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, Institute of Meteorology and Climate Research, 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)

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 re-run of the latest ECMWF reanalysis ERA5. By that, we correct for model biases and drifts and improve the spatial resolution from 36 km to 0.1°. This is exemplary performed over 4 predominately semi-arid study domains across the world, which include the river basins of the Karun (Iran), the São Francisco (Brazil), the Tekeze-Atbara and Blue Nile (Sudan, Ethiopia and Eritrea), and the Catamayo-Chira (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 days 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 5 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: Tekeze-Atbara and Blue Nile Basins (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.

1 Introduction

Since the launch of seasonal hydrometeorological forecasts, it is widely agreed that sub-seasonal 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) or driving hydrological models (Thober et al., 2015), for 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 for 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. 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), and 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 seven 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 is computationally highly expensive (e.g. Manzanas et al., 2018b), empirical-statistical techniques usually require reliable reference data,

5   which is 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, 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

10   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

15   to preserve observed local rainfall characteristics 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).

20   Despite all these efforts, most studies focus on selected regions and only short periods of time. Furthermore, the corrected data is 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 and Blue Nile) as well as Iran (Karun basin). Our study regions are marked by a strong

25   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 significant changes due to the construction of the Grand Ethiopian Renaissance Dam (GERD) near the Ethiopian-Sudanese boarder, has been referred to in the worldwide media as the filling and operation of the GERD will have huge implications on the whole

30   Nile-region. This underlines the urgent need for longer-term forecasts to mitigate the impacts of climatically extreme events, to improve the regions' disaster preparedness, and to better plan the distribution of water resources from the coming season particularly in transboundary river basins.

Reanalysis data of the ERA5-Land (ECMWF, 2019) re-run of the land component of ERA5 climate reanalyses (Hersbach et al., 2018) is used as reference for applying the BCSD on raw forecasts from the seasonal forecasting system SEAS5 (Johnson

35   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 data set 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 deci-
5 sion support for the regional water management, or as a comprehensive repository for teaching the dealing with 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 above mentioned constraints of seasonal forecasts, and hence provide a significant contribution towards improving the usefulness of such infor-
10 mation and praxis-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
15 resolution of 0.1° is obtained from the ERA5-Land offline land surface re-run 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
20 days after 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).
25 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.

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 four month 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 (one month prior to and after the main 4 months) is also provided.

Basin	Area [km^2]	Annual rainfall [mm]	Temperature [K]	Rainy season	Seasonal precipitation [%]
Karun (KA, D01)	67313	640	289 (265/310)	DJFM	73 (94)
Extended São Francisco (SF, D02)	740820	858	298 (289/307)	DJFM	60 (85)
Blue Nile (BN, D03)	308197	1336	297 (287/308)	JJAS	73 (90)
Tekeze-Atbara (TA, D03)	205097	727	298 (286/310)	JJAS	83 (92)
Catamayo-Chira (CC, D04)	17761	1666	293 (287/301)	JFMA	65 (81)

2.2 SEAS5 seasonal forecasts

The fifth generation of ECMWF's seasonal forecasting system is operational since November 2017. The modelled 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
5   1981 till 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).

10   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, predominantly positive predictive skill during the first two forecast months is achieved. Recently, Gubler et al. (2019) did a comprehensive performance analysis of bias-corrected SEAS5-forecasts against homogenized station data across South America. They found that, in general, prediction skill of temperature forecasts is higher than the skill of precipitation forecasts and that particularly regions
15   which are influenced by Niño-3.4 show higher skills. Highest forecast skills 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
20   and orographic complexity which contain five semi-arid river basins: the Karun Basin (Domain D01, Iran), the Extended São Francisco Basin (D02, Brazil), the Tekeze-Atbara and Blue-Nile Basins (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. 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 with altitudes of up to 4000 m (except for the Extended São Francisco basin) and relatively high seasonal precipitation amounts (e.g., due to convective effects), while the downstream conditions are mainly

5   arid. Particularly in the Karun and Catamayo-Chira basins, there is a very strong climatic and elevation gradient from the head- to the tailwater within only a few hundred kilometers. Therefore, it is crucial to take these elevations and strong gradients into account to obtain realistic estimates of precipitation and temperature for the mountainous headwaters. 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.

10   This combination of dependency from incoming water resources and lack of observations is particularly dangerous 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. 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

15   amounts of precipitation, which then led to water shortages and even societal unrest during the coming months. Only one 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

20   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 *transbound-*

25   *ary* river basins. The headwaters of the Tekeze-Atbara and Blue Nile are located in Ethiopia and contribute, together with the Baro-Akobo (Sabot) River, about 85% of the Nile Water (Yitayew and Melesse, 2011). Both rivers cross the Ethiopian-Sudanese boarder 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 originates in Ecuador,

30   is dammed in the Poechos reservoir right after the Ecuadorian-Peruvian boarder, 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 build 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

5   The bias-correction and spatial disaggregation approach (BCSD, Wood, 2002) is a widely-used two-step technique for calibrating of, e.g., climate forecasts towards any kind of reference data. First, a quantile-mapping 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 and these anomalies are interpolated to a higher-resolution grid. Finally, the high-resolution climatology is added back to the interpolated anomalies to obtain a bias-corrected

10   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

15   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

20   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 (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

25   interpolation technique to the *full* precipitation, temperature and radiation forecasts and not, as in most other studies, to the anomalies.

– The 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 ±15 day window around the forecasted day, respectively.

30   – 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 allowing 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 spatial 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 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 exemplary the 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. Fig. 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
5 and radiation biases can reach values of up to ±5 mm/day (CC), ±2 K (KA) and ±30 W/m² (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
10 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 K 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 behaviour with maximum
15 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 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
20 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 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
25 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. Overall, the RMSE of SEAS5 BCSD is much lower compared to the raw forecasts. As an example, the RMSE of the precipitation and
30 radiation forecasts for February over the CC-basin are reduced from around 6 mm/day and 30 W/m² and to less than 4 mm/day and 20 W/m², respectively. The minimum temperature 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 forecasts for the rainy season of the SF-Basin are much lower after applying the BCSD, the RMSE remains almost

unchanged. 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 forecasts over the SF-Basin.

