn.* 38, 613–634. doi:10.1007/s00382-011-1055-8

Decharme, B., and Douville, H. (2006). Uncertainties in the GSWP-2 Precipitation Forcing and Their Impacts on Regional and Global Hydrological Simulations. *Clim. Dyn.* 27, 695–713. doi:10.1007/s00382-006-0160-6

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Quart. J. Roy. Meteorol. Soc.* 137, 553–597. doi:10.1002/qj.828

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., et al. (2018). Validation of the CHIRPS Satellite Rainfall Estimates over Eastern Africa. *Quart. J. Roy. Meteorol. Soc.* 144, 292–312. doi:10.1002/qj.3244

Domínguez-Castro, F., García-Herrera, R., and Vicente-Serrano, S. M. (2018). Wet and Dry Extremes in Quito (Ecuador) since the 17th Century. *Int. J. Climatol.* 38, 2006–2014. doi:10.1002/joc.5312

Dudhia, J. (1989). Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimension Model. *J. Atmos. Sci.* 46, 3077–3107. doi:10.1175/1520-0469(1989)046<3077:nsocod>2.0.co;2

ECMWF (2019). ERA5-Land Hourly Data from 1981 to Present. *Tech. Rep.* 11, 29. doi:10.24381/cds.e2161bac

Efstathiou, G. A., Zoumakis, N. M., Melas, D., Lolis, C. J., and Kassomenos, P. (2013). Sensitivity of WRF to Boundary Layer Parameterizations in Simulating a Heavy Rainfall Event Using Different Microphysical Schemes. Effect on Large-Scale Processes. *Atmos. Res.* 132-133, 125–143. doi:10.1016/j.atmosres.2013.05.004

Fersch, B., and Kunstmann, H. (2014). Atmospheric and Terrestrial Water Budgets: Sensitivity and Performance of Configurations and Global Driving Data for Long Term Continental Scale WRF Simulations. *Clim. Dyn.* 42, 2367–2396. doi:10.1007/s00382-013-1915-5

Flaounas, E., Bastin, S., and Janicot, S. (2011). Regional Climate Modelling of the 2006 West African Monsoon: Sensitivity to Convection and Planetary Boundary Layer Parameterisation Using WRF. *Clim. Dynam.* 36, 1083–1105. doi:10.1007/s00382-010-0785-3

Food and Agriculture Organization of the United Nations (2015). http://www.fao.org/3/a-i4560e.pdf

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al. (2015). The Climate Hazards Infrared Precipitation with Stations–A New Environmental Record for Monitoring Extremes. *Sci. Data* 2, 150066. doi:10.1038/sdata.2015.66

Gan, M. A., Kousky, V. E., and Ropelewski, C. F. (2004). The South America Monsoon Circulation and its Relationship to Rainfall over West-Central Brazil. *J. Clim.* 17, 47–66. doi:10.1175/1520-0442(2004)017<0047:TSAMCA>2.0.CO;2

Gbode, I. E., Dudhia, J., Ogunjobi, K. O., and Ajayi, V. O. (2019). Sensitivity of Different Physics Schemes in the WRF Model during a West African Monsoon Regime. *Theor. Appl. Climatol.* 136, 121. doi:10.1007/s00704-018-2538-x

Grell, G. A., and Dévényi, D. (2002). A Generalized Approach to Parameterizing Convection Combining Ensemble and Data Assimilation Techniques. *Geophys. Res. Lett.* 29, 1–38. doi:10.1029/2002GL015311

Grell, G. A., and Freitas, S. R. (2014). A Scale and Aerosol Aware Stochastic Convective Parameterization for Weather and Air Quality Modeling. *Atmos. Chem. Phys.* 14, 5233–5250. doi:10.5194/acp-14-5233-2014

Grimm, A. M. (2003). The El Niño Impact on the Summer Monsoon in Brazil: Regional Processes versus Remote Influences. *J. Clim.* 16, 263–280. doi:10.1175/1520-0442(2003)016<0263:TENIOT>2.0.CO;2

Harris, I., Jones, P., Osborn, T., and Lister, D. (2014). Updated High-Resolution Grids of Monthly Climatic Observations - the CRU TS3.10 Dataset. *Int. J. Climatol.* 34, 623–642. doi:10.1002/joc.3711

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz–Sabater, J., et al. (2020). The ERA5 Global Reanalysis. *Q. J. R. Meteorol. Soc.* 13, 133. doi:10.1002/qj.3803

Hersbach, H. (2000). Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather Forecast.* 15, 559–570. doi:10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2

Hong, S.-Y., Dudhia, J., and Chen, S.-H. (2004). A Revised Approach to Ice Microphysical Processes for the Bulk Parameterization of Clouds and Precipitation. *Mon. Wea. Rev.* 132, 103. doi:10.1175/1520-0493(2004)132<0103:aratim>2.0.co;2

Hong, S.-Y., Noh, Y., and Dudhia, J. (2006). A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Mon. Wea. Rev.* 134, 2318. doi:10.1175/mwr3199.1

Hong, S. Y., and Lim, J. O. J. (2006). The WRF Single-Moment 6-Class Microphysics Scheme (WSM6). *J. Korean Meteorol. Soc.* 42, 129–151. doi:10.1109/candar.2017.58

Hsiao, L.-F., Yang, M.-J., Lee, C.-S., Kuo, H.-C., Shih, D.-S., Tsai, C.-C., et al. (2013). Ensemble Forecasting of Typhoon Rainfall and Floods over a Mountainous Watershed in Taiwan. *J. Hydrol.* 506, 55–68. doi:10.1016/j.jhydrol.2013.08.046

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D. (2008). Radiative Forcing by Long-Lived Greenhouse Gases: Calculations with the AER Radiative Transfer Models. *J. Geophys. Res.* 113, 29. doi:10.1029/2008JD009944

Janjic, Z. I. (2000). Development and Evaluation of a Convection Scheme for Use in Climate Models. *J. Atmos. Sci.* 57, 573. doi:10.1175/1520-0469(2000)057<3686:codaeo>2.0.co;2

Janjic, Z. I. (1994). The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer and Turbulence Closure Schemes. *Mon. Wea. Rev.* 122, 927.

Jerez, S., López–Romero, J. M., Turco, M., Lorente–Plazas, R., Gómez–Navarro, J. J., Jiménez–Guerrero, P., et al. (2020). On the Spin–Up Period in WRF Simulations over Europe: Trade–Offs between Length and Seasonality. *J. Adv. Model. Earth Sy.* 12. doi:10.1029/2019MS001945

Jeworrek, J., West, G., and Stull, R. (2019). Evaluation of Cumulus and Microphysics Parameterizations in WRF across the Convective Gray Zone. *Wea. Forecast.* 34, 1097–1115. doi:10.1175/WAF-D-18-0178.1

Kain, J. S., and Fritsch, J. M. (1990). A One-Dimensional Entraining/Detraining Plume Model and its Application in Convective Parameterization. *J. Atmos. Sci.* 47, 2784–2802. doi:10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2

Kain, J. S. (2004). The Kain-Fritsch Convective Parameterization: An Update. *J. Appl. Meteorol.* 43, 170–181. doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2

Klein, C., Heinzeller, D., Bliefernicht, J., and Kunstmann, H. (2015). Variability of West African Monsoon Patterns Generated by a WRF Multi-Physics Ensemble. *Clim. Dyn.* 45, 2733–2755. doi:10.1007/s00382-015-2505-5

Liu, C., Ikeda, K., Thompson, G., Rasmussen, R., and Dudhia, J. (2011). High-Resolution Simulations of Wintertime Precipitation in the Colorado Headwaters Region: Sensitivity to Physics Parameterizations. *Mon. Wea. Rev.* 139, 3533–3553. doi:10.1175/MWR-D-11-00009.1

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N., and Riegger, J. (2014). Large-Scale Runoff from Landmasses: A Global Assessment of the Closure of the Hydrological and Atmospheric Water Balances. *J. Hydrometeorol.* 15, 2111–2139. doi:10.1175/JHM-D-13-0157.1

Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H. (2020). Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-Term Reference Forecast Product for the Water Sector in Semi-arid Regions. *Earth Syst. Sci. Data* 17, 131. doi:10.5194/essd-2020-177

Marengo, J. A., Alves, L. M., Alvala, R. C., Cunha, A. P., Brito, S., and Moraes, O. L. (2018). Climatic Characteristics of the 2010-2016 Drought in the Semiarid Northeast Brazil Region. *Anais da Academia Brasileira de Ciências* 90, 1973–1985. doi:10.1590/0001-3765201720170206

Martins, E. S. P. R., Coelho, C. A. S., Haarsma, R., Otto, F. E. L., King, A. D., Jan van Oldenborgh, G., et al. (2018). A Multimethod Attribution Analysis of the Prolonged Northeast Brazil Hydrometeorological Drought (2012–16). *Bull. Amer. Meteorol. Soc.* 99, S65–S69. doi:10.1175/BAMS-D-17-0102.1

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A. (1997). Radiative Transfer for Inhomogeneous Atmospheres: RRTM, a Validated Correlated K Model for the Longwave. *J. Geophys. Res.* 102, 16663–16682. doi:10.1029/97JD00237

Moya-Álvarez, A. S., Martínez-Castro, D., Flores, J. L., and Silva, Y. (2018). Sensitivity Study on the Influence of Parameterization Schemes in WRF-ARW Model on Short- and Medium-Range Precipitation Forecasts in the Central Andes of Peru. *Adv. Meteorol.* 2018, 1–16. doi:10.1155/2018/1381092

Müller, O. V., Lovino, M. A., and Berbery, E. H. (2016). Evaluation of WRF Model Forecasts and Their Use for Hydroclimate Monitoring over Southern South America. *Wea. Forecast.* 31, 1001–1017. doi:10.1175/WAF-D-15-0130.1

Ochoa, A., Campozano, L., Sánchez, E., Gualán, R., and Samaniego, E. (2016). Evaluation of Downscaled Estimates of Monthly Temperature and Precipitation for a Southern Ecuador Case Study. *Int. J. Climatol.* 36, 1244–1255. doi:10.1002/joc.4418

Parkes, B., Higginbottom, T. P., Hufkens, K., Ceballos, F., Kramer, B., and Foster, T. (2019). Weather Dataset Choice Introduces Uncertainty to Estimates of Crop Yield Responses to Climate Variability and Change. *Environ. Res. Lett.* 14, 124089. doi:10.1088/1748-9326/ab5ebb

Pleim, J. E. (2007). A Combined Local and Nonlocal Closure Model for the Atmospheric Boundary Layer. Part I: Model Description and Testing. *J. Appl. Meteorol. Climatol.* 46, 1383–1395. doi:10.1175/JAM2539.1

Pohl, B., Crétat, J., and Camberlin, P. (2011). Testing WRF Capability in Simulating the Atmospheric Water Cycle over Equatorial East Africa. *Clim. Dyn.* 37, 1357–1379. doi:10.1007/s00382-011-1024-2

