Seasonal Drought Prediction for Semiarid Northeast Brazil: Verification of Six Hydro-Meteorological Forecast Products

José Miguel Delgado[1], Sebastian Voss[1], Gerd Bürger[1], Klaus Vormoor[1], Aline Murawski[2], Marcelo Rodrigues[3], Eduardo Martins[3], Francisco Vasconcelos Júnior[3], and Till Francke[1]

[1]Institute of Earth and Environmental Sciences, University of Potsdam, Germany
[2]German Research Centre of Geosciences GFZ Potsdam, Germany
[3]Research Institute for Meteorology and Water Resources - FUNCEME, Fortaleza, Brazil

Correspondence to: José Miguel Delgado (martinsd@uni-potsdam.de)

Abstract. A set of seasonal drought forecast models was assessed and verified for the Jaguaribe River in semiarid northeast Brazil. Meteorological seasonal forecasts were provided by the operational forecasting system used at FUNCEME (Ceará's research foundation for meteorology) and by the European Centre for Medium-Range Weather Forecasts (ECMWF). Three downscaling approaches were tested and combined with the models in hindcast mode for the period 1981 to 2014. The forecast

5   issue time was January and the forecast period was January to June. Hydrological drought indices were obtained by fitting a generalized linear model to observations. In short, it was possible to obtain forecasts for *a)* monthly precipitation, *b)* meteorological drought indices, and *c)* hydrological drought indices.

    The skill of the forecasting systems was evaluated with regard to root mean square error (RMSE) and the relative operating characteristic (ROC) skill score. Forecasts of monthly precipitation had little or no skill considering RMSE. Still, the forecast

10   of extreme events of low monthly precipitation showed skill for the rainy season (ROC skill score of 0.24 to 0.33). A similar picture was seen when forecasting meteorological drought indices: low skill regarding RMSE and significant skill when forecasting drought events of e.g. $SPEI_{01}$ (ROC skill score of 0.53 to 0.61). Similar results were obtained for low regional reservoir storage forecasts. Regarding the skill in the forecasted months, it was greater for April, when compared to February and March (the remaining months of the rainy season).

15   This work showed that a multimodel ensemble can forecast drought events of time scales relevant to water managers in northeast Brazil. But no or little skill could be found in the forecasts of the whole range of monthly precipitation or drought indices (e.g. forecasting average years). Both this work and those here revisited showed that major steps forward are needed in forecasting the rainy season in northeast Brazil.

1 Introduction

20   Northeastern Brazil has historically been the epicenter of major drought events. Fioreze et al. (2012) identified 100 severe droughts since the 16th century in this region, while Marengo et al. (2016) identified 68 major events for the same period. Within this region, the state of Ceará has been in the frontline of the fight against this natural hazard. This has been both due to the impacts suffered in the past and to the measures taken to improve its resilience.

Droughts in Ceará reflect a meteorological anomaly over the tropical Atlantic Ocean. Dry years are generally related to a positive sea surface temperature (SST) anomaly on the tropical North Atlantic, associated with a negative anomaly on the tropical South Atlantic and over the equator. This forces a northward shift of the intertropical convergence zone, taking the rainbelt to northern latitudes. The causes for this anomaly are linked to the occurrence of the El Niño Southern Oscillation and

5   to the North Atlantic Oscillation (Hastenrath, 2012).

Past famines and mass migrations triggered large investments in infrastructure in recent decades. These investments brought hundreds of strategic reservoirs and thousands of small dams to a semi-arid landscape, which are being managed according to a transparent water allocation process (Formiga-Johnsson and Kemper, 2005). In order to support water allocation and management, the state runs a seasonal drought forecasting system and issues annual quantitative and qualitative forecasts of

10 the magnitude of the rainy season. These predictions can support decisions ranging from agricultural management (choice of crop, planning of seeding time) to water distribution and reservoir operation.

Currently, the forecasting system in Ceará is based on the general circulation model ECHAM4.6 (Roeckner et al., 1992). It runs from January to August on persisted SSTs (observed SSTs which are assumed invariant), covering each year's rainy season (February to April). The forecasts produced by this model are generally downscaled with the NCEP regional spectral

15 model (Juang et al., 1997), in order to resolve the spatial variability of Ceará. Verification and the current forecast can be retrieved under http://www3.funceme.br/previsao-climatica/.

In this study we intend to evaluate and extend this prediction system by employing

1. an additional underlying GCM,

2. a statistical approach based on the classification of weather patterns,

20 3. empirical-statistical downscaling methods to increase the spatial resolution and temporal fidelity of the predictions, and

4. drought indices as powerful integrative descriptors for the description of drought severity.

By these means, we aim to address the following questions:

What skill do the seasonal *meteorological drought* forecasts have?

While the term *meteorological drought* focuses on the atmospheric forcing causing water shortage, its effective implications

25 for society are more specifically accounted for by the term *hydrological drought* (Araújo and Bronstert, 2016). Since the aim of the prediction system is to support water management, we sharpen the previous question in this regard:

Can we forecast *hydrological* droughts in Ceará based on these seasonal meteorological forecasts?
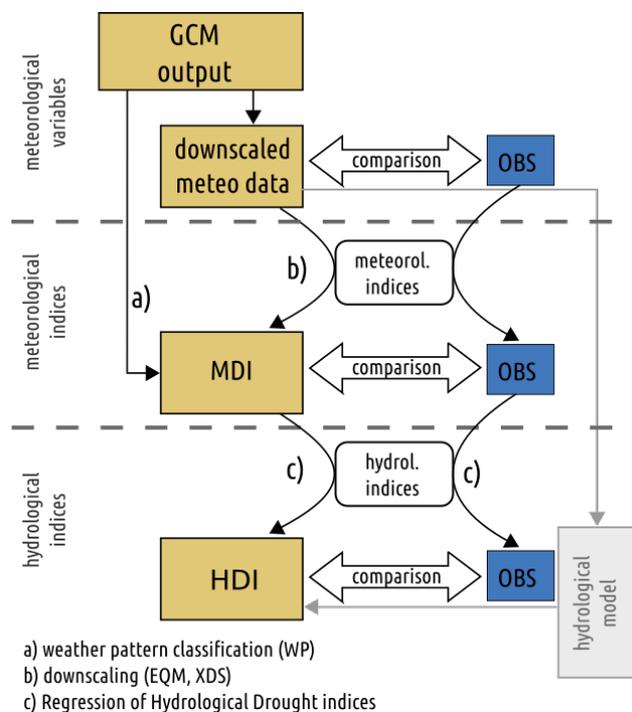
Figure 1. Flowchart explaining the methodology used for predicting meteorological data, meteorological drought indices (MDI) and hydrological drought indices (HDI).

