Climatology

User-oriented global predictions of the GPCC drought index for the next decade

Andreas Paxian[1*], Markus Ziese[1], Frank Kreienkamp[2], Klaus Pankatz[1], Sascha Brand[1], Alexander Pasternack[3], Holger Pohlmann[4], Kameswarrao Modali[4] and Barbara Früh[1]

[1]Business Area of Climate and Environment, Deutscher Wetterdienst, Offenbach (Main), Germany
[2]Business Area of Climate and Environment, Deutscher Wetterdienst, Stahnsdorf, Germany
[3]Institute of Meteorology, Freie Universität Berlin, Berlin, Germany
[4]Department of Ocean in the Earth System, Max Planck Institute for Meteorology, Hamburg, Germany

Abstract

Multi-year droughts strongly impact food production and water management. Thus, predictions for the next decade are required for decision makers. This study analyzes the decadal prediction skill of the Global Precipitation Climatology Centre Drought Index (GPCC-DI) and its components, namely the Standardized Precipitation Index (SPI-DWD) adapted by the German Meteorological Service (Deutscher Wetterdienst, DWD) and the Standardized Precipitation Evapotranspiration Index (SPEI) within the German global decadal prediction system. The decadal predictions are recalibrated. The prediction skills of the two prediction types ensemble mean predictions and probabilistic predictions are evaluated against those of the commonly applied reference predictions observed climatology and uninitialized simulations. The evaluation of 4-year mean droughts for the lead-year period 1–4 at 5° spatial resolution shows high prediction skills for the SPEI in the tropics, especially northern Africa, and several heterogeneously distributed hot spots for the SPI-DWD. The advantage of GPCC-DI is its global coverage, but it hardly enhances the SPI-DWD and SPEI skills. The recalibration clearly enhances ensemble mean prediction skills in slightly improving correlations and in strongly reducing standard deviations as well as large conditional biases in decadal predictions. For probabilistic predictions, impacts of conditional biases and recalibration are less prominent. To meet user requirements decadal drought predictions with higher temporal and spatial resolutions are analyzed. 1-year mean droughts for lead year 1 mostly show smaller prediction skills than 4-year means because of larger small-scale noise, but some regions reveal improved skills due to regional processes predictable at the 1-year time scale, e.g. over the western United States. Drought predictions at 2° resolution show similar spatial skill patterns with enhanced fine-scale structures mostly without losing prediction skill. A user-oriented evaluation of the decadal GPCC-DI prediction for the severe North African drought of 2008–2011 reproduces most observed drought index tendencies in both prediction types, but intensities are often underestimated. Finally, the decadal GPCC-DI prediction for 2018–2021 presents a drought over North Africa and Arabia and wetting over the Northern Hemisphere in both prediction types. For 2018, predicted patterns are similar but with smoothed intensities. In summary, decadal drought prediction skill depends on the indices, time periods, and areas considered. However, the analyzed drought indices can provide skillful high-resolution information for several future time periods and regions meeting user needs for decadal drought predictions.

Keywords: decadal climate prediction, drought index, probabilistic prediction, bias correction, evaluation, user need

1 Introduction

Multi-year large-scale droughts have devastating impacts on society due to their strong influence on water resources and food security as well as economic development (Benson and Clay, 1998), e.g. in the United States in the 1930s (Schubert et al., 2004), in the Sahel in the 1970–1980s (Janicot et al., 1996) or in south-western Asia, southern Europe and the United States in 1998–2002 (Hoerling and Kumar, 2003). Local stakeholders and political decision makers need precise information on extreme events to implement necessary adaptation measures

(Changnon, 2003). Several user workshops held at the German Meteorological Service (Deutscher Wetterdienst, DWD) have revealed that government agencies and international organizations from the forestry, agriculture, humanitarian disaster risk reduction and water management sectors have a strong interest in multi-year drought information on the 1–10-year time scale (https://www.dwd.de/EN/climate_environment/climateresearch/climateprediction/decadalprediction/start_decadalprediction.html).

The still rather new research area of decadal climate predictions (Murphy et al., 2010) focuses on this multi-year to decadal time period, bridging seasonal forecasts and centennial climate projections, which is of particular interest for medium-term planners, managers, and policy

*Corresponding author: Andreas Paxian, Deutscher Wetterdienst, Frankfurter Str. 135, 63067 Offenbach (Main), Germany, email: andreas.paxian@dwd.de

makers (MEEHL et al., 2009). Decadal predictability is thought to be generated from both boundary conditions, such as greenhouse gases and aerosols (VAN OLDEN-BORGH et al., 2012), and the initialization of ocean (MATEI et al., 2012), sea ice and land surface data (BEL-LUCCI et al., 2015). The German research project MiKlip (Decadal Climate Prediction, Mittelfristige Klimaprognosen) develops the German global decadal climate prediction system which offers prediction skill e.g. for European summer temperatures, extra-tropical cyclone tracks, oceanic carbon uptake, and the Quasi-Biennial Oscillation (MAROTZKE et al., 2016). The DWD aims at providing operational decadal climate predictions for the next 1–10 years by the end of 2019, enabling, for instance, multi-year global drought predictions as described in this paper.

Droughts are characterized by a shortage of available water (HEIM, 2002). They can be quantified by means of drought indices describing the anomalies of water availability compared to a long-term climatology (PALMER, 1965). Various indices have been developed based on different purposes and datasets: the Palmer Drought Severity index (PDSI, PALMER, 1965) makes use of many input data, such as precipitation, evapotranspiration, runoff, soil water storage capacity and recharge, and water loss from soil, some of which are not easy to obtain for the whole globe (LLOYD-HUGHES and SAUNDERS, 2002). The PDSI is highly standardized but uses parameterizations for evapotranspiration and several empirical relationships. Another example is the Reconnaissance Drought Index (RDI, TSARIKIS and VANGELIS, 2005) which describes the ratio of precipitation and potential evapotranspiration (PET) and can also be standardized. Yet, PET has to be parameterized and the RDI fails if PET equals zero. The Standardized Precipitation Index (SPI) is defined by dividing the precipitation anomaly by its standard deviation (MCKEE et al., 1993). It cannot consider enhanced temperature and evapotranspiration in a changing climate and fails in arid areas (LLOYDS-HUGHES and SAUNDERS, 2002). However, this problem can be reduced using the SPI-DWD as adapted by DWD (PIETZSCH and BISSOLLI, 2011). The Standardized Precipitation Evapotranspiration Index (SPEI, VICENTE-SERRANO et al., 2010) standardizes the climatic water balance, i.e. precipitation minus PET. Again, the determination of PET is challenging. When applying THORNTHWAITE (1948) for PET parameterization, the SPEI fails in cold areas. As both SPI and SPEI are recommended by the World Meteorological Organization (WMO, 2009), ZIESE et al. (2014) have developed the Global Precipitation Climatology Centre Drought Index (GPCC-DI), which combines SPI-DWD and SPEI. As a result, this new drought index enables past and present-day drought monitoring with nearly global coverage.

The drought indices described are used for operational drought monitoring and generation of outlooks for the next months, e.g. the North American Drought Monitor (https://gis.ncdc.noaa.gov/maps/ncei/drought/na), the SPEI Global Drought Monitor (http://sac.csic.es/spei/) or the GPCC Drought Index Product (ftp://ftp.dwd.de/pub/data/gpcc/html/gpcc_di_doi_download.html). Concerning decadal drought predictions, no operational outlooks exist until now. However, several research studies outline prediction skills for droughts or drought-related variables in certain regions often related to teleconnections to sea surface temperatures (SSTs): RAMESH et al. (2017) find predictability of periods of cool tropical Pacific SSTs that last for several years up to a decade and are linked to prolonged North American droughts, such as the Dust Bowl in the 1930s. Skillful decadal predictions are also possible for soil water storage over North America if soil conditions are properly initialized (CHIKAMOTO et al., 2015). However, limited skill is found in decadal precipitation predictions of nine coarse general circulation models (GCMs) over the continental United States. This can be improved by statistical downscaling for high resolution impact assessments (SALVI et al., 2017). Rainfall in the Sahel is connected to Atlantic, Indian Ocean and Eastern Mediterranean SSTs (DIATTA and FINK, 2014; PAXIAN et al., 2016; PAETH et al., 2017). However, various decadal prediction skills are found in an ensemble of initialized GCMs because of differing ocean initializations and teleconnection patterns (GAETANI and MOHINO, 2013; MARTIN and THORNCROFT, 2014). SHEEN et al. (2017) present skillful predictions of Sahel rainfall on inter-annual to multi-year time scales. For Europe, IONITA et al. (2017) find valuable inter-annual to decadal prediction skills for summer droughts due to lagged relationships to Mediterranean SSTs; REASON et al. (2006) show strong teleconnections between tropical Pacific SSTs and dry spells for some areas in southern Africa.

In light of all this, this study aims at investigating the skill of global decadal drought predictions in applying the GPCC-DI approach and its components SPI-DWD and SPEI to decadal predictions of the German decadal prediction system. The drought predictions are evaluated for 4-year means at 5° spatial resolution. This is the standard configuration the MiKlip community has defined for decadal prediction evaluation. Additionally, higher temporal and spatial resolutions (1-year means and 2° resolution) are investigated for the purpose of meeting user needs for decadal drought predictions. Observed climatology and uninitialized simulations are chosen as reference predictions for evaluation because they are most commonly applied by climate data users on the decadal time scale. Furthermore, recalibration is applied to improve decadal prediction skills. For this reason, Section 2 of the paper presents the model and observational datasets used. Section 3 describes the GPCC-DI method, the recalibration approach and the evaluation strategy applied for decadal predictions. Section 4 shows the skill results for decadal predictions of the input variables and components of the GPCC-DI in standard configuration. This is needed to understand the sources of the GPCC-DI skill (chapter 4.1) as well as the GPCC-DI

skill itself, including the impacts of standard deviation, conditional bias and recalibration (chapter 4.2). The section also presents user-oriented skill analyses in higher resolution and a user-oriented evaluation of a single event (chapter 4.3) as well as the decadal GPCC-DI prediction for the next years (chapter 4.4). Finally, Section 5 summarizes the results, draws major conclusions and gives a short outlook.

2 Data

In order to obtain and evaluate global decadal GPCC-DI predictions simulations from the German global decadal climate prediction system, corresponding uninitialized climate simulations and different observational datasets have been used.

2.1 Decadal climate predictions

The global decadal predictions result from the German decadal climate prediction system in the 'baseline 1' configuration, which means the Max Planck Institute for Meteorology Earth System Model Low Resolution (MPI-ESM-LR) is combined with the 'baseline 1' procedure to initialize the model with observations (Müller et al., 2012; Pohlmann et al., 2013; Marotzke et al., 2016). MPI-ESM-LR is based on the coupled GCM ECHAM6/MPI-OM. Its atmospheric component features a horizontal resolution of T63 (∼ 1.9°) with 47 vertical levels, whereas the ocean has a GR15 (∼1.5°) resolution with 40 levels (Jungclaus et al., 2013; Pohlmann et al., 2013; Stevens et al., 2013). The 'baseline 1' initialization procedure combines full-field initialization in the atmosphere and anomaly initialization in the ocean. It nudges a MPI-ESM-LR assimilation run to temperature, surface pressure, divergence, and vorticity fields from ERA40 atmospheric reanalyses (Uppala et al., 2005) before 1990 and ERA-Interim reanalyses (Dee et al., 2011) after 1990. In the oceanic component, temperature and salinity anomalies from ORAS4 reanalyses (Balmaseda et al., 2013) of the European Centre for Medium-Range Weather Forecasts (ECMWF) are used. Global decadal predictions have been initialized from this MPI-ESM-LR assimilation run on 1 January in every year from 1961 until 2018, i.e. 58 decades, for a simulation period of 10 years. For each decade, a ten member ensemble is simulated using 1-day-lagged initialization for each member (Pohlmann et al., 2013). External forcing has been derived from observations and the RCP4.5 scenario (Representative Concentration Pathway, Moss et al., 2010) before and after 2005, respectively.

2.2 Uninitialized climate simulations

For evaluating global decadal predictions, corresponding uninitialized climate simulations applying the same model configuration and external forcing and covering the same time period are used. In this way the effect of initialization on decadal prediction skill is investigated (Goddard et al., 2013). The uninitialized MPI-ESM-LR simulations combine the 'historical' and 'rcp45' experiments of the fifth phase of the Coupled Model Intercomparison Project (CMIP5, Taylor et al., 2012) before and after 2005, respectively. Ten MPI-ESM-LR ensemble members have been simulated to fit the ensemble size of the decadal climate predictions.