4.2 Topographic and resolution-dependent effects

To evaluate the impact of the improved spatial resolution, Fig. 6 shows the total accumulated precipitation during the four
5   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 precipitation pattern along the Andes in the D04-domain and
10   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.

4.3 Lead-time dependent effects

15   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. 4 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.

20   One reason for these drifts is a southward shift of the locations with higher temperatures and higher radiations with increasing lead time. This is visible in Fig. 5, which shows the July-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. 4 also show the largest deviations in the Northern part of D03.

25   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. 4). The remaining differences for precipitation, temperature and radiation between the low- and higher-lead forecasts are below 0.5 mm/day, 0.5 K and 10 W/m^2, respectively. Similarly, as depicted in Fig. 4, the lead-0 and lead-5 forecasts for precipitation, temperature, and radiation forecasts in July
30   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 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 with respect to ERA5-Land is evaluated with the *Continuous Ranked Probability Skill Score* (CRPSS 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 above mentioned 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 four 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.

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.

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 Fig. 5 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

5 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-

10 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

15 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

20 forecasts which show a negative bias of the raw forecasts (Fig. 2), high CRPSS values (Fig. 8) and reduced 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 precip-

25 itation 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

30 dry season (April). We assume that this is caused by the application of a 31-day-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

35 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 can not fill the complete 31-day-window for estimating the forecast CDF. As an example, the reference CDF for the first of January is based on the values from the 17th of December till the 16th of January, while the CDF of the January forecast for the 1st of January only uses the values from the 1st to the 16th of January. If there are strong temporal climate gradients or heteroscedasticity, e.g., during the transition from a cold to a warm period,

5 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.

For the small basins of KA and CC with large altitude differences, the added effect of spatial disaggregation within the BCSD

10 approach is 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. 6). 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

15 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.

Finally, we would 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 computational demanding and can

20 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

25 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   Conclusions

In this study we present a comprehensive dataset of bias-corrected and spatially disaggregated seasonal forecasts for four

30 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 re-run 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 till 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 days. Currently, the data covers 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
5 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 $36\,\mathrm{km}$ 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
10 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 to, e.g., global approaches. Therefore, future works have to further examine methods and approaches to account for such strong gradients in the reference climatologies.

15 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 an 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
20 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, are therefore already adopted by several weather services and other higher-level authorities in the study domains.

We are convinced that such products have the potential to significantly improve the regional water management particularly
25 in semi-arid regions, in which a careful management of available freshwater resources is crucial. Furthermore, especially in the context of global climate change with increasing risks of climatic extreme events, it is crucial that longer-term forecasts like our SEAS5 BCSD product are adopted by authorities and included in the water resources planning to ensure a sustainable regional water management and improve disaster preparedness.

7 Data availability

30 The bias-corrected and spatially disaggregated seasonal forecasts are published via the World Data Center for Climate (WDCC), which is hosted by the 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). 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
5   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)

10   – 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
15   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), 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 days (daily) and 6 months (monthly), respectively.

20   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/.
25   This system is currently under joint development with the company GAF AG (https://www.gaf.de, 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 Server. In contrast to the products available via the WDCC, the operational forecasts are only available via the THREDDS-
30   Server. These are published with a delay of about 1 day 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.

Appendix A: Empirical quantile mapping

A1   Empirical quantile mapping

We follow the classical empirical quantile mapping approach as depicted in, e.g., Wood (2002); Voisin et al. (2010). The bias-correction is performed separately for each single pixel. For each forecasted day, we select a window of $\pm$ 15 days 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 the 1st of January from the January forecast is based on 16 (days in January) $\times$ 25 (ensemble members) $\times$ 36 (years) = 14400 values. The CDF for the 31st of January from the same issue month is based on 27900 values (as we use the 15 days before and after the 31st of January). Obviously, during the first and last 15 days of a forecast, we have to use less 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) $\times$ 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 θ, the bias-correction of each forecasted value $X_{i,\theta}$ is performed through

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

This is shown exemplarily 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 exemplary

demonstrated 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.

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

$$Bias \quad = \tfrac{1}{N}\left(X_p - X_o\right) \tag{B1}$$

and

$$RMSE \quad = \sqrt{\tfrac{1}{N}\left(X_p - X_o\right)^2}, \tag{B2}$$

respectively, where X_p and X_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 is 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

5   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 $> 10\,\mathrm{mm/day}$, the CRPSS extends this to *all* possible events. As a skillscore, comparing the prediction skill of different forecasts, the CRPSS is defined as

$$CRPSS \quad = 1 - \frac{CRPS_{\mathrm{forecast}}}{CRPS_{\mathrm{reference\ forecast}}}. \tag{B3}$$

10   Here, the Continuous Ranked Probability Score ($CRPS$) for a forecasted quantity x is defined as

$$CRPS(F_p, x) \quad = \int\limits_{-\infty}^{+\infty} \left(F_p\left(x\right) - F_o\left(x\right)\right)^2 dx \tag{B4}$$

where $F_p\left(x_p\right)$ and $F_o\left(x_o\right)$ 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

15   $$F_o\left(x\right) \quad = H\left(x - x_o\right) \tag{B5}$$

where $H\left(x\right)$ 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

20   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. C.L. 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 Server and implemented the operationalization of the presented approach; C.L. and T.P. conducted the
25   evaluation of the BCSD forecasts and wrote the first draft of the paper; C.L., T.P., P.L. and H.K. 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) project. We would also like to thank the ECMWF and the Copernicus Climate Service for providing the SEAS5 seasonal forecasts as

5   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 DRKZ and WDCC teams and, particularly, Mr. Heinrich Widmann and Dr. Andrea Lammert for uploading, publishing, and supporting our product.

References

Abatzoglou, J. T. and Brown, T. J.: A comparison of statistical downscaling methods suited for wildfire applications, International Journal of Climatology, 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
5   of climate model outputs for climate change impact assessment in the U.S. northeast, Global and Planetary 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?, Hydrology and Earth System Sciences, 22, 3515–3532, https://doi.org/10.5194/hess-22-3515-2018, 2018.

10   Alidoost, F., Stein, A., Su, Z., and Sharifi, A.: Multivariate copula quantile mapping for bias correction of reanalysis air temperature data, Journal of Spatial Science, pp. 1–17, 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.

