Hydrol. Earth Syst. Sci. Discuss., 9, 11093–11129, 2012 www.hydrol-earth-syst-sci-discuss.net/9/11093/2012/ doi:10.5194/hessd-9-11093-2012 © Author(s) 2012. CC Attribution 3.0 License.



This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

Seasonal forecasts of drought indices in African basins

E. Dutra¹, F. Di Giuseppe¹, F. Wetterhall¹, and F. Pappenberger^{1,2}

¹European Centre for Medium-Range Weather Forecasts, Reading, UK ²College of Hydrology and Water Resources, Hohai University, Nanjing, China

Received: 10 September 2012 – Accepted: 19 September 2012 – Published: 28 September 2012

Correspondence to: E. Dutra (emanuel.dutra@ecmwf.int)

Published by Copernicus Publications on behalf of the European Geosciences Union.



Abstract

Vast parts of Africa rely on the rainy season for livestock and agriculture. Droughts can have a severe impact in these areas which often have a very low resilience and limited capabilities to mitigate their effects. This paper tries to assess the predictive capabilities

- of an integrated drought monitoring and forecasting system based on the Standard precipitation index (SPI). The system is firstly constructed by temporally extending near real-time precipitation fields (ECMWF ERA-Interim reanalysis and the Climate Anomaly Monitoring System-Outgoing Longwave Radiation Precipitation Index, CAMS-OPI) with forecasted fields as provided by the ECMWF seasonal forecasting system and then is
- evaluated over four basins in Africa: the Blue Nile, Limpopo, Upper Niger, and Upper Zambezi. There are significant differences in the quality of the precipitation between the datasets depending on the catchments, and a general statement regarding the best product is difficult to make. All the datasets show similar patterns in the South and North West Africa, while there is a low correlation in the tropical region which makes it difficult
- to define ground truth and choose an adequate product for monitoring. The Seasonal forecasts have a higher reliability and skill in the Blue Nile, Limpopo and Upper Niger in comparison with the Zambezi. This skill and reliability depends strongly on the SPI time-scale, and more skill is observed at larger time-scales. The ECMWF seasonal forecasts have predictive skill which is higher than using climatology for most regions.
- In regions where no reliable near real-time data is available, the seasonal forecast can be used for monitoring (first month of forecast). Furthermore, poor quality precipitation monitoring products can reduce the potential skill of SPI seasonal forecasts in two to four months lead time.



1 Introduction

Most of Africa relies on the rainy season for water supply for livestock and agriculture (IWMI, 2010). Therefore, rain shortage can have a severe impact over this continent which in many areas have a very low resilience and limited capabilities to mitigate

- ⁵ drought effects. A long sequence of droughts can also occur, such as the mega sub-Sahel drought in the 1970s and 1980s (e.g. Nicholson et al., 1998). A more recent example is the 2010/2011 drought that afflicted the Horn of Africa (Dutra et al., 2012). Monitoring and forecasting both the length and geographical extension of droughts is a key component of increasing resilience.
- ¹⁰ Droughts are typically classified in four types: meteorological, hydrological, agricultural and socioeconomic, and there are many drought indicators associated to each drought type (e.g Keyantash and Dracup, 2002). In this work we focus on the meteorological drought using the Standardized Precipitation Index (SPI) initially developed by Mckee et al. (1993) and recommended by the World Meteorological Organization as
- a standard to characterize meteorological droughts (WMO, 2009). The SPI is based on the probability of an observed precipitation deficit occurring over a given prior accumulated time period. This time period (also referred to time-scale) is defined accordingly to the particular application, with typical values of 3, 6 and 12 months. The flexibility of the accumulation in different time periods allows a range of meteorological, agricul-
- ²⁰ tural and hydrological applications. For short accumulations periods, e.g. 3-month, the SPI will be associated with short- and medium-term soil moisture conditions, which are potentially useful for agricultural production in rain-fed regions. However, a relatively normal 3-month period (or even wet) can occur during prolonged droughts. The 6-month SPI is associated with medium-term precipitation anomalies affecting stream
- flows and reservoir levels. The 12-month SPI reflects long-term precipitation anomalies which are associated with changes to water resources, such as large reservoirs and groundwater. Despite the general association of SPI time-scale with soil moisture, stream-flow and ground water changes, such relations are regionally dependent on the



precipitation regime and geological characteristics, among others. In each particular region/application, a detailed study should ideally be carried out to relate the different SPI time scales to the target variable, such as soil moisture available to crops or natural reservoirs. The SPI calculation is only dependent on monthly means of precipitation,

- which are usually available on near real-time from observations (in-situ and/or satellite) and also from seasonal forecasts (in both cases generally associated with large uncertainties). The data availability, and simplicity of the calculation, makes the SPI a drought index with potential capabilities for combining monitoring and forecasting, that has been widely demonstrated over Africa in other sectoral studies such as health,
 water management, energy and agriculture (Ingram et al., 2002; Jones et al., 2000;
- Lamb et al., 2011; Millner and Washington, 2011; Thomson et al., 2006).

15

The monitoring component relies on near real-time observations of surface variables such as precipitation, temperature and soil moisture. These observations can either be derived through the merging of ground observations and remote sensing information or by using numerical weather prediction (NWP) models such as re-analysis tools.

The forecasting component can rely on climatological forecasts or in recent years on numerical weather prediction (NWP). As the quality of forecast models steadily improves over the monthly to seasonal lead time there is in fact an increasing interest to test their potential in sectorial applications (Simmons and Hollingsworth, 2002). For ex-20 ample, Yoon et al. (2012) recently proposed a methodology to forecast 3- and 6-month SPI for the prediction of meteorological drought over the Contiguous United States

based on the NCEP climate forecast system (CFS) seasonal forecasts of precipitation (Saha et al., 2006).

The latest version of the ECMWF seasonal forecasting systems, system 4 (S4), was released in November 2011 (Molteni et al., 2011). Despite the recent model improvements, predicted fields such as temperature, and to a higher extent precipitation, can be biased and in some areas have little or no skill. This is particularly the case in some regions in Africa, where in-situ observations are scarce and models often show persistent systematic errors. One such example is the prediction of the West Africa monsoon



system, where S4 is able to reproduce the progression of the West Africa monsoon, but shows persistent biases in terms of a southerly shift of the precipitation in the main monsoon months of July and August (Agustí-Panareda et al., 2010; Tompkins and Feudale, 2009).