2.3 Observational data

Since this study relies on the GPCC-DI approach developed by Ziese et al. (2014), the same sources for gridded global observations of monthly temperature means and precipitation totals are used. Land-surface precipitation for the time period 1901–2013 at 1° spatial resolution is taken from GPCC Full Data Reanalysis Version 7 (Schneider et al., 2015). This dataset is the most accurate precipitation reanalysis of GPCC; it supports climate model validation analyses and is based on 75,000 global stations with at least 10 years of record length. The precipitation station anomalies have been interpolated to a regular 1° grid, and the corresponding GPCC Climatology V2015 (Meyer-Christoffer et al., 2015) has been added. Concerning temperature, the Climate Prediction Center (CPC) of the National Centers for Environmental Prediction (NCEP) has combined the station data from the Global Historical Climatology Network version 2 (GHCN, Peterson and Vose, 1997) and the Climate Anomaly Monitoring System (CAMS, Ropelewski et al., 1984). For station interpolation, a background climatology and an objective analysis scheme based on Cressman (1959) applying several iterations through the grid at stepwise smaller radii of influence and implemented into the Grid Analysis and Display System (GrADS) have been used. The resulting NOAA CPC GHCN_CAMS dataset (Fan and Van den Dool, 2008) of land-surface temperatures from 1948 to the present at 0.5° spatial resolution is used in this study.

3 Methods

The following sections present the theoretical definition of the GPCC-DI, its application to decadal predictions as well as the recalibration and evaluation approaches used in this study for decadal predictions.

3.1 Definition of the GPCC-DI

As defined by Ziese et al. (2014), the GPCC-DI is determined in two steps (Fig. 1): first, the SPI-DWD and SPEI are calculated per grid box applying Gamma and Log Logistic distribution functions for parameter estimation, respectively. However, even the calculation of the adjusted SPI-DWD (Pietzsch and Bissolli, 2011) fails in very dry areas with monthly precipitation near
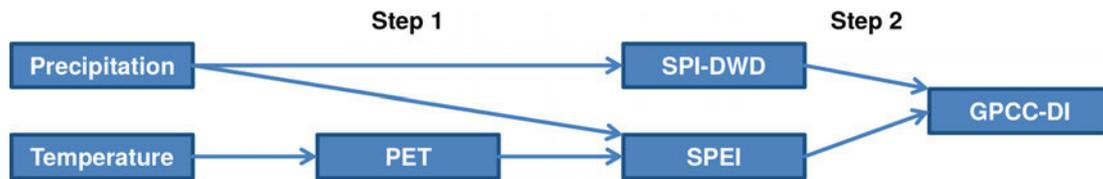
Figure 1: Concept of calculating the GPCC-DI in two steps: determination of SPI-DWD and SPEI from monthly temperature and precipitation (step 1) and averaging of GPCC-DI from monthly SPI-DWD and SPEI (step 2).

0 mm because the maximum of the Gamma distribution is not larger than zero, as required for the determination of the SPI (WU et al., 2007). The PET for determining the SPEI is calculated following THORNTH-WAITE (1948) based on temperature and astronomical data because the availability of detailed global observational input data for other parameterizations, e.g. the FAO Penman-Monteith equation (ALLEN et al., 1998), is restricted (ZIESE et al., 2014). However, the applied algorithm defines PET as 0 mm in cold regions with monthly temperatures close to or below 0 °C, and the calculation of SPEI fails.

Thus, in a second step, the GPCC-DI is computed for each grid box by averaging both indices if available and applying one index if only one is available in order to reach nearly global coverage. This drought index has been developed to enable past and present-day near global drought monitoring (ZIESE et al., 2014) extending other regionally limited approaches. It is transferred to decadal predictions to facilitate homogeneous drought predictions for global comparison studies. It describes the standardized anomaly of available water relating to the standard deviation σ of a standardized normal distribution (similar to SPI and SPEI). Its general interpretation follows that of the SPI (LLOYD-HUGHES and SAUNDERS, 2002): positive values correspond to wet conditions, values between -1 and 1 to normal conditions, and negative values to droughts (ZIESE et al., 2014). Due to differing definitions of different drought indices (see introduction section), the GPCC-DI combination of two indices, however, makes the resulting drought information more robust and less sensitive to outliers: positive values of the GPCC-DI denote the physical conditions causing large amounts of both rainfall and climatic water balance (precipitation minus PET), i.e. high rainfall and low PET; values between -1 and 1 indicate both normal conditions and those of discordant indices as well as negative values those of low rainfall and high PET. This means that the impact of rainfall in defining the GPCC-DI is twice as large as that of PET.

3.2 Calculation of the GPCC-DI for decadal predictions

Since decadal predictability arises from slowly varying components of the climate system, e.g. the ocean, spatial and temporal smoothing is useful in skill analy-

sis in order to reduce (the influence of) unpredictable small-scale noise (RÄISÄNEN and YLHÄISI, 2011, GODDARD et al., 2013). On the other hand, climate data users need high spatial and temporal resolution for climate impact studies. To fulfil these contrasting requirements, the global GPCC-DI decadal predictions are calculated for the MiKlip standard configuration, i.e. 4-year means (January–December) on a 5° spatial grid. In addition, further skill analyses are also performed for 1-year means and on a 2° spatial grid.

For parameter calculation, the following reference time periods are chosen: 1967–1970 to 2010–2013 for 4-year means (Fig. 2) and 1970 to 2013 for 1-year means (Figure S1), which provide the maximum evaluation time periods possible. 1967–1970 and 1970 represent the first time periods available for all possible lead-year periods, i.e. year 1–4 until 7–10 for 4-year means and year 1 until 10 for 1-year means, when decadal predictions start in 1961. 2010–2013 and 2013 are the last time periods available for the evaluation of 4-year and 1-year means because 2013 is the last year in the series of high-quality precipitation observations used in this study.

The input values for parameter calculation are monthly precipitation and monthly PET data computed from temperature and astronomical data following THORNTHWAITE (1948). Since the chosen drought indices are standardized values, all input data originating from decadal predictions, uninitialized simulations, and observations have to be aggregated before running parameter calculation. This aggregation may include temporal smoothing, spatial interpolation via first order conservative remapping, and ensemble averaging of all ten members. The parameters of the SPI and SPEI distribution functions are computed from precipitation and climatic water balance (precipitation minus PET) values of all years in the reference time period, but separately for each dataset (decadal predictions, uninitialized simulations, observations), temporal smoothing (1-year, 4-year means), spatial resolution (5°, 2°), ensemble averaging (all single members for the probabilistic forecast, ensemble mean for the ensemble mean forecast), and grid box. Finally, the estimated SPI and SPEI parameters are used to compute time series of SPI-DWD adapted by DWD following PIETZSCH and BISSOLLI (2011), SPEI, and GPCC-DI for the whole time period of all three datasets.
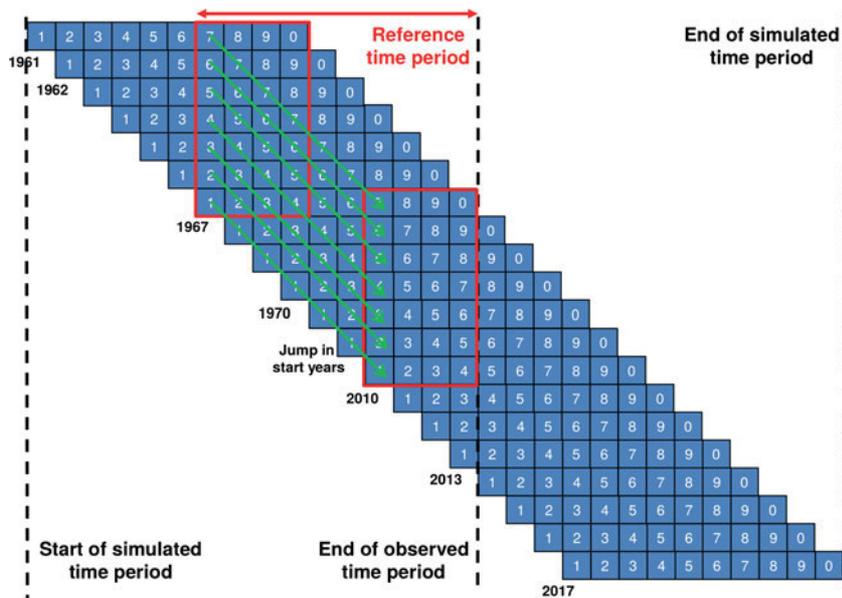
Figure 2: Concept of determining the reference time period for parameter calculation (here 1967–1970 to 2010–2013 for 4-year means, see **Figure S1** for 1-year means). The horizontal bars denote decadal predictions from lead year 1 to lead year 10 (written as 0). The first (1961, 1962, . . . , 1967, . . . , 1970) and last start years (2010, . . . , 2013, . . . , 2017) are highlighted. A jump in start years is indicated in between. The red boxes denote the start and end of the reference time period, and the green arrows describe all years considered for lead year dependent evaluation.

3.3 Recalibration of decadal predictions

For analyzing the prediction skill of temperature, precipitation and PET of all three datasets, the same pre-processing steps to SPI, SPEI, and GPCC-DI have been taken concerning temporal smoothing (1-year and 4-year means), spatial resolution (5° and 2°), ensemble averaging of ten members, and anomaly calculation with respect to the reference time periods for parameter calculation. For the evaluation of decadal predictions (see below), anomalies are calculated separately for model and observations as well as for different lead-year periods. Thus, the model drift from initialization to model climatology, i.e. the lead time dependent bias between model and observations, is implicitly considered (GOD-DARD et al., 2013; BOER et al., 2016) following the suggestion by the International Clivar Project Office (ICPO, 2011) for lead time dependent bias adjustment over all start dates.

Additionally, the Decadal Climate Forecast Recalibration Strategy (DeFoReSt) of PASTERNACK et al. (2018) has been developed within the MiKlip community: It applies the parametric drift correction of KRUSCHKE et al. (2015) considering the drift along lead-year periods by means of a third order polynomial (GANGSTØ et al., 2013) and a linear trend in polynomial parameters over start dates to allow for non-stationary model drifts (KHARIN et al., 2012). This mean or unconditional bias and drift adjustment has been extended by a similar third order parametric adjustment of the conditional bias, i.e. the bias depending on the magnitude, and a quadratic parametric recalibration of the ensemble spread. A cross validation approach leaves out all

decadal predictions initialized within the 10-year prediction period as training data for the estimation of the correction parameters of a certain 10-year decadal prediction. The method reveals improved prediction skill compared to the simple lead time dependent bias adjustment (PASTERNACK et al., 2018).

This recalibration approach has been applied to all considered variables and indices. However, the approach is univariate and cannot assure the consistency of corrected precipitation and temperature (or PET) values as input variables of the GPCC-DI. Thus, the SPI-DWD, SPEI, and GPCC-DI have been calculated from consistent original model data and corrected after index calculation. Note that the correction of the ensemble spread only influences the probabilistic forecasts but has no impact on the ensemble mean forecasts because the ensemble mean is not changed.

3.4 Evaluation of decadal predictions

For each variable and index, the skill of decadal predictions in reproducing past observations is assessed in comparison to the reference forecasts observed climatology in the reference time period (see Section 3.2), i.e. the mean without trend extrapolation, and uninitialized simulations which are both commonly applied by climate data users on the decadal time scale. The evaluation time period equals the reference time period of parameter estimation. Two different prediction types are evaluated: the ensemble mean prediction and the probabilistic prediction of the ensemble distribution.