Que, L.-J., Que, W.-L., and Feng, J.-M. (2016). Intercomparison of Different Physics Schemes in the WRF Model over the Asian Summer Monsoon Region. *Atmos. Oceanic Sci. Lett.* 9, 169–177. doi:10.1080/16742834.2016.1158618

Radanovics, S., Vidal, J.-P., and Sauquet, E. (2018). Spatial Verification of Ensemble Precipitation: An Ensemble Version of SAL. *Wea. Forecast.* 33, 1001–1020. doi:10.1175/WAF-D-17-0162.1

Ruiz, J. J., Saulo, C., and Nogués-Paegle, J. (2010). WRF Model Sensitivity to Choice of Parameterization over South America: Validation against Surface Variables. *Mon. Wea. Rev.* 138, 3342–3355. doi:10.1175/2010MWR3358.1

Sikder, S., and Hossain, F. (2016). Assessment of the Weather Research and Forecasting Model Generalized Parameterization Schemes for Advancement of Precipitation Forecasting in Monsoon Driven River Basins. *J. Adv. Model. Earth Sy.* 8, 1210–1228. doi:10.1002/2016MS000678

Simmons, A. J., Poli, P., Dee, D. P., Berrisford, P., Hersbach, H., Kobayashi, S., et al. (2014). Estimating Low-Frequency Variability and Trends in Atmospheric Temperature Using ERA-Interim. *Quart. J. Roy. Meteorol. Soc.* 140, 329–353. doi:10.1002/qj.2317

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, M., Duda, M. G., et al. (2008). A Description of the Advanced Research WRF Version 3. *Tech. Note NCAR* 113. doi:10.5065/D68S4MVH

Soares, P. M. M., Cardoso, R. M., Miranda, P. M. A., de Medeiros, J., Belo-Pereira, M., and Espirito-Santo, F. (2012). WRF High Resolution Dynamical Downscaling of ERA-Interim for Portugal. *Clim. Dyn.* 39, 2497–2522. doi:10.1007/s00382-012-1315-2

Tariku, T. B., and Gan, T. Y. (2017). Sensitivity of the Weather Research and Forecasting Model to Parameterization Schemes for Regional Climate of Nile River Basin. *Clim. Dyn.* 50, 4231–4247. doi:10.1007/s00382-017-3870-z

Taylor, K. E. (2001). Summarizing Multiple Aspects of Model Performance in a Single Diagram. *J. Geophys. Res. Atmos.* 106, 7183–7192. doi:10.1029/2000JD900719

Tiedtke, M. (1989). A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. *Monthly Weather Rev.* 117, 1779. doi:10.1175/1520-0493(1989)117<1779:acmfsf>2.0.co;2

Ulate, M., Dudhia, J., and Zhang, C. (2014). Sensitivity of the Water Cycle over the Indian Ocean and Maritime Continent to Parameterized Physics in a Regional Model. *J. Adv. Model. Earth Sy.* 6, 1095–1120. doi:10.1002/2014MS000313

von Sperling, E. (2012). Hydropower in Brazil: Overview of Positive and Negative Environmental Aspects. *Energ. Proced.* 18, 110–118. doi:10.1016/j.egypro.2012.05.023

Wernli, H., Hofmann, C., and Zimmer, M. (2009). Spatial Forecast Verification Methods Intercomparison Project: Application of the SAL Technique. *Wea. Forecast.* 24, 1472–1484. doi:10.1175/2009WAF2222271.1

Wernli, H., Paulat, M., Hagen, M., and Frei, C. (2008). SAL - A Novel Quality Measure for the Verification of Quantitative Precipitation Forecasts. *Mon. Wea. Rev.* 136, 4470–4487. doi:10.1175/2008MWR2415.1

Yang, B., Qian, Y., Lin, G., Leung, R., and Zhang, Y. (2012). Some Issues in Uncertainty Quantification and Parameter Tuning: A Case Study of Convective Parameterization Scheme in the WRF Regional Climate Model. *Atmos. Chem. Phys.* 12, 2409–2427. doi:10.5194/acp-12-2409-2012

Yang, Q., Yu, Z., Wei, J., Yang, C., Gu, H., Xiao, M., et al. (2021). Performance of the WRF Model in Simulating Intense Precipitation Events over the Hanjiang River Basin, China – A Multi-Physics Ensemble Approach. *Atmos. Res.* 248, 105206. doi:10.1016/j.atmosres.2020.105206

Zhang, C., Wang, Y., Lauer, A., and Hamilton, K. (2012). Configuration and Evaluation of the WRF Model for the Study of Hawaiian Regional Climate. *Mon. Wea. Rev.* 140, 3259–3277. doi:10.1175/MWR-D-11-00260.1

Zhang, C., and Wang, Y. (2017). Projected Future Changes of Tropical Cyclone Activity over the Western North and South Pacific in a 20-Km-Mesh Regional Climate Model. *J. Clim.* 30, 5923–5941. doi:10.1175/JCLI-D-16-0597.1

Zheng, Y., Alapaty, K., Herwehe, J. A., Genio, A. D. D., and Niyogi, D. (2016). Improving High-Resolution Weather Forecasts Using the Weather Research and Forecasting (WRF) Model with an Updated Kain-Fritsch Scheme. *Mon. Wea. Rev.* 144, 833–860. doi:10.1175/MWR-D-15-0005.1

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# *Supplementary Material*

## 1 SUPPLEMENTARY FIGURE

The general domain precipitation bias, feature-based simulated location and volume of the WRF WSM3 and WSM6 sub-ensembles precipitation patterns of all three study domains are analyzed in the ensemble Structure-Amplitude-Location analysis (eSAL) of Fig. S1. Little differences between the performance of WSM3 and WSM6 sub-ensembles are found for EP9 and EP3 with respect to amplitude, structure and location errors. NB9 shows slightly smaller amplitude errors for WSM3 than for WSM6.

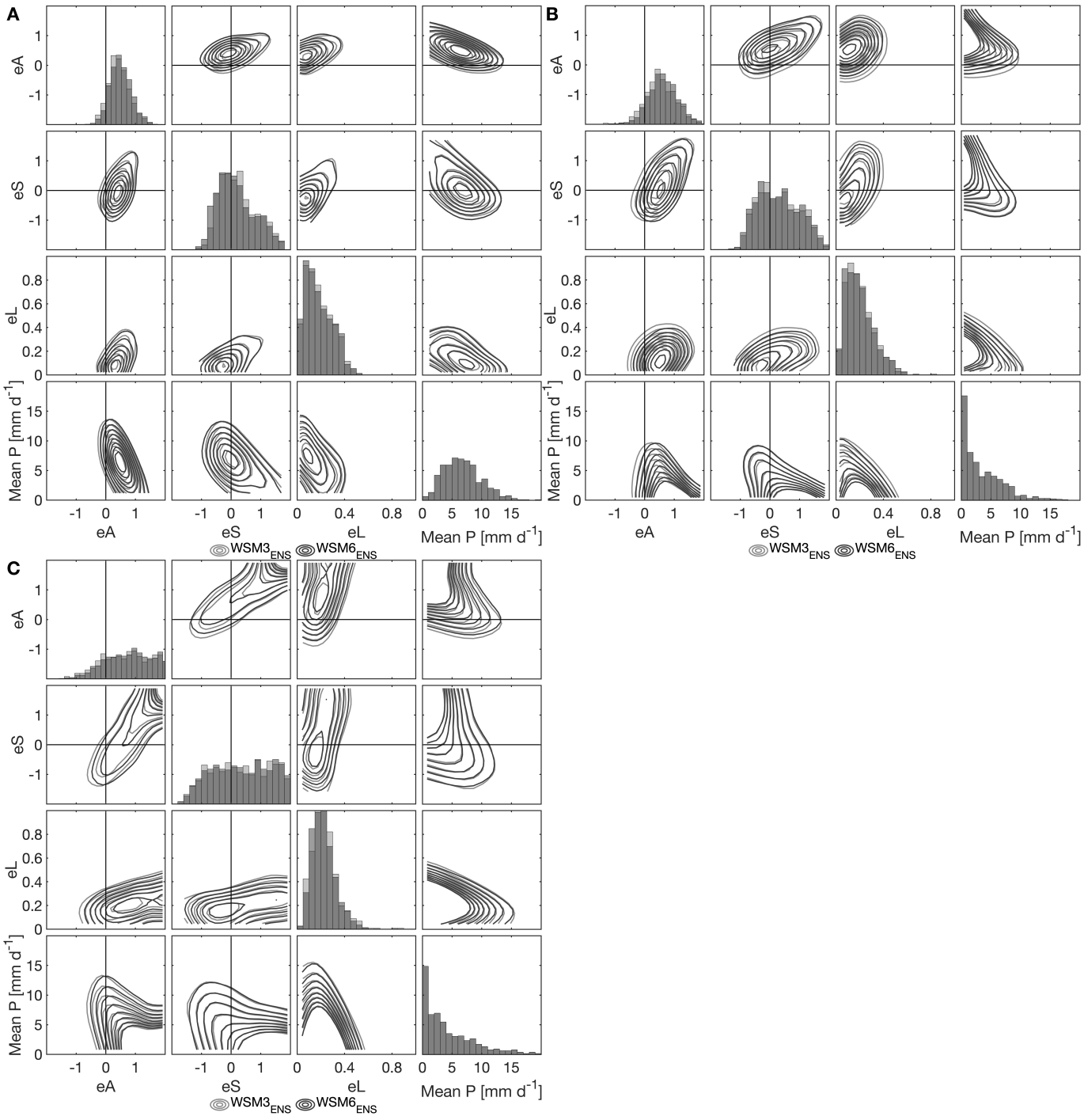


Figure S1: Ensemble Amplitude-Structure-Location (eSAL) analysis for (A) EP9, (B) NB9 and (C) EP3 with respect to CHIRPS for the physical parameterization groups of WSM3 and WSM6. For each physics ensemble, $eA$, $eS$, $eL$ and the reference domain average precipitation (Mean $P$) are juxtaposed in density contour plots of daily values from Jul 2006 till Jun 2008. The diagonal shows the probability histograms of $eA$, $eS$, $eL$ of the respective physics ensemble and of the reference domain average precipitation (Mean $P$). When the transparently colored probability histograms of $eA$, $eS$, or $eL$ overlap, darker colors are produced. As all different parameterization groups use CHIRPS as reference, histograms of reference mean $P$ always overlap.