- In this paper an integrated monitoring and forecasting drought system for four African river basins has been designed to explore the current capability of ECMWF products to provide drought information over the Africa continent. This has been done by combining globally available precipitation monitoring products with the forecast from S4. The drought has been derived from precipitation fields and defined as an anomaly
- on the rainfall accumulation over different time-scales using the SPI. The four basins were chosen to represent different synoptic conditions typical of the African continent. The drought monitoring and forecasting system is described in Sect. 2 followed by the evaluation of the system in Sect. 3 and the main conclusions in Sect. 4.

2 Data and methods

15 2.1 Precipitation datasets

2.1.1 Precipitation monitoring

Several datasets could be used for drought monitoring. However, there are few technical requirements a dataset should fulfill to be suitable for an operational monitoring and forecasting chain employing the SPI which will be described in the following section.

Firstly, it should be long enough (at least 30 yr, as recommended by Mckee et al., 1993) and statistically homogeneous. This means that observations should as much as possible (i) avoid changes in rain-gauges location and measuring equipments; (ii) use similar techniques to derive precipitation from remote sensing data, even when using different platforms. When employing dynamical model outputs, the model should have the same spatial and temporal structure (as in reanalysis or global/regional climate)



models) to avoid disruptions due to model changes, such as a change in resolution or parametrization schemes. Changes in the observation systems and/or models can produce artificial signals, such as trends, that will be reflected in the drought indicators. Secondly, the dataset needs to be available in near real-time, meaning no more than 1 5 month delay.

The long-term homogeneity and near real-time update are two criteria difficult to achieve, especially on a global scale. Two freely available products that partially fulfill the requirements: the re-analysis produced by the dynamical ECMWF model ERA-Interim (Dee et al., 2011) and the observational based product Climate Anomaly Monitoring System-Outgoing Longwave Radiation Precipitation Index (CAMS-OPI, Janowiak and Xie, 1999). These datasets have a global coverage, span 30 yr time and are available in near real-time.

ERA-Interim (ERAI) starts the 1 January 1979 and is extended forward in near realtime (Dee et al., 2011). It has a spectral T255 horizontal resolution, which corresponds

- to approximately 79 km in the grid-point space and employs a sequential 4-D-var data assimilation scheme which ensures the optimal consistency between available observations and the model background. Full 3-D fields are stored 6-hourly while a large number of surface parameters, including precipitation are provided 3-hourly. In the following, monthly means of precipitation taken from the daily forecasts starting at 00:00 UTC and 12:00 UTC with lead times +24 h to +48 h (2nd day of forecast) are
- processed. The forecast delay should reduce the initial spin-up to a minimum.

10

The CAMS-OPI is a merged dataset produced by the NOAA Climate Prediction Centre (CPC) combining satellite rainfall estimates from the Outgoing Longwave Radiation (OLR) Precipitation Index (OPI) with ground-based rain gauge observations from the

²⁵ Climate Anomaly Monitoring System (CAMS). The OPI estimates are computed from NOAA polar-orbiting IR window channel data using the technique developed by Xie and Arkin (1998). While it is recognized that the OPI has significant limitations for many climate applications, the merged CAMS-OPI dataset is provided to whom requires near real-time applications. For example, Sohn et al. (2011) found that CAMS-OPI was



reliable for monitoring large-scale precipitation variations over the East Asian region. Janowiak and Xie (1999) provide a complete description of the CAMS-OPI merged dataset, which is available from January 1979 to present in a $2.5^{\circ} \times 2.5^{\circ}$ lat/lon regular resolution.

For research purposes, CPC encourage users to use instead either GPCP or CMAP (Xie and Arkin, 1997) merged climate rainfall datasets, both of which have better quality control measures and include satellite passive microwave rain estimates. Therefore, the Global Precipitation Climatology Project (GPCP) version 2.2 (Huffman et al., 2011) monthly precipitation is used in the following as a benchmark. It is available for the period January 1979 to December 2010 on a 2.5° × 2.5° lat/lon regular resolution.

2.1.2 Precipitation forecasting

Forecasted precipitation is the second required input to construct the drought forecasting system. In the implementation presented here this is provided by the most recent seasonal forecasting system at ECMWF (S4) which became operational in November

- ¹⁵ 2011 issuing 51 ensemble members with six months lead time. S4 has the same horizontal resolution of ERAI (about 79 km), and is fully coupled with an ocean model with a horizontal resolution of 1°. The initial perturbations are defined with a combination of atmospheric singular vectors and an ensemble of ocean analyses. Atmosphere model uncertainties are simulated using the 3-time level stochastically perturbed parameter-
- ized tendency (SPPT) scheme and the stochastic back-scatter scheme (SPBS) that are also operational in the current ECMWF medium-range ensemble prediction system. The hindcast set is provided for calibration, covering a period of 30 yr (1981 to 2010) with the same configuration as the operational forecasts but only with 15 ensemble members. Molteni et al. (2011) present an overview of S4 model biases and forecast performance.



2.2 Drought forecasting system

2.2.1 Drought metric

Drought is predicted by means of the Standardized Precipitation Index (SPI) Mckee et al. (1993). In the calculation of the SPI for a specific k time-scale the total precipitation for a certain month m (m = 1, ..., N, where N is the number of months in the time 5 serie) is the sum of the precipitation for the period [m - k + 1 to m]. For each calendar month (i.e. all Januaries, Februaries, etc) the accumulated precipitation distribution (N/12 samples) is fitted to a probability distribution. The resultant cumulative probability distribution (CDF) is then transformed into the standard normal distribution (mean zero with one standard deviation), resulting in the SPI. It is common to select a parametric 10 probability distribution to fit the precipitation. Different statistical distributions can be used, such as the gamma distribution (Lloyd-Hughes and Saunders, 2002) or Pearson III (Vicente-Serrano, 2006), and there are statistical tests (e.g. Kolmogorov-Smirnov) to assess the suitability of the selected distribution. In this study the gamma distribution was chosen since it is commonly used. The fitting of the distribution uses the approxi-15 mation proposed by Greenwood and Durand (1960).