Concerning the ensemble mean prediction, the mean squared error skill score (MSESS, MURPHY, 1988; GOD-DARD et al., 2013; KADOW et al., 2016) compares the

mean squared error $MSE_{P,O}$ between decadal predictions P_j and observations O_j over $j = 1, n$ years or start dates to the error $MSE_{R,O}$ between reference forecast R_j and observations O_j (Eq. 3.1):

$$MSESS_{P,R,O} = 1 - \frac{MSE_{P,O}}{MSE_{R,O}},$$

$$\text{with } MSE_{P,O} = \frac{1}{n} \sum_{j=1}^{n} (P_j - O_j)^2 \qquad (3.1)$$

Following MURPHY (1988) and GODDARD et al. (2013) the MSESS with the reference forecast observed climatology can be decomposed into the square of the correlation coefficient $r_{p,o}$ between predictions and observations and the square of the conditional bias which is based on $r_{p,o}$ and the quotient between predicted and observed standard deviations s_p and s_o, respectively (Eq. 3.2). The decomposition of the MSESS calculated with the reference forecast uninitialized simulations is more complex applying correlations and conditional biases between both decadal predictions and observations as well as uninitialized simulations and observations:

$$MSESS_{P,R,O} = r_{P,O}^2 - \left(r_{P,O} - \frac{s_P}{s_O}\right)^2 \qquad (3.2)$$

For the probabilistic prediction, the anomalies of the three input datasets are separated into three terciles or categories of equal frequency describing above normal, normal and below normal conditions with respect to the reference time period. The ranked probability score $RPS_{P,O}$ defines the squared error between predicted and observed cumulative probabilities $P_{j,k}$ and $O_{j,k}$, respectively, over $j = 1, n$ years or start dates and $k = 1$, K categories. $P_{j,k}$ results from the distribution of ensemble members, $O_{j,k} = 0$ if observations lie within a category higher than k and $O_{j,k} = 1$ else. The ranked probability skill score (RPSS, FERRO et al., 2008; WILKS, 2011; KRUSCHKE et al., 2014) compares the $RPS_{P,O}$ of decadal predictions to the $RPS_{R,O}$ of a reference prediction (Eq. 3.3):

$$RPSS_{P,R,O} = 1 - \frac{RPS_{P,O}}{RPS_{R,O}},$$

$$\text{with } RPS_{P,O} = \frac{1}{n} \sum_{j=1}^{n} \sum_{k=1}^{K} (P_{j,k} - O_{j,k})^2 \qquad (3.3)$$

Both skill scores give values larger/smaller than 0 if the decadal prediction is more/ less skillful in reproducing past observations than the reference prediction. If the skill score is 0, both are equally skillful. If it is 1, the decadal prediction perfectly matches past observations. Please note that the two reference predictions may show strong differences, e.g. the uninitialized simulations reveal a large temperature trend in the reference time period due to rising atmospheric greenhouse gas concentrations fitting well to real observations. The averaged observed climatology does not show such a temperature

trend and is less skillful. Since decadal predictions contain information on both long-term trend and initialization, their temperature skill score compared to observed climatology is larger than that with respect to uninitialized simulations which are more skillful due to the included trend. Indeed, the skill score compared to climatology reveals the benefit of long-term trend and initialization and that with respect to uninitialized simulations the advantage of initialization only (see further below).

Significance of the skill scores is tested via bootstrapping, i.e. 500 random samples of equal size are chosen from the reference time period with replacement and evaluated applying a significance level of 95 %. The MSESS and RPSS have been calculated with the MurCSS and ProblEMS software tools of the Central Evaluation System of the MiKlip project for standardized evaluation of decadal predictions, respectively (ILLING et al., 2014).

4   Results

In the results section the decadal prediction skills of the input variables and components of the GPCC-DI are verified in MiKlip standard configuration to clarify the sources of the GPCC-DI skill. Then, the GPCC-DI skill itself and the influences of standard deviation, conditional bias and recalibration are presented. Furthermore, the user-oriented skill analyses in higher temporal and spatial resolution and a user-oriented evaluation of a single event are shown. Finally, the decadal prediction of the GPCC-DI for the next years is given.

4.1   Decadal prediction skill of the GPCC-DI input variables in standard configuration

The decadal prediction skill of recalibrated ensemble mean and probabilistic predictions of the GPCC-DI input variables and components is evaluated in MiKlip standard configuration, i.e. 4-year means for the lead-year period 1–4 at 5° spatial resolution. This evaluation is the background analysis to understand the sources of the decadal prediction skill of the GPCC-DI (see below). Prediction skill often reveals similar patterns for ensemble mean and probabilistic predictions, but strong differences are found in comparison to different reference predictions.

First, recalibrated precipitation predictions (Fig. 3) reveal rather heterogeneous skill patterns, but several hot spot regions can be identified. Both prediction types show some skill with respect to the reference prediction observed climatology in western Africa, northern Europe, and some parts of Asia and North America. Compared to uninitialized simulations skill is even higher, especially in southern Africa, northern Europe, Central Asia, and the Arctic regions, highlighting the impact of initialization on precipitation prediction skill for the next four years. However, prediction skill for precipitation is much lower than for temperature or PET (see below) because precipitation is not clearly linked to the long-term
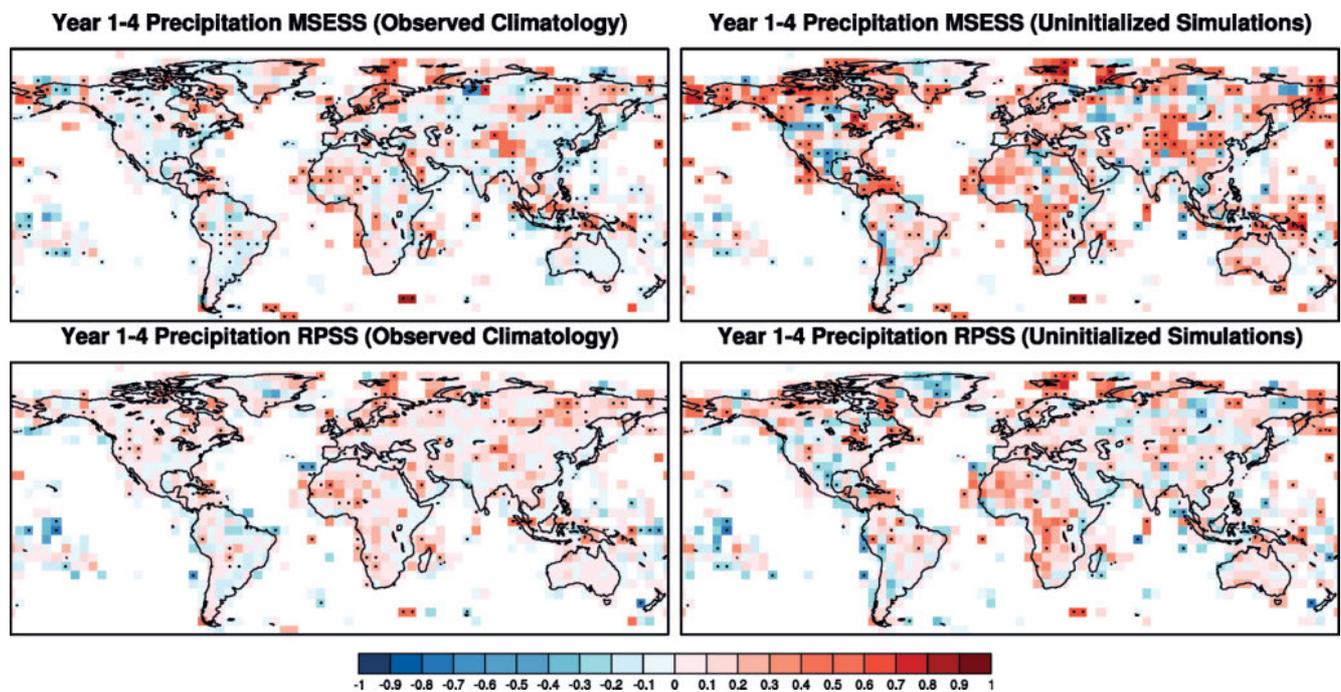
Year 1-4 Precipitation MSESS (Observed Climatology)   **Year 1-4 Precipitation MSESS (Uninitialized Simulations)**

Year 1-4 Precipitation RPSS (Observed Climatology)   **Year 1-4 Precipitation RPSS (Uninitialized Simulations)**

-1 -0.9 -0.8 -0.7 -0.6 -0.5 -0.4 -0.3 -0.2 -0.1 0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1

Figure 3: Decadal prediction skill of MPI-ESM-LR for 4-year mean precipitation predictions for lead-year period 1–4 and 5° spatial resolution: MSESS (top) and RPSS (bottom) compared to the reference forecasts observed climatology (left) and uninitialized simulations (right). Precipitation observations have been taken from the GPCC Full Data Reanalysis Version 7. Dots denote significant skill scores at a significance level of 95 %.

greenhouse gas trend. Furthermore, precipitation reveals small-scale variations in time and space which are less predictable by slow decadal climatic processes, e.g. in the ocean, than more homogeneous temperature variations. Thus, large-scale precipitation variations might be predictable but small-scale noise can hardly be foreseen. Limited decadal prediction skill of global GCMs for precipitation has already been stated in former studies (GAETANI and MOHINO, 2013; MARTIN and THORNCROFT, 2014) and can be improved by higher resolution, e.g. in statistical downscaling approaches (SALVI et al., 2017).

For temperature predictions (Fig. 4, top), high positive skill is widely spread over the whole globe. Highest values are found in the tropics and Arctic regions in comparison to observed climatology. Skill is still high in these regions but disappears in northern Asia and parts of Europe, North America, and Africa with respect to uninitialized simulations, especially for probabilistic predictions. This confirms that a large part of decadal temperature prediction skill is explained by the long-term greenhouse gas trend.

Decadal PET predictions (Fig. 4, bottom) inherit high skill over the whole tropics from their temperature input dataset. PET cannot be calculated for several regions on the northern Hemisphere due to low monthly temperatures. Skill is especially high in whole northern Africa, East Asia, and Brazil compared to observed climatology as well as north-eastern Africa, India, and Brazil with respect to uninitialized simulations. As for temperature, skill strongly decreases for probabilistic

predictions in comparison to the latter reference forecast.

For probabilistic predictions of the SPI-DWD (Fig. 5, top), similar skill patterns to precipitation are found but with slightly less significance. This also holds for ensemble mean predictions compared to observed climatology, but more negative MSESS values can be stated than for precipitation. This is because the standardization of the SPI-DWD increases the standard deviation of the smoothed ensemble mean precipitation to high observed values. This also increases the conditional bias which can be partly reduced by recalibration. Ensemble mean prediction skill compared to uninitialized simulations reveals large positive values nearly over the whole globe because the conditional biases of uninitialized simulations are much larger than those of recalibrated decadal predictions. The impacts of conditional bias and recalibration on prediction skill will be further explained in the evaluation of the GPCC-DI decadal prediction skill.

Probabilistic SPEI predictions (Fig. 5, bottom) inherit skillful patterns from PET but with less intensity, especially in northern Africa and Central and South America compared to observed climatology as well as parts of Africa and America with respect to uninitialized simulations. The ensemble mean predictions compared to observed climatology reveal slightly more negative skill values than PET, but high significant skill over northern Africa, Arabia, and America remains. The comparison to uninitialized simulations shows again large areas of high positive skill but partly shifted com-
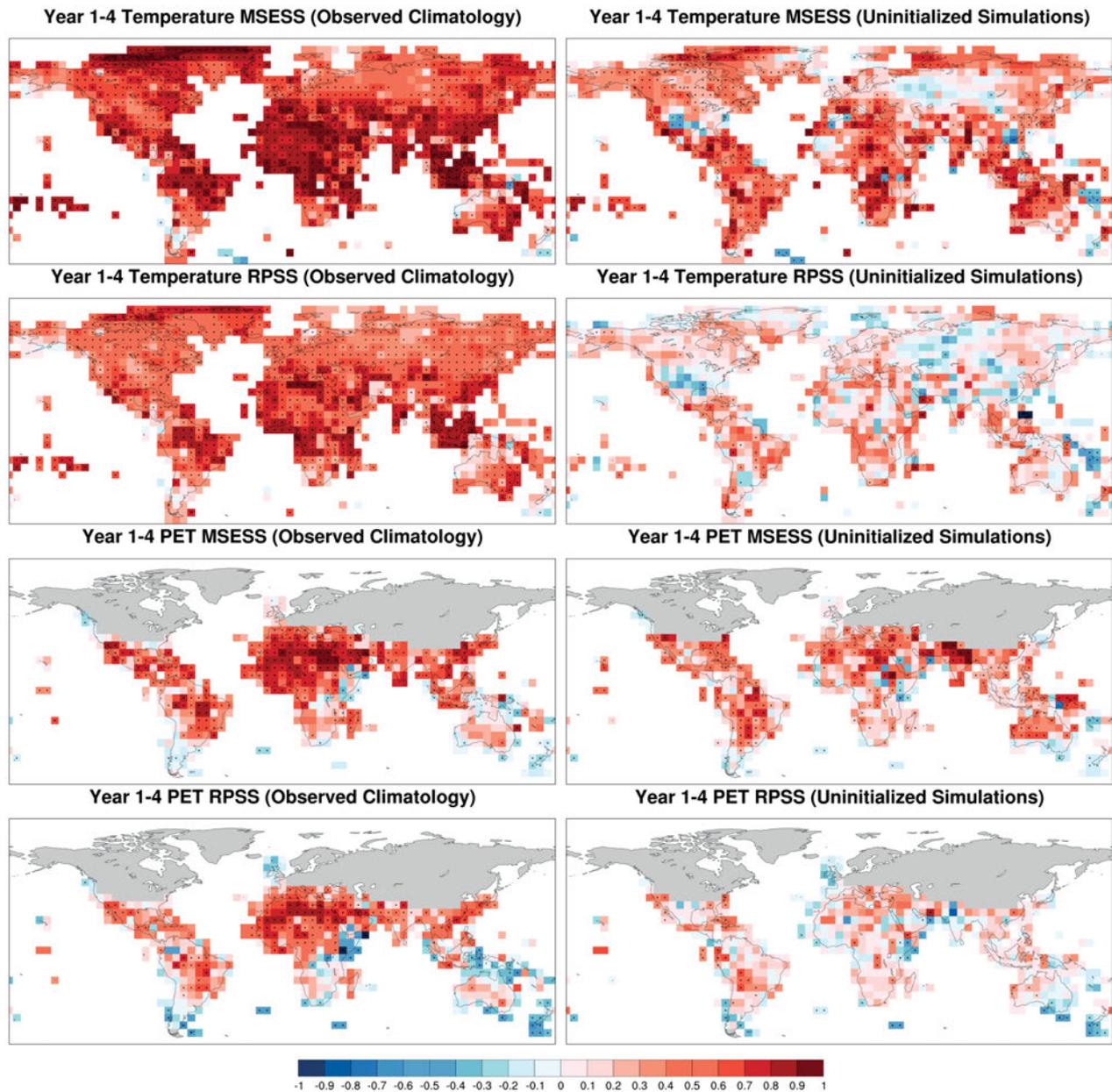
Figure 4: Decadal prediction skill of MPI-ESM-LR for 4-year mean temperature (top) and PET predictions (bottom) for lead-year period 1–4 and 5° spatial resolution: MSESS (first and third row) and RPSS (second and forth row) compared to the reference forecasts observed climatology (left) and uninitialized simulations (right). Temperature observations have been taken from the NOAA CPC GHCN_CAMS dataset. Dots denote significant skill scores at a significance level of 95 %.