Anghileri, D., Monhart, S., Zhou, C., Bogner, K., Castelletti, A., Burlando, P., and Zappa, M.: The Value of Subseasonal Hydrometeoro-
15   logical Forecasts to Hydropower Operations: How Much Does Preprocessing Matter?, Water Resources Research, 55, 10 159–10 178, https://doi.org/10.1029/2019WR025280, 2019.

Block, P.: Tailoring seasonal climate forecasts for hydropower operations, Hydrology and Earth System Sciences, 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
20   studies, International Journal of Climatology, 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, Regional Environmental 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, Monthly Weather Review, 78, 1–3, https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2, 1950.

25   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, Theoretical and Applied Climatology, 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, Climate Dynamics, 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.:
30   Forecast verification: current status and future directions, Meteorological Applications, 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 Resources Research, 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, Hydrology and Earth System Sciences, 20, 3601–3618, https://doi.org/10.5194/hess-20-3601-2016, 2016.

35   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, International Journal of Climatology, 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, Hydrology and Earth System Sciences, 17, 2359–2373, https://doi.org/10.5194/hess-17-2359-2013, 2013.

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

Elagib, N. A. and Elhag, M. M.: Major climate indicators of ongoing drought in Sudan, Journal of Hydrology, 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?, Environmental Research Communications, 1, 031 002, https://doi.org/10.1088/2515-7620/ab114e, 2019.

10 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 and Forecasting, 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 Resources Research, 50, 7167–7186,
15 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, Bulletin of the American Meteorological Society, 83, 683–698, https://doi.org/10.1175/1520-
20 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, Climate Dynamics, 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 and Forecasting, 15,
25 559–570, 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:
30 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, 2019.

Hwang, S. and Graham, W. D.: Development and comparative evaluation of a stochastic analog method to downscale daily GCM precipita-
35 tion, Hydrology and Earth System Sciences, 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, Geoscientific Model Development, 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, Climate Dynamics, 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, Theoretical and Applied Climatology, 118, 211–221, https://doi.org/10.1007/s00704-013-1055-1, 2014.

5 Lafon, T., Dadson, S., Buys, G., and Prudhomme, C.: Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods, International Journal of Climatology, 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, Hydrological Processes, 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

10 from Northeast Brazil, Climate 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, Journal of Hydrometeorology, 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, Journal of Hydrometeorology, 15, 2111–2139,

15 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

20 spatially disaggregated seasonal forecasts for the Catamayo-Chira Basin (Ecuador/Peru), 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), World Data Center for Climate (WDCC) at DKRZ, https://doi.org/10.26050/WDCC/SaWaM_D01_SEAS5_BCSD, 2020b.

25 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), 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), World Data Center for

30 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, Climate Dynamics, 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?, Journal of Geophysical Research: Atmospheres, 124, 9423–9441, https://doi.org/10.1029/2019JD031155, 2019.

35 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?, Climate Dynamics, 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, Climate Dynamics, 53, 1287–1305, https://doi.org/10.1007/s00382-019-04640-4, 2019.

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), Bulletin of the American Meteorological Society, 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, Hydrology and Earth System Sciences, 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, Journal of 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 Environmental 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, Proceedings of the National Academy of Sciences, 102, 12 623–12 628, 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), Journal of Applied Meteorology and Climatology, 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, Climatic 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, Australian Journal of Agricultural and Resource Economics, 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 sub-seasonal to seasonal rainfall forecasts from global climate models and evaluation for 12 Australian catchments, Hydrology and Earth System Sciences, 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, Journal of Geophysical Research: Atmospheres, 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, International Journal of Geophysics, 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?, Hydrology and Earth System Sciences, 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), Journal of Hydrometeorology, 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, Hydrology and Earth System Sciences, 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.

Tryhorn, L. and DeGaetano, A.: A comparison of techniques for downscaling extreme precipitation over the Northeastern United States, International Journal of Climatology, 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 Energy, 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, Scientific Reports, 9, 1464, https://doi.org/10.1038/s41598-018-38071-8, 2019.

Vandal, T., Kodra, E., and Ganguly, A. R.: Intercomparison of machine learning methods for statistical downscaling: the case of daily and extreme precipitation, Theoretical and Applied Climatology, 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 and Forecasting, 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, Bulletin of the American Meteorological Society, 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, Journal of Geophysical Research: Solid Earth, 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, Journal of Hydrometeorology, 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 Resources Research, 54, 9224–9254, https://doi.org/10.1029/2017WR022149, 2018.

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, Meteorological Applications, 24, 315–325, https://doi.org/10.1002/met.1654, 2017.

Wood, Andrew, M. E. K. A. L. D. P.: Long-range experimental hydrologic forecasting for the eastern United States, Journal of Geophysical Research, 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, Journal of 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.