2.2.2 Drought forecasts

20

25

The extension of the SPI from the monitoring period, i.e. past, to the seasonal forecast range has been performed by merging the seasonal forecasts of precipitation with the monitoring product and is somehow a delicate step of the whole system.

Firstly both the forecasts and monitoring products have to be interpolated to a common resolution. This step will depend on the available data (resolution of monitoring and seasonal forecasts of precipitation) and final application of the drought forecasts. Two options are available: (i) downscale the forecasts to a higher resolution using a simple (for example bilinear interpolation) or more advance methods (for example



statistical downscaling) as it was proposed by Yoon et al. (2012); or (ii) upscale the

forecasts and monitoring to a coarser resolution or to a region. The second option has the advantage of reducing the amount of data to process, and filters some of the intrinsic noise of grid-point precipitation from the dynamical models (Lander and Hoskins, 1997) and has been preferred in this exercise. The precipitation was therefore spatially aggregated (mass conserving) to a basin scale (the basin definitions are described in the end of this section).

5

Secondly some care needs to be taken on the seasonal forecast biases. Uncertainties in S4 precipitation forecasts are mainly controlled by model biases (Di Giuseppe et al., 2011). These biases can drift over time, i.e. change with lead time, and can jeopardize the prediction skill. Moreover, since the merging procedure involved two different precipitation datasets, care has to be taken to ensure that the two dataset have the same mean climate. This is achieved by performing a simple bias correction of monthly precipitation in the form: $P'_{m,l} = \alpha_{m,l}P_{m,l}$ where *P* and *P'* are the original and corrected seasonal forecast of precipitation, respectively, α a multiplicative correc-

¹⁵ tion factor and the subscripts *m* and *l* the calendar month (1 to 12) and the forecast lead time (0 to five months), respectively. The correction factor is given by the ratio: $\alpha_{m,l} = \bar{P}_m^{\text{mon}} / \bar{P}_{m,l}$, where \bar{P}_m^{mon} is the multi-annual mean of precipitation of the monitoring dataset for a particular calendar month, and $\bar{P}_{m,l}$ the multi-annual and ensemble mean of the forecasts for a particular calendar month *m* and lead time *l*. This simple bias correction does not address inter-annual variability and ensemble spread. More sophisticated bias correction methods (e.g. Yoon et al., 2012) are possible, but being mostly focused on spatially integrated quantities this is not believed necessary.

To create a seamless transition from the monitoring to the forecast fields, the interpolated and bias-corrected S4 precipitation data were merged with the monitoring fields.

²⁵ This merge is a simple concatenation of the precipitation time series, performed for the entire S4 hindcast period. For a particular initial forecast date in month m (m = 1 (January 1981),..., 360 (December 2010)) the accumulated precipitation for SPI-k with lead time l is given by:



Discussion Paper **HESSD** 9, 11093–11129, 2012 Seasonal forecasts of drought indices in African basins **Discussion** Paper E. Dutra et al. **Title Page** Introduction Abstract Conclusions References **Discussion** Paper Tables **Figures** 14 Back Close **Discussion** Paper Full Screen / Esc **Printer-friendly Version** Interactive Discussion

(1)

$$\begin{cases} \sum_{i=\max(0,l-k+1)}^{i=l} P'_{m,i}, & l-k+1 \ge 0\\ \sum_{i=\max(0,l-k+1)}^{i=l} P'_{m,i} + \sum_{j=m+l-k+1}^{j=m-1} P_j^{mon}, & l-k+1 < 0 \end{cases}$$

The application of Eq. (1) to all years and ensemble members for a specific calendar month (for example for the forecasts starting in February all months (m = 2, 13, 25, ..., 350) creates a sample of 450 (30 yr × 15 ensemble members) values of accumulated precipitation that are transformed to the normal space following the SPI calculation procedure described before.

2.3 Verification metrics

The validation of the monitoring and forecasts where done with Anomaly Correlation Coefficient (ACC), Continuous Rank Probability Skill Scores (CRPSS), Relative Operating Characteristic (ROC) diagrams and, Reliability (REL) diagrams. The ACC is cal-10 culated as the ordinary correlation coefficient on the anomalies, i.e. removing the mean annual cycle. The continuous rank probability score (CRPS, see Hersbach, 2000) can be interpreted as the integral of the Brier Score over all possible threshold values of the parameter under consideration. Since the CRPS is not a normalized measure, this skill score (CRPSS) was used. In the skill score calculation the reference forecast 15 was taken from the observational dataset to produce a climatological forecast (CLM) with the same ensemble size as the forecast, by randomly sampling different years. The ROC diagram (Mason and Graham, 1999) displays the false alarm rate (FAR) as a function of hit rate (HR) for different thresholds (i.e. fraction of ensemble members detecting an event) identifying whether the forecast has the attribute to discriminate 20 between an event or not. The area under the ROC curve is a summary statistic repre-

senting the skill of the forecast system. The area is standardized against the total area

along the diagonal (no information) has an area of 0.5. The reliability diagrams (Hamill, 1997) measure the consistency between predicted probabilities of an event and the observed relative frequencies and are used to assess the reliability and confidence of the forecasts.

5 2.4 Selection of the basins

The drought forecasting system was tested in four river basins of the African continent: Blue Nile (NB), Limpopo (LP), Upper Niger (NG) and Zambezi (ZB) (Table 1 and Fig. 1). The catchment definitions were taken from the river network and basins created by Yamazaki et al. (2009). The location of the basins was selected to sample different climatic regions of Africa, with different precipitation distributions along the year. The 10 regions are defined as hydrological basins instead of lat/lon limits since basins have a geographical meaning draining to the same river. All the basins have a similar size (see Table 1) of about 3.5×10^5 km², corresponding to approximately 60 grid points of ERAI or S4 and 4.5 grid points of GPCP and CAMS-OPI. The possible ranges of basins sizes are not addressed in this study. This will mainly depend on prior knowledge of 15 the region (precipitation patterns and variability), and underlying skill of the seasonal forecasts (avoiding merging region with different skill). Very small basins will not allow for the spatial filtering that reduces precipitation noise, while very large basins (e.g. entire Nile, Niger) might account for different climatic regions with different forecast skill. 20