pared to PET, e.g. from northern to southern Africa. This can be explained by impacts of the precipitation skill because SPEI is defined by the standardization of precipitation minus PET.

4.2 Decadal prediction skill of the GPCC-DI in standard configuration

The GPCC-DI predictions (Fig. 6) show a global coverage in averaging SPI-DWD and SPEI or taking the one index existing. On the northern Hemisphere GPCC-DI skill is mostly equal to SPI-DWD skill. However, in

northern Africa GPCC-DI and SPEI skills do not exactly match because the SPI-DWD exists for some time periods (but not for all) and influences the GPCC-DI skill. Thus, the GPCC-DI skill compared to observed climatology is more restricted to north-eastern Africa and Arabia. The skill with respect to uninitialized simulations reveals often higher significance and some improvements over north-western Africa. However, the GPCC-DI does not clearly improve the single decadal prediction skills of SPI-DWD and SPEI.

To explain these skill results of recalibrated GPCC-DI predictions, Fig. 7 presents the decomposition of the MSESS. For the comparison to observed
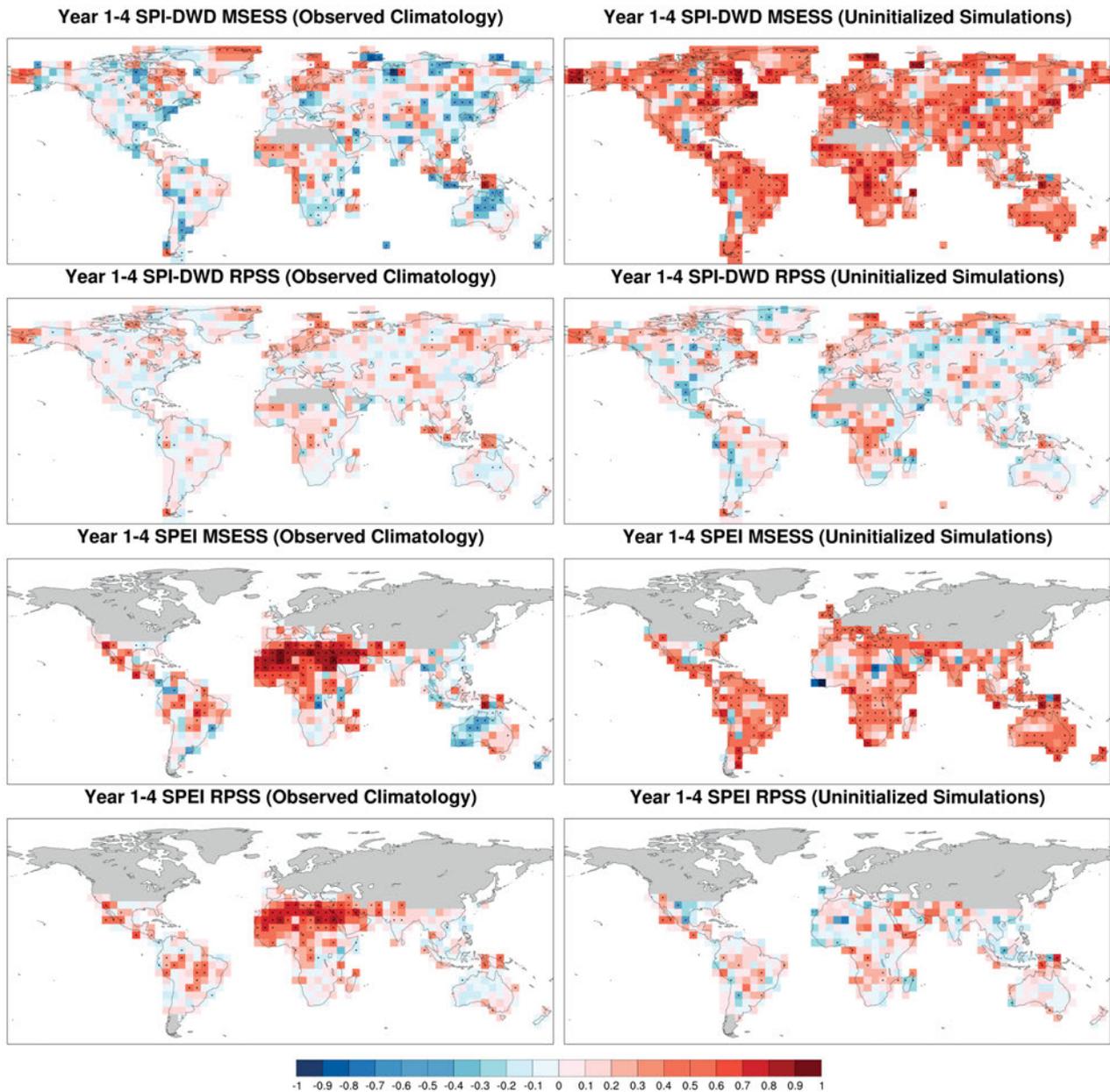
Figure 5: Decadal prediction skill of MPI-ESM-LR for 4-year mean SPI-DWD (top) and SPEI predictions (bottom) for lead-year period 1–4 and 5° spatial resolution: MSESS (first and third row) and RPSS (second and forth row) compared to the reference forecasts observed climatology (left) and uninitialized simulations (right). Precipitation and temperature observations have been taken from the GPCC Full Data Reanalysis Version 7 and the NOAA CPC GHCN_CAMS dataset, respectively. Dots denote significant skill scores at a significance level of 95 %.

climatology, the correlations and conditional biases between decadal predictions and observations are shown. GPCC-DI shows quite high correlations over the whole globe, especially over northern Africa, northern Europe, Central Asia, and Brazil. Conditional biases are slightly negative over most regions because recalibration is used for conditional bias correction (see below). Thus, the MSESS is only slightly reduced compared to correlations (see Eq. (3.2), Fig. 6). For the comparison to uninitialized simulations, the differences between the correlations and conditional biases described before and those between uninitialized simulations and observations are

shown, i.e. the gain of decadal predictions with respect to uninitialized simulations. Decadal predictions reveal mostly better correlations to observations than uninitialized simulations, e.g. over Canada, northern Europe, Central Africa, and Central Asia. The gain in conditional biases is very high because recalibrated decadal predictions show only small biases compared to strongly negative conditional biases of uninitialized simulations. Thus, the MSESS is strongly positive over the whole globe (Fig. 6).

Generally, the recalibration of GPCC-DI predictions following PASTERNACK et al. (2018) slightly improves
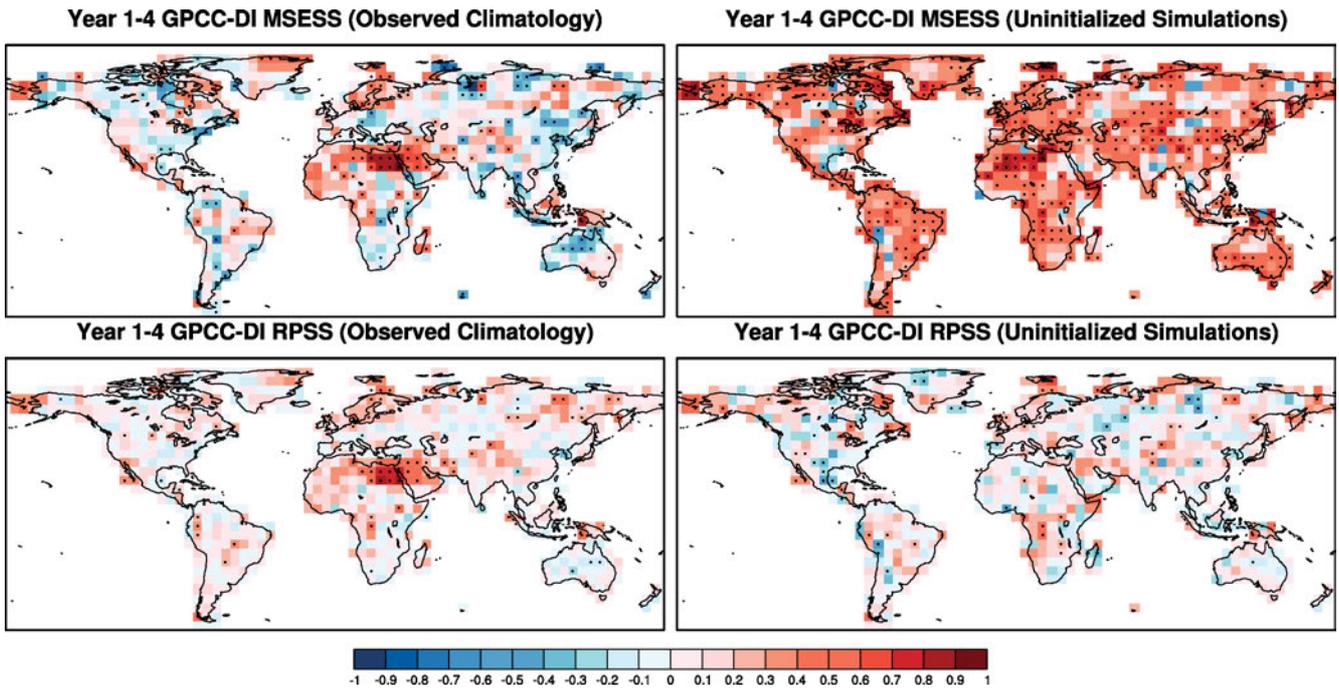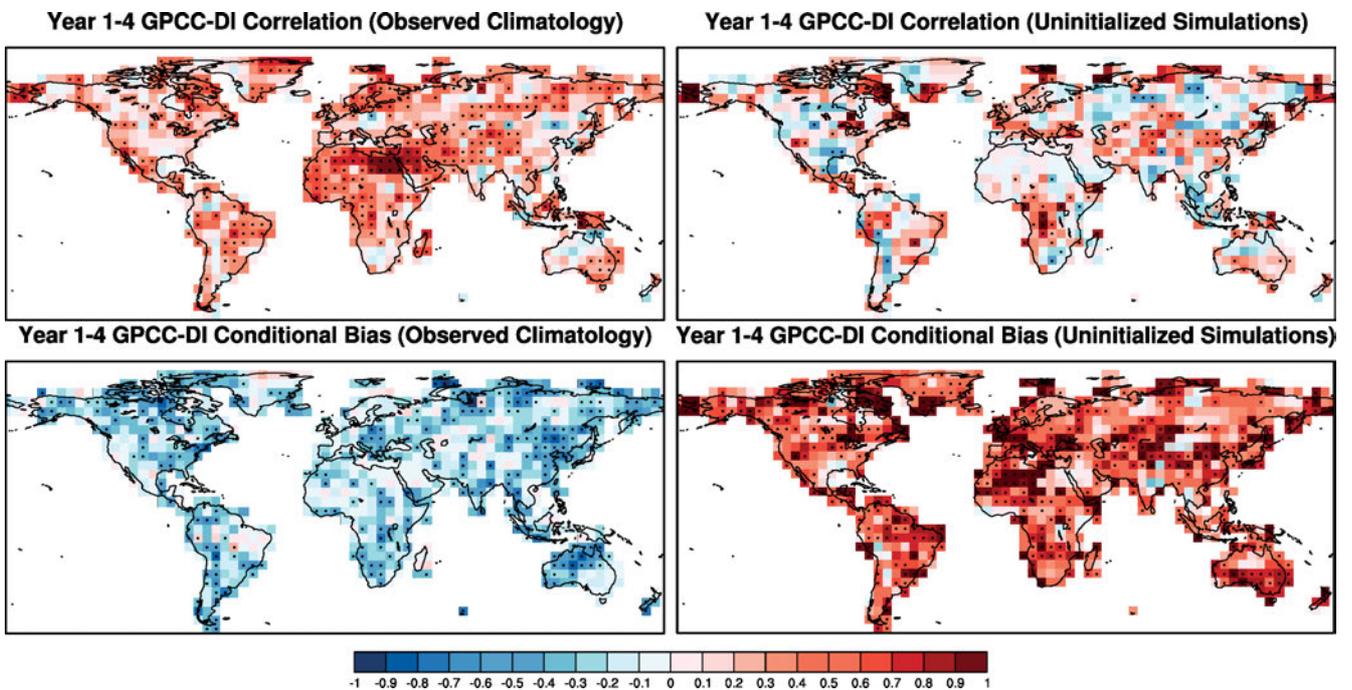
Figure 6: Decadal prediction skill of MPI-ESM-LR for 4-year mean GPCC-DI predictions for lead-year period 1–4 and 5° spatial resolution: MSESS (top) and RPSS (bottom) compared to the reference forecasts observed climatology (left) and uninitialized simulations (right). Precipitation and temperature observations have been taken from the GPCC Full Data Reanalysis Version 7 and the NOAA CPC GHCN_CAMS dataset, respectively. Dots denote significant skill scores at a significance level of 95 %.