5   Yuan, X., Wood, E. F., Luo, L., and Pan, M.: A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction, Geophysical Research Letters, 38, L13 402, 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?, Journal of Climate, 30, 3185–3196, https://doi.org/10.1175/JCLI-D-16-0652.1, 2017.
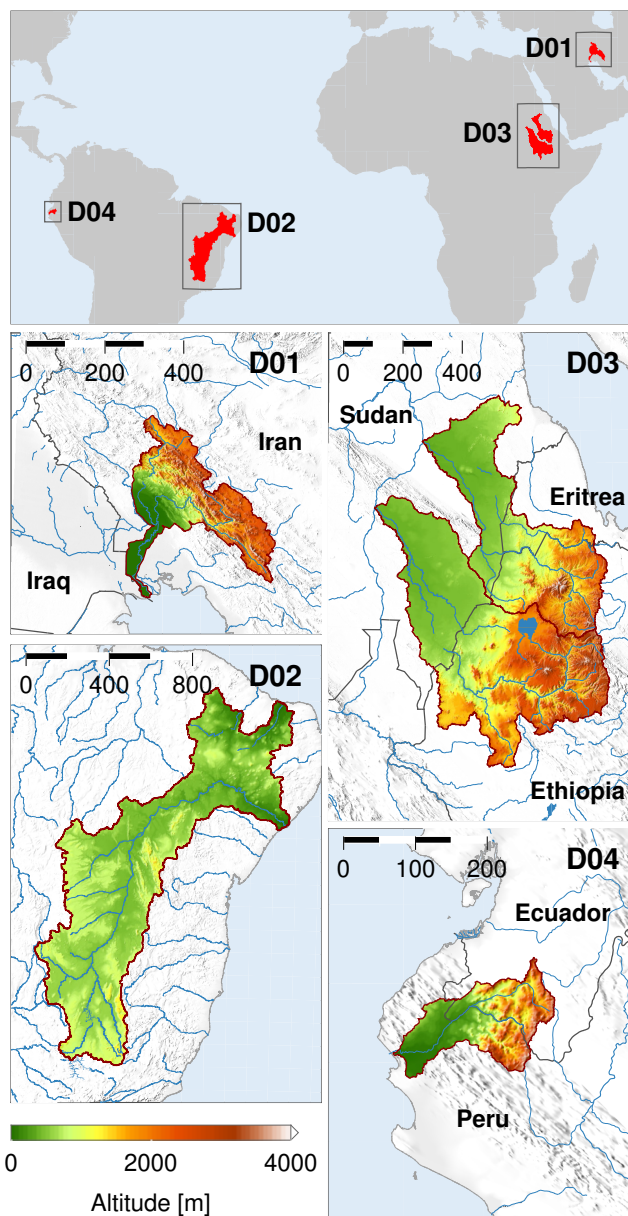
Figure 1. Overview of the four study domains: the Karun Basin (KA, Iran, D01, top left), the Extended São Francisco Basin (SF, Brazil, D02, bottom left), the Tekeze-Atbara (to the North) and Blue Nile (to the South) Basins (TA and BN, Sudan, Ethiopia and Eritrea, D03, top right), and the Catamayo-Chira 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.