3 Results

25

3.1 Quality of observations

Throughout the paper GPCP is assumed to be the ground truth and is used as benchmark for drought monitoring. However, the quality of this large-scale dataset is significantly influenced by stations count and changes in the number of stations in time. Since



all basin have a similar area (see Table 1), the differences in the stations count and its change in time can potentially compromise the reliability of GPCP and its temporal homogeneity (essential for drought monitoring). The analysis of the temporal evolution of the number of stations present in the Global Precipitation Climatology Centre (GPCC,

- ⁵ Rudolf and Schneider, 2005; Schneider et al., 2011), the underlying data used in GPCP over land, along with the error estimates provided by GPCP (Fig. 2) provides a qualitative overview of possible challenges in GPCP over each basins. In the NG there is a significant drop in the stations count from the early 80's to late 90's of about 50 % to a very low number of stations during the last decade. This is reflected in an increase
- of the error estimates during the last decade. In both BN and ZB basins, the number of stations is lower than in NG, being much lower (around 10) in the ZB. LP is the basin with more and stable number of stations, expect for a drop in the last 5 yr of the dataset. The number of stations present in CAMS-OPI (Fig. 2) is much lower than in GPCP over the selected basins, especially in NG, NG and ZB. This will impact its potential use of CAMS-OPI for real time monitoring, that will be addressed in the next section.

3.2 Drought monitoring

Precipitation over the Niger (NG) and the Blue Nile (BN) is controlled by the south to north and back progression of the West Africa Monsoon. Peak rainfall occurs in the boreal summer (June–September, Fig. 3a, b), when the ITCZ moves to its far north²⁰ ern limit producing disturbances that are dynamically linked to the African Easterly Jet. These are the first cause for the large scale precipitation observed in the region at the monsoon onset. Westward propagating mesoscale disturbances generate the dominant convective systems. They feed into the large scale disturbance only during late boreal summer (when enough moisture is available) changing the rainfall regime

from frontal precipitation (June–July) to convective (August–September). The Limpopo and Zambezi (ZB) river basins are instead located in East Africa and have their peak precipitation occurring during boreal winter (Fig. 3c, d). The variability is therefore generally out of phase with that in West Africa (i.e. dry (wet) West Africa corresponds to



wet (dry) East Africa). Although wave activity has not been identified, rainfall tends to be organized into mesoscale convective systems analogous to those in Sahelian West Africa.

In the BN and NG basins, the rainy season (June–September) is captured by all datasets, with an overestimation in the BN by ERAI (Fig. 3b). S4 forecasts overestimate precipitation in both BN and NG in the first month of forecasts with a reduction of the peak rainfall with lead time. This is an example of model drift with lead time, justifying the applied bias correction for each calendar month (initial forecast date) and lead time. In the LP and ZB basins all datasets show a reasonable agreement with GPCP, and S4 has a reduced drift in the mean cycle with lead time.

The temporal correlation of the SPI-3, 6 and 12 months calculated with ERAI, CAMS-OPI and S4 first month of forecasts (S4L0) against the SPI calculated with GPCP (Table 2, and Fig. 4 for the SPI-12 time series) gives an overview of the potential quality of each dataset for drought monitoring in the regions. Both ERAI and CAMS-OPI have

- ¹⁵ a good agreement with GPCP in LP for the different time-scales, while in the remaining three basins the correlations are lower. In the NG and BN, the SPI derived from the first month of forecasts from S4 has a better agreement with GPCP than ERAI or CAMS-OPI, while in ZB all datasets display low correlations (Table 2). It should be noted that S4L0 has a better intra-seasonal to inter-annual variability of precipitation anomalies
- than ERAI or CAMS-OPI in the NG and BN regions. ERAI overestimates the decrease in rainfall in the Central African region which is likely to be associated with a substantial warm bias in the model due to an underestimation of aerosol optical depth in the region, as well as changes in the data entering the data assimilation system (Dee et al., 2011). This resulted in an unrealistic model drying that penalizes the SPI scores. The poor
- ²⁵ performance of CAMS-OPI, when compared with GPCP, in the NG, BN and ZB basins is most likely linked with the low number of stations used (Fig. 2), mainly due to the near real-time restriction, since not all stations report in near real-time.

An independent dataset, the monthly mean anomalies of river discharge, available from the Global Runoff Data Centre (GRDC) were compared with the SPI. The river



discharge accounts for the upstream drainage area that corresponds to the basin definition, and was only available in the Niger (GRDC station Dire, ID: 1134700) and Zambezi (GRDC station KATIMA MULILO, ID: 1291100). There is a reasonable agreement between the GPCP SPI-12 and river discharge in the NG basin with a temporal correla-

- tion of 0.65 (Table 3, Fig. 4) while the correlations with the remaining datasets are much lower. This can be interpreted as an indirect assessment of the precipitation datasets. In ZB the temporal correlations are lower for GPCP. This could be associated with the reduced number of stations entering GPCP. Even if precipitation would be very close to reality, timing delays in the river network, interaction with groundwater, among other
- ¹⁰ processes, will introduce non-linear effects in the river discharge making its comparison with the SPI difficult. Furthermore, such relation might be only present in a sub-set of calendar months/ SPI time scales, as Vicente-Serrano et al. (2012) exemplified for the Congo and Orange river basins when comparing with the Standardized Precipitation Evapotranspiration Index (SPEI, Vicente-Serrano et al., 2010).
- The spatial patterns of the temporal correlations of the SPI-3 and 12 calculated from the different products are in agreement with GPCP in Southern and North-West Africa, while in Central Africa (between the 20° North/South parallels) ERAI and CAMS-OPI have low or non-significant correlations, especially for the SPI-12 (Fig. 5). However, CAMS-OPI tends to have a better performance than ERAI. S4L0 has in general a lower variability than ERAI or CAMS-OPI, except over a latitudinal band South of Sahel (in-
- cluding the NG and BN basins), being a good candidate for drought monitoring in those regions, considering the poor performance of ERAI and CAMS-OPI.