2 Methods

2.1 General approach

This work employed a cascade of models and algorithms ranging from two general circulation models (one atmospheric and one coupled) at the top to hydrological indices at the bottom (Fig. 1). Each step involved different types of target variables being

5   forecasted: The *meteorological forecasts* (Fig. 1, top) refer to meteorological variables ("meteo data") from GCM-forecasts and the subsequent downscaling and bias correction to match the spatial and temporal resolution. The *meteorological indices* (same figure, centre), refer to the indices that were used to describe the magnitude of the forecasted meteorological drought. Finally, the *hydrological indices* (same figure, bottom) were calculated based on meteorological indices in an attempt to infer the magnitude of a hydrological drought characterized by meteorological and hydrological properties. To allow for the comparison

10   with observations, we use results of GCM hindcast, i.e. a model that has been run with data only known until the specified time in the past. As these are supposed to represent and technically resemble true forecasts, they are referred to as "forecasts" henceforward. All results and computations after the statistical downscaling have a monthly time step. Similarly, all results and computations here presented were aggregated to selected subbasins (Fig. 2).
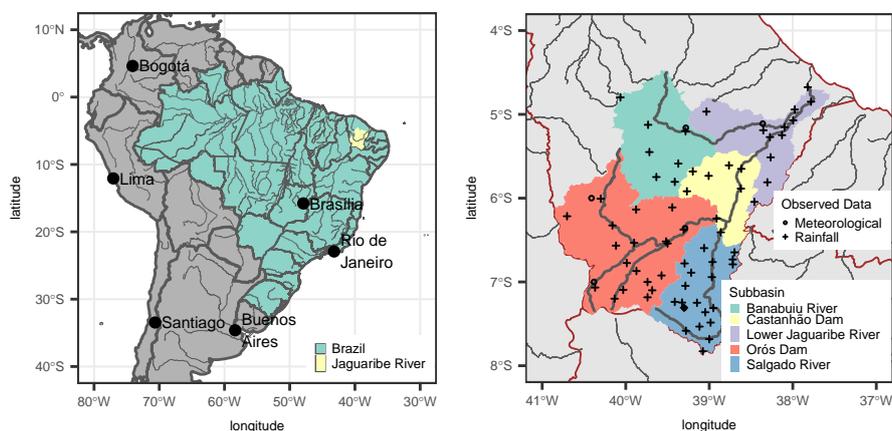
Figure 2. Left panel shows the location of the Jaguaribe River basin in South America. Right panel shows the Jaguaribe River together with its main tributaries, division into subcatchments used in this work, and meteorological and rainfall observations stations.

2.2 Study area

The spatial domain chosen for this analysis is the Jaguaribe river basin. Due to the river's regional importance, a lot has been written about its hydrology and development (see e.g. Araújo and Bronstert, 2016; Araújo, 1990). The Jaguaribe is the most important river in Ceará. Its catchment has an area of $70000 \mathrm{km}^2$ and is home to about 2.7 million people (IPECE, 2017). Annual precipitation amounts to $755 \mathrm{mm}$, of which about 90% fall in the months January to June. Average potential evapotranspiration is estimated to $2100 \mathrm{mm}$. Due to its dominant geology composed of a crystalline complex, aquifers in the region are unproductive. Runoff is practically the only source of drinking water for people and animals as well as irrigation. To that end, most of the water is stored in thousands of reservoirs of all scales across the watershed.

The main tributaries are the Banabuiú river in central Ceará and the Salgado river in southern Ceará. We aggregated the results of this research into five subcatchments: the aforementioned tributaries Banabuiú and Salgado, the upper (upstream of Orós Reservoir), middle (upstream of Castanhão Reservoir) and lower (downstream of Castanhão Reservoir) Jaguaribe. An overview of the state and location of these catchments and tributaries is given in Fig. 2.

2.3 Seasonal Forecast Models ("GCM output")

To address the first research question we employed different combinations of dynamical and statistical models and a weather pattern classification methodology to produce meteorological drought indices. The dynamical seasonal forecast models were provided by FUNCEME and ECMWF in the form of hindcasts for the period 1981 to 2014. Details like resolution, reference and short description are given in Table 1.

The ECMWF operational seasonal forecasting system S4 has 51 ensemble members and six months lead time. It is fully coupled to an ocean circulation model. The system has been systematically verified (Vitart, 2013; Molteni et al., 2011; Richardson

et al., 2012). The hindcast version of the system has the same specifications of the operational model but only 15 ensemble members.

The seasonal forecasting system used by Ceará's meteorological agency are based on the general circulation model ECHAM4.6. Details on this model can be found in Table 1. The operational and hindcast version have 20 ensemble members and are run

5 on initial sea surface temperatures (SSTs) persisted during the forecasting period (8 months). The forecast model is initialized in January with the conditions of the atmosphere modeled by an AMIP-type run for December 31st. The AMIP-type run starts in 1961 and is forced by monthly observed SSTs (NOAA Optimum Interpolation SST V2). This means that only modeled and no observed atmospheric conditions are used to initialize the forecast model.

2.4 Downscaling of GCM output

10 In order to predict precipitation over particular locations it is necessary to downscale the GCM forecasts. Three statistical downscaling approaches were employed: expanded downscaling (XDS), empirical quantile mapping (EQM) and weather pattern classification (WP, see Table 1 for details and references). To differenciate between two fundamentally different downscaling approaches, weather pattern classification will not be refered to as downscaling approach/method throughout the text.

The downscaling approaches used here yielded a full set of meteorological variables distributed across the catchment at

15 points where observations were available (daily mean temperature, relative humidity, wind speed and daily total precipitation and radiation). The forecasting products obtained from the combinations of GCM and downscaling will be named after their components: *XDS:ECHAM*, *XDS:ECMWF*, *EQM:ECHAM*, *EQM:ECMWF*, *WP:ECHAM*, and *WP:ECMWF*.

Weather patterns were classified using the SANDRA methodology described in Philipp et al. (2016). The selection of the *optimal* classification was done visually in respect to the explained variation of the observed meteorological drought indices.

20 The classification itself was independent of the MDIs, so that no artificial skill was to be expected from forecasting the stations.

2.5 Drought Quantification using Drought Indices

Meteorological droughts were quantified in magnitude and temporal scale using meteorological drought indices (MDI). After careful appraisal regarding data demand and current conventions, the following indices were selected: SPI_{01}, SPI_{12}, SPI_{36} (Svoboda et al., 2012; Mckee et al., 1993), $SPEI_{01}$, $SPEI_{12}$ and $SPEI_{36}$ (Vicente-Serrano et al., 2009). The numbers in the

25 index name (e.g. SPI_{01}) refer to the temporal scale in months for which the index was computed.

The forecast always takes place in January for the period from January until June. Indices obtained by downscaling forecasts with temporal scale greater than the lead time of the forecast will include values from the observation set. SPI_{12}, for example, will contain at least six months of measured precipitation and a maximum of six months forecasted precipitation. The skill of a SPI_{12} forecast is therefore expected to be greater than the skill of a SPI_{01} forecast beforehand. This feature does not apply

30 to WP classification.