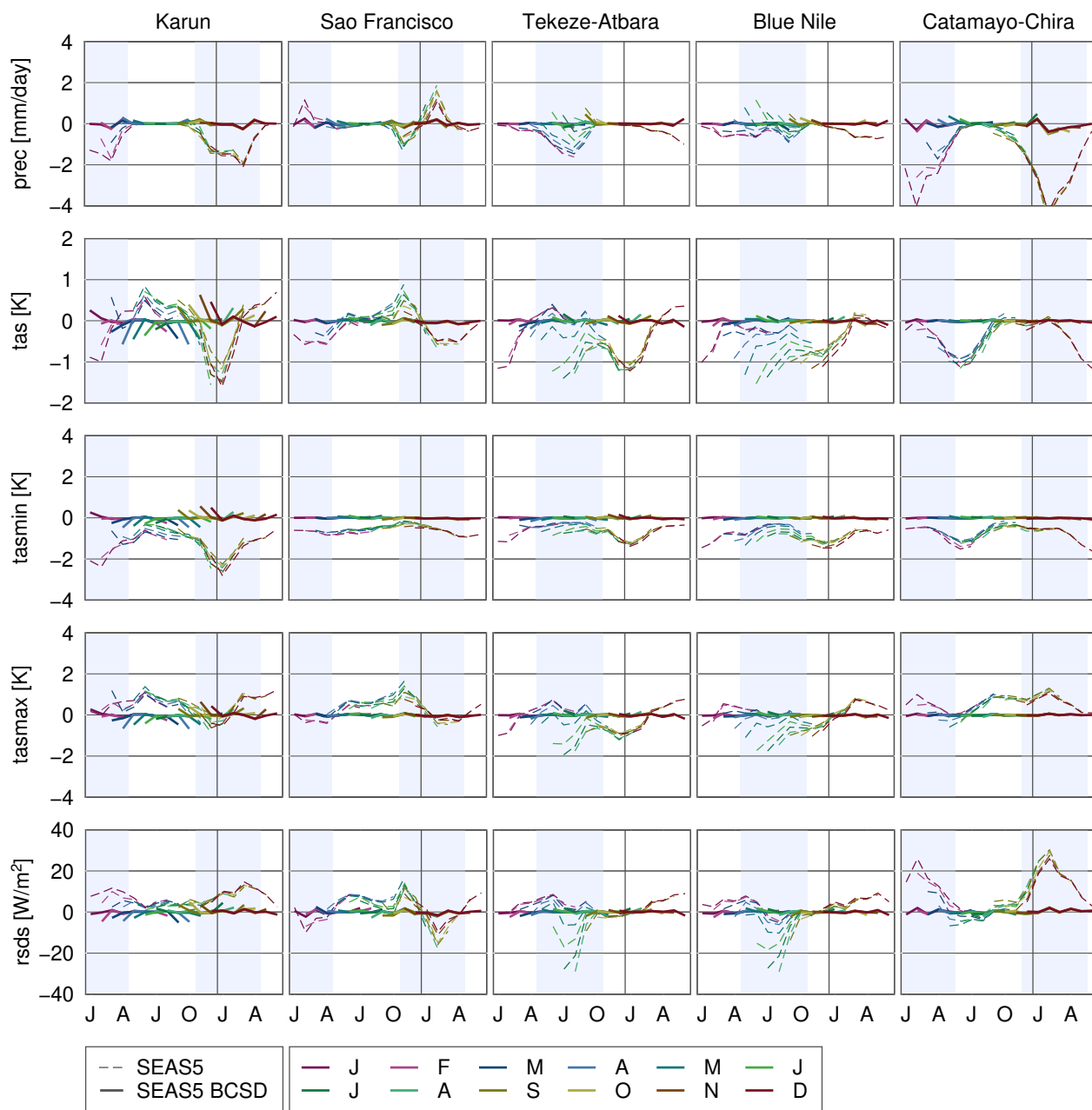
Figure 2. Bias of monthly (from top to bttom) precipitation (pr), average temperature (tas), minimum temperature (tasmin), maximum temperature (tasmax) and radiation (rsds) forecasts from SEAS5 (dashed lines) and SEAS5 BCSD (solid lines) with respect to ERA5 Land during the period 1981 to 2016 for all five study basins. The colors of the lines indicate the twelve 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 till December. The blue-shaded area depicts the six months of the rainy season.
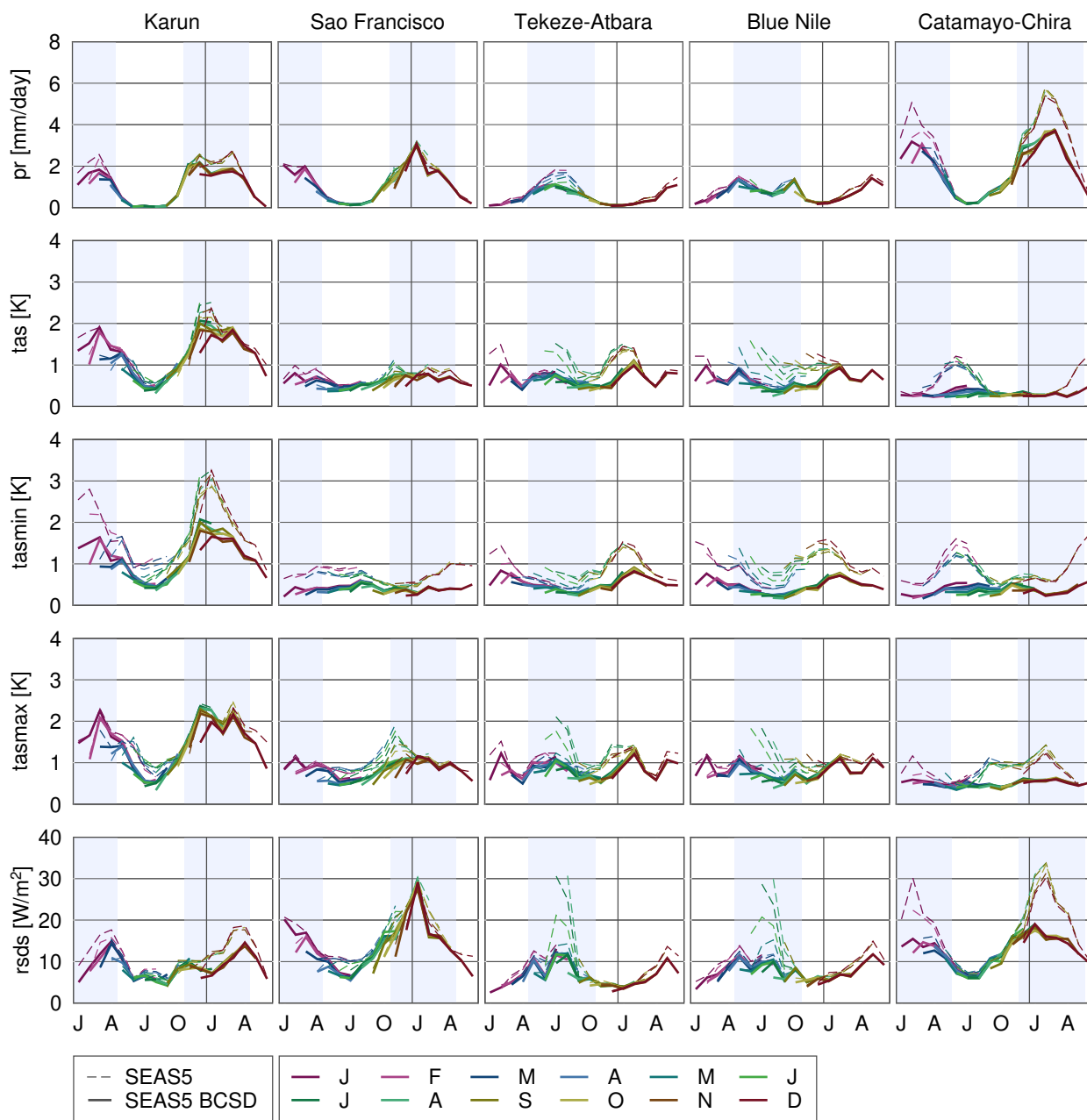
Figure 3. Same as Fig. 2 but for the RMSE of SEAS5 and SEAS5 BCSD compared to ERA5-Land.