3.3 Drought forecasting

The skill of the seasonal forecasts of SPI will depend on the skill of the underlying seasonal forecasts of precipitation and on the quality of the monitoring (for long SPI time scales and short lead times, where the monitoring dominates over the forecast). These two components of the skill can be separately analyzed, by (i) accessing the



skill of the seasonal forecasts of precipitation and (ii) evaluating the potential skill of the SPI seasonal forecast, i.e. using a perfect monitoring product.

3.3.1 Precipitation monitoring skill

Over LP, NG and BN, S4 has skill in first month of forecasts, explaining the good performance of the SPI calculated using S4L0 when compared with ERAI and CAMS-OPI, especially in the NG and BN basins (Fig. 6). This can be primarily attributed to the predictability coming from the land-atmosphere initial conditions that will dominate the first days of the forecast. With increasing lead time, there is a general drop in skill that is only present in regions/seasons associated with large-scale climate forcings that can be captured by the coupled atmosphere-ocean modeling system. In both NG and BN, S4 has skill up to two/three months lead time for the forecasts valid between June to September, which is also the main rainy season, while in the LP a similar skill with lead time is also found during November to February, also the rainy season. In the ZB,

S4 has a reduced skill (only 3 months at 0 lead time), which is also visible in ERAI and CAMS-OPI. However, ZB was also the basin with lower number of rain-gauges included in GPCP, therefore being the most uncertain in terms of verification.

3.3.2 Forecast skill of the benchmark

The potential skill of the SPI forecasts was evaluated by merging the S4 precipitation with GPCP to create a benchmark of the different SPI time-scales for the seasonal forecasts described above. This method isolates the contribution of the seasonal forecasts of precipitation to the SPI skill, avoiding the problems of the different monitoring products. On a regional scale, this can be adapted by using local information, such as long-term rain gauges and/or gridded precipitation datasets. The SPI seasonal forecasts using GPCP+S4 were benchmarked against forecasts using the same monitor ²⁵ ing merged with climatological forecasts (CLM) created by randomly sampling different 15 yr (same ensemble size as S4) of GPCP.



The ACC of the SPI-12 is very close to 1 in all basins for 0 and 1 months lead time (Fig. 7e-h). In this case, the SPI-12 is built from 11 or 10 months of the monitoring and 1 or 2 months of the seasonal forecasts for the 0 and 1 months lead time, respectively. For the short lead times, the monitoring dominates the ACC of the SPI-12 which yields scores close to 1 since the verification is done against the same dataset used for 5 monitoring. In the SPI-3 the ACC for the 0 lead time is already lower than in the SPI-12, and rapidly drops to low values or not significant in regions/periods with low or no precipitation predictability. For long lead times, there is a drop in the SPI-12 skill in particular for the verification in the calendar months after the rainy season. This is associated with the different weight of the monitoring and forecast in regions with 10 a pronounced annual cycle. The SPI-12 forecasts valid before the rainy season will tend to have a higher skill, since the core precipitation information comes from the monitoring (year before), while the forecasts valid right after the rainy season will rely on the S4 seasonal prediction. The CRPSS identifies the verification months and lead time where the SPI forecasts using S4 outperform a simple climatological forecast 15

(Fig. 7i–p). Those periods are consistent with higher ACC of S4 compared with CLM (with symbols in Fig. 7a–h), and reflect the underlying skill of S4 precipitation (Fig. 6). The previous skill analysis, based on ACC and CRPSS, considered the full range of

SPI forecasts. For drought detection/early warning, ROC and REL diagrams (among

- others like the Brier Score) are a useful tool, by testing categorical forecasts, i.e. event or no event. A drought event is defined as SPI < -1. The ROC diagrams in Fig. 8 of the SPI-3 and SPI-6 represent the skill of using only precipitation forecasts (no monitoring), while in the SPI-12 6 months of monitoring are merged with 6 months of seasonal forecasts. The ROC scores of CLM are close to 0.5 (no information) in all basins for
- the SPI-3 and SPI-6 at 2 and 5 months lead time, since these are just a random climatological sampling (Fig. 8). On the other hand, S4 has skill in drought detection in the NG, BN and LP, and no skill in ZB. For the SPI-12 at 5 months lead time, the ROC of CLM is always around 0.8 and are outperform by S4 in the NG, BN and LP.



The reliability diagrams (Fig. 9) further support the previous results showing that SPI-3 and SPI-6 with 2 and 5 months lead time, respectively, are reliable in the NG, BN and LP and tend to be over-confident (reliability curves with a slope < 1; Fig. 9). While the ACC and ROC evaluation indicated a clear difference between S4 and CLM forecasts for the SPI-12 at 5 months lead time, the reliability diagrams show similar results, with 5 slopes of the reliability curves close to 1 with S4 being under-confident (slopes > 1) in the BN and LP. The variation of ROC and ROC skill score (ROCSS) with lead time resume the above results (Fig. 10): (i) in the ZB S4 is similar to a climatological forecast, i.e. no skill, while it outperforms CLM in the NG, BN and LP; (ii) in the SPI-3 the 2 months lead time (using the first 3 months of the seasonal forecast) and in the SPI-10 6 the 5 months lead time (using the first 6 months of the seasonal forecast) has the highest skill scores; (iii) the skill score of the SPI-12 is reduced, i.e. it is difficult to beat a climatology based forecast for long SPI time scales, where the monitoring dominates; and (iv) SPI S4 forecasts are never worst than CLM, and their skill are only driven by

the accumulated skill of the S4 precipitation forecasts.

3.3.3 Seasonal forecast skill

20

The potential skill allows a clear understanding of the importance and impact of the skill of S4 precipitation, but for a near real-time operational implementation GPCP is not available. Therefore, a similar analysis to the previous section was performed using other precipitation products that have long-term records and are available in near real-time to assess the actual predictive skill of the merged forecast.