Although time scales greater than 6 months are in strict terms of no value for the verification of the forecasting system, they put the system into perspective. Droughts are long, creeping phenomena that must be quantified on large temporal scales. A

Table 1. Output variables of each prediction model used in this paper

Model/Method	Short description	Reference	Spatial Scale
FUNCEME Seasonal Forecast System	A 20 member ECHAM4.6 ensemble. Atmospheric circulation model, initial SSTs persisted for 6 months. Initial conditions of the atmosphere modeled by AMIP-type run (starting in 1961). AMIP run is forced by monthly observed SSTs (NOAA Optimum Interpolation SST V2).	Roeckner et al. (1992)	approx. 2.8 degree longitude/latitude
ECMWF Seasonal Forecast System	A multimodel 15 member ensemble including ocean circulation. Initial conditions coming from ERA Interim.	e.g. Stockdale et al. (1998)	approx. 0.7 degree latitude/longitude
expanded downscaling	Simulates local events consistent with prevailing atmospheric circulation while preserving local covariability	Bürger (1996)	network of monitoring stations
empirical quantile mapping	Improves systematic biases throughout the statistical distribution by mapping the empirical cumulative distributions of the observed and modelled variable	e.g. Wetterhall et al. (2012)	network of monitoring stations
weather pattern classification	Including pre-selection of variables, variable combinations and spatial domain.	e.g. Murawski et al. (2016); Philipp et al. (2007)	network of monitoring stations

forecasting system should at least be able to improve the ability to predict well the onset of a drought given the previous months rainfall deficit.

Regarding hydrological droughts, various hydrological drought indices (HDI) were reviewed and three were considered suitable for this work. All other indices reviewed either *a)* require consumptive data for water use, which is impractical for

5   the given settings or *b)* focus on streamflow, which misses the most important features (ephemeral rivers, role of reservoirs) of the hydrological system of Ceará and many other semi-arid regions. The indices chosen were the surface water supply index (SWSI) as formulated in Doesken et al. (1991) with a weight of 0.5 for precipitation within the reservoir catchment and 0.5 for reservoir volume; the regional reservoir volume divided by the maximum regional reservoir volume; and the monthly variation of the regional reservoir volume. Drought events were here considered to be the periods that fall below the 30th percentile of

10   each index.

2.6 Regression of Hydrological Drought indices

Forecasts of hydrological drought indices were obtained by searching and fitting a generalized linear model (GML) to observations of hydrological drought indices, using meteorological drought indeces as predictors. Model parsimony was enforced by predictor selection comprising a heuristic search for the best Akaike information criterion (AIC) under the constraint of
5   checking the predictors for multi-collinearity.

2.7 Forecast Verification

At each level of Fig. 1, a verification of the forecast was performed. The first metric used to calculate forecast skill for the hindcast period was the root mean square error (RMSE). RMSE was computed for each member, ensemble mean and climatology. Climatology was considered the reference forecast (Wilks, 2005). The mean square error was computed for
10   monthly values in the forecast period (1981-2014, January-June) and averaged over the entire period. The root square of this measure is the RMSE. It shows the capability of the model to correctly forecast monthly precipitation, but it does not quantify the skill to predict particular events of water scarcity. January to June precipitation represents over 90% of the annual precipitation in he Jaguaribe basin.

The other metric employed was the relative operating characteristic (ROC) skill score. The relative operating characteristic
15   (Wilks, 2005) was applied to each forecasted month with a threshold of -1 for defining an event. The respective hit rate and false alarm rate (two concepts underlying the ROC diagram) refer to the above-mentioned event. The threshold -1 captures dry spells of moderate to extreme magnitude. For precipitation a threshold based on the 30th percentile of the series of observed monthly precipitation was used. The threshold for defining HDI drought events was based on the 30th percentile of the series of observed monthly HDI.
20   ROC skill score was calculated as

$$ROCSS = 2 \cdot AUC - 1 \tag{1}$$

as in Wilks (2005), where AUC is the area under the ROC curve. The ROCSS can have values between -1 and 1, where anything below zero means *no skill*. A ROCSS of 0 corresponds to the skill of a referrence random forecast.

3 Results and Discussion

25   ## 3.1 Forecasting precipitation

The RMSE of the precipitation forecast are presented in Fig. 3. ECMWF ranks better than ECHAM, while EQM:ECMWF results in the lowest RMSEs and XDS:ECHAM in the greatest. Still, the best results in terms of RMSE are comparable to the climatology, meaning that there is limited skill in forecasting monthly precipitation. The spatial distribution of RMSE of the ensemble mean in April shows a concentration of high RMSEs in the lower Jaguaribe catchment for EQM and in the Salgado
30   catchment for XDS.
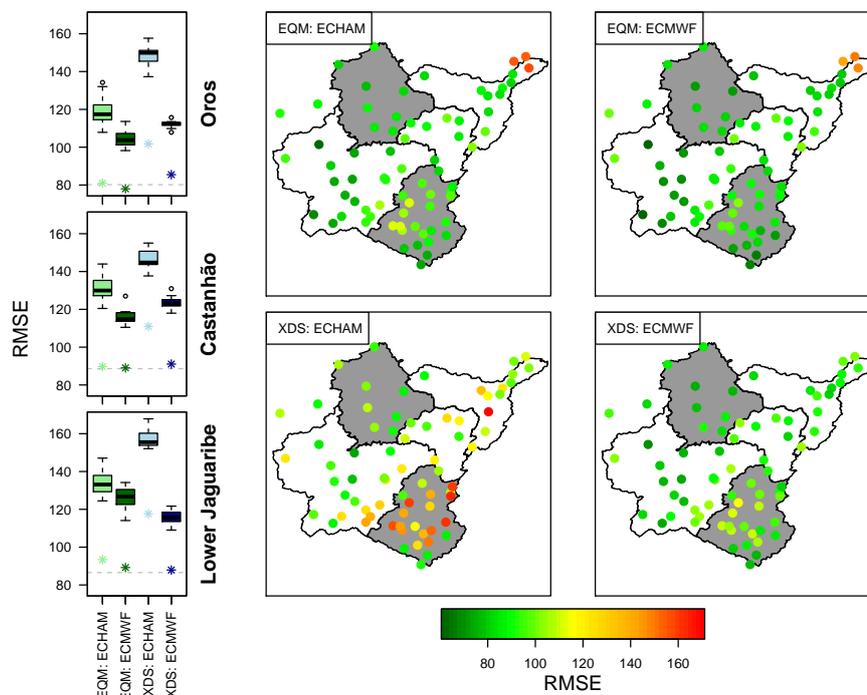
Figure 3. Root mean square error of the forecast of monthly precipitation. On the left panels, boxplots show the spread of the RMSE of each model. The asterisks (*) shows the RMSE of the ensemble mean. The RMSE of using climatology as a forecasting product is given by the grey dashed line. The four panels on the right show RMSE for each individual station for each model. Note that in general the ensemble mean ranks better than the best of the ensemble members.