# 5 Synthesis

## 5.1 Summary and Discussion

Seasonal forecasts offer the timescales that fill the gap between medium-range weather forecasts and decadal climate projections. These are exactly the timescales on which mitigation of hydrometeorological climate extremes can be provided and which are decisive for the planning and development of seasonal water resources management. Facing climate change, particularly in semi-arid regions with high population growth rates and high variability of water resources, there is increased urgency to develop and apply methods allowing for the planning even several months ahead. With the **Article I:** *Seasonal Forecasts offer Economic Benefit for Hydrological Decision Making in Semi-Arid Regions* (Portele et al., 2021b), we provide evidence that between 1981 and 2018, the frequency of very warm months and drought months increased significantly from 10 to 60 % and from 10 to 30 %, respectively, in the semi-arid study regions in South America, West and Northeast Africa and West Asia. To aid climate proofing in the face of these increased frequencies of droughts and warm extreme events, the potential use and value of global, unprocessed seasonal forecasts in hydrological decision-making was demonstrated. It was shown that switching towards seasonal-forecast based preemptive actions in hydrological decision making offers a clear economic benefit. This was accomplished by evaluating the performance of the latest SEAS5 global seasonal forecasting system for decision support in water management and dam operations in a user-oriented way based on relative economic savings (potential economic value, PEV). This method allows stakeholders to understand the forecast skill from an economic perspective, offering a broader range of possible applications of seasonal forecasts. The results revealed that high potential economic values occur not only for short forecast horizons, but up to seven months in advance. For droughts, potential economic savings up to 70 % of those from optimal early action could be achieved when basing preemptive action on seasonal forecasts. Even for forecast horizons of several months, savings of at least 20 % occur for the warm extreme and droughts. The direct, real-case application of the underlying economic cost-loss model was demonstrated for hydropower production at a typical representative of dams in semi-arid regions, i.e., for the Upper Atbara Dam Complex in Sudan. In one example year, avoidable losses of 16 Mio US$ were achieved for early-action based drought reservoir operation. The first article (Portele et al., 2021b) thus created a paradigm for the relevance and beneficial use of ensemble-based seasonal forecasts in an interdisciplinary context and their potential use for an user-oriented transfer to regional decision-making. Both the credibility and probabilistic nature of the seasonal forecasts were taken into account, and as little manipulation as possible was applied to the raw forecast data. The applied approach allowed the use of unprocessed forecasts and only implicitly incorporated bias correction by defining the extreme events separately on the distributions of reference and forecast data (and per lead time), respectively. The good performance in forecast value even for several

months ahead confirmed the suitability of SEAS5 seasonal forecasts with a horizontal resolution of 36 km for regional applications and statements at the catchment scale.

Limitations of this are obvious when absolute precipitation or temperature values are actually needed or distributed gridcell information is requested. At that point, post-processing of the global seasonal forecasts enters the scene, correcting for biases and model drift with lead time, as well as refining the included information regionally. In particular, the horizontal resolutions of 36 km need to be increased in the areas with complex topography, e.g., to capture enhanced precipitation at mountains. The **Article II:** *Bias-corrected and Spatially Disaggregated Seasonal Forecasts: a Long-term Reference Forecast Product for the Water Sector in Semi-arid Regions* (Lorenz et al., 2021) therefore delivered the corresponding next step of regional refinement and post-processing. As reference for the empirical statistical approach, ERA5-Land - the offline land surface re-run of the latest ECMWF reanalysis ERA5 - was used. With that, an increase in resolution of the seasonal forecasts SEAS5 from 36 km to 0.1° was achieved for four study regions in Ecuador/Peru, Northeast Brazil, Sudan/Ethiopia and Iran. The bias correction was performed with the method of quantile mapping on daily values, including corrections for extreme values in the forecasts and for precipitation intermittency. The analysis showed that the bias correction and spatial disaggregation (BCSD) approach is capable of achieving a better match of spatial patterns with ERA5-Land than the raw forecasts and reducing lead-dependent model drift effects. Monthly precipitation forecast biases of up to 4 mm $d^{-1}$ were successfully removed, and temperature biases were reduced from about 2 K to about 0.5 K or were entirely resolved. Shortwave radiation errors of up to 30 W $m^{-2}$ were further reduced to 5 W $m^{-2}$ or less. The agreement of raw SEAS5 and of SEAS5 BCSD with ERA5-Land was moreover compared in the Continuous Ranked Probability Skillscore quality measure, which indicates the agreement of the entire distribution of values with the reference. Substantial improvements in distributions of values were obtained for all considered variables of precipitation, temperature, and radiation, with the largest improvements for the minimum 2-m temperature. Deficiencies of the approach were evident for some months with large gradients, e.g., at the beginning or end of the rainy season, for which the inclusion of distributions of values 15 days before and after the current correction day in the quantile mapping approach reduced forecast performance. The BCSD approach was shown to perform well in terms of its ability to indirectly correct for altitude and to better represent small-scale features according to the higher-resolved reference. The free provision of the whole set of post-processed variables from seasonal forecasts across several domains for the long study period from 1981 to 2019 through the widely used portal World Data Centre for Climate (WDCC; Lorenz et al., 2020a,b,c,d) makes it very practicable and convenient for potential users to a) analyze the suitability (and, more general, the overall performance) of seasonal forecasts for the regional water management and b) use this post-processed and refined product for driving subsequent models. By post-processing the entire available (re-)forecast period of SEAS5, this article paves the way for opera-

tionalizing the provision of bias-corrected and spatially disaggregated skillful seasonal forecasts tailored to the study regions.

The refinement and correction of global fields using statistical methods for subsequent impact modeling is always tied to the presence of a high-resolution, multi-year and multi-variable data set. In this regard, any refinement and correction, e.g., of topographic effects, can only be as good as their representation in the reference data set. Likewise, the final spatial and temporal resolution is limited to this reference data set and cannot be further increased at will. Furthermore, it is possible that forecasts and reference data map different processes due to their different resolution (e.g. convective precipitation events). Since empirical statistical approaches do not describe the underlying physics and only correct for the long-term bias, the error of individual events may even be increased after bias correction. Therefore, their ability of correcting such scale-dependent inconsistencies is limited.

Here, dynamical downscaling with regional climate models (RCMs) like the Weather Research and Forecast (WRF) model can substantially contribute to the increasing need for reliable high-resolution regional hydrometeorological information. Under the constraints of available computing power and model stability with the physical processes involved, any desired horizontal resolution from several kilometers to a few hundred meters or any temporal resolution as high as a few minutes can be achieved. According to their model resolution, they are able to represent physical processes or at least to parameterize them accordingly. In order to minimize model bias, however, RCMs must be adapted in their physical setup to the region of interest. For the processes of precipitation generation, in particular the parameterizations of convection, cloud microphysics, radiation, and turbulent mixing are relevant. For the two study regions in South America, the **Article III:** *Ensemble-tailored Pattern Analysis of High-resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America* (Portele et al., 2021a) therefore involved an extensive sensitivity experiment with 32 ensemble simulations evaluating the effect of different parameterization schemes on the simulated WRF precipitation in two 9-km domains and one 3-km domain. The 32 runs were evaluated according to the distributional, temporal and spatial patterns of simulated precipitation and this pattern analysis was tailored on parameterization sub-ensembles to better understand the effects on precipitation when combining one fixed scheme with others. It was shown that there can be enormous spatial and temporal ranges of variation in precipitation, with up to four times the precipitation amounts of three different reference datasets (CHIRPS, MSWEP, ERA5-Land). For the 9-km domains as well as for the convection-resolving 3-km domain, the applied schemes, in particular, for convection (although only used in the parent domain for the 3-km domain) and for radiation dominated the sub-ensemble tendencies by the largest deviations from the overall ensemble mean and by the largest differences in the value distributions or structure of the precipitation features. Not a single best performing physical setup could be identified considering all distributional, temporal and spatial evaluation metrics. Nevertheless, with the ensemble-tailored pattern analysis methods, suited parameterizations were

narrowed down, allowing a choice according to application needs and reference data. This computationally expensive experiment showed that, depending on the physics setup chosen, there can be huge gains, but also losses, in model performance. For subsequent impact modeling or application to seasonal forecasting, it is therefore essential to undertake this effort of an extensive parameterization sensitivity study.

In the context of seasonal forecasting, however, the need for a preceding extensive parameterization sensitivity study for dynamical downscaling with WRF implies an extra, indispensable, but non-negligible computational effort that comes on top of the computational demands of dynamically downscaling ensemble forecasts - for SEAS5 with 51 members - for each month and for seven months lead time. 2.5 years for each of the 32 parameterization combinations already resulted in 960 simulation months and required 18.5 Mio CPU hours including several repeated runs due to model instabilities and incomplete output writing. For each new issue month, 357 simulations months arise from the 51 ensemble members and the seven months lead time. Of course, national weather services would only simulate their respective domain, and not, as in this sensitivity study two Ecuadorian-Peruvian and one Brazilian domain at once. Also the outermost parent domain at 27 km horizontal resolution necessary with the ERA-Interim input data would be skipped as SEAS5 is already available at 36 km. However, for a grid and time step ratio of 1:3, and about equally sized domains in terms of gridcell number, the nest domain requires about four times (from the ratio and some nesting overhead) the computing time of the parent domain. Thus, not too great savings in computational effort are to be expected from this. In fact, an enormous amount of computational effort still remains for the dynamical downscaling of the seasonal forecasts for each area that exceeds the limited computational resources in many developing countries.

The sensitivity study of Article III (Portele et al., 2021a) revealed another drawback of the dynamical downscaling approach with respect to seasonal forecasts: Depending on the final choice of the physics setup, there still remain smaller or larger biases in the probability distribution or discrepancies in the spatial precipitation patterns compared to the reference. Therefore, bias correction or some additional post-processing is often still required, although a choice of setup would be done according to the least necessary manipulation afterwards. To perform such a bias correction, several years of simulations and corresponding evaluations would be required. For seasonal forecasts, this involves long re-forecast periods with ensemble simulations to correct the remaining mean forecast errors relative to the reference data. In particular, the lead-dependent effects have to be taken into account, since the global input fields suffer from model drifts that cannot be reduced by using dynamical downscaling. An example would be the shift of the location of atmospheric circulation patterns with lead time in the global forecast data (as was the case for Sudan/Ethiopia in Article II (Lorenz et al., 2021) with a southward shift of higher temperatures and higher radiation with increasing lead times) which would result in the wrong model physics to apply during the dynamical downscaling. Apart from post-processing the WRF downscaled seasonal forecasts, the use

of lead-dependent correction of global input data for the study domain might also be promising (following the a-priori bias-correction as in Pontoppidan et al., 2018; Bruyère et al., 2014).

Regarding the transferability of the identified setup with ERA-Interim to the downscaling of the seasonal forecasts SEAS5, another result from Article III (Portele et al., 2021a) should be kept in mind: The 3-km experiment showed the large dependence of the performance of the simulated precipitation in the convection-resolving domain on the cumulus parameterization of the parent domain. Here, the consistency of the ECMWF Integrated Forecast System (IFS) cycle should allow this transferability: All IFS cycles for SEAS5, ERA5, and ERA-Interim include a convection parameterization based primarily on the Tiedtke (1989) cumulus scheme (Johnson et al., 2019; Hersbach et al., 2020; ECMWF, 2007), but with ongoing developments and improvements in each cycle. Additional studies would be required for other prediction systems using different cumulus convection physics schemes. Another approach, as in Mori et al. (2021), would be to perform the sensitivity analysis directly on the seasonal forecasts, but again this raises the issue of computationally expensive ensemble simulations. Mori et al. (2021) here only tested the different parameterization schemes with the input of one SEAS5 ensemble member.