The ROC scores for the near real-time forecasts are equal from 2 months lead time onwards for the SPI-3, and for the 5 months lead time in the SPI-6 since these do not include precipitation from the monitoring (Figs. 10 and 11). In the NG and BN ERAI and S4L0 were similar, outperforming CAMS-OPI however having a similar skill with

and S4L0 were similar, outperforming CAMS-OPI however having a similar skill with 0 months lead time to using GPCP as monitoring at 2 months lead time. This means that the problems identified in those datasets (Sect. 3.2) leads to a reduction of skill of 2 months in the NG and BN and 1 month in LP for the SPI-3. For the SPI-6 the skill



reduction is between 3 to 4 months, while for the SPI-12 only CAMS-OPI is able to reach similar skill to GPCP at 5 months lead time in the LP basin. These results high-light the role of the precipitation monitoring quality for SPI seasonal forecasts, showing that significant gains in skill can be obtained by using good quality observation/modeled precipitation.

4 Conclusions

5

10

25

In this paper the use of different observational (GPCP and CAMS-OPI) and reanalysis datasets (ERA-Interim) were evaluated concerning their value as monitoring tools for droughts in four African basins. Furthermore, the skill of the new seasonal forecast (S4) was tested in its ability to forecast droughts on a seasonal scale, in combination with reanalysis/observations as well as a stand-alone tool.

There is a clear difference in skill of monitoring precipitation anomalies, and thereby also droughts, depending on the region. In general, monitoring is difficult in the tropical convergence zone, and the different datasets show the highest divergence in these ¹⁵ regions. It is therefore important to carefully assess the performance of the monitoring dataset for the specific region of interest. GPCP shows the highest correlation on longer time-scales with normalized river discharge in the Niger and Zambezi basins although the low number of stations used and its change over time. However, GPCP is discontinued and cannot serve as a near real-time monitoring tool in the future, but serve as a benchmark observational tool. In this study it was also used to bias-correct the reanalysis data. The main conclusions from the monitoring component are:

- drought monitoring in Africa with ERA-Interim is mainly possible outside the tropical region;
- the usefulness of near real-time monitoring tools has to be carefully selected depending on region;



 in regions where no reliable near real-time data is available (in this study Niger, Blue Nile and Zambezi) it is better to use just the first month of the seasonal forecasts for drought monitoring.

The new ECMWF seasonal forecast system 4 shows skill as a forecasting tool in most basins, and is in no basin performing worse than climate. However at longer time scales, where a merge with observational datasets are needed, the selection of the best observation dataset is paramount. The main conclusions of the drought forecasting system applied to the Africa basins are:

- The system 4 seasonal forecast has predictive skill in comparison with climatology
- in the Niger, Blue Nile and Limpopo and no skill in the Zambezi basin;
- Poor quality monitoring products can reduce the potential skill of SPI seasonal forecasts with 2 to 4 months lead time.

The methodology presented in the paper to merge a monitoring and seasonal fore-casts of precipitation in a regional-scale can be adapted by using other sources of
precipitation for the monitoring (e.g. in-situ rain gauges, gridded precipitation datasets, remote sensing estimates) and seasonal forecasts (e.g. other systems, multi-model approaches, statistical methods). This methodology can be also applied on a grid-point basis, following downscaling methods as proposed by Yoon et al. (2012), but care has to be taken when interpreting how seasonal scale predictions of precipitation can be
reliable on local scales. Furthermore, the role, quality and skill of other drought indicators (e.g. based on soil moisture, river discharge, evaporation) has to be established, but such work will be highly dependent on reliable monitoring networks.

Acknowledgements. River discharge observations were provided by the Global Runoff Data Centre (GRDC). This work was funded by the FP7 EU projects DEWFORA http://www.dewfora. net and GLOWASIS (http://www.glowasis.eu).



10

25

References

30

- Agustí-Panareda, A., Balsamo, G., and Beljaars, A.: Impact of improved soil moisture on the ECMWF precipitation forecast in West Africa, Geophys. Res. Lett., 37, L20808, doi:10.1029/2010gl044748, 2010.
- ⁵ Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Ros-
- ¹⁰ nay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.
 - Di Giuseppe, F., Molteni, F., and Tompkins, A. M.: A rainfall calibration methodology for impacts modelling based on spatial mapping, Q. J. Roy. Meteor. Soc., in press, 2012
- ¹⁵ Dutra, E., Magnunson, L., Wetterhall, F., Cloke, H. L., Balsamo, G., Boussetta, S., and Pappenberger, F.: The 2010–11 drought in the Horn of Africa in the ECMWF reanalysis and seasonal forecast products, Int. J. Climatol., online first: doi:10.1002/joc.3545, 2012.
 - Greenwood, J. A. and Durand, D.: Aids for Fitting the Gamma Distribution by Maximum Likelihood, Technometrics, 2, 55–56, 1960.
- Hamill, T. M.: Reliability diagrams for multicategory probabilistic forecasts, Weather Forecast., 12, 736–741, doi:10.1175/1520-0434(1997)012<0736:rdfmpf>2.0.co;2, 1997.
 - Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather Forecast., 15, 559–570, doi:10.1175/1520-0434(2000)015<0559:dotcrp>2.0.co;2, 2000.
- ²⁵ Huffman, G. J., Bolvin, D. T., and Adler, R. F.: GPCP Version 2.2 Combined Precipitation Data set, WDC-A, NCDC, Asheville, NC., available at: http://www.ncdc.noaa.gov/oa/wmo/ wdcamet-ncdc.html (last access: September 2012), 2011.
 - Ingram, K. T., Roncoli, M. C., and Kirshen, P. H.: Opportunities and constraints for farmers of West Africa to use seasonal precipitation forecasts with Burkina Faso as a case study, Agr. Syst., 74, 331–349, 2002.
 - IWMI: Managing water for rainfed agriculture, International Water Management Institute Water Issue Brief, 10, 4 pp., doi:10.5337/2010.223, 2010.



Janowiak, J. E. and Xie, P.: CAMS-OPI: a global satellite-rain gauge merged product for realtime precipitation monitoring applications, J. Climate, 12, 3335–3342, doi:10.1175/1520-0442(1999)012<3335:coagsr>2.0.co;2, 1999.

Jones, J. W., Hansen, J. W., Royce, F. S., and Messina, C. D.: Potential benefits of climate forecasting to agriculture, Agr. Ecosyst. Environ., 82, 169–184, 2000.

Keyantash, J. and Dracup, J. A.: The quantification of drought: An evaluation of drought indices, B. Am. Meteorol. Soc., 83, 1167–1180, 2002.