The ensemble mean of the forecast, shown by the asterisks in Fig. 3 as well as in other figures below, always displayed a lower RMSE than *any of the ensemble members*. This happens because the ensemble mean "smoothes out unpredictable detail and presents the more predictable elements of the forecast" ((WMO) World Meteorological Organization, 2012). Despite its usefulness, the ensemble mean is not entirely appropriate for predicting drought events. Ensemble means do not provide any
5   information on the probability of an extreme event.

Other than RMSE, which does not provide any information on the skill of extreme event forecasts, the ROCSS is explicitly suited for that purpose, as shown in Fig. 4. The ROC curve is built based on the hit rate and false alarm rate of occurrence of a predefined event. The variation of the ROCSS over time can be attributed to lead time (skill decreasing with increasing lead time) and to low or no precipitation in the months before and after the rainy season. Months of typically low precipita-
10   tion showed poor ROCSS (Fig. 4: January, May, June). When comparing downscaling techniques and GCMs, EQM mostly outperformed XDS, while the skill was less affected by using different GCMs.

To put our results into context, we could find three reports with a statement of verification concerning precipitation forecast in Ceará. Castro et al. (2013) presents a RMSE of between 120 and 130 mm for the *Sertão Interior de Inhamuns*, using an
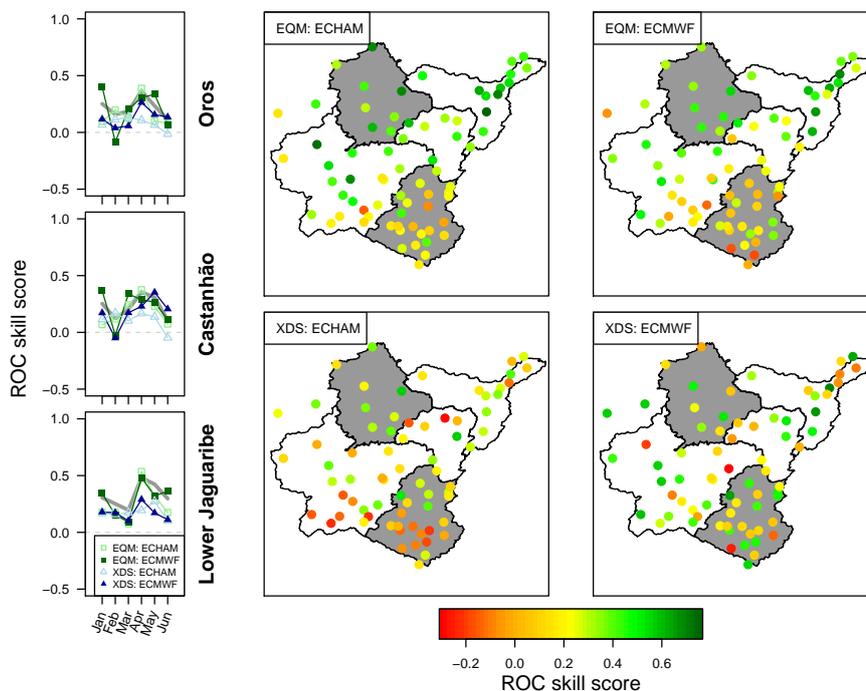
Figure 4. ROCSS of the forecast of monthly low precipitation event. On the left side the ROCSS is shown for each model/downscaling combination and for the forecasting months averaged over the respective subcatchments. The four panels on the right show ROCSS averaged over all forecasting months for each station.
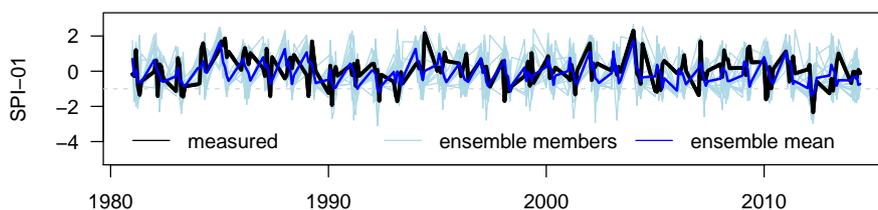
Figure 5. Time series presentation of the seasonal forecast of SPI_{01} in the Castanhão subcatchment given by ECMWF:EQM. Only periods from January to June are shown. the threshold "moderate drought event" is given by the grey dashed line.

empirical model with forecasts issued in January for the period February to June. Moura and Hastenrath (2004), with a forecast issued in end of February for the period of March to June, i.e. with shorter lead-time than our work, shows a RMSE of 50 to 70 mm (Hastenrath and Greischar (1993) obtained similar results).
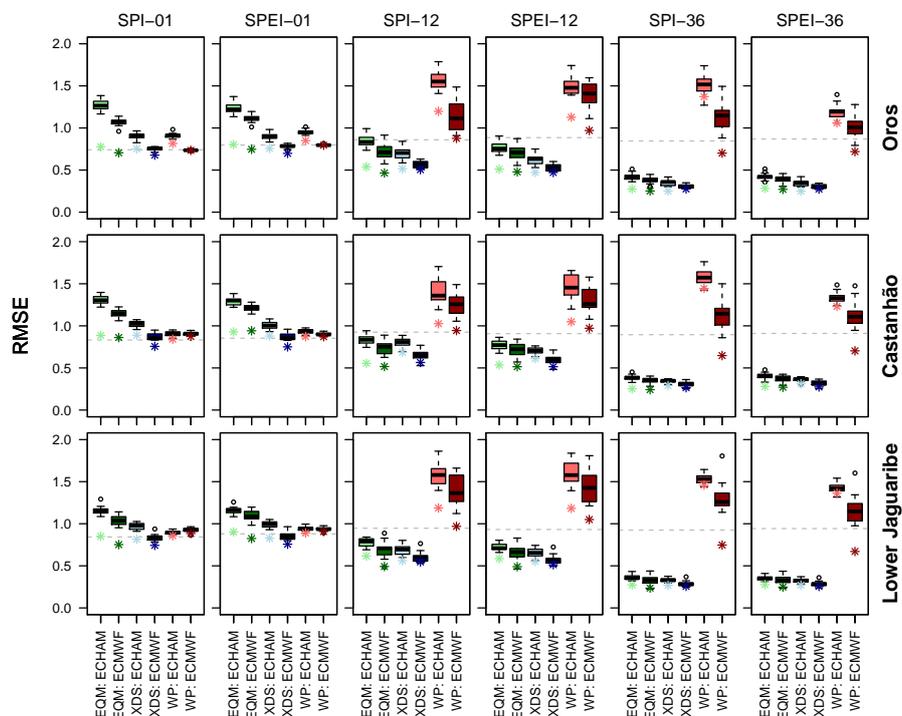
Figure 6. Boxplots of the root mean square error of forecasted meteorological drought index. Asterisk (*) shows the RMSE of the ensemble mean and box plots show the spread of the individual members. Note that in general the ensemble mean ranks better than the best of the ensemble members.