For the application of a time-consuming and computationally demanding dynamical downscaling in the framework of meteorological services, it would be desirable to develop a technique to reduce the ensemble size while conserving the specific characteristics of an ensemble forecast. However, the complete randomness in generating the SEAS5 ensemble due to the two step algorithm including initial condition perturbations and stochastic model perturbations (i.e., no control forecast exists, Johnson et al., 2019) does not allow the selection of the "best" ensemble members for further downscaling. In this regard, the ensemble generation of the Climate Forecast System, Version 2 (CFSv2) of the National Centers for Environmental Prediction (NCEP; Saha et al., 2014) in turn allows a selection of sub-ensembles, since the numbering of ensemble members is fixed and refers to the lagged start dates (every 5th day up to four weeks before the release in case of the reforecasts from 1982 to 2010, daily for forecasts since 2011) and times of day (4 times per start date) for each monthly forecast release. This motivated and allowed the study of, e.g., Siegmund et al. (2015), who chose a 22-member sub-ensemble of CFSv2 to dynamically downscale one study rainy season in West Africa. Their analysis also showed that post-processing of WRF output in seasonal forecasts would still be required for, e.g., the treatment of model drizzle rain, or for lead-dependent effects. As another example, Yuan and Liang (2011) simulated 5 retrospective CFS forecast ensemble members for 27 seasons and therewith allowed a more climatological analysis of seasonal forecast improvement with WRF dynamical downscaling. For SEAS5, however, the selection of ensemble members would be purely random and may therefore include members that capture the true variance well or represent only parts of the entire ensemble. Therefore, for the final application of operational WRF seasonal

forecasts, approaches to deal with the large, randomly generated SEAS5 ensemble still need to be developed.

Although dynamical downscaling of the seasonal forecasts SEAS5 does not appear to be operationally feasible at that time, dynamically downscaled reanalysis data can still provide useful information on current water resources needed for sustainable water management. The conducted effort of the sensitivity experiment, especially with the identification of Tiedtke or BMJ and RRTM-Dudhia as suitable schemes in the 9-km domains with high similarity of spatial precipitation patterns with CHIRPS, can finally help local authorities to run their own high-resolution simulations. Furthermore, it contributes to increase their knowledge of the hydrometeorological situation, as well as of hydrological or ecosystem conditions of past and current years through the subsequent application of impact modeling. In particular, with the latest reanalysis ERA5, which is updated daily with a latency of about 5 days (Hersbach et al., 2020), the current situation for water management could be monitored even better with high-resolution WRF simulations. The comparable high-resolution ERA5 land product, in turn, is only available monthly with a lag of about three months relative to the current date (Muñoz-Sabater et al., 2021).

In general, statistical post-processing methods range from very simple (e.g., linear scaling) to highly complex multivariate statistical approaches. The BCSD approach, as a quite simple approach including the empirical quantile mapping was chosen due to several reasons: It is widely accepted in the community and has been extensively applied for post-processing all sorts of weather and climate projections (e.g., Thrasher et al., 2013; Ning et al., 2015; Briley et al., 2017; Nyaupane et al., 2018). While there are approaches that perform better in some aspects (e.g., Abatzoglou and Brown, 2012; Schepen et al., 2020), a highly robust and computationally relatively cheap method was needed that performs consistently across a very large dataset on the gridpoint-scale. To improve the bias correction approach, spatial, temporal and inter-variable correlations (e.g., Schepen et al., 2020) should be considered in future studies. In all downscaling techniques, whether statistical or dynamical approaches, the performance of the posterior forecasts always depends on the performance of the raw global forecast. Once the global model poorly predicts the large-scale circulation, the refined forecast can hardly capture the anomaly correctly (e.g., Haerter et al., 2011; Yuan et al., 2015) - the statistical post-processed forecasts because of the tight constraint on the global values, and the RCM because of the transfer of the wrong signal to local processes (e.g., precipitation) and scales according to the globally forced lateral and lower boundary conditions. Therefore, conditional bias correction methods (Yuan et al., 2015, for a broad overview), i.e., the posterior forecasts becomes very close to the observation or to the climatology depending on the model skill, are worth considering. Haerter et al. (2011) further suggest a cascade bias correction approach that includes the correction of several timescales, i.e., hourly, daily, monthly, depending on the final application. This could further improve the monthly evaluated - but daily corrected (with a moving 30-day window) - SEAS5 BCSD forecasts.

As a robust, easy to implement, and performant approach, BCSD allows the step to the operationalization of hydrometeorological forecasts. It is comparatively little computationally intensive, yet its application to gridded data requires consistent parallelization of processes, user-friendly and application-oriented storage strategies, and, notwithstanding, adequate high-performance computing facilities. The users in the study regions often lack the computational power or expert knowledge for applying very advanced and complex methods. The BCSD-post-processing, however, can be easily transferred to stakeholders in the respective regions so that local experts can continue to process the seasonal forecasts. Once the historical statistical relationship of forecast and reference is available, all current 51 forecast ensemble members of the hydrometeorological variables of precipitation, 2-m temperature, and radiation can be bias-corrected and spatially disaggregated within just one day after the official release of the SEAS5 forecasts (5th of each month). Thus, the high benefit and good performance of the lowest lead month can still be used in the water management context to stimulate or update decisions for the remaining three weeks of the current month. For global approaches, first tests already showed that BCSD can be applied to a global gridded study area (land areas) with reasonable computational effort when disaggregated to the 0.25° resolution of ERA5.

The final dissemination of the post-processed SEAS5 BCSD product is achieved by releasing the monthly SEAS5 BCSD forecasts in an online decision support system (DSS, `https://sawam.gaf.de`). This provides access to high-resolution corrected seasonal forecasts to local experts who often do not have the opportunity to apply the ECMWF high-resolution operational products. Other products (e.g., CFSv2 from NCEP) are only available in a very coarse spatial resolution and all of these data sets need some sort of post-processing. In addition, the acquisition, analysis and application of seasonal forecasting data is often not straightforward which hinders potential users to explore and use such products. Thus, the SEAS5 BCSD product might currently provide the most easy "point of entry" for potential users of seasonal forecasts. In the DSS, it is important not only to present the ensemble mean values, but to show the entire forecast range with corresponding probabilities. Only then can additional credibility and confidence in the forecasts be established. In this context, the presentation of derived forecast quantities such as categorical forecasts plays a major role for the user-friendliness and the understanding of probabilistic seasonal forecasts. For those, the long reforecast period offered with the BCSD approach is decisive. Depending on the specific application in water management, both gridded and regionally averaged data are of interest and need to be provided in the DSS.

The SEAS5 BCSD forecasts were not only generated to provide high-resolution and corrected seasonal hydrometeorological (re-)forecasts, they were further intended to offer the opportunity for spatially distributed hydrological and ecosystem modeling for the coming season. Proper calibration and reliable functioning of hydrological or ecosystem models require long time series and a (spatially and temporally) consistent, bias-corrected system of long-term reference data and a seasonal forecasting product. Long

spin-up periods for the ecosystem or hydrological models are required and can be performed with ERA5-Land. With the spun-up model, the consistent, bias-corrected SEAS5 BCSD are used to finally provide seasonal forecasts of hydrological and ecosystem variables. For initial testing with (re-)forecasts of SEAS5 BCSD, this model chain already proved to be successful. However, for actual operational hydrological or ecosystem seasonal forecasts, high-resolution reference data are still missing that reach up to the present day. ERA5-Land is provided with a latency of three months (Muñoz-Sabater et al., 2021) and thus does not provide a seamless link to the seasonal forecasts. To close this gap, seamless approaches are still lacking and future work is needed.

In the three different articles (Portele et al., 2021b; Lorenz et al., 2021; Portele et al., 2021a), the performance analysis of seasonal forecasts or simulated hydrometeorological fields was a central issue. Depending on the specific requirements and application, very different performance measures were required and implemented. In general, forecast performance is evaluated primarily in terms of quality, but only the valuation with potential economic benefits for forecast-based early action can provide clear decision support, especially for probabilistic seasonal forecasts. The provision of the economic value of seasonal forecasts is able to bridge between science and decision making (e.g., Bruno Soares et al., 2018; Lopez et al., 2018), which can only rarely be achieved by the provision of statistical quality skillscores, like the CRPSS (e.g., Crochemore et al., 2016). In turn, for the assessment of performance gain of the post-processed BCSD forecast compared to the raw forecast exactly the CRPSS, bias, and RMSE provided the desired information on value distribution and mean forecast error that had to be corrected. For the analysis of WRF output, classical approaches based on correlation, standard deviation, and RMSD were complemented by approaches that specifically consider the local displacement and spatial structure errors of precipitation patterns. This spatial evaluation becomes increasingly important as the horizontal resolution improves, since classical point verification scores can double penalize high-resolution fields (double-penalty problem, Jolliffe and Stephenson, 2011): for missing the actual precipitation in the gridcell, and for predicting it in another gridcell where no precipitation was present in the reference. Thus, this work presented key components for the evaluation of hydrometeorological fields from very different perspectives, taking into account the final transfer to practice, the assessment of forecast improvement, and the spatial similarity of simulated and reference fields at high resolutions.