Figure 7: Decomposition of the MSESS of MPI-ESM-LR for 4-year mean GPCC-DI predictions for lead-year period 1–4 and 5° spatial resolution: Correlation coefficients (top) and conditional biases (bottom) between decadal predictions and observations for the comparison to the reference forecast observed climatology (left). For the comparison to the reference forecast uninitialized simulations (right), the differences between the correlation coefficients and conditional biases on the left side and those between uninitialized simulations and observations are shown, i.e. the gain of decadal predictions w.r.t. uninitialized simulations. Precipitation and temperature observations have been taken from the GPCC Full Data Reanalysis Version 7 and the NOAA CPC GHCN_CAMS dataset, respectively. Dots denote significance at a significance level of 95 %.

correlations but clearly reduces the large negative conditional biases of original predictions in reducing the standard deviations (Figure S2). This results in strongly improved MSESS values with respect to both reference forecasts without changing spatial skill patterns. Improvements of probabilistic predictions are less pronounced because impacts of conditional biases on tercile probabilities are less prominent when terciles are built separately for predictions and observations (Figure S3). The effect of the ensemble spread correction is found to be rather negligible because the original spread of the MiKlip decadal prediction ensemble is already quite appropriate. In addition, Figure S4 in the supplement presents the corresponding correlation coefficients for all considered input variables and components of the GPCC-DI.

Finally, the GPCC-DI skills of lead-year periods 4–7 and 7–10 have been calculated to analyze the skill development over the full 10-year decadal time period (Figure S5). Prediction skills remain at similar size in most regions and do not clearly drop down until the end of the decade as expected from decreasing impacts of initialization, except for ensemble mean predictions with respect to observed climatology.

4.3 Decadal prediction skill of the GPCC-DI in user-oriented evaluation

User workshops showed that climate data users need decadal predictions with much higher temporal and spatial resolution (https://www.dwd.de/EN/climate_environment/climateresearch/climateprediction/decadalprediction/start_decadalprediction.html). Thus, Fig. 8 reveals the 1-year mean GPCC-DI prediction skills for lead year 1. Compared to lead-year period 1–4, the skill for 1-year means drops down over many regions in both prediction types, but negative skill scores are reduced as well. This is because correlations mostly decrease due to higher impacts of small-scale noise, but conditional biases are reduced because of smaller standard deviations. Some regions reveal larger prediction skills and higher correlations caused by regional-scale processes predictable only for 1-year means, e.g. the western United States due to prediction skill of the 1-year mean Pacific Decadal Oscillation (not shown), southern Africa and Australia reaching significant positive values. Ensemble mean predictions compared to uninitialized simulations reveal positive skill for nearly the whole globe. Thus, user-oriented GPCC-DI predictions with higher temporal resolution are possible. Skill decreases in some areas, but new skillful regions emerge.

Furthermore, Fig. 9 presents 4-year mean GPCC-DI prediction skills for lead-year period 1–4 and a higher spatial resolution of 2°. For all prediction types and reference forecasts, the highly-resolved skill patterns reveal similar spatial distributions to the 5° resolution, but fine-scale features emerge, even if no orographic structures are evident. Prediction skills do not drop down at

higher resolution but remain at a similar level. So do as well conditional biases and correlations. However, significance is reached in some small-scale areas where no significance is found at low resolution, e.g. in some parts of Africa, Asia, and eastern Europe. The zoom over Europe underlines that skillful regional decadal predictions are possible. For the DWD focus area of Germany, ensemble mean predictions compared to uninitialized simulations are skillful over the whole area and probabilistic predictions over the northern and eastern parts. Thus, GPCC-DI predictions at higher spatial resolution for user applications can reveal improved spatial patterns without losing prediction skill.

Again, the corresponding correlation coefficients for the GPCC-DI at different lead-year periods and higher temporal and spatial resolution are presented in the supplement (Figure S6).

Finally, user workshops revealed that climate data users also appreciate the evaluation of prominent single events in addition to the robust statistical skill analysis over many years to build trust in predictions (https://www.dwd.de/EN/climate_environment/climateresearch/climateprediction/decadalprediction/start_decadalprediction.html), even if evaluating single events is strongly impacted by chance. Thus, Fig. 10 presents the evaluation of an exemplary decadal prediction of MPI-ESM-LR for the 4-year mean GPCC-DI of lead-year period 1–4 and 2° spatial resolution with observations. The time period 2008–2011 has been chosen revealing the strongest drought of the whole reference period in north-eastern Africa where decadal prediction skill is very high. The observed GPCC-DI shows strong drought patterns over northern Africa, the Arabian Peninsula, Madagascar, Argentina, Mexico, and northern Canada (Fig. 10, top left). Wetter conditions have been observed over most of North America, north-eastern Europe, northern and central Asia, Indonesia, Australia, and Namibia.

The recalibrated ensemble mean GPCC-DI prediction of MPI-ESM-LR (initialized in 2008) reveals similar drought patterns over north-eastern Africa and the Arabian Peninsula as expected from the high prediction skill in this region (Fig. 10, top right). Most global dry and wet tendencies of observations are captured by the predictions, but intensities of events are often underestimated. This results from high correlations of GPCC-DI tendencies as well as decreased conditional biases and standard deviations due to recalibration. Thus, more skillful GPCC-DI predictions are achieved if the standard deviation of the standardized index is reduced to about 50 % of the observed one (not shown). Major discrepancies in GPCC-DI tendencies can be found over southern Europe, western Russia, Namibia, and Australia.

The recalibrated probabilistic GPCC-DI prediction shows the probabilities of the three terciles below-normal, normal, and above-normal based on the distribution of ensemble members. The tercile boundaries of the observed GPCC-DI equal those of the standardized
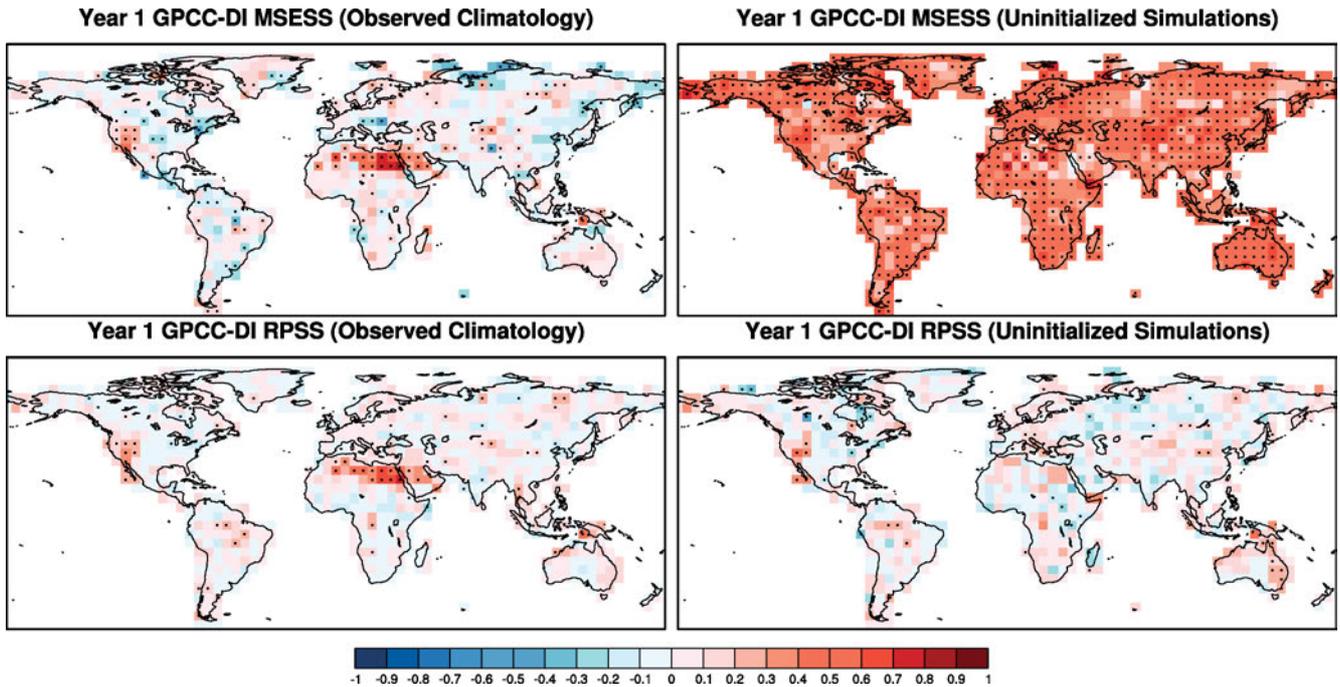
Figure 8: Decadal prediction skill of MPI-ESM-LR for 1-year mean GPCC-DI predictions for lead year 1 and 5° spatial resolution: MSESS (top) and RPSS (bottom) compared to the reference forecasts observed climatology (left) and uninitialized simulations (right). Precipitation and temperature observations have been taken from the GPCC Full Data Reanalysis Version 7 and the NOAA CPC GHCN_CAMS dataset, respectively. Dots denote significant skill scores at a significance level of 95 %.

normal distribution, i.e. around ±0.44. Thus, Fig. 10 (top left) approximately displays the three observed terciles as brown, white and green sectors of the color bar. The tercile boundaries of the probabilistic prediction are slightly shifted away from standardized values due to recalibration. However, the probabilities of terciles can be compared because terciles are calculated separately for observations and predictions. The probabilistic prediction reveals a similar distribution of global dry and wet patterns to the ensemble mean prediction, thus, capturing most observed tendencies (Fig. 10, bottom). However, it shows stronger drying than the ensemble mean prediction over western North America and northwestern Africa, thus, overestimating observed droughts. It also show stronger wetting over northern Alaska and north-eastern Greenland than the ensemble mean prediction often improving the agreement to observations. Thus, the prediction types reveal rather similar results, but both should be investigated because differences are prominent over some regions as already found in the skill analysis.