3.2 Forecasting meteorological drought indices

A time series of seasonal MDI forecasts was plotted to illustrate the forecast spread given by model EQM:ECMWF (Fig. 5). The improvement provided by the ensemble mean, when compared to each member, is clearly visible. Also visible are several observed events of moderate to severe drought (below the dashed grey line). The ensemble mean is able to forecast at least a

5   few of these events. The balance between hit rate and false alarm rate can be seen in the form of ROCSS in Fig. 7 below.

The RMSE of MDI forecasts is shown in Fig. 6. With the exception of the predictions produced by the WP approach for SPI_{01}, the general ranking of the approaches is quite consistent among the three subbasins. As with precipitation, the RMSE of SPI_{01} and $SPEI_{01}$ generally does not differ from that of the climatology and is greatest for ECHAM and EQM. EQM:ECMWF and XDS:ECMWF show consistently lowest RMSE and XDS:ECMWF performs better than the climatology.

10   For longer time scales, i.e. to the right of SPI_{12} in Fig. 6, the WP approach shows notably poorer performance than the other methods. A reason could be the way MDIs of longer time-scales are computed: when using GCM-downscaling approaches, observed data has to be used for computing indices of longer time scales. Still, here all models except WP classification show an improvement regarding climatology.
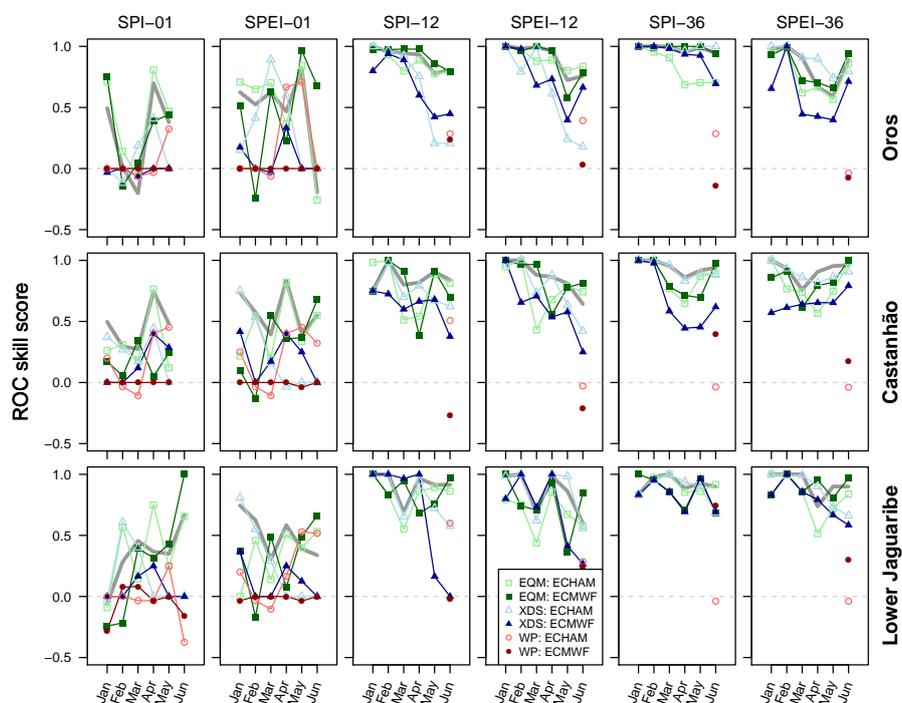
Figure 7. Boxplots of the ROCSS of forecasted meteorological drought event based on an event of index lower than -1. The grey line shows the result of the multimodel ensemble mean.

RMSE reflects the prediction skill for the whole range of the indices, including wet spells and dry spells/droughts. When aiming primarily at forecasting drought events, this verification may be misleading. Nevertheless, this metric shows which models are most appropriate for this domain and confirms the plausibility of the forecasting system also for wet years.

The ROCSS for the different months of the forecasting period shows a slightly different picture than the RMSE previously presented. Fig. 7 shows ROCSS for time scales of 1, 12 and 36 months in three regions of the Jaguaribe river. For a time scale of one month, there is no clear pattern concerning the relationship between lead time and skill for none of the forecasting models. MDIs with time scales of 12 and 36 months show a clear decrease in skill with increasing lead time.

Contrary to the results for RMSE, EQM:ECHAM show comparably good ROCSS in forecasting $SPI_{01}/SPEI_{01}$ drought events for January to May in all three regions. Still, the comparably low skill of the March forecast is problematic, March being the month of greatest precipitation in most of the catchment. April features a maximum ROCSS for EQM:ECHAM within the main months of the rainy season.

The good results of the $SPI_{36}/SPEI_{36}$ forecasts are not surprising, given that by definition they are computed with monthly precipitation from the previous 36 months, where most of the contributing months are coming from past observations and not within the forecast period. More interesting are the results for $SPI_{12}/SPEI_{12}$, in particular for June. Here, six wet months contributing to the index are forecasted and the remaining six dry months (July to December of the previous year) come from

Table 2. Regression used for predicting HDI from MDI. Only the predictant *Reservoir Volume* is shown. *Reservoir Volume* stands for regional reservoir volume as a percentage of regional storage capacity.

Region	Response	Method	Formula	R^2
Orós	*Reservoir Volume*	GLM, 2nd order	$68.41 - 26.76\text{SPEI}_{36} + 53.49\text{SPI}_{36} - 29.34\text{SPEI}_{36}\text{SPI}_{12} - 9.98\text{SPEI}_{36}\text{SPI}_{36} + 37.50\text{SPI}_{36}\text{SPI}_{12}$	0.64
Castanhão	*Reservoir Volume*	GLM, 2nd order	$58.88 + 12.61 * \text{SPEI}_{36} + 3.41 * \text{SPI}_{12}/\text{SPEI}_{12} - 3.03\text{SPI}_{01}\text{SPI}_{12}/\text{SPEI}_{36} + 7.68\text{SPEI}_{12}\text{SPI}_{36}/\text{SPEI}_{36} + 0.85\text{SPI}_{12}\text{SPI}_{36}/\text{SPEI}_{36}^2$	0.39
Lower Jaguaribe	*Reservoir Volume*	GLM, 2nd order	$31.43 + 16.31\text{SPEI}_{36} - 4.07\text{SPI}_{01}^2/\text{SPEI}_{01}$	0.49

observations. Since the variability of precipitation in July to December is very low (mostly close to zero), this index measures well the skill of the forecasting system concerning long term droughts. Here, a greater differenciation is visible, in particular EQM:ECHAM and EQM:ECMWF standing out. WP:ECHAM also compares relatively well to other models.

5   The multi-model ensemble skill, shown by the gray line is generally close to the upper envelope formed by that of the individual models. For $SPEI_{01}$ in the months January to May (rainy season) the skill of the multi-model ensemble is always positive and oscillates around 0.5. An interesting result is the improvement in skill when SPI_{01} is replaced by $SPEI_{01}$. The gray line, which shows the ROCSS for the multi-model ensemble, indicates an increase in skill during the rainy season.