All articles (Portele et al., 2021b; Lorenz et al., 2021; Portele et al., 2021a) showed that a reliable reference dataset is the basis for all performance evaluations. In particular, the need for a multi-year and consistent multi-variable reference data set with the resolution of SEAS5 for analyses at the catchment scale, or higher resolution for statistical bias correction and spatial disaggregation posed an additional challenge. For the data-sparse study regions, where access to reliable local station data in a dense measuring network is limited or where available station data often have errors and gaps, only reanalysis data can meet these requirements. For the economic value analysis of the first article (Portele et al., 2021b), certainly other reference data sets could have been used instead of

ECMWF's latest reanalysis product ERA5. The requirements for the analysis of the first article were a global gridded data set with a continuous timeseries from 1981 to 2018 including at least precipitation and mean 2-m temperature on monthly timesteps with roughly the horizontal resolution of SEAS5. Here, the CRU data set at 0.5° (CRU TS v4, Harris et al., 2020) would have been a valuable alternative for ERA5, however, station coverage especially in South America and Africa in the most recent decades was rather poor. The search for a suitable reference data set for the bias correction and spatial disaggregation of the second article (Lorenz et al., 2021), however, turned out to be much more challenging. To allow for a regional refinement and a later application of the forecasts in impact modeling, the reference data set should provide a high spatial resolution suitable for regional applications (10-20 km), at least a daily timestep and the minimum set of variables necessary for impact modeling, i.e., precipitation, mean, maximum and minimum 2-m temperature, as well as shortwave radiation. Moreover, a long timeseries was essential to derive the statistical relationship. Before the release of ERA5-Land in the second half of 2019, no data set was able to fulfill all these requirements. ERA5 lacked the high spatial resolution; CRU additionally lacked the daily timestep. For precipitation alone, higher-resolution data sets existed (like CHIRPS), however, a consistent multi-variate data set comprising all study regions was not available. First tests of the BCSD approach were therefore conducted with a data set that even needed to be created specifically for that purpose: ERA5 was spatially disaggregated to 0.1° and corrected for the climatological mean using the high-resolution climatology of WorldClim 2 (at 1 km, Fick and Hijmans, 2017) for temperature and radiation, and using CHIRPS for precipitation. The statistical downscaling therefore showed all the more how dependent it is on the existence of a long-term, spatiotemporally suitably resolved and multi-variate reference data set. Finally, the evaluation of the WRF output in the third article (Portele et al., 2021a) introduced new issues, particularly for the validation of the high-resolution 3-km domain. WRF is capable of producing higher resolved fields, especially in topographically highly complex terrain, but sound reference data at this resolution are needed to evaluate its performance. For temperature, no such high-resolution global, gridded dataset exists at all. In turn, CHIRPS is so far the only data set providing near-global gridded precipitation at around 5 km spatial resolution over a long time period. However, the analysis revealed potential deficiencies of this reference product when focusing on precipitation in high mountain areas, such as the South American Andes. The lack of local station data in the regions where CHIRPS and WRF output differed most also made it difficult to draw definitive conclusions about WRF performance in these areas. Nonetheless, it seems plausible that the underlying physics in WRF could at least capture more details of the small-scale atmospheric conditions and variability there. Any approaches that rely on sparse or inadequate station or space-borne observational data will have even more issues in these areas. In general, global data sets purport to provide a consistent level of uncertainty, but their performance in some regions - despite state-of-the-art model and remote sensing data - can be substantially worse than in other regions. These challenges regarding a suitable, reliable reference data set remind us that a dense network of local station data that is quality-checked, standard-

ized, and freely available would still be required. Only then can high-resolution global data sets be further calibrated and improved. The assessment of other secondary fields such as seasonal forecasts with these data sets certainly assumes fidelity to reality, but correspondingly, these secondary fields can also only be as good as the reference used.

## 5.2 Answers to the Research Questions

1. Can seasonal forecasts support decision-making and provide economic benefit for the regional water management in semi-arid regions?

Decision makers in the water resources management are often still hesitant to use seasonal forecasts as a decision tool. But indeed, this work (Portele et al., 2021b) demonstrated that for seven hot spot drought-prone basins in South America, Northeast Africa, West Africa and West Asia, switching towards seasonal-forecast based preemptive actions in hydrological decision making offers a clear economic benefit and aids climate proofing. This work proved that high potential economic values do not only occur for short forecast horizons, but up to seven months ahead: Seasonal drought forecasts achieve economic savings of up to 70 % of what would have been possible with optimal early action; for several months ahead, savings of at least 20 % are achieved for very warm months and drought periods. The direct, real-case application of the underlying economic cost-loss model is demonstrated for hydropower production at a typical representative of large dams in semi-arid regions, for the Upper Atbara Dam Complex in Sudan. For an example drought year, losses of 16 Mio US$ could have been saved there. This work therewith creates a paradigm for the relevance and beneficial use of ensemble-based seasonal forecasts in an interdisciplinary context and their potential use for an user-oriented transfer to regional decision-making.

2. How can empirical-statistical regionalization and post-processing techniques improve the regional quality of global seasonal forecasts?

Due to their relatively coarse resolution of 36 km at best, and due to model biases and drifts, global seasonal forecasts often cannot be used directly for high-resolution applications and impact studies. This work (Lorenz et al., 2021) showed that an applied bias-correction and spatial-disaggregation (BCSD) of global seasonal forecasts based on daily values successfully corrects for model biases of monthly mean precipitation of up to 4 mm $d^{-1}$, reduces mean monthly temperature biases from about 2 K to about 0.5 K or eliminates them altogether, and decreases biases in monthly mean shortwave radiation of up to 30 W $m^{-2}$ to 5 W $m^{-2}$ or less. For all considered variables - precipitation, temperature, and radiation - considerable improvements of the value distributions were obtained, with the largest improvements for the minimum 2-m temperature. Lead-dependent bias correction further allowed to remarkably reduce model drift effects. In addition to that, the BCSD approach was shown to improve the regional quality of global seasonal forecasts as it is able to indirectly correct for elevation and to better represent small-scale features consistent with the higher-resolution reference. Weaknesses

of the applied empirical-statistical regionalization and post-processing procedure were found in relation to the used sliding window for the correction of daily values and the limitations given the availability of a high-resolution multivariable reference dataset. Nevertheless, the computationally cost-effective BCSD approach can provide an entire 39-years reforecast period and an operationally generated high-resolution, high-accuracy seasonal forecast dataset for regional applications and impact studies.

3. How robust are dynamically downscaled hydrometeorological fields with respect to the applied physical model setup?

The refinement of global hydrometeorological information is not only possible with statistical methods, but also with dynamical downscaling using regional climate models. However, regional climate models need to be adapted in their physical setup to the region of interest. Especially the parameterizations of convection, cloud microphysics, radiation, and turbulent mixing determine the processes of precipitation generation. This work (Portele et al., 2021a) showed that especially the applied schemes for convection and radiation dominate the tendencies of all runs with one particular parameterization scheme (sub-ensemble). It was also found that there can be enormous precipitation ranges in space and time, with precipitation amounts up to four times that of reference data. This implies that a high uncertainty is introduced into dynamically downscaled hydrometeorological fields by different applied physical model setups, i.e. a low robustness is given. It is therefore concluded that it is absolutely essential to carry out such computationally expensive parameterization analyses before forecasts or input data for impact studies can be produced. The applied ensemble-tailored methods for analyzing temporal, distributional, and spatial patterns allow for the narrowing down of appropriate and more robust parameterization schemes according to the application and reference data, ultimately enabling the use of the dynamically downscaled information for, e.g., hydrometeorological monitoring strategies of water resources.

## 5.3 Conclusions

This thesis substantially improved the understanding of seasonal forecasts and demonstrated the clear advantage of integrating them into hydrological decision making. It offered the tools to finally overcome the hurdle of actually incorporating seasonal forecasts into water management operations, and ultimately provides integral components to support sustainable regional water management. First, the study of Portele et al. (2021b) tackled the perceived unreliability and reluctance towards probabilistic forecasts. The specifically implemented approach of the potential economic benefit of forecast-based early action demonstrated the potential of the global raw seasonal forecasts SEAS5: For the semi-arid study regions, economic savings of up to 70 % of those from optimal early action could be achieved in the case of drought forecasts. In general, decision-makers could especially benefit from seasonal drought forecasts and forecasts for very warm months even several months ahead. The potential economic benefit includes

the forecast probability and the user's cost-loss situations, and therefore proved to be a strong tool for user-oriented, economic decision-related verification of probability forecasts.

Classical performance analyses revealed the deficiencies in the quality of seasonal forecasts in terms of absolute value distributions and mean forecast error, and demonstrated the need, in particular for spatially distributed impact modeling of hydrology or ecosystems, to provide post-processed, i.e., bias-corrected and regionally refined, seasonal forecasts. Therewith, the second tool offered was the development of an empirical statistical bias-correction and spatially disaggregation approach (Lorenz et al., 2021). The global seasonal forecasts were thus made usable in a broader context for local stakeholders and for impact studies by providing the required spatiotemporal time scales, and by correcting biases against reference data and model drifts with increasing forecast lead time. In the final step of operationalization, region-tailored seasonal forecasts of hydrometeorological variables relevant to water management are made available in an online decision support system. The online tool gives local stakeholders and experts from the four study domains the opportunity to integrate high-resolution, corrected seasonal forecasts in their decisions on water resources. This allows them to investigate the potential of the SEAS5 BCSD forecasts for regional water management, drought forecasting and irrigation planning, to build trust in the performance of these forecast, and to finally refine decision calendars and dates for management strategies of the coming months.

The final tool was the elaboration of suitable physics parameterization schemes in the regional climate model WRF to reproduce the temporal, distributional and spatial patterns of precipitation over two semi-arid study regions (Portele et al., 2021a). This setup should provide the entry point to be able to dynamically downscale seasonal forecasts, and therewith provide a physically based, tailor-made alternative of regional refinement with the freely available, community model WRF to the statistical approach. However, the tremendous computational demand of dynamically downscaling 51 ensemble members of SEAS5 seasonal forecasts for each month with seven months lead time and further downscaling a long period of probabilistic reforecasts for correction purposes, makes it at the current stage not feasible to provide operational WRF SEAS5 seasonal forecasts. Nevertheless, regionally refined information on currently available water resources, i.e., applying the WRF setup (identified according to the performed sensitivity analysis and specific for the application) to the near-present day reanalyses ERA5, could improve the hydrometeorological data basis in the data-sparse semi-arid regions and would allow the high-resolution monitoring of extreme events.

With all tools provided, this thesis actively contributes to a sustainable water management in the semi-arid study regions. Tools were chosen to be user-friendly and to provide profound decision support (potential economic benefit), computationally feasible and operationally implementable (bias correction and regional refinement) in the study regions, but also scientifically challenging (dynamical downscaling). All types of ap-

proaches to evaluating the performance of seasonal forecasts are important and justified to raise awareness and advance the use of seasonal forecasts. They show that even with long lead times, information can be obtained for decision making and economic value can still be achieved. Moreover, the tools raise the awareness of challenges facing with the application of seasonal forecasts. Often, they cannot be treated as everyday medium-range weather forecasts, and especially the communication of their probabilistic nature needs to be reinforced. Long (re-)forecast periods with several months lead time need to be considered for performance assessments and for common statements about “above" or “below normal" conditions, highly increasing the storage and processing requirements. The application of certain computationally demanding model approaches, in particular at the gridpoint scale, may be unfeasible at that time. Differences between forecasting systems, e.g., SEAS5 and CFSv2, need to be understood to be able to choose the proper tools, such as for further refinement. Dynamical downscaling may be computationally too demanding for SEAS5 due to the large ensemble number and randomness in their ensemble generation. However, systems like CFSv2 allow the selection of fixed ensemble members and therewith ease the application of dynamical downscaling.