5

20

25

- Lamb, P. J., Timmer, R. P., and Lélé, M. I.: Professional development for providers of seasonal climate prediction, Clim. Res., 47, 57–75, doi:10.3354/cr00949, 2011
- Lander, J. and Hoskins, B. J.: Believable Scales and Parameterizations in a Spectral Transform Model, Mon. Weather Rev., 125, 292–303, doi:10.1175/1520-0493(1997)125<0292:bsapia>2.0.co;2, 1997.

Lloyd-Hughes, B. and Saunders, M. A.: A drought climatology for Europe, Int. J. Climatol., 22, 1571–1592, 2002.

- ¹⁵ Mason, S. J. and Graham, N. E.: Conditional probabilities, relative operating characteristics, and relative operating levels, Weather Forecast., 14, 713–725, doi:10.1175/1520-0434(1999)014<0713:cproca>2.0.co;2, 1999.
 - McKee, T. B. N., Doesken, J., and Kleist, J.: The relationship of drought frecuency and duration to time scales, Eight Conf. On Applied Climatology, Anaheim, CA, Am. Meteorol. Soc., 179– 184, 1993.
 - Millner, A. and Washington, R.: What determines perceived value of seasonal climate forecasts? A theoretical analysis, Global Environ. Change, 21, 209–218, 2011.
 - Molteni, F., Stockdale, T., Balmaseda, M., BALSAMO, G., Buizza, R., Ferranti, L., Magnunson, L., Mogensen, K., Palmer, T., and Vitart, F.: The new ECMWF seasonal forecast system (system 4), ECMWF Tech. Memo., 656, 49 pp., 2011.
 - Nicholson, S. E., Tucker, C. J., and Ba, M. B.: Desertification, drought, and surface vegetation: an example from the West African Sahel, B. Am. Meteorol. Soc., 79, 815–829, doi:10.1175/1520-0477(1998)079<0815:ddasva>2.0.co;2, 1998.

Rudolf, B. and Schneider, U.: Calculation of gridded precipitation data for the global land-

³⁰ surface using in-situ gauge observations, in: Proceeding of the 2nd Workshop of the International Precipitation Working Group IPWF, Monterey, October 2004, EUMETSAT, ISBN 92-9110-070-6, ISSN 1727-432X, 231–247, 2005.



- Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., Van den Dool, H. M., Pan, H. L., Moorthi, S., Behringer, D., Stokes, D., Peña, M., Lord, S., White, G., Ebisuzaki, W., Peng, P., and Xie, P.: The NCEP climate forecast system, J. Climate, 19, 3483–3517, doi:10.1175/jcli3812.1, 2006.
- Schneider, U., Becker, A., Meyer-Christoffer, A., Ziese, M., and Rudolf, B.: Global precipitation analysis products of the GPC C., Global Climatology Centre (GPCC), DWD, Internet Publikation, 1–13, 2011.
 - Simmons, A. J. and Hollingsworth, A.: Some aspects of the improvement in skill of numerical weather prediction, Q. J. Roy. Meteor. Soc., 128, 647–677, doi:10.1256/003590002321042135, 2002.
- Sohn, S.-J., Tam, C.-Y., Ashok, K., and Ahn, J.-B.: Quantifying the reliability of precipitation datasets for monitoring large-scale East Asian precipitation variations, Int. J. Climatol., 32, 1520–1526, doi:10.1002/joc.2380, 2011.

10

Thomson, M. C., Doblas-Reyes, F. J., Mason, S. J., Hagedorn, R., Connor, S. J., Phindela, T.,

¹⁵ Morse, A. P., and Palmer, T. N.: Malaria early warnings based on seasonal climate forecasts from multi-model ensembles, Nature, 439, 576–579, 2006.

Tompkins, A. M. and Feudale, L.: Seasonal ensemble predictions of West African monsoon precipitation in the ECMWF system 3 with a focus on the AMMA special observing period in 2006, Weather Forecast., 25, 768–788, doi:10.1175/2009waf2222236.1, 2009.

- Vicente-Serrano, S.: Differences in spatial patterns of drought on different time scales: an analysis of the Iberian Peninsula, Water Resour. Manag., 20, 37–60, 2006.
 - Vicente-Serrano, S. M., Begueria, S., and Lopez-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, 2010.
- Vicente-Serrano, S. M., Beguería, S., Gimeno, L., Eklundh, L., Giuliani, G., Weston, D., El Kenawy, A., López-Moreno, J. I., Nieto, R., Ayenew, T., Konte, D., Ardö, J., and Pegram, G. G. S.: Challenges for drought mitigation in Africa: the potential use of geospatial data and drought information systems, Appl. Geogr., 34, 471–486, doi:10.1016/j.apgeog.2012.02.001, 2012. WMO: Press release, December 2009 No. 872, 2009.
- Xie, P. and Arkin, P. A.: Global Precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, B. Am. Meteorol. Soc., 78, 2539–2558, doi:10.1175/1520-0477(1997)078<2539:gpayma>2.0.co;2, 1997.



Xie, P. and Arkin, P. A.: Global Monthly Precipitation Estimates from Satellite-Observed Outgoing Longwave Radiation, J. Climate, 11, 137–164, doi:10.1175/1520-0442(1998)011<0137:gmpefs>2.0.co;2, 1998

Yamazaki, D., Oki, T., and Kanae, S.: Deriving a global river network map and its sub-grid

- topographic characteristics from a fine-resolution flow direction map, Hydrol. Earth Syst. Sci., 13, 2241–2251, doi:10.5194/hess-13-2241-2009, 2009.
 - Yoon, J.-H., Mo, K., and Wood, E. F.: Dynamic-model-based seasonal prediction of meteorological drought over the contiguous United States, J. Hydrometeorol., 13, 463–482, doi:10.1175/jhm-d-11-038.1, 2012.