In both RMSE and ROCSS, WP classification ranks comparably well for time scales of one month ($SPI_{01}/SPEI_{01}$), but poorly for longer time scales. There are two methodological reasons for that. First, the variables for each WP classification

10   are selected from a pool of candidate variables according to their skill in predicting SPEI or SPI. The selected variables are therefore different among different time scales of SPI/SPEI. In other words, the SPI/SPEI with longer time scales are not predicted simply as an accumulation of previous SPI/SPEI states, but as a fundamentally distinct WP classification. The second reason is that the EQM/XDS downscaled forecasts of SPI/SPEI of longer time scales (12 and 36 months) are based in part on observed values, since they include data belonging to the period before the forecast issue time.

15   A similar forecast assessment has been reported by e.g. Dutra et al. (2013). Events were defined by a SPI_{03} lower than -1, with a lead time of 3 months. ROCSS obtained were in the order of 0.6 for the Blue Nile, which is comparable with the results presented in this paper, but much lower for other rivers e.g. Zambezi.

3.3   Forecasting hydrological drought indices

The GLM equations obtained and their respective R^2 are shown in Table 2. Long scale MDIs (like SPI_{12} or SPI_{36}) prevail as

20   predictors of HDI. This reflects the time scale of reservoir storage variations. At a given moment in time, the reservoir storage reflects several months of inflow. Similarly, the effect of a month of high inflow in the reservoir storage level is likely to be only residual, which is given by the comparably low factors associated to SPI_{01} and $SPEI_{01}$ in the Jaguaribe region Table 2.

Hydrology and
Earth System
Sciences
Discussions
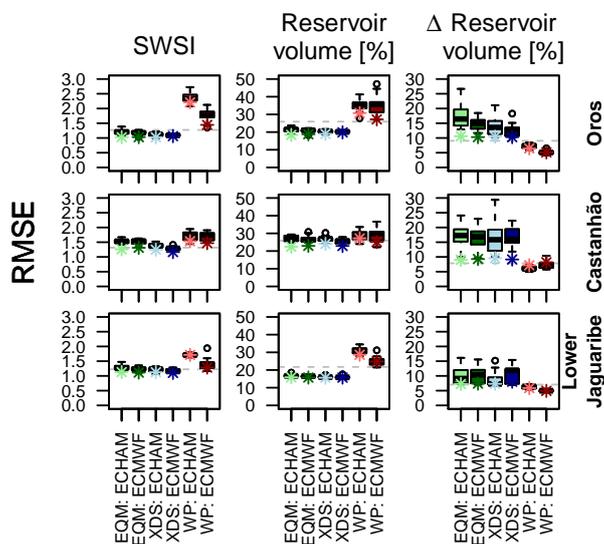
Figure 8. Root mean square error of the forecast of SWSI, regional reservoir volume and regional reservoir volume month-to-month variation. The forecast period is January to June. Three regions are presented: Lower Jaguaribe, Orós and Castanhão. The horizontal grey dashed line shows the RMSE of the climatology.
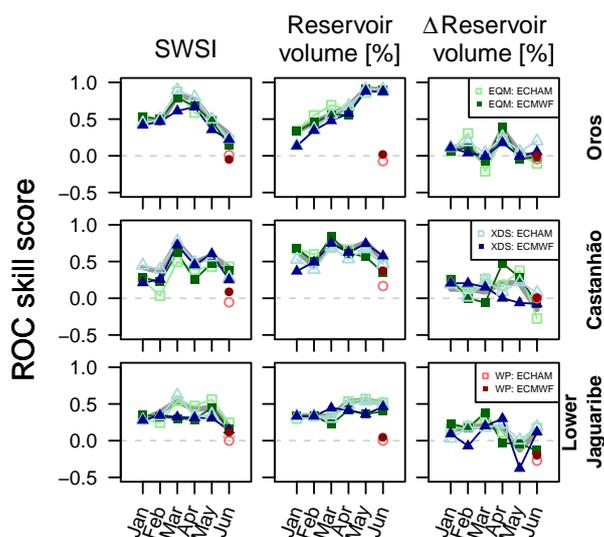
Figure 9. ROCSS of the forecast of SWSI, regional reservoir volume and regional reservoir volume month-to-month variation. The forecast months are January to June. Three regions are presented: Lower Jaguaribe, Orós and Castanhão.

The forecast of the three HDIs shows notable differences between downscaling techniques EQM/XDS and the WP classification (Fig. 8). EQM/XDS have more ROC skill and lower RMSE. Notably, WP classification has lower RMSE when forecasting

Table 3. ROCSS of January-May multimodel ensemble forecast. The ensemble includes ECHAM and the ECMWF seasonal forecast model, as well as the EQM and XDS downscaling techniques. The ROCSS are averaged over each region. Columns show different indices used for the forecast: P is seasonal precipitation, SPI_{01} and $SPEI_{01}$ are standardized precipitation indices with scale 1 month, and *Reservoir Volume* stands for regional reservoir volume in percentage of regional storage capacity. The ROC refers to an event described in Sect. 2

region	ROCSS					
	P	SPI_{01}	$SPEI_{01}$	$SPEI_{12}$	*Reservoir Volume*	Δ *Reservoir Volume*
Orós	0.24	0.28	0.61	0.93	0.59	0.11
Castanhão	0.25	0.46	0.59	0.91	0.65	0.15
Lower Jaguaribe	0.33	0.28	0.53	0.90	0.42	0.11

reservoir volume change, which reflects the results found in Sect. 3.2, i.e. greater skill in forecasting SPI_{01} and $SPEI_{01}$ than longer time scales.

Forecast quality is only weakly affected by different GCMs or downscaling techniques. SWSI and reservoir volume can be forecasted with less error using the downscaling approach. The WP approach apparently ranks better when forecasting HDI with high variation in short timescales as the monthly variation of the regional reservoir volume. This especially applies to Orós and lower Jaguaribe catchments. In the Castanhão catchment the RMSE is similar for all approaches.

Again, WP classification considers by design only a range of discrete MDIs, which can affect RMSE. MDIs were limited to nine values, of which -0.75, 0 and 0.75 are the closest to zero. The continuous values of MDI derived by downscaling are problematic, because the multilinear regression also considers division by the meteorological drought index. When the MDI are close to zero, outliers arise and skew the RMSE.

This latter issue affecting the RMSE of the hydrological drought index is not felt by the ROCSS, since it considers only the distinction of an event between *dry* and *not dry*. The ROCSS shows small differences between GCMs or downscaling methods (Fig. 9). Dry conditions can be most skillfully forecasted for the SWSI and the reservoir volume. The skill is mostly constant during the season and differs between the regions mainly for the reservoir volume with very high skill for May and June in Orós. Except for the reservoir volume in Orós, WP classification is least skillful in forecasting droughts.