However, the mere provision of tools to improve regional decision-making on water resources is not enough. Close cooperation with local stakeholders, capacity building and knowledge transfer through on-site workshops, including methods training in the study regions are required for the final transfer of developed tools into practice. Starting with complicated drought indicators or skill score measures, it was necessary to use appropriate and easy-to-understand indicators and performance measures and to demonstrate the advantages of seasonal forecasts in practical applications. Different political, societal and administrative hierarchies needed to be acknowledged in the study regions to be able to support the regional water management decisions. In general, local knowledge is required to properly identify the challenges and constraints (e.g., politically induced anthropogenic drought), to address the demands (e.g., transboundary water management without upstream data), and to be able to provide reasonable solutions according to the local conditions (e.g., refined seasonal forecasts for entire transboundary basins). On the other hand, elaborate and in-depth studies and results were in turn only possible with the support of local partners and with the knowledge of regional challenges. A commensurate workflow and effective exchange with local authorities had to be established, including the provision of (sparse) local station data. Especially the in-depth case study at the Upper Atbara Dam was only possible with the close cooperation and data exchange with local authorities. Despite all global approaches and global data sets, water management issues can thus only be solved by profound expertise in the local systems and close exchange with local partners.

Finally, the detected increase in the frequency of drought months from 10 to 30 % in semi-arid regions of Brazil, West and Northeast Africa, and Iran in recent decades (Portele et al., 2021b) calls for consistent action against the impacts of climate change. Global changes in precipitation and water availability in the context of climate change

are noticeable on all spatial and temporal scales, and lead to a radical change in the basis for action for sustainable water management. In this work, seasonal predictions proved to be a strong tool in the management of climate extremes, and with proven value and skill even months ahead they are suited as adaptation and mitigation strategies to climate change. Therewith, seasonal forecasts can be viewed as the product that water managers have been waiting for and there is increasing importance of seasonal forecasts in decision support for a variety of socioeconomic applications. It was shown that reservoir management in water-scarce regions can highly benefit from integrating early action plans for, e.g., droughts (Portele et al., 2021b). The stored volume of water in a reservoir can thus be managed more efficiently, increasing annual electricity production in hydropower plants, and in a second step enhancing the sustainability of providing water resources for irrigation or drinking water supply. This is of urgent importance, as the adverse impacts of droughts on ecosystem productivity and the associated increasing pressure on global food security are well known (e.g., Gampe et al., 2021). The ability to store water resources more sustainably in large reservoirs can thus help to better link water, energy, and food in a local-global understanding of water resources at all levels of decision-making. This research therefore leads to a considerable increase in competence in the field of operational reservoir management. There is a great need worldwide and thus also a great potential for transferring the results and methods to other semi-arid regions. The larger implication of this study can be perceived when considering that there are more than 58,000 large dams worldwide (International Commission on Large Dams (ICOLD-CIGB), 2021; Mulligan et al., 2020) for which the use of seasonal forecasts could improve dam operations, and as a further step ensure water supply in regions of water scarcity. This work thereby contributes towards Sustainable Development Goal 6, that is to “ensure availability and sustainable management of water and sanitation for all".

In conclusion, for locally to globally consistent, scientifically sound, coordinated decisions leading to the sustainable use of water, the gathered scientific knowledge and technological progress in seasonal forecasting must be integrated as an integral factor in all decisions related to the management of water resources. To this end, viable, user-friendly solutions are key to bridging the gap from science to decision-making. Operationalizing the developed tools and disseminating them in a decision support system are further essential to reinforce the integration of these advanced solutions. Science must therefore not limit itself to research on water resources or seasonal forecasts. Rather, it must empower society to also use the gained knowledge and actively promote its transfer into practice. Finally, coordination across all levels of government and all relevant sectors and stakeholder groups must be enhanced to integrate scientific knowledge and technological progress into decision making at all levels.

## 5.4 Outlook

Future work should be dedicated to the advancement and increased application of the presented tools and methods. Bias correction methods for seasonal forecasts could be further optimized by using multivariate (e.g. Schepen et al., 2020), skill-dependent (e.g. Yuan et al., 2015), or temporal cascade correction approaches (e.g. Haerter et al., 2011). Other approaches, e.g., linking the large- scale synoptic weather (circulation) patterns with station precipitation and temperature (e.g. Laux et al., 2021a), should also be considered. In this regard, the developed SEAS5-BCSD seasonal forecasts (Lorenz et al., 2021) can serve as a long-term reference product for new correction and regional refinement approaches.

Post-processing and correction methods may not only include classical approaches, but artificial intelligence (AI) and/or machine learning (ML) techniques should also come into play (Haupt et al., 2021). Different algorithms can be tested here with the aim of learning a pattern between predictor and predictand to optimize the seasonal forecasts, and to downscale the global fields (e.g., Xu et al., 2019). AI/ML approaches have the advantage that multi-variable, multi-source, and multi-lead-time data from the pixel to the global scale can be easily incorporated into forecast improvement (e.g., Feng et al., 2020). Nonetheless, feasibility in terms of Random-Access Memory (RAM) always plays a role here as well, if many input data are included. But the general advantage of ML models in terms of computing power, i.e., they run faster with much lower computer resources than dynamical models, underlines the importance of considering these new tools also for improving seasonal forecast quality in the future (Anochi et al., 2021; Cohen et al., 2019).

Regarding the economic benefits of seasonal forecasts, the cost-loss analysis at the Upper Atbara Dam could be extended to its multipurpose nature, including costs and losses related to, e.g., irrigation. In addition, with the coupling of a hydrological model, future seasonal inflow to the reservoir can be estimated, and simulation of, e.g., drought events to derive real drought management plans incorporating the seasonal forecasts should be considered. For the final operational hydrological seasonal forecasts, the temporal gap between the reference spin-up data and the start of the seasonal forecasts needs to be closed. Multivariable, long-term reference input data for impact modeling are at best provided up to a few days (ERA5) or a few months (e.g., ERA5 land) before the present day. Future work here should address the seamless link from spin-up to seasonal forecasting.

With the chosen WRF setup according to application and reference data, a high-resolution monitoring product could be established in the study regions providing hydrometeorological information close to present day with ERA5 reanalysis (Hersbach et al., 2020) input. For the use of WRF with seasonal forecasts SEAS5, the bias-correction of large-scale input fields (e.g., Pontoppidan et al., 2018) may provide solutions to minimize the effects of SEAS5 model drifts on dynamical downscaling. Hence, future stud-

ies should also focus on the remaining bias compared with the reference and investigate whether the expected additional post-processing with long reforecast periods is still required.

From the promising results for the semi-arid regions, the regional seasonal forecasting system with forecast horizons of up to seven months should be extended to other target regions, such as Central Europe and Germany, which have been severely affected by prolonged periods of extreme drought in recent years (e.g., Schuldt et al., 2020). In Central Europe, climate change will increasingly call for longer-term irrigation strategies for agriculture (e.g., Riediger et al., 2014). But well-founded indications of other extreme weather events to be expected, such as most recently the July 2021 flood in parts of Central Europe, are also needed. Here, the seamless forecasts from several days to weeks and months, including sub-seasonal forecasts up to six weeks in advance, issued weekly, are also essential. For this, special techniques for seamless forecast construction and validation from daily to seasonal time scales need to be developed and applied (e.g., Dirmeyer and Ford, 2020). The ability to predict not only long-term drought or above-average precipitation events at monthly scales, but also to provide indications of flooding at the weekly scale, facilitates early warnings and preventive action for a wide range of extreme weather events.

Another overarching aim for the future should be to achieve independence from currently institutional computing resources, to make the prediction system publicly and freely available in the long term and on a sustainable basis. BCSD and also other correction methods require high-performance computing environments, which are not available in many countries. The prediction system should therefore be further developed using cloud-based computing and storage power (e.g., Google Cloud, 2021; Helmholtz Federated IT Services (HIFIS), 2021), so that it can be easily transferred to other IT environments and made available to other users. This would even more allow the development of climate change mitigation strategies with seasonal forecasts in climate-sensitive regions in the future.

Despite these methodological and technical advances, we should also emphasize the improvement of quality-controlled reference data to be available, reliable and consistent over a long period of time, as the lack of such information is still one of the major limiting aspects in many parts of the world. We must therefore continue to expand regional station networks in climatically sensitive regions such as Africa and ensure their maintenance. In addition, the final transfer into practice is only possible with good and long-standing regional contacts. Therefore, close and trusting cooperation with regional partners is urgently needed in order to actually reach the long-term sustainability goals, to achieve the implementation of scientific results in practice and to further improve the acquired knowledge. It is important that we work together to combat the advancing climate change, and it is precisely such regional and practice-oriented research that makes an important contribution to this end.

# References


Abatzoglou, J. T. and Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *Int. J. Climatol.* 32, 772–780. doi:10.1002/joc.2312

AghaKouchak, A., Feldman, D., Hoerling, M., Huxman, T., and Lund, J. (2015). Water and climate: Recognize anthropogenic drought. *Nature* 524, 409–411. doi:10.1038/524409a

Ahmadalipour, A., Moradkhani, H., Castelletti, A., and Magliocca, N. (2019). Future drought risk in Africa: Integrating vulnerability, climate change, and population growth. *Sci. Total Environ.* 662, 672–686. doi:10.1016/j.scitotenv.2019.01.278

Alves, O., Balmaseda, M. A., Anderson, D., and Stockdale, T. (2004). Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Quart. J. Roy. Meteorol. Soc.* 130, 647–667. doi:10.1256/qj.03.25

Anochi, J. A., de Almeida, V. A., and de Campos Velho, H. F. (2021). Machine Learning for Climate Precipitation Prediction Modeling over South America. *Remote Sens.* 13, 2468. doi:10.3390/rs13132468

Arnell, N. W., van Vuuren, D. P., and Isaac, M. (2011). The implications of climate policy for the impacts of climate change on global water resources. *Glob. Env. Change* 21, 592–603. doi:10.1016/j.gloenvcha.2011.01.015

Ashraf, S., AghaKouchak, A., Nazemi, A., Mirchi, A., Sadegh, M., Moftakhari, H. R., et al. (2019). Compounding effects of human activities and climatic changes on surface water availability in Iran. *Climatic Change* 152, 379–391. doi:10.1007/s10584-018-2336-6

Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature* 525, 47–55. doi:10.1038/nature14956

Bell, C. J., Gray, L. J., Charlton-Perez, A. J., Joshi, M. M., and Scaife, A. A. (2009). Stratospheric Communication of El Niño Teleconnections to European Winter. *J. Climate* 22, 4083–4096. doi:10.1175/2009JCLI2717.1

Briley, L. J., Ashley, W. S., Rood, R. B., and Krmenec, A. (2017). The role of meteorological processes in the description of uncertainty for climate change decision-making. *Theor. Appl. Climatol.* 127, 643–654. doi:10.1007/s00704-015-1652-2

Brunner, M. I., Slater, L., Tallaksen, L. M., and Clark, M. (2021). Challenges in modeling and predicting floods and droughts: A review. *WIREs Water* e1520. doi:10.1002/wat2.1520

Bruno Soares, M., Daly, M., and Dessai, S. (2018). Assessing the value of seasonal climate forecasts for decision-making. *WIREs: Clim. Chang.* 9, e523. doi:10.1002/wcc.523