Table 1. Bas	ins definitions	S. See also Fig. 1.	
--------------	-----------------	---------------------	--

Basin	Outlet	Area ($\times 10^5$ km ²)
Upper Niger (NG) Blue Nile (BN) Upper Zambezi (ZB) Limpopo (LP)	16.24° N, -3.39° E 15.50° N, 32.68° E -17.46° N, 24.25° E -24.25° N, 32.79° E	3.62 3.17 3.34 3.45



Discussion Pa	HESSD 9, 11093–11129, 2012								
per Discussior	Seasonal of drought African E. Dutr	forecasts indices in basins a et al.							
n Pap	Title Page								
er	Abstract	Introduction							
	Conclusions	References							
)iscuss	Tables	Figures							
ion P	14	►I.							
aper	•								
_	Back	Close							
Discus	Full Scre	en / Esc							
sion	Printer-frien	Printer-friendly Version							
Pap	Interactive	Discussion							
er									

Table 2. Temporal correlation of the 3, 6 and 12-month SPI from between GPCP and ERAI, CAMS-OPI and S4L0 (each column) for the different basins.

	ERAI		CAMS-OPI			S4L0				
	3	6	12	3	6	12		3	6	12
NG	0.60	0.48	0.34	0.49	0.36	0.22		0.57	0.54	0.56
BN	0.57	0.50	0.41	0.62	0.54	0.41		0.65	0.68	0.70
ZB	0.47	0.47	0.48	0.42	0.48	0.47		0.37	0.42	0.39
LP	0.83	0.86	0.90	0.87	0.89	0.89		0.53	0.59	0.68

Discussion Pa	HES 9, 11093–1	SSD 1129, 2012
aper Discussion	Seasonal of drought African E. Dutr	forecasts indices in basins ra et al.
n Pape	Title	Page
r _	Abstract	Introduction References
Discuss	Tables	Figures
sion Pa	I	۶I
Iper	•	•
—	Back	Close
Discussion	Full Scree Printer-frier	een / Esc ndly Version
n Paper	Interactive	Discussion

Table 3. Temporal correlation of the 12-month SPI from ERAI, CAMS-OPI, S4 and GPCP with the normalized monthly discharge in the NG and ZB basins.

	ERAI	CAMS-OPI	S4L0	GPCP
NG	0.36	0.27	0.49	0.65
ZB	0.45	0.46	0.37	0.54



Fig. 1. Basins definition (dark gray), and the full basin (dark and light gray). See also Table 1.





Fig. 2. Temporal evolution of the mean annual number of stations present in GPCC (dashed lines) and CAMS-OPI (circle symbols) in each basin and GPCP error estimates normalized by the mean precipitation (solid line). CAMS-OPI number of stations was multiplied by 5.





Fig. 3. Mean annual cycle of precipitation over the selected basins. The shaded area represents the range (±1 standard deviation) of observed (GPCP-red) and modeled (S4, gray) precipitation in the hindcast period (1981–2010), comparing with ERAI (blue), CAMS-OPI (green). The time series for the seasonal forecasts uses the first (S4-L0,thick black) and last (S4-L6, tick gray) month of forecasts and the 15 ensemble members.





Fig. 4. Evolution of the 12-month SPI in the different basins given by S4 (first forecast month), CAMS-OPI, GPCP and ERAI precipitation, and normalized monthly discharge. The horizontal ticks represent January of each year. In the discharge rows, the symbols identify missing data.





Fig. 5. Temporal correlation of **(a–c)** 3-month and **(d–f)** 12 month SPI between **(a,d)** GPCP and ERAI, **(b,e)** GPCP and CAMS-OPI and **(c,f)** GPCP and S4.





Fig. 6. Anomaly correlation coefficient (ACC) of 3-montlhy mean precipitation as a function of verification season (horizontal axis) and dataset (EI: ERAI, CS: CAMS-OPI) and S4 lead time. For example, the color associated with column OND of line ERAI corresponds to the ACC of ERAI versus GPCP mean October–December precipitation (over 30 yr), while the column OND of line SL4L2 corresponds to the ACC of S4 forecasts initialized in August valid for October–December (two months lead time) compared with October–December precipitation of GPCP. Only ACC significant at p < 0.05 are displayed. The forecast and ERAI and CAMS-OPI are verified against GPCP for the period 1981 to 2010.







Fig. 7. Anomaly correlation coefficient (ACC) the seasonal forecasts of SPI-3 (**a**–**d**), and SPI-12 (**e**–**h**) and continuous rank probability skill score (CRPSS) of SPI-3 (**i**–**I**) and SPI-12 (**m**–**p**). In the color matrix, the horizontal axis represents the verification month and the vertical axis the lead time (months). In the ACC S4 forecasts are compared with GPCP, and the white circles indicate that the S4 ACCB,> CLM ACC by at least 0.05. In the CRPSS panels S4 CRPS is benchmarked against the CPRS of CLM.



Fig. 8. Relative operating characteristic (ROC) diagram representing false alarm rate versus hit rate for the SPI-3 for 2 months lead time (a-d), the SPI-6 for 5 months lead time (e-h) and the SPI-12 for 5 months lead time (i-I), given by S4 (black) and CLM (gray) in the different basins (columns). Calculations based on 20 thresholds (fraction of ensemble members below -1), from 1 (symbols closer to 0,0) to 0 (symbols closer to 1, 1). Both S4 and CLM seasonal forecasts were merged with GPCP for the SPI calculation and the forecasts are verified against the SPI calculated with GPCP. ROC values are given in the legend of each panel.





Fig. 9. Reliability diagrams (CLM, S4) and frequency histograms (fCLM, fS4) for SPI < -1 forecasts produced by S4 (black lines and white bars) and CLM (gray lines and bars). For perfect reliability the curves should fall on top of the dashed diagonal line. The thin solid lines (CLM* and S4*) are the weighted least-squares regression lines of the reliability curves, and the slope of each curve is displayed in each panel. Each panel represents a particular basin (column) and SPI timescale (lines), with the same organization has in Fig. 8.











Fig. 11. Relative operating characteristic (ROC) of the S4 SPI forecasts as a function of lead time (months) for different time-scales (lines) and for the four basins (columns). Each panel compares the ROC of S4 forecasts using GPCP (square symbols), CAMS-OPI (triangle up symbols), ERAI (circle symbols) and S4L0 (triangle left symbols) as monitoring. Note that for the SPI-3 all ROC scores are the same from 2 month lead time onwards and for the SPI-6 for the 5 month lead time since in those lead times only S4 precipitation is used.