We could not find reports on streamflow/reservoir forecasting systems for the region of Ceará stating ROCSS, RMSE or other verification measure. Still, for other semi-arid regions of the world, similar skill values could be found in the literature. Trambauer et al. (2015) forecasted events of standardized runoff index of 6 months lower than -0.5 with variable lead times. Their best catchment point to a ROCSS of 0.7 with a lead time of 5 months. Seibert et al. (2017) forecasted events with a standardized streamflow index (Vicente-Serrano et al., 2012) below -0.5, reporting ROCSS of 0 at the outlet of a large river (the Limpopo in southern Africa) to close to 1 in its headwater catchments.

3.4 Multimodel ensemble forecast

Finally, we present the skill score of the multimodel ensemble forecast in Table 3. Each type of index considered (precipitation, meteorological drought index and hydrological drought index) is presented. Results of the WP classification were excluded from the multimodel ensemble, because they contributed very negatively to the overall skill.

5   The forecasts of low precipitation events (given in column P), as well as the forecasts of drought defined by the SPI_{01} show comparably low skill. Forecasts of $SPEI_{01}$ and reservoir storage show greatest skill. This points to a possible bias in the forecasting that is compensated by introducing temperature in the equation of SPEI. Still it is possible to forecast events of low relative reservoir storage with similar skill as for $SPEI_{01}$.

    The good skill of the reservoir storage forecast is likely related to the long memory of the reservoir system. The forecasted

10 precipitation will affect the reservoir only marginally, since most of its storage is accumulated throughout several years. In fact, the multilinear regression used for predicting reservoir storage employs as predictors $SPEI_{12}$ as well as other MDIs of shorter time scales. In this sense, the reservoir storage forecast has less skill than would be expected, knowing that the skill of $SPEI_{12}$ was very high.

    Since the $SPEI_{12}$ forecasts are calculated based on a combination of past observed data and the forecasted months, it is

15 clear that the reservoir storage forecast is not a good indicator for streamflow. Reservoir inflow depends more on short time scales and contrary to reservoir storage is not affected by interannual memory. Note the comparably lower skill of the reservoir storage variation (between 0.11 and 0.15), a variable which is expected to closely reflect reservoir inflow.

    Improvements in the skill of hydrological drought forecasting can still be obtained. Both reservoir storage's and reservoir inflow's forecast skill showed considerable differences to the skills of the corresponding meteorological forecasts. The multi-

20 linear regression employed for estimating hydrological drought is likely the reason for such a drop in forecast skill. A way of circumventing this multilinear regression would be to add a process-based hydrological model to the forecasting system.

4 Conclusions

The plausibility and skill of a set of drought forecasting models was presented. Different types of drought events were considered: a rainfall anomaly during the rainy season, standardized precipitation indices below a given threshold and anomalies

25 in regional reservoir storage. The forecast products considered were combinations of two models, ECHAM and the ECMWF seasonal forecast, two downscaling techniques, XDS and EQM, and a weather pattern classification approach.

    Each model provided an ensemble of predictions, so deterministic and probabilistic measures of skill could be used. Results showed that models with little to no skill by a deterministic measure (RMSE) showed skill under a probabilistic skill score. This underlines the importance of having at least one deterministic and one probabilistic measure of skill. The deterministic

30 measure also allowed the see the significant improvement introduced by the ensemble mean: the ensemble mean had in most cases a greater root mean square error than the climatology. The RMSE of the ensemble mean however was comparable to the climatology and in some cases lower. Still, no model had an RMSE that significantly departed from the RMSE of the climatology.

It was possible to provide a probabilistic forecast based on a multi-model ensemble. The probabilistic forecast was obtained by binding all members of all models into one product. The skill of this forecast is given in Figs. 4, 7, 9, and Table 3. These can be considered to be our best guess of a probabilistic drought forecast, since it is consistently among the best forecast skills provided by the individual models. Individual members surpassed the multi-model ensemble skill only occasionally, for

5 particular combinations of regions, months and indices.

The skill of the hydrological drought forecast, namely the relative reservoir storage, was comparable to the skill of the $SPEI_{01}$ forecast, around 0.6 for the regions of Castanhão and Orós and 0.5 for lower Jaguaribe. The skill obtained for the hydrological drought forecast is likely inflated by the long memory of the reservoir system and by the dependence on meteorological drought indices of long time scales. The hydrological drought index that most resembles streamflow, i.e. reservoir

10 storage variation, was forecasted with much lower skill. Improvements are expected by coupling a process-based hydrological model to the seasonal forecasting system.

This work showed that a multimodel ensemble can forecast drought events of time scales relevant to water managers in northeast Brazil with skill. But none or little skill could be found in the forecasts of the whole range of monthly precipitation or drought indices (e.g. forecasting average years). Both this work and others here revisited showed that major steps forward

15 are needed in forecasting the rainy season in northeast Brazil.

Hydrology and
Earth System
Sciences
Discussions

Open Access

EGU

References

Araújo, J. A. d. A.: Barragens no Nordeste do Brasil, DNOCS, Fortaleza, josé amaury de aragão araújo edn., 1990.

Araújo, J. C. d. and Bronstert, A.: A method to assess hydrological drought in semi-arid environments and its application to the Jaguaribe River basin, Brazil, Water International, 41, 213–230, doi:10.1080/02508060.2015.1113077, http://dx.doi.org/10.1080/02508060.2015.1113077, 2016.

Bürger, G.: Expanded downscaling for generating local weather scenarios, Climate Research, 7, 111–128, http://www.int-res.com/abstracts/cr/v07/n2/p111-128/, 1996.

Castro, T. d., Souza, F, B, J. M., Alves, Pontes, R. S. T., Reis, D., L.n, L., and Daher, S.: Neo-fuzzy neuron model for seasonal rainfall forecast: A case study of Ceara's eight homogenous regions, Journal of Intelligent & Fuzzy Systems, 25, 389–394, doi:10.3233/IFS-2012-0645, http://content.iospress.com/articles/journal-of-intelligent-and-fuzzy-systems/ifs645, 2013.

Doesken, N., McKee, T., and Kleist, J.: Development of a Surface Water Supply Index for the Western United States, Tech. Rep. Climatology Report 91-3, Colorado Climate Center, Department of Atmospheric Science, Colorado State University, http://climate.colostate.edu/pdfs/climo_rpt_91-3.pdf, 1991.

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, doi:10.5194/hess-17-2359-2013, https://www.hydrol-earth-syst-sci.net/17/2359/2013/, 2013.

Fioreze, A. P., Bubel, A. P. M., Callou, A. P., Mendonça, B. C. d. S., Nunes, C. M., Pinto, C. G., Viana, C. F. G., Junior, D. S. R., Martins, E. S. P. R., Rodrigues, F. S. F., Filho, F. d. A. d. S., Teixeira, F. J. C., Viana, F. L., Nascentes, J. C. d. M., Filho, J. G. C. G., Júnior, J. A. d. L., Campos, J. N. B., Carvalho, J. O. d., Gonçalves, J. Y. d. B., Burte, J., Silva, L. M. C. d., Azevedo, L. G. T. d., Bursztyn, M., Cerqueira, M. R. S., Coimbra, Thomas, P., Nobre, P., Vieira, R. F., Alves, R. F. F., Chacon, S. S., and Paulino, W. D.: A questão da água no Nordeste, Ministério da Ciência e Tecnologia (MCT), http://livroaberto.ibict.br/handle/1/669, 2012.