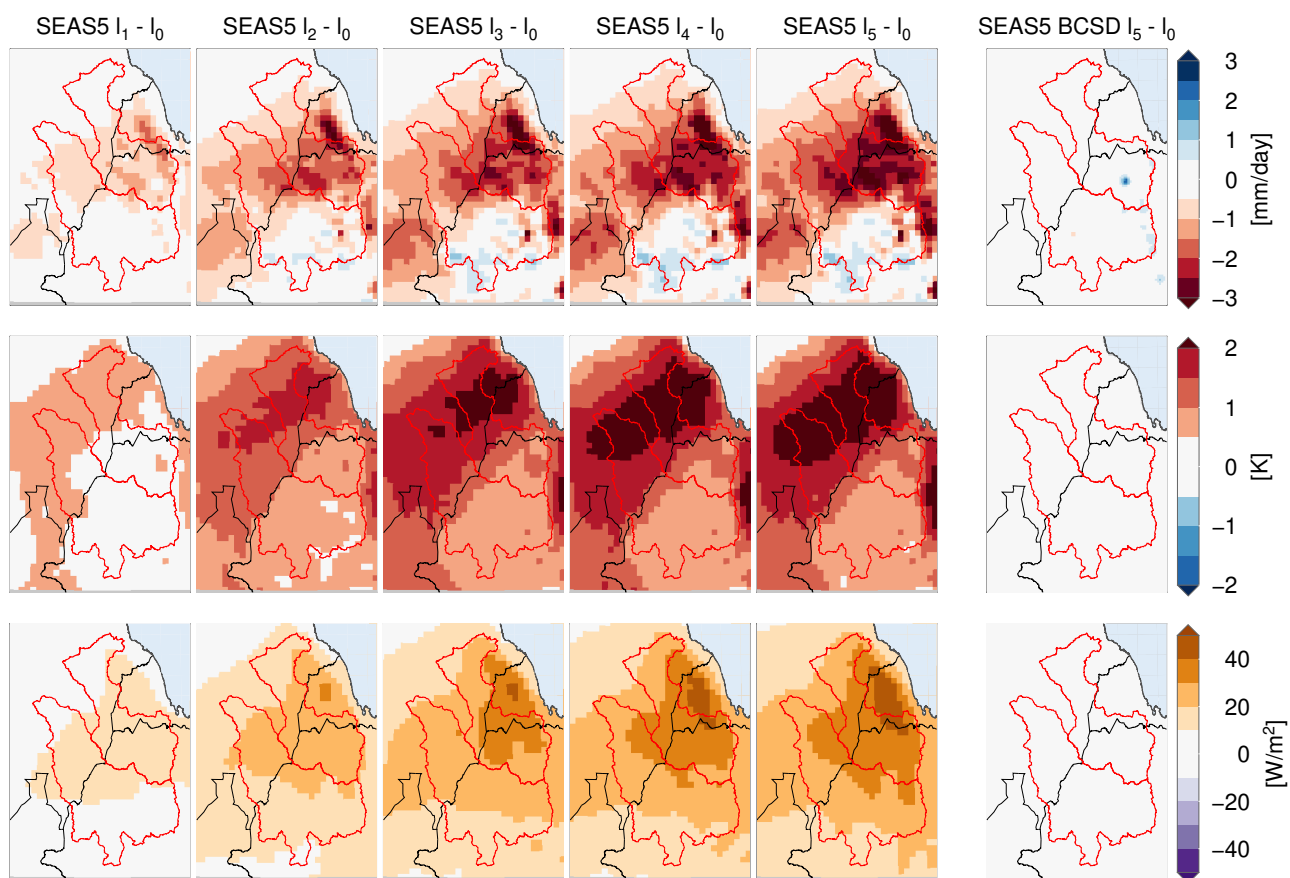
Figure 4. Differences between the July-forecasts from different issue months for precipitation (top row), temperature (middle row) and radiation (bottom row) for the African domain D03.
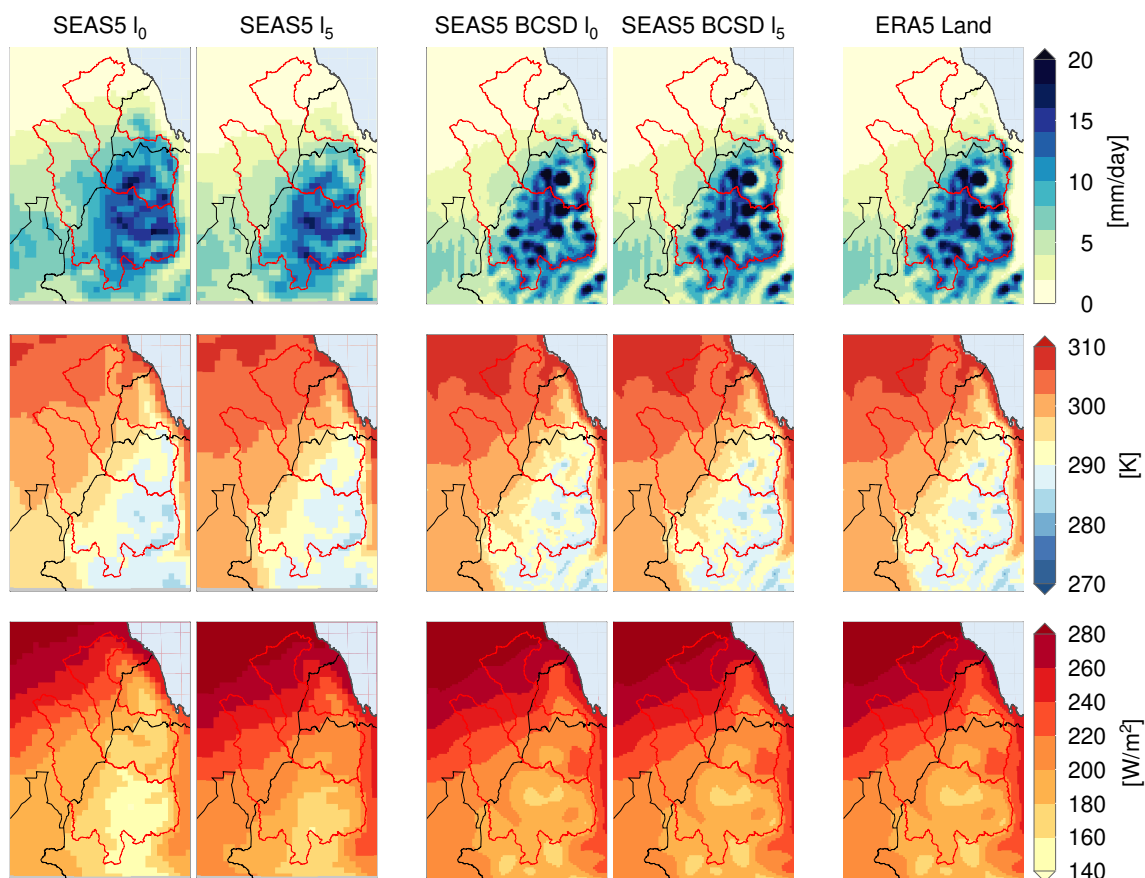
Figure 5. Average July predictions from SEAS5 lead 0 and 5 (first two columns), SEAS5 BCSD lead 0 and 5 (center columns) and reference values from ERA5-Land (right column) for precipitation (top row), average temperature (middle row) and radiation (bottom row) for the African domain D03.
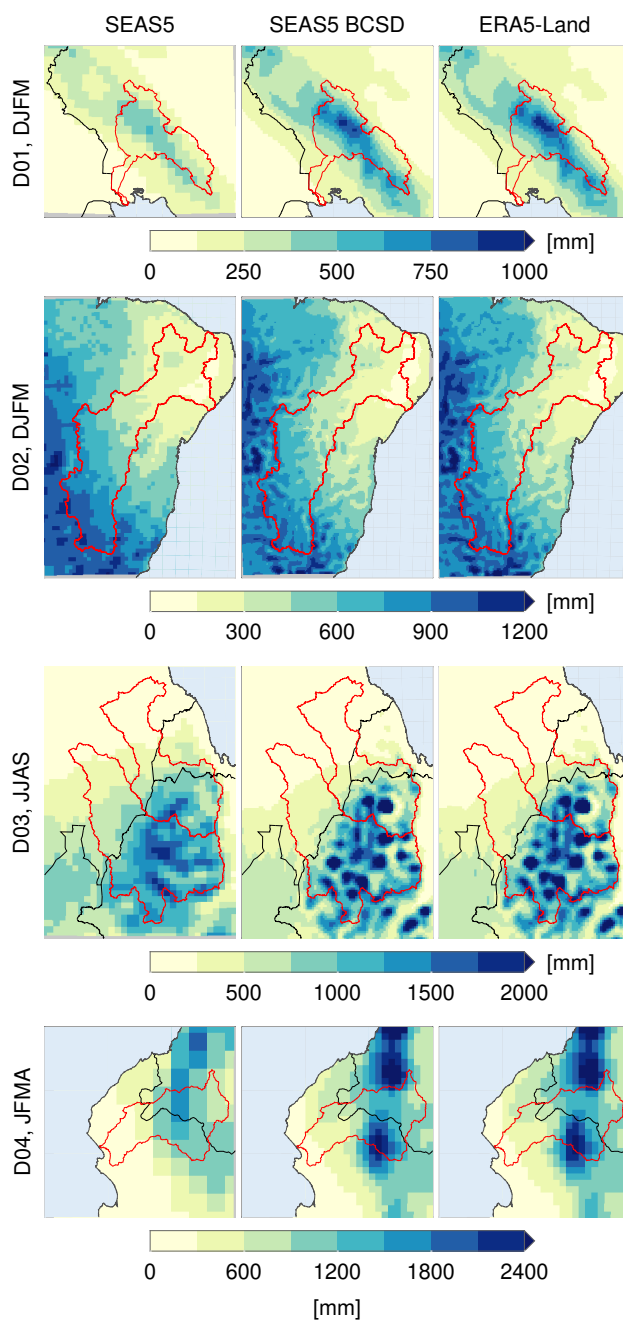