4.4 Decadal prediction of the GPCC-DI for next years

The final results section presents a decadal prediction of MPI-ESM-LR (initialized in 2018) for the 4-year mean GPCC-DI of lead-year period 1–4, i.e. years 2018–2021, and 2° spatial resolution (Fig. 11). Recalibrated ensemble mean predictions show a large drought over northeastern Africa and the Arabian Peninsula and small drying patterns over West and Central Africa as well as

South America. Wetting is predicted for many northern Hemispheric regions, especially for Greenland and parts of Siberia. The recalibrated probabilistic predictions generally confirm these findings, however, reveal stronger drying probabilities over north-western and Central Africa, South America and India as well as more intense wetting tendencies over northern Europe, Siberia and Indonesia.

A decadal prediction for the 1-year mean GPCC-DI of lead year 1, i.e. year 2018, is shown in Figure S7. Similar drying and wetting patterns to 2018–2021 are found but more smoothed with less intensities or probabilities of occurrence. The only prominent feature is the drying over northern Africa and the Arabian Peninsula. The observed GPCC Drought Index Product (ftp://ftp.dwd.de/pub/data/gpcc/html/gpcc_di_doi_download.html) averaged over January to June 2018 indicates the predicted drought over northern Africa and Arabia as well as wetting over many northern Hemispheric regions (not shown). In contrast to predictions, stronger drying has been observed over eastern Australia and wetting over the United States, the Mediterranean area, and East Africa. The prominent summer drought over Central Europe could not be found in both observed and predicted GPCC-DI. However, please note that until now only the mean of six observed months of the year 2018 is available.

Generally, we recommend climate data users to apply decadal drought predictions only for those time periods and regions where significant positive skill is found in the past.
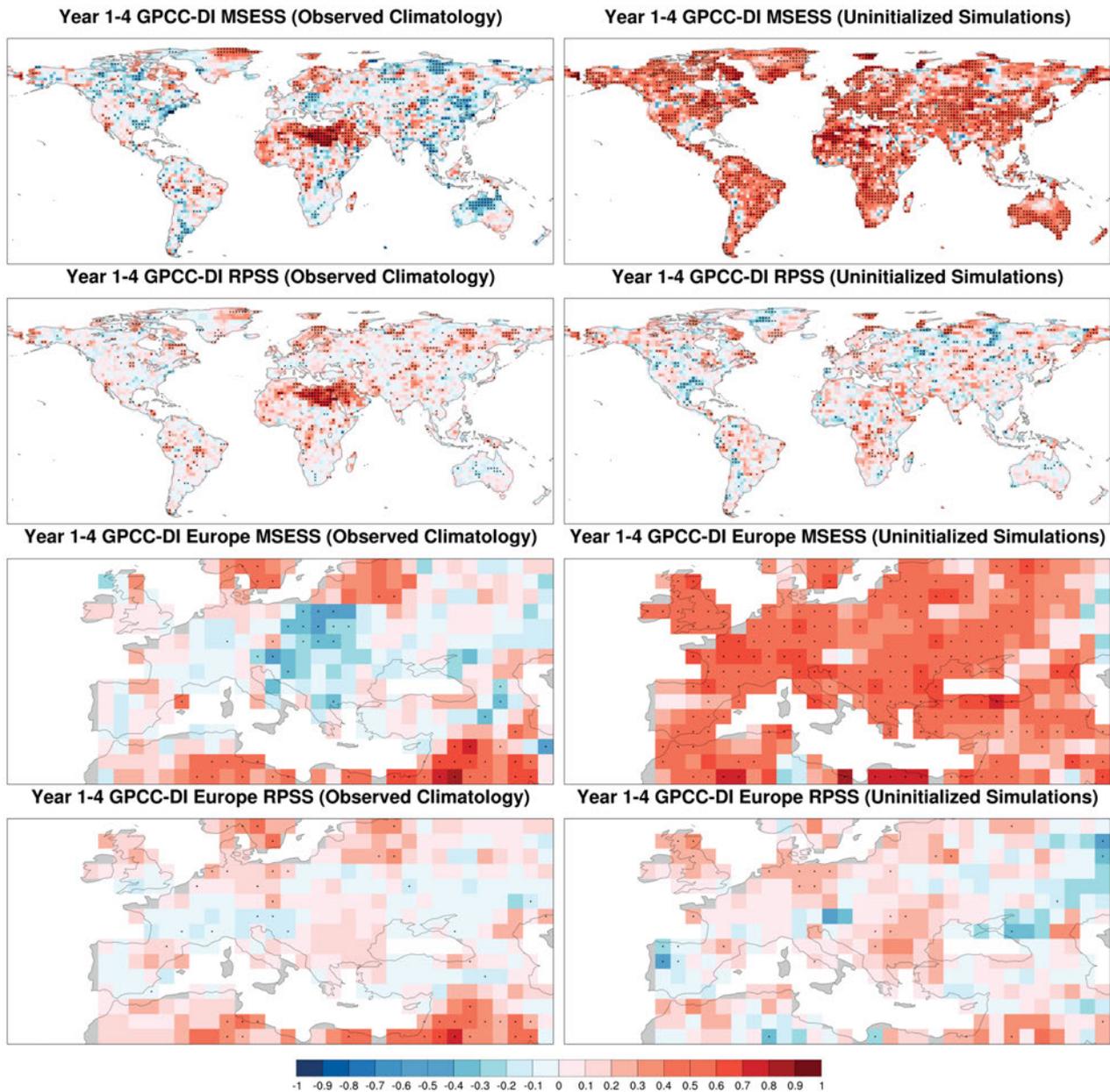
Figure 9: Decadal prediction skill of MPI-ESM-LR for 4-year mean GPCC-DI predictions for lead year 1–4 and 2° spatial resolution at global scale (top) and zoomed-in over Europe (bottom): MSESS (first and third row) and RPSS (second and forth row) compared to the reference forecasts observed climatology (left) and uninitialized simulations (right). Precipitation and temperature observations have been taken from the GPCC Full Data Reanalysis Version 7 and the NOAA CPC GHCN_CAMS dataset, respectively. Dots denote significant skill scores at a significance level of 95 %.

5   Summary and conclusions

This study has analyzed the skill of global decadal drought predictions in determining the GPCC-DI (ZIESE et al., 2014) and its components SPI-DWD and SPEI for decadal predictions of MPI-ESM-LR. Ensemble mean and probabilistic drought predictions have been evaluated via MSESS and RPSS, respectively. Uninitialized MPI-ESM-LR simulations and the observed climatology based on GPCC Full Data Reanalysis Version 7 precipitation and NOAA CPC GHCN_CAMS temperature data have been selected as reference predictions.

For both ensemble mean and probabilistic predictions, the recalibration of PASTERNACK et al. (2018) has clearly improved the standard ICPO (2011) correction. Major results of the decadal drought prediction skill are summarized in the following:

The evaluation in MiKlip standard configuration, i.e. 4-year means for lead-year period 1–4 at 5° spatial resolution, reveals high decadal prediction skill for temperatures on the whole globe and for PET in the tropics, especially in northern Africa, as well as several heterogeneously distributed skill hot spots for precipitation. SPI-DWD and SPEI show similar skill pat-
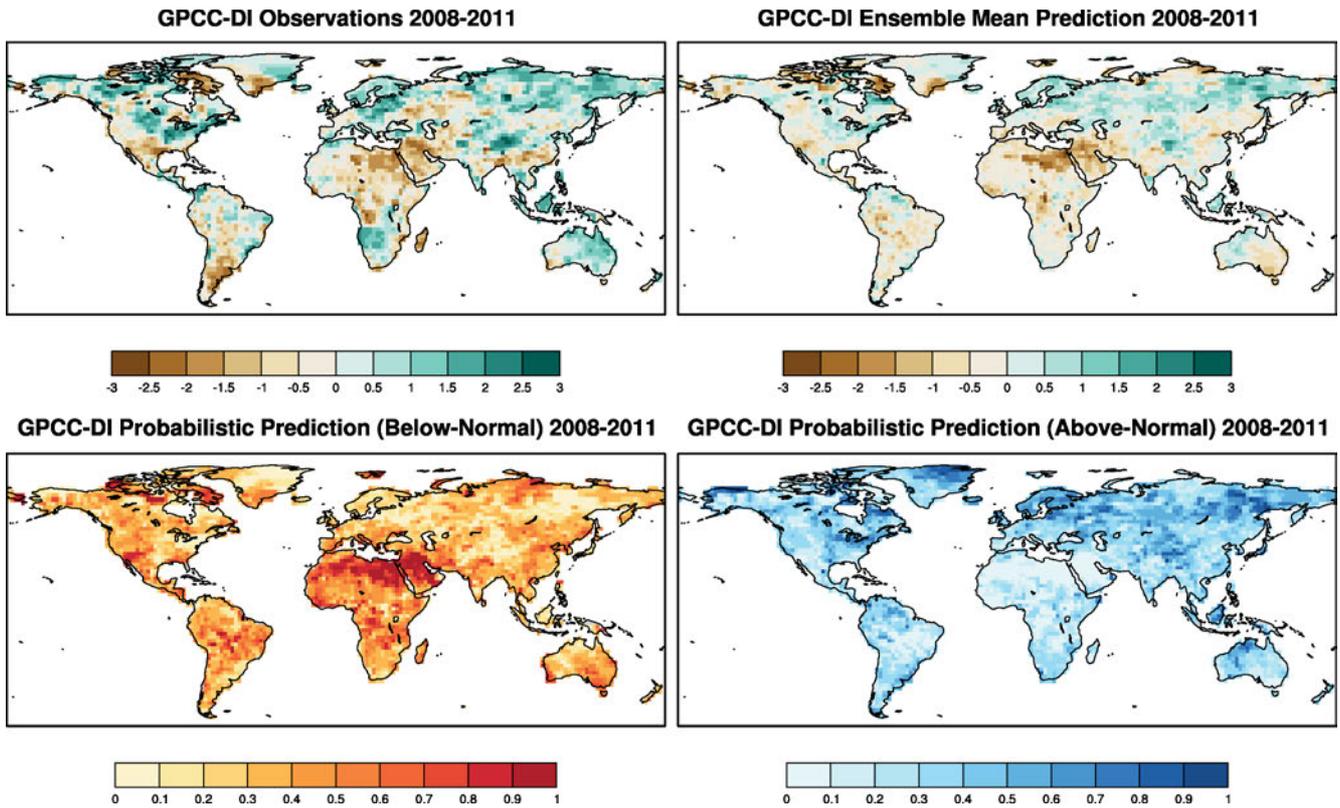
Figure 10: Observed GPCC-DI for 2008–2011 (top left) and decadal prediction of MPI-ESM-LR (initialized in 2008) for 4-year mean GPCC-DI for lead-year period 1–4 and 2° spatial resolution: ensemble mean prediction of GPCC-DI values (top right) and probabilistic predictions of below-normal (bottom left), normal (not shown) and above-normal conditions (bottom right) describing the probabilities of terciles based on the distribution of ensemble members. Precipitation and temperature observations have been taken from the GPCC Full Data Reanalysis Version 7 and the NOAA CPC GHCN_CAMS dataset, respectively.

terns to precipitation and PET, respectively. GPCC-DI presents a global coverage but hardly improves the skill of its components SPI-DWD and SPEI. The recalibration strongly improves prediction skills in slightly enhancing correlations and clearly reducing standard deviations and large negative conditional biases of original predictions. The MSESS compared to uninitialized simulations is strongly positive over many regions over the globe because recalibrated decadal predictions reveal much smaller conditional biases than uninitialized simulations. Impacts of conditional biases are less prominent for probabilistic predictions. In most areas prediction skills remain at similar size until the end of the decade, except for ensemble mean predictions with respect to observed climatology.

Additionally, user-oriented decadal GPCC-DI predictions with higher temporal and spatial resolutions are evaluated. Prediction skills of 1-year means for lead year 1 are smaller than those of 4-year means over many areas due to increased small-scale noise, but some regions reveal reduced negative skills or higher positive skills due to regional processes relevant at the 1-year time scale, e.g. in the western United States. Prediction skills at 2° resolution show similar spatial patterns to the 5° resolution but with fine-scale structures, e.g. over the focus area of Germany. They do not decrease at higher resolution but remain mostly at a similar level.

Furthermore, a user-oriented evaluation of the decadal GPCC-DI prediction of a prominent single event, i.e. the strong north-eastern African drought in 2008–2011, reveals that most observed global drought and wet tendencies are reproduced, but intensities are often underestimated. This results from high correlations but reduced standard deviations. Ensemble mean and probabilistic predictions reveal mostly similar tendencies, except for some regions in North America and Africa.

Finally, a decadal GPCC-DI prediction for 2018–2021 presents a large drought over north-eastern Africa and the Arabian Peninsula and strong wetting over Greenland and parts of Siberia in both ensemble mean and probabilistic predictions. The prediction for 2018 reveals similar patterns to 2018–2021 but more smoothed with less intensities and probabilities of occurrence and agrees well with the already available GPCC-DI observations of January to June in 2018.