Formiga-Johnsson, R. M. and Kemper, K.: Institutional and Policy Analysis of River Basin Management: The Jaguaribe River Basin, Ceara, Brazil, SSRN Scholarly Paper ID 757424, Social Science Research Network, Rochester, NY, https://papers.ssrn.com/abstract=757424, 2005.

Hastenrath, S.: Exploring the climate problems of Brazil's Nordeste: a review, Climatic Change, 112, 243–251, doi:10.1007/s10584-011-0227-1, 2012.

Hastenrath, S. and Greischar, L.: Further Work on the Prediction of Northeast Brazil Rainfall Anomalies, J. Climate, 6, 743–758, doi:10.1175/1520-0442(1993)006<0743:FWOTPO>2.0.CO;2, 1993.

IPECE: Anuário Estatístico do Ceará, http://www.ipece.ce.gov.br/index.php/anuario-estatistico-do-ceara, 2017.

Juang, H.-M. H., Hong, S.-Y., and Kanamitsu, M.: The NCEP Regional Spectral Model: An Update, Bull. Amer. Meteor. Soc., 78, 2125–2143, doi:10.1175/1520-0477(1997)078<2125:TNRSMA>2.0.CO;2, 1997.

Marengo, J. A., Torres, R. R., and Alves, L. M.: Drought in Northeast Brazil—past, present, and future, Theor Appl Climatol, pp. 1–12, doi:10.1007/s00704-016-1840-8, https://link.springer.com/article/10.1007/s00704-016-1840-8, 2016.

Mckee, T. B., Doesken, N. J., and Kleist, J.: The relationship of drought frequency and duration to time scales, http://www.citeulike.org/group/15400/article/14027580, 1993.

Molteni, F., Stockdale, T., Balmaseda, M., Balsamo, G., Buizza, R., Ferranti, L., Magnusson, L., Mogensen, K., Palmer, T., and Vitart, F.: The new ECMWF seasonal forecast system (System 4), European Centre for Medium-Range Weather Forecasts, 2011.

Moura, A. D. and Hastenrath, S.: Climate Prediction for Brazil's Nordeste: Performance of Empirical and Numerical Modeling Methods, J. Climate, 17, 2667–2672, doi:10.1175/1520-0442(2004)017<2667:CPFBNP>2.0.CO;2, 2004.

Murawski, A., Bürger, G., Vorogushyn, S., and Merz, B.: Can local climate variability be explained by weather patterns? A multi-station evaluation for the Rhine basin, Hydrol. Earth Syst. Sci., 20, 4283–4306, doi:10.5194/hess-20-4283-2016, https://www.hydrol-earth-syst-sci.net/20/4283/2016/, 2016.

Philipp, A., Della-Marta, P. M., Jacobeit, J., Fereday, D. R., Jones, P. D., Moberg, A., and Wanner, H.: Long-Term Variability of Daily North Atlantic–European Pressure Patterns since 1850 Classified by Simulated Annealing Clustering, J. Climate, 20, 4065–4095, doi:10.1175/JCLI4175.1, http://journals.ametsoc.org/doi/abs/10.1175/JCLI4175.1, 2007.

Philipp, A., Beck, C., Huth, R., and Jacobeit, J.: Development and comparison of circulation type classifications using the COST 733 dataset and software, Int. J. Climatol., 36, 2673–2691, doi:10.1002/joc.3920, http://onlinelibrary.wiley.com/doi/10.1002/joc.3920/abstract, 2016.

Richardson, D. S., Bidlot, J., Ferranti, L., Ghelli, A., Haiden, T., Hewson, T., Janousek, M., Prates, F., and Vitart, F.: Verification statistics and evaluations of ECMWF forecasts in 2011-2012, European Centre for Medium-Range Weather Forecasts, 2012.

Roeckner, E., Arpe, K., Bengtsson, L., Brinkop, S., D{\"u}menil, L., Esch, M., Kirk, E., Lunkeit, F., Ponater, M., Rockel, B., Sausen, R., Schlese, U., Schubert, S., and Windelband, M.: Simulation of the present-day climate with the {ECHAM-3} model: impact of model physics and resolution, 1992.

Seibert, M., Merz, B., and Apel, H.: Seasonal forecasting of hydrological drought in the Limpopo Basin: a comparison of statistical methods, Hydrol. Earth Syst. Sci., 21, 1611–1629, doi:10.5194/hess-21-1611-2017, https://www.hydrol-earth-syst-sci.net/21/1611/2017/, 2017.

Stockdale, T. N., Anderson, D. L. T., Alves, J. O. S., and Balmaseda, M. A.: Global seasonal rainfall forecasts using a coupled ocean–atmosphere model, Nature, 392, 370–373, doi:10.1038/32861, http://www.nature.com/nature/journal/v392/n6674/abs/392370a0.html, 1998.

Svoboda, M., Hayes, M., Wood, D. A., and World Meteorological Organization, W.: Standardized Precipitation Index User Guide, WMO, WMO, http://library.wmo.int/pmb_ged/wmo_1090_en.pdf, 2012.

Trambauer, P., Werner, M., Winsemius, H. C., Maskey, S., Dutra, E., and Uhlenbrook, S.: Hydrological drought forecasting and skill assessment for the Limpopo River basin, southern Africa, Hydrol. Earth Syst. Sci., 19, 1695–1711, doi:10.5194/hess-19-1695-2015, wOS:000353877000007, 2015.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I.: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index, J. Climate, 23, 1696–1718, doi:10.1175/2009JCLI2909.1, 2009.

Vicente-Serrano, S. M., López-Moreno, J. I., Beguería, S., Lorenzo-Lacruz, J., Azorin-Molina, C., and Morán-Tejeda, E.: Accurate Computation of a Streamflow Drought Index, Journal of Hydrologic Engineering, 17, 318–332, doi:10.1061/(ASCE)HE.1943-5584.0000433, 2012.

Vitart, F.: Evolution of ECMWF sub-seasonal forecast skill scores over the past 10 years, ECMWF Research Department Technical Memorandum, p. 28, 2013.

Wetterhall, F., Pappenberger, F., He, Y., Freer, J., and Cloke, H. L.: Conditioning model output statistics of regional climate model precipitation on circulation patterns, Nonlin. Processes Geophys., 19, 623–633, doi:10.5194/npg-19-623-2012, https://www.nonlin-processes-geophys.net/19/623/2012/, 2012.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, google-Books-ID: _b4R9j2Iy7EC, 2005.

(WMO) World Meteorological Organization: Guidelines on Ensemble Prediction Systems and Forecasting, WMO, WMO, http://library.wmo.int/pmb_ged/wmo_1091_en.pdf, 2012.