Figure 6. Total precipitation during the four main months of the rainy season for the (from top to bottom) TABN, SF, KA, and CC-domain from raw (left) SEAS5 and BCSD (right) lead-0 forecasts, averaged over the period 1981 to 2016.
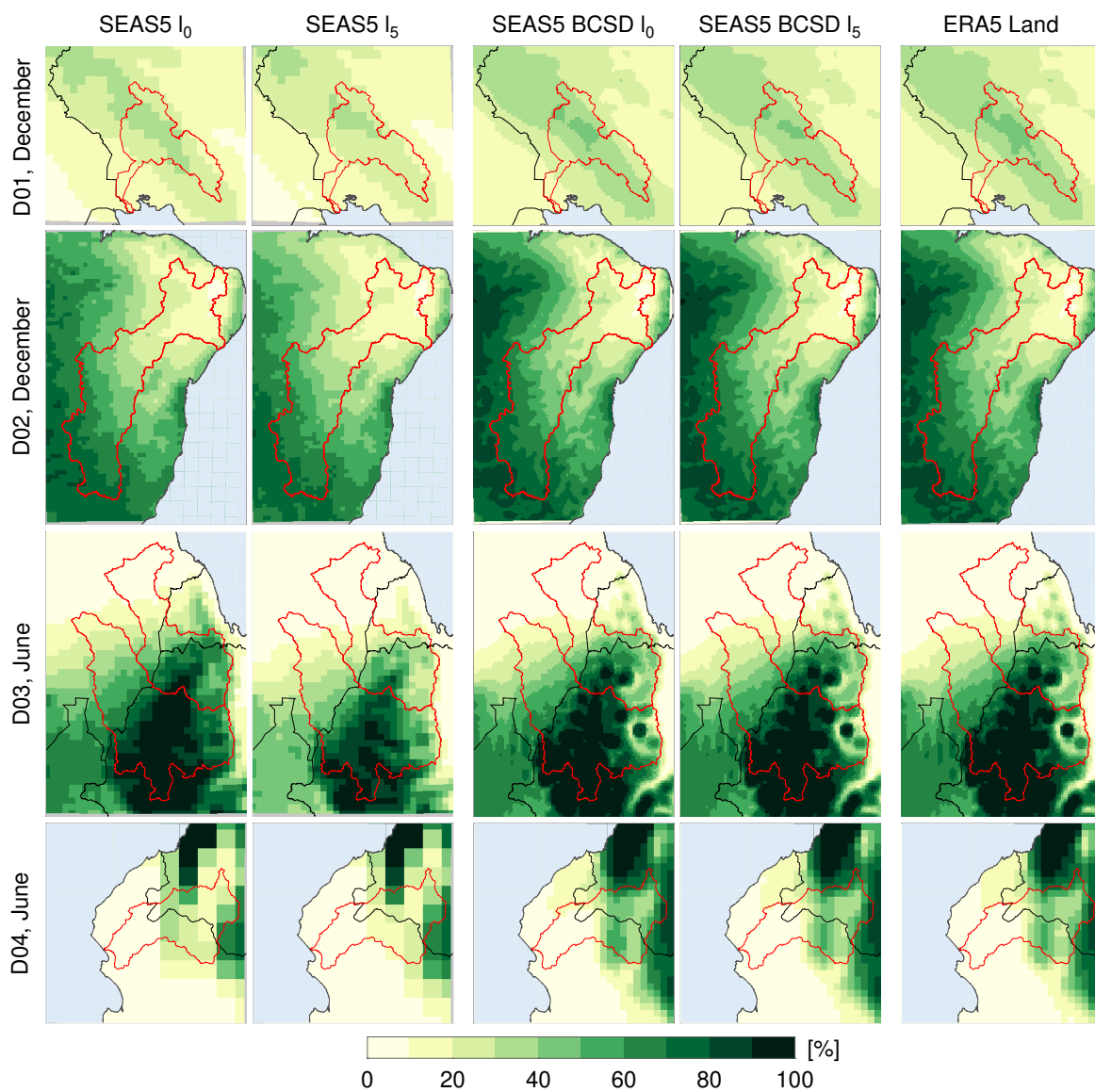
Figure 7. Wet day (> 1 mm/day) probability for the four study domains for a single month (D01: December, D02: December, D03: June, D04: May) from SEAS5, SEAS5 BCSD for both lead-0 and lead-5, respectively, and the reference ERA5-Land.
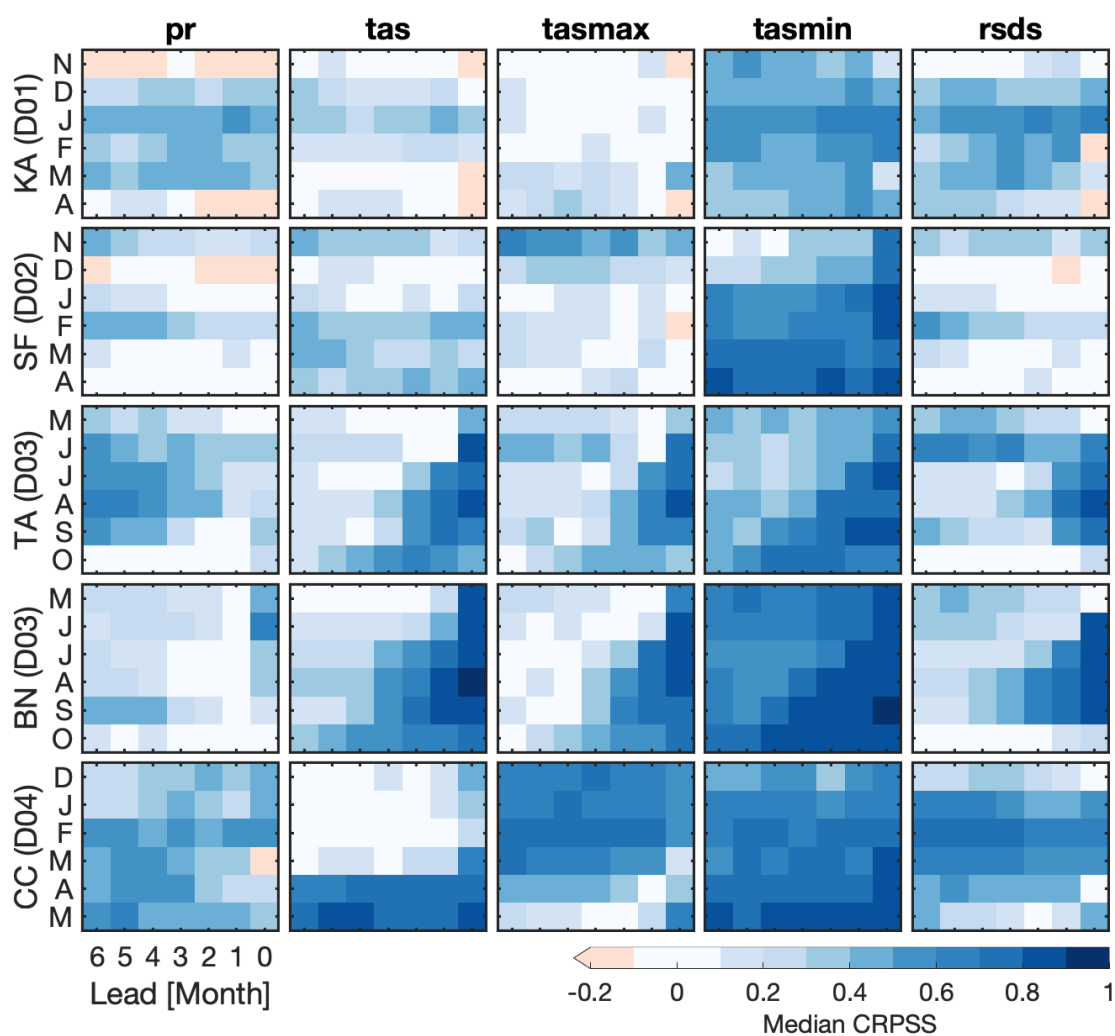
Figure 8. Median Continuous Ranked Probability Skill Score (CRPSS) of area averages over the five basins (from top to bottom) of SEAS5-BCSD against raw SEAS5 forecasts with respect to the reference ERA5-Land. The