Several conclusions can be drawn from these findings: (1) the skill of decadal drought predictions strongly depends on the considered evaluation criteria: drought index, reference prediction, prediction type, time period, and region. SPI-DWD and SPEI show strongly differing results due to various input variables. So do the different reference predictions observed climatology and uninitialized simulations. Ensemble mean and probabilistic predictions often reveal similar skills, but differences
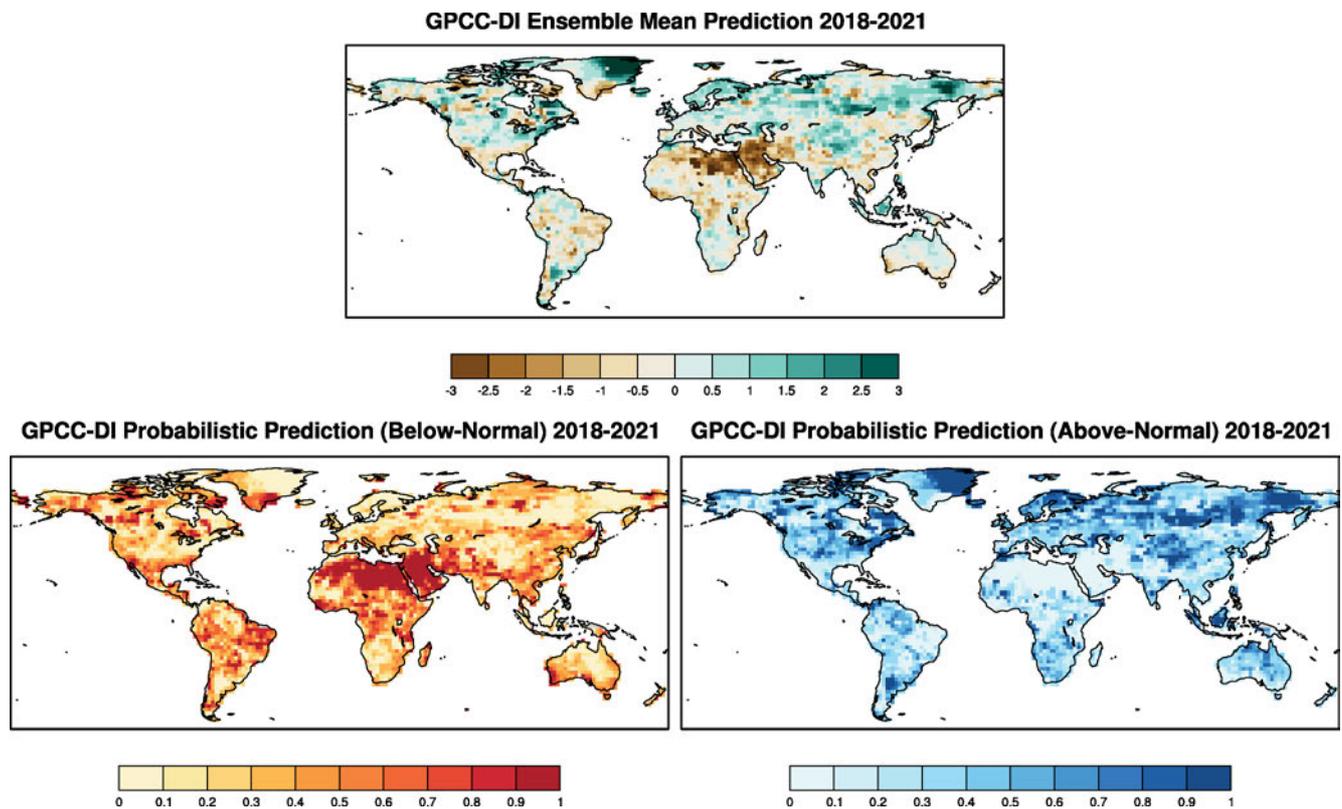
Figure 11: Decadal prediction of MPI-ESM-LR (initialized in 2018) for 4-year mean GPCC-DI for lead-year period 1–4 and 2° spatial resolution: ensemble mean prediction of GPCC-DI values (top) and probabilistic predictions of below-normal (bottom left), normal (not shown) and above-normal conditions (bottom right) describing the probabilities of terciles based on the distribution of ensemble members. Precipitation and temperature observations have been taken from the GPCC Full Data Reanalysis Version 7 and the NOAA CPC GHCN_CAMS dataset, respectively.

emerge with respect to the reference prediction uninitialized simulations which often show higher conditional biases than recalibrated decadal predictions. Drought prediction skill features large variability in time, concerning temporal smoothing, and space but mostly remains at similar size when shifting from the beginning to the end of a decade and from lower to higher spatial resolution. The described results are calculated by robust statistical approaches involving 58 decades, 10 model ensemble members, cross validation, and significance tests via bootstrapping in order to exclude random impacts. Thus, the presented prediction skills just reflect the state of the art of global decadal drought predictions: The applied MPI-ESM-LR model is able to reproduce several physical processes enabling predictions for certain variables in certain time periods and regions, e.g. the PDO prediction skill for lead year 1 related to precipitation and SPI-DWD prediction skill in the western United States. Other physical processes enabling additional prediction skill are not captured by this global decadal prediction system until now. Limited or differing decadal prediction skills of GCMs have already been stated, e.g. for precipitation over the continental United States (Salvi et al., 2017) or in the Sahel (Gaetani and Mohino, 2013; Martin and Thorncroft, 2014).

(2) The drought prediction skill of decadal MPI-ESM-LR predictions is able to significantly improve the commonly used reference predictions observed climatology and uninitialized simulations for several time periods and regions. Thus, based on these findings a user-oriented skill matrix can be built to decide if drought prediction skill for a certain user need (time period, region) with respect to a certain reference prediction applied until now is found and for which evaluation criteria (drought index, prediction type). However, not every single user need can be fulfilled because prediction skill is limited as described. Promising hot spots of drought prediction skill are central South America, northern and central Africa, and northern Europe for 4-year means as well as the western United States, northern Africa, and Australia for 1-year means. In other regions precipitation (Central Asia and the Arctic regions) and PET predictions (whole tropics) are skillful. Some of these skill hot spots confirm recent research studies: Interannual to decadal prediction skill has already been found for North American droughts (Ramesh et al., 2017) and soil water storage (Chikamoto et al., 2015), Sahel rainfall (Sheen et al., 2017), European summer droughts (Ionita et al., 2017), and some southern African dry spells (Reason et al., 2006).

(3) The GPCC-DI has been developed to enable past and present-day drought monitoring with global coverage (ZIESE et al., 2014) and transferred to decadal predictions to facilitate as well near global predictions. However, due to its concept of averaging SPI-DWD and SPEI values its decadal prediction skill hardly improves that of its components. SPI-DWD and SPEI are generally more appropriate to reach high decadal prediction skills in certain regions. The GPCC-DI presents a homogenous drought index for global comparison studies with mostly lower prediction skill. Thus, the choice of considered drought index strongly depends on the kind of application.

(4) The recalibration of PASTERNACK et al. (2018) is able to strongly enhance the ensemble mean prediction skill of decadal predictions in slightly improving correlations and clearly reducing conditional biases and thus, standard deviations. Original variables like precipitation already reveal strongly smoothed ensemble mean standard deviations compared to observations. However, the standard deviation of standardized drought indices is as high as the observed one producing large conditional biases and making the reduction of standard deviations for skill improvement especially necessary. Thus, the user-oriented standardization of variables in order to reach observed high standard deviations is only recommended if correlations are high enough to assure that skill is not deteriorated by large conditional biases. If correlations are low, standardization is not recommended because larger standard deviations would then even worsen the skill. The improvements of the probabilistic prediction skill due to recalibration are less pronounced because impacts of conditional biases and standard deviations are smaller.

Finally, some aspects of this study might motivate further research to reduce uncertainties and improve decadal drought predictions: Concerning the methodical concept, some tests have recently been performed to improve the calculation of SPI-DWD and SPEI by optimizing applied probability density functions and fitting approaches. The SPI-DWD adaptation term could be adapted to large model ensembles to avoid too strong adjustments towards negative indices. The calculation of PET could be improved in applying more satisfied approaches, e.g. the FAO Penman-Monteith equation (ALLEN et al., 1998), and more complex drought indices than SPI and SPEI could be applied to decadal predictions, such as hydrological indices. However, prediction skill of the needed complex input variables has to be found. Concerning spatial resolution, further skill analyses are planned for global decadal predictions of the MPI-ESM-HR (High Resolution, MÜLLER et al., 2018) system at 1° resolution and the downscaled predictions of the regional climate model COSMO-CLM (CCLM, ROCKEL et al., 2008) at 0.5° resolution over Europe and the empirical-statistical downscaling method EPISODES at 0.11° resolution over Germany (KREIENKAMP et al., 2018). Comparisons to other global decadal prediction systems as done in the Decadal Climate Prediction Project (DCPP, BOER et al., 2016) might also be of interest. Finally, the skill of multi-year seasonal drought predictions might be investigated to analyze if seasonally defined droughts are more skillful in certain regions. The major aim is to provide user-oriented decadal drought predictions at high temporal and spatial resolution as an operational climate service at DWD.

Abbrevation list

BMBF	Federal Ministry of Education and Research in Germany (Bundesministerium fuer Bildung und Forschung)
CAMS	Climate Anomaly Monitoring System
CCLM	COSMO-CLM
CMIP5	Fifth Phase of the Coupled Model Intercomparison Project
CPC	Climate Prediction Center
DCPP	Decadal Climate Prediction Project
DeFoReSt	Decadal Climate Forecast Recalibration Strategy
DWD	German Meteorological Service (Deutscher Wetterdienst)
ECMWF	European Centre for Medium-Range Weather Forecasts
GCM	General Circulation Model
GHCN	Global Historical Climatology Network
GPCC-DI	Global Precipitation Climatology Centre Drought Index
ICPO	International Clivar Project Office
MiKlip	Decadal Climate Prediction (Mittelfristige Klimaprognosen)

MPI-ESM-HR	Max Planck Institute for Meteorology Earth System Model High Resolution
MPI-ESM-LR	Max Planck Institute for Meteorology Earth System Model Low Resolution
MSESS	Mean Squared Error Skill Score
NCEP	National Centers for Environmental Prediction
PET	Potential Evapotranspiration
PDO	Pacific Decadal Oscillation
PDSI	Palmer Drought Severity index
RCP	Representative Concentration Pathway
RDI	Reconnaissance Drought Index
RPSS	Ranked Probability Skill Score
SPEI	Standardized Precipitation Evapo-transpiration Index
SPI	Standardized Precipitation Index
SPI-DWD	Standardized Precipitation Index adapted by DWD
WMO	World Meteorological Organization

References

ALLEN, R., L. PEREIRA, D. RAES, M. SMITH, 1998: Crop evapotranspiration – Guidelines for computing crop water requirements. – FAO Irrigation and drainage paper **56**. – FAO – Food and Agriculture Organization of the United Nations, Rome. http://www.fao.org/docrep/X0490E/x0490e00.htm#Contents (accessed at 15.12.2017), ISBN: 92-5-104219-5, 1998.

BALMASEDA, M.A., K. MOGENSEN, A.T. WEAVER, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4. – Quart. J. Roy. Meteor. Soc. **139**, 1132–1161. DOI:10.1002/qj.2063.

BELLUCCI, A., R. HAARSMA, N. BELLOUIN, B. BOOTH, C. CAGNAZZO, B. van den HURK, N. KEENLYSIDE, T. KOENIGK, F. MASSONNET, S. MATERIA, M. WEISS, 2015: Advancements in decadal climate predictability: The role of nonoceanic drivers. – Rev. Geophys. **53**, 165–202. DOI:10.1002/2014RG000473.

BENSON, C., E.J. CLAY, 1998: The impact of drought on sub-Saharan economies. – World Bank Tech. Paper No. **401**, Overseas Development Institute, Regent's College, London, U.K., DOI:10.1596/0-8213-4180-4.

BOER, G.J. , D.M. SMITH, C. CASSOU, F. DOBLAS-REYES, G. DANABASOGLU, B. KIRTMAN, Y. KUSHNIR, M. KIMOTO, G.A. MEEHL, R. MSADEK, W.A. MUELLER, K.E. TAYLOR, F. ZWIERS, M. RIXEN, Y. RUPRICH-ROBERT, R. EADE, 2016: The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. – Geosci. Model Dev. **9**, 3751–3777. DOI:10.5194/gmd-9-3751-2016.