Bruyère, C. L., Done, J. M., Holland, G. J., and Fredrick, S. (2014). Bias corrections of global models for regional climate simulations of high-impact weather. *Clim. Dyn.* 43, 1847–1856. doi:10.1007/s00382-013-2011-6

Cohen, J., Coumou, D., Hwang, J., Mackey, L., Orenstein, P., Totz, S., et al. (2019). S2S reboot: An argument for greater inclusion of machine learning in subseasonal to seasonal forecasts. *WIREs Clim. Change* 10. doi:10.1002/wcc.567

Crochemore, L., Ramos, M.-H., Pappenberger, F., Andel, S. J. v., and Wood, A. W. (2016). An Experiment on Risk-Based Decision-Making in Water Management Using Monthly Probabilistic Forecasts. *B. Am. Meteorol. Soc.* 97, 541–551. doi:10.1175/BAMS-D-14-00270.1

Dirmeyer, P. A. and Ford, T. W. (2020). A Technique for Seamless Forecast Construction and Validation from Weather to Monthly Time Scales. *Mon. Wea. Rev.* 148, 3589–3603. doi:10.1175/MWR-D-19-0076.1

Distefano, T. and Kelly, S. (2017). Are we in deep water? Water scarcity and its limits to economic growth. *Ecol. Econom.* 142, 130–147. doi:10.1016/j.ecolecon.2017.06.019

Domínguez-Castro, F., García-Herrera, R., and Vicente-Serrano, S. M. (2018). Wet and dry extremes in Quito (Ecuador) since the 17th century. *Int. J. Climatol.* 38, 2006–2014. doi:10.1002/joc.5312

Echevin, V., Colas, F., Espinoza-Morriberon, D., Vasquez, L., Anculle, T., and Gutierrez, D. (2018). Forcings and Evolution of the 2017 Coastal El Niño Off Northern Peru and Ecuador. *Frontiers Marine Sci.* 5, 367. doi:10.3389/fmars.2018.00367

ECMWF (2007). IFS Documentation CY31R1 - Part IV: Physical Processes. In *IFS Documentation CY31R1* (ECMWF). 155. doi:10.21957/orqi6pqsa

Environment Management Group (2011). *Global Drylands: A UN system-wide response*. Tech. rep., UN Environment Programme WCMC

Feng, P., Wang, B., Liu, D. L., Ji, F., Niu, X., Ruan, H., et al. (2020). Machine learning-based integration of large-scale climate drivers can improve the forecast of seasonal rainfall probability in Australia. *Environ. Res. Lett.* 15, 084051. doi:10.1088/1748-9326/ab9e98

Fick, S. E. and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37, 4302–4315. doi:10.1002/joc.5086

Folland, C. K., Scaife, A. A., Lindesay, J., and Stephenson, D. B. (2012). How potentially predictable is northern European winter climate a season ahead? *Int. J. Climatol.* 32, 801–818. doi:10.1002/joc.2314

Frankignoul, C. and Sennéchael, N. (2007). Observed Influence of North Pacific SST Anomalies on the Atmospheric Circulation. *J. Climate* 20, 592–606. doi:10.1175/JCLI4021.1

Gampe, D., Zscheischler, J., Reichstein, M., O'Sullivan, M., Smith, W. K., Sitch, S., et al. (2021). Increasing impact of warm droughts on northern ecosystem productivity over recent decades. *Nat. Clim. Chang.* 11, 772–779. doi:10.1038/s41558-021-01112-8

Google Cloud (2021). Cloud Computing Services. *https://cloud.google.com*

Gosling, S. N. and Arnell, N. W. (2016). A global assessment of the impact of climate change on water scarcity. *Climatic Change* 134, 371–385. doi:10.1007/s10584-013-0853-x

Haerter, J. O., Hagemann, S., Moseley, C., and Piani, C. (2011). Climate model bias correction and the role of timescales. *Hydrol. Earth Syst. Sci.* 15, 1065–1079. doi:10.5194/hess-15-1065-2011

Hansen, J. W. (2002). Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges. *Agricult. Syst.* 74, 309–330. doi:10.1016/S0308-521X(02)00043-4

Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data* 7, 109. doi:10.1038/s41597-020-0453-3

Hartmann, H. C., Pagano, T. C., Sorooshian, S., and Bales, R. (2002). Confidence Builders: Evaluating Seasonal Climate Forecasts from User Perspectives. *Bull. Amer. Meteorol. Soc.* 83, 683–698. doi:10.1175/1520-0477(2002)083<0683:CBESCF>2.3.CO;2

Haupt, S. E., Chapman, W., Adams, S. V., Kirkwood, C., Hosking, J. S., Robinson, N. H., et al. (2021). Towards implementing artificial intelligence post-processing in weather and climate: proposed actions from the Oxford 2019 workshop. *Phil. Trans. R. Soc. A* 379, 20200091. doi:10.1098/rsta.2020.0091

Helmholtz Federated IT Services (HIFIS) (2021). Helmholtz Cloud Services. *https://www.helmholtz.de/forschung/information-data-science/helmholtz-federated-it-services-hifis/*

Hermanson, L., Ren, H.-L., Vellinga, M., Dunstone, N. D., Hyder, P., Ineson, S., et al. (2018). Different types of drifts in two seasonal forecast systems and their dependence on ENSO. *Clim. Dyn.* 51, 1411–1426. doi:10.1007/s00382-017-3962-9

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quart. J. Roy. Meteorol. Soc.* , qj.3803doi:10.1002/qj.3803

[Dataset] Hoogeveen, J. (2009). Global map of aridity - 10 arc minutes, FAO-UN. http://www.fao.org/geonetwork/srv/en/main.home?uuid=221072ae-2090-48a1-be6f-5a88f061431a

International Commission on Large Dams (ICOLD-CIGB) (2021). World Register of Dams - General Synthesis. *https://www.icold-cigb.org/GB/world_register/general_synthesis.asp*

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geosci. Model Dev.* 12, 1087–1117. doi:10.5194/gmd-12-1087-2019

Jolliffe, I. T. and Stephenson, D. B. (eds.) (2011). *Forecast Verification: A Practitioner's Guide in Atmospheric Science* (John Wiley & Sons, Ltd)

Klein, C., Heinzeller, D., Bliefernicht, J., and Kunstmann, H. (2015). Variability of West African monsoon patterns generated by a WRF multi-physics ensemble. *Clim. Dyn.* 45, 2733–2755. doi:10.1007/s00382-015-2505-5

Kummu, M., Ward, P. J., de Moel, H., and Varis, O. (2010). Is physical water scarcity a new phenomenon? Global assessment of water shortage over the last two millennia. *Environ. Res. Lett.* 5, 034006. doi:10.1088/1748-9326/5/3/034006

Lamb, P. J. and Peppler, R. A. (1987). North Atlantic Oscillation: Concept and an Application. *Bull. Amer. Meteorol. Soc.* 68, 1218–1225. doi:10.1175/1520-0477(1987)068<1218:NAOCAA>2.0.CO;2

Laux, P., Böker, B., Martins, E. S., Vasconcelos Junior, F., Moron, V., Portele, T., et al. (2021a). A semi-objective circulation pattern classification scheme for the semi-arid Northeast Brazil. *Int. J. Climatol.* 41, 51–72. doi:10.1002/joc.6608

Laux, P., Dieng, D., Portele, T. C., Wei, J., Shang, S., Zhang, Z., et al. (2021b). A High-Resolution Regional Climate Model Physics Ensemble for Northern Sub-Saharan Africa. *Front. Earth Sci.* 9, 792. doi:10.3389/feart.2021.700249

Lopez, A., Coughlan de Perez, E., Bazo, J., Suarez, P., van den Hurk, B., and van Aalst, M. (2018). Bridging forecast verification and humanitarian decisions: A valuation approach for setting up action-oriented early warnings. *Wea. Climate Extremes* , 100167doi:10.1016/j.wace.2018.03.006

Lopez, A. and Haines, S. (2017). Exploring the Usability of Probabilistic Weather Forecasts for Water Resources Decision-Making in the United Kingdom. *Wea. Climate Soc.* 9, 701–715. doi:10.1175/WCAS-D-16-0072.1

Lorenz, C. and Kunstmann, H. (2012). The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis. *J. Hydrometeorol.* 13, 1397–1420. doi:10.1175/JHM-D-11-088.1

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N., and Riegger, J. (2014). Large-Scale Runoff from Landmasses: A Global Assessment of the Closure of the Hydrological and Atmospheric Water Balances. *J. Hydrometeorol.* 15, 2111–2139. doi:10.1175/JHM-D-13-0157.1

[Dataset] Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H. (2020a). Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Catamayo-Chira Basin (Ecuador/Peru). World Data Center for Climate (WDCC) at DKRZ. doi:10.26050/WDCC/SaWaM_D04_SEAS5_BCSD

[Dataset] Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H. (2020b). Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Karun Basin (Iran). World Data Center for Climate (WDCC) at DKRZ. doi:10.26050/WDCC/SaWaM_D01_SEAS5_BCSD

[Dataset] Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H. (2020c). Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Rio São Francisco Basin (Brazil). World Data Center for Climate (WDCC) at DKRZ. doi:10.26050/WDCC/SaWaM_D02_SEAS5_BCSD

[Dataset] Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H. (2020d). Seasonal Water Resources Management for Semiarid Areas: Bias-corrected and spatially disaggregated seasonal forecasts for the Tekeze-Atbara and Blue Nile Basins (Ethiopia/Eritrea/Sudan). World Data Center for Climate (WDCC) at DKRZ. doi:10.26050/WDCC/SaWaM_D03_SEAS5_BCSD

Lorenz, C., Portele, T. C., Laux, P., and Kunstmann, H. (2021). Bias-corrected and spatially disaggregated seasonal forecasts: a long-term reference forecast product for the water sector in semi-arid regions. *Earth Syst. Sci. Data* 13, 2701–2722. doi:10.5194/essd-13-2701-2021

Loucks, D. P. and van Beek, E. (2017). *Water Resource Systems Planning and Management* (Cham: Springer International Publishing). doi:10.1007/978-3-319-44234-1

Luo, J.-J., Behera, S., Masumoto, Y., Sakuma, H., and Yamagata, T. (2008). Successful prediction of the consecutive IOD in 2006 and 2007. *Geophys. Res. Lett.* 35, L14S02. doi:10.1029/2007GL032793

Manzanas, R. (2020). Assessment of Model Drifts in Seasonal Forecasting: Sensitivity to Ensemble Size and Implications for Bias Correction. *J. Adv. Model. Earth Sy.* 12. doi:10.1029/2019MS001751

Marengo, J. A., Alves, L. M., Alvala, R. C., Cunha, A. P., Brito, S., and Moraes, O. L. (2018). Climatic characteristics of the 2010-2016 drought in the semiarid Northeast

Brazil region. *Anais da Academia Brasileira de Ciências* 90, 1973–1985. doi:10.1590/0001-3765201720170206