CRPSS values are derived for precipitation (pr), mean (tas), maximum (tasmax), and minimum (tasmin) temperature and shortwave radiation (rsds) as monthly medians for each of the six 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.
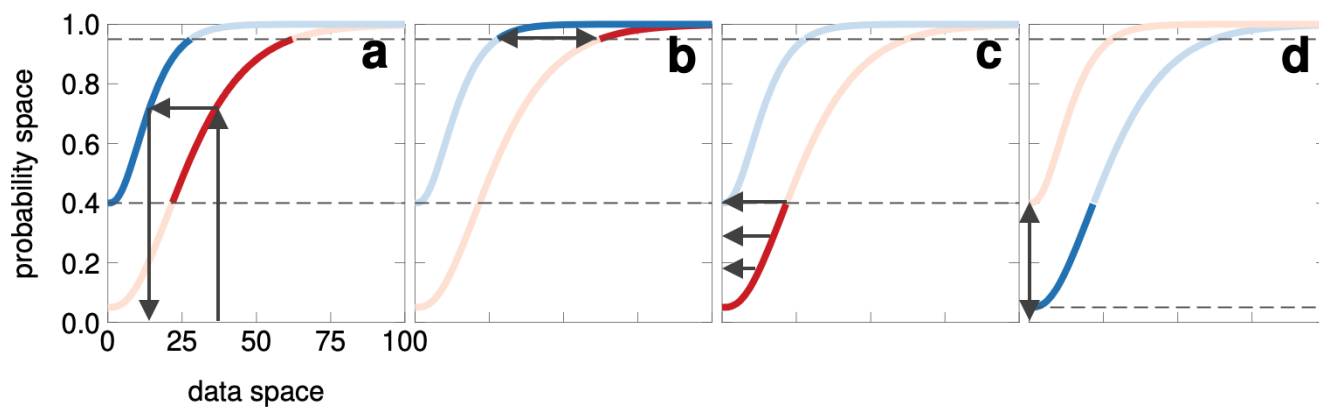
Figure A1. a) Empirical quantile mapping between model-based (red) and reference (blue) data; b) delta-approach for correcting extreme values above the maximum Weibull quantiles; c) correction of precipitation intermittency when the dry-day probability of the reference (lower dashed line) is higher; d) correction of precipitation intermittency when the dry-day probability of the reference is lower