Marengo, J. A., Chou, S. C., Kay, G., Alves, L. M., Pesquero, J. F., Soares, W. R., et al. (2012). Development of regional future climate change scenarios in South America using the Eta CPTEC/HadCM3 climate change projections: climatology and regional analyses for the Amazon, São Francisco and the Paraná River basins. *Clim. Dyn.* 38, 1829–1848. doi:10.1007/s00382-011-1155-5

Martins, E. S. P. R., Coelho, C. A. S., Haarsma, R., Otto, F. E. L., King, A. D., Jan van Oldenborgh, G., et al. (2018). A Multimethod Attribution Analysis of the Prolonged Northeast Brazil Hydrometeorological Drought (2012–16). *Bull. Amer. Meteorol. Soc.* 99, S65–S69. doi:10.1175/BAMS-D-17-0102.1

Masih, I., Maskey, S., Mussá, F. E. F., and Trambauer, P. (2014). A review of droughts on the African continent: a geospatial and long-term perspective. *Hydrol. Earth Syst. Sci.* 18, 3635–3649. doi:10.5194/hess-18-3635-2014

Mekonnen, M. M. and Hoekstra, A. Y. (2016). Four billion people facing severe water scarcity. *Sci. Adv.* 2, e1500323. doi:10.1126/sciadv.1500323

Mori, P., Schwitalla, T., Ware, M. B., Warrach-Sagi, K., and Wulfmeyer, V. (2021). Downscaling of seasonal ensemble forecasts to the convection-permitting scale over the Horn of Africa using the <scp>WRF</scp> model. *Int. J. Climatol.* 41. doi:10.1002/joc.6809

Mulligan, M., van Soesbergen, A., and Sáenz, L. (2020). GOODD, a global dataset of more than 38,000 georeferenced dams. *Sci. Data* 7, 31. doi:10.1038/s41597-020-0362-5

Munia, H. A., Guillaume, J. H. A., Wada, Y., Veldkamp, T., Virkki, V., and Kummu, M. (2020). Future Transboundary Water Stress and Its Drivers Under Climate Change: A Global Study. *Earth's Future* 8. doi:10.1029/2019EF001321

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., et al. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data Discuss.* doi:10.5194/essd-2021-82

Ning, L., Riddle, E. E., and Bradley, R. S. (2015). Projected Changes in Climate Extremes over the Northeastern United States. *J. Climate* 28, 3289–3310. doi:10.1175/JCLI-D-14-00150.1

Nyaupane, N., Thakur, B., Kalra, A., and Ahmad, S. (2018). Evaluating Future Flood Scenarios Using CMIP5 Climate Projections. *Water* 10, 1866. doi:10.3390/w10121866

OECD/FAO (2016). *OECD-FAO Agricultural Outlook 2016-2025*. OECD-FAO Agricultural Outlook (Paris: OECD Publishing). doi:10.1787/agr_outlook-2016-en

Patt, A. and Gwata, C. (2002). Effective seasonal climate forecast applications: Examining constraints for subsistence farmers in Zimbabwe. *Global Environ. Change* 12, 185–195. doi:10.1016/S0959-3780(02)00013-4

Peyravi, M., Peyvandi, A. A., Khodadadi, A., and Ahmadi Marzaleh, M. (2019). Flood in the South-West of Iran in 2019; Causes, Problems, Actions and Lesson Learned. *Bull. Emerg. Trauma* 7, 199–200. doi:10.29252/beat-070219

Pontoppidan, M., Kolstad, E., Sobolowski, S., and King, M. P. (2018). Improving the Reliability and Added Value of Dynamical Downscaling via Correction of Large-Scale Errors: A Norwegian Perspective. *J. Geophys. Res. Atmos.* 123, 875–11. doi:10.1029/2018JD028372

Portele, T. C., Laux, P., Lorenz, C., Janner, A., Horna, N., Fersch, B., et al. (2021a). Ensemble-Tailored Pattern Analysis of High-Resolution Dynamically Downscaled Precipitation Fields: Example for Climate Sensitive Regions of South America. *Front. Earth Sci.* 9, 669427. doi:10.3389/feart.2021.669427

Portele, T. C., Lorenz, C., Dibrani, B., Laux, P., Bliefernicht, J., and Kunstmann, H. (2021b). Seasonal forecasts offer economic benefit for hydrological decision making in semi-arid regions. *Sci. Rep.* 11, 10581. doi:10.1038/s41598-021-89564-y

Privé, N. C. and Errico, R. M. (2013). The role of model and initial condition error in numerical weather forecasting investigated with an observing system simulation experiment. *Tellus A: Dyn. Meteorol. Ocean.* 65, 21740. doi:10.3402/tellusa.v65i0.21740

Rayner, S., Lach, D., and Ingram, H. (2005). Weather Forecasts are for Wimps: Why Water Resource Managers Do Not Use Climate Forecasts. *Climatic Change* 69, 197–227. doi:10.1007/s10584-005-3148-z

Richardson, D. S. (2000). Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteorol. Soc.* 126, 649–667. doi:10.1002/qj.49712656313

Riediger, J., Breckling, B., Nuske, R. S., and Schröder, W. (2014). Will climate change increase irrigation requirements in agriculture of Central Europe? A simulation study for Northern Germany. *Environ. Sci. Europe* 26, 18. doi:10.1186/s12302-014-0018-1

Roshan, G., Grab, S. W., and Najafi, M. S. (2020). The role of physical geographic parameters affecting past (1980–2010) and future (2020–2049) thermal stress in Iran. *Nat. Hazards* 102, 365–399. doi:10.1007/s11069-020-03930-z

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2014). The NCEP Climate Forecast System Version 2. *J. Climate* 27, 2185–2208. doi:10.1175/JCLI-D-12-00823.1

Schepen, A., Everingham, Y., and Wang, Q. J. (2020). Coupling forecast calibration and data-driven downscaling for generating reliable, high-resolution, multivariate seasonal climate forecast ensembles at multiple sites. *Int. J. Climatol.* 40, 2479–2496. doi:10.1002/joc.6346

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., et al. (2014). Multimodel assessment of water scarcity under climate change. *Proceedings of the National Academy of Sciences* 111, 3245–3250. doi:10.1073/pnas.1222460110

Scholes, R. J. (2020). The Future of Semi-Arid Regions: A Weak Fabric Unravels. *Climate* 8, 43. doi:10.3390/cli8030043

Schuldt, B., Buras, A., Arend, M., Vitasse, Y., Beierkuhnlein, C., Damm, A., et al. (2020). A first assessment of the impact of the extreme 2018 summer drought on Central European forests. *Basic Appl. Ecol.* 45, 86–103. doi:10.1016/j.baae.2020.04.003

Siegmund, J., Bliefernicht, J., Laux, P., and Kunstmann, H. (2015). Toward a seasonal precipitation prediction system for West Africa: Performance of CFSv2 and high--resolution dynamical downscaling. *J. Geophys. Res. Atmos.* 120, 7316–7339. doi:10.1002/2014JD022692

Sillmann, J., Thorarinsdottir, T., Keenlyside, N., Schaller, N., Alexander, L. V., Hegerl, G., et al. (2017). Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities. *Wea. Climate Extremes* 18, 65–74. doi:10.1016/j.wace.2017.10.003

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, M., Duda, M. G., et al. (2008). A description of the advanced research WRF version 3. *Tech. Note NCAR* , 113doi:10.5065/D68S4MVH

Takahashi, K. (2004). The atmospheric circulation associated with extreme rainfall events in Piura, Peru, during the 1997–1998 and 2002 El Niño events. *Ann. Geophys.* 22, 3917–3926. doi:10.5194/angeo-22-3917-2004

Thrasher, B., Xiong, J., Wang, W., Melton, F., Michaelis, A., and Nemani, R. (2013). Downscaled Climate Projections Suitable for Resource Management. *Eos Transact. Amer. Geophys. Union* 94, 321–323. doi:10.1002/2013EO370002

Tiedtke, M. (1989). A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. *Mon. Wea. Rev.* 117, 1779

Torres, R. R., Lapola, D. M., and Gamarra, N. L. R. (2017). Future Climate Change in the Caatinga. In *Caatinga*, eds. J. M. C. Silva, I. R. Leal, and T. M (Cham: Springer International Publishing). 383–410. doi:10.1007/978-3-319-68339-3_15

Troccoli, A. (2010). Seasonal climate forecasting. *Meteorol. Appl.* 17, 251–268. doi:10.1002/met.184

United Nations (2015). *Transforming our World: The 2030 Agenda for Sustainable Development*. Tech. rep., A/RES/70/1

Vicente-Serrano, S. M., Aguilar, E., Martínez, R., Martín-Hernández, N., Azorin-Molina, C., Sanchez-Lorenzo, A., et al. (2017). The complex influence of ENSO on droughts in Ecuador. *Clim. Dyn.* 48, 405–427. doi:10.1007/s00382-016-3082-y

Viviroli, D., Kummu, M., Meybeck, M., Kallio, M., and Wada, Y. (2020). Increasing dependence of lowland populations on mountain water resources. *Nature Sustain.* 3, 917–928. doi:10.1038/s41893-020-0559-9

von Sperling, E. (2012). Hydropower in Brazil: Overview of Positive and Negative Environmental Aspects. *Energy Procedia* 18, 110–118. doi:10.1016/j.egypro.2012.05.023

Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B. (2000). Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science* 289, 284–288. doi:10.1126/science.289.5477.284

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. *Nature* 467, 555–561. doi:10.1038/nature09440

Wheeler, K. G., Jeuland, M., Hall, J. W., Zagona, E., and Whittington, D. (2020). Understanding and managing new risks on the Nile with the Grand Ethiopian Renaissance Dam. *Nat. Commun.* 11, 5222. doi:10.1038/s41467-020-19089-x

Wilks, D. S. (2001). A skill score based on economic value for probability forecasts. *Meteorol. Appl.* 8, 209–219. doi:10.1017/S1350482701002092

Wood, A. W. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.* 107, 4429. doi:10.1029/2001JD000659

World Economic Forum (2021). *The Global Risks Report 2021, 16th Edition*. Tech. rep., ISBN: 978-2-940631-24-7

World Meteorological Organization (WMO) (2021). Global Producing Centres of Long-Range Forecasts. *https://public.wmo.int/en/programmes/global-data-processing-and-forecasting-system/global-producing-centres-of-long-range-forecasts*

Xu, L., Chen, N., Zhang, X., Chen, Z., Hu, C., and Wang, C. (2019). Improving the North American multi-model ensemble (NMME) precipitation forecasts at local areas using wavelet and machine learning. *Clim. Dyn.* 53, 601–615. doi:10.1007/s00382-018-04605-z

Yuan, X. and Liang, X.-Z. (2011). Improving cold season precipitation prediction by the nested CWRF-CFS system. *Geophys. Res. Lett.* 38, L02706. doi:10.1029/2010GL046104

Yuan, X., Wood, E. F., and Ma, Z. (2015). A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development. *WIREs Water* 2, 523–536