CHANGNON, S.D., 2003: Measure of Economic Impacts of Weather Extremes. – Bull. Amer. Meteor. Soc. **84**, 1231–1235. DOI:10.1175/BAMS-84-9-1231.

CHIKAMOTO, Y., A. TIMMERMANN, S. STEVENSON, P. DINEZIO, S. LANGFORD, 2015: Decadal predictability of soil water, vegetation, and wildfire frequency over North America. – Climate Dyn. **45**, 2213–2235. DOI:10.1007/s00382-015-2469-5.

CRESSMAN, G.F., 1959: An operational objective analysis system. – Mon. Wea. Rev. **87**, 367–374. DOI:10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

DEE, D.P., S.M. UPPALA, A.J. SIMMONS, P. BERRISFORD, P. POLI, S. KOBAYASHI, U. ANDRAE, M.A. BALMASEDA, G. BALSAMO, P. BAUER, P. BECHTOLD, A.C.M. BELJAARS, L. van de BERG, J. BIDLOT, N. BORMANN, C. DELSOL, R. DRAGANI, M. FUENTES, A.J. GEER, L. HAIMBERGER, S.B. HEALY, H. HERSBACH, E.V. HÓLM, L. ISAKSEN, P. KÅLLBERG, M. KÖHLER, M. MATRICARDI, A.P. MCNALLY, B.M. MONGE-SANZ, J.-J. MORCRETTE, B.-K. PARK, C. PEUBEY, P. DE ROSNAY, C. TAVOLATO, J.-N. THÉPAUT, F. VITART, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. – Quart. J. Roy. Meteor. Soc. **137**, 553–597. DOI:10.1002/qj.828.

DIATTA, S., A.H. FINK, 2014: Statistical relationship between remote climate indices and West African monsoon variability. – Int. J. Climatol. **34**, 3348–3367. DOI:10.1002/joc.3912.

FAN, Y., H. van den DOOL, 2008: A global monthly land surface air temperature analysis for 1948–present. – J. Geophys. Res. **113**, D01103. DOI:10.1029/2007JD008470.

FERRO, C.A.T., D.S. RICHARDSON, A.P. WEIGEL, 2008: On the effect of ensemble size on the discrete and continuous ranked probability scores. – Meteor. Appl. **15**, 19–24. DOI:10.1002/met.45.

GAETANI, M., E. MOHINO, 2013: Decadal prediction of the Sahel precipitation in CMIP5 simulations. – J. Climate **26**, 7708–7719. DOI:10.1175/JCLI-D-12-00635.1.

GANGSTØ, R., A.P. WEIGEL, M.A. LINIGER, C. APPENZELLER, 2013: Methodological aspects of the validation of decadal predictions. – Climate Res. **55**, 181–200. DOI:10.3354/cr01135.

GODDARD, L., A. KUMAR, A. SOLOMON, D. SMITH, G. BOER, P. GONZALEZ, V. KHARIN, W. MERRYFIELD, C. DESER, S.J. MASON, B.P. KIRTMAN, R. MSADEK, R. SUTTON, E. HAWKINS, T. FRICKER, G. HEGERL, C.A.T. FERRO, D.B. STEPHENSON, G.A. MEEHL, T. STOCKDALE, R. BURGMAN, A.M. GREENE, Y. KUSHNIR, M. NEWMAN, J. CARTON, I. FUKUMORI, T. DELWORTH, 2013: A verification framework for interannual-to-decadal predictions experiments. – Climate Dyn. **40**, 245–272. DOI:10.1007/s00382-012-1481-2.

HEIM, R.R., 2002: A review of twentieth-century Drought Indices used in the United States. – Bull. Amer. Meteor. Soc. **83**, 1149–1165. DOI:10.1175/1520-0477(2002)083<1149:AROTDI>2.3.CO;2.

HOERLING, M., A. KUMAR, 2003: The perfect ocean for drought. – Science **299**, 691–694. DOI:10.1126/science.1079053.

ICPO, 2011: Data and bias correction for decadal climate predictions. – In: International CLIVAR Project Office: CLIVAR Publication Series No. **150**, 5 pp. https://eprints.soton.ac.uk/171975/1/ICPO150_Bias.pdf (accessed at 15.12.2017).

ILLING, S., C. KADOW, O. KUNST, U. CUBASCH, 2014: MurCSS: A Tool for Standardized Evaluation of Decadal Hindcast Systems. – J. Open Res. Softw.'2, p.e24. DOI:10.5334/jors.bf.

IONITA, M., L.M. TALLAKSEN, D.G. KINGSTON, J.H. STAGGE, G. LAAHA, H.A.J. van LANEN, P. SCHOLZ, S.M. CHELCEA, K. HASLINGER, 2017: The European 2015 drought from a climatological perspective. – Hydrol. Earth Sys. Sci. **21**, 1397–1419. DOI:10.5194/hess-21-1397-2017.

JANICOT, S., V. MORON, B. FONTAINE, 1996: Sahel drought and ENSO dynamics. – Geophys. Res. Lett. **23**, 515–518. DOI:10.1029/96GL00246.

JUNGCLAUS, J.H., N. FISCHER, H. HAAK, K. LOHMANN, J. MAROTZKE, D. MATEI, U. MIKOLAJEWICZ, D. NOTZ, J.S. VON STORCH, 2013: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-Earth sys-

tem model. – J. Adv. Model. Earth Syst. **5**, 422–446. DOI: 10.1002/jame.20023.

KADOW, C., S. ILLING, O. KUNST, H.W. RUST, H. POHLMANN, W.A. MÜLLER, U. CUBASCH, 2016: Evaluation of forecasts by accuracy and spread in the MiKlip decadal climate prediction system. – Meteorol. Z. **25**, 631–643. DOI: 10.1127/metz/2015/0639.

KHARIN, V.V., G.J. BOER, W.J. MERRYFIELD, J.F. SCINOCCA, W.S. LEE, 2012: Statistical adjustment of decadal predictions in a changing climate. – Geophys. Res. Lett. **39**, L19705. DOI: 10.1029/2012GL052647.

KREIENKAMP, F., A. PAXIAN, B. FRÜH, P. LORENZ, C. MATULLA, 2018: Evaluation of the Empirical-Statistical Downscaling method EPISODES. – Climate Dyn. DOI:10.1007/s00382-018-4276-2.

KRUSCHKE, T., H.W. RUST, C. KADOW, G.C. LECKEBUSCH, U. ULBRICH, 2014: Evaluating decadal predictions of northern hemispheric cyclone frequencies. – Tellus A **66**, 22830. DOI:10.3402/tellusa.v66.22830.

KRUSCHKE, T., H.W. RUST, C. KADOW, W.A. MÜLLER, H. POHLMANN, G.C. LECKEBUSCH, U. ULBRICH, 2015: Probabilistic evaluation of decadal prediction skill regarding Northern Hemisphere winter storms. – Meteorol. Z. **25**, 721–738. DOI:10.1127/metz/2015/0641.

LLOYD-HUGHES, B., M.A. SAUNDERS, 2002: A drought climatology for Europe. – Int. J. Climatol. **22**, 1571–1592. DOI: 10.1002/joc.846.

MAROTZKE, J., W. MÜLLER, F. VAMBORG, P. BECKER, U. CUBASCH, H. FELDMANN, F. KASPAR, C. KOTTMEIER, C. MARINI, I. POLKOVA, K. PRÖMMEL, H. RUST, D. STAMMER, U. ULBRICH, C. KADOW, A. KÖHL, J. KRÖGER, T. KRUSCHKE, J. PINTO, H. POHLMANN, M. REYERS, M. SCHRÖDER, F. SIENZ, C. TIMMRECK, M. ZIESE, 2016: MiKlip – a National Research Project on Decadal Climate Prediction. – Bull. Amer. Meteor. Soc. **97**, 2379–2394. DOI:10.1175/BAMS-D-15-00184.1.

MARTIN, E.R., C. THORNCROFT, 2014: Sahel rainfall in multi-model CMIP5 decadal hindcasts. – Geophys. Res. Lett. **41**, 2169–2175. DOI:10.1002/2014GL059338.

MATEI, D., H. POHLMANN, J. JUNGCLAUS, W. MÜLLER, H. HAAK, J. MAROTZKE, 2012: Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model. – J. Climate **25**, 8502–8523. DOI: 10.1175/JCLI-D-11-00633.1.

MCKEE, T.B., N.J. DOESKEN, J. KLIEST, 1993: The relationship of drought frequency and duration to time scales. – In: Proceedings of the 8th Conference of Applied Climatology, 17–22 January, Anaheim. American Meteorological Society, Boston, 179–184.

MEEHL, G.A., L. GODDARD, J. MURPHY, R.J. STOUFFER, G. BOER, G. DANABASOGLU, K. DIXON, M.A. GIORGETTA, A.M. GREENE, E. HAWKINS, G. HEGERL, D. KAROLY, N. KEENLYSIDE, M. KIMOTO, B. KIRTMAN, A. NAVARRA, R. PULWARTY, D. SMITH, D. STAMMER, T. STOCKDALE, 2009: Decadal prediction – Can it be skilful? – Bull. Amer. Meteor. Soc. **90**, 1467–1485. DOI:10.1175/2009BAMS2778.1.

MEYER-CHRISTOFFER, A., A. BECKER, P. FINGER, B. RUDOLF, U. SCHNEIDER, M. ZIESE, 2015: GPCC Climatology Version 2015 at 1.0°: Monthly Land-Surface Precipitation Climatology for Every Month and the Total Year from Rain-Gauges built on GTS-based and Historic Data. – DWD GPCC, DOI: 10.5676/DWD_GPCC/CLIM_M_V2015_100.

MOSS, R.H., J.A. EDMONDS, K.A. HIBBARD, M.R. MANNING, S.K. ROSE, D.P. VAN VUUREN, T.R. CARTER, S. EMORI, M. KAINUMA, T. KRAM, G.A. MEEHL, J.F.B. MITCHELL, N. NAKICENOVIC, K. RIAHI, S.J. SMITH, R.J. STOUFFER, A.M. THOMSON, J.P. WEYANT, T.J. WILBANKS, 2010: The next generation of scenarios for climate change research and assessment. – Nature **463**, 747–756. DOI:10.1038/nature08823.

MÜLLER, W.A., J. BAEHR, H. HAAK, J.H. JUNGCLAUS, J. KRÖGER, D. MATEI, D. NOTZ, H. POHLMANN, J.S. VON STORCH, J. MAROTZKE, 2012: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology. – Geophys. Res. Lett. **39**, L22707. DOI:10.1029/2012GL053326.

MÜLLER, W.A., J.H. JUNGCLAUS, T. MAURITSEN, J. BAEHR, M. BITTNER, R. BUDICH, F. BUNZEL, M. ESCH, R. GHOSH, H. HAAK, T. ILYINA, T. KLEINE, L. KORNBLUEH, H. LI, K. MODALI, D. NOTZ, H. POHLMANN, E. ROECKNER, I. STEMMLER, F. TIAN, J. MAROTZKE, 2018: A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). – J. Adv. Model. Earth Syst. **10**, 1383–1413. DOI:10.1029/2017MS001217.

MURPHY, A.H., 1988: Skill scores based on the mean squared error and their relationships to the correlation coefficient. – Mon. Wea. Rev. **116**, 2417–2424. DOI:10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2.

MURPHY, J., V. KATTSOV, N. KEENLYSIDE, M. KIMOTO, G. MEEHL, V. MEHTA, H. POHLMANN, A. SCAIFE, D. SMITH, 2010: Towards prediction of decadal climate variability and change. - Proc. Environ. Sci. **1**, 287–304. DOI:10.1016/j.proenv.2010.09.018.

PAETH, H., A. PAXIAN, D. SEIN, D. JACOB, H.-J. PANITZ, M. WARSCHER, A. FINK, H. KUNSTMANN, M. BREIL, T. ENGEL, A. KRAUSE, J. TOEDTER, B. AHRENS, 2017: Decadal and multi-year predictability of the West African monsoon and the role of dynamical downscaling. – Meteorol. Z. **26**, 363–377. DOI:10.1127/metz/2017/0811.

PALMER, W.C., 1965: Meteorological Drought. – Research paper no. **45**, U.S. Weather Bureau, Washington D.C., 65 pp. http://www.ncdc.noaa.gov/temp-and-precip/drought/docs/palmer.pdf (accessed at 15.12.2017).

PASTERNACK, A., J. BHEND, M.A. LINIGER, H.W. RUST, W.A. MÜLLER, U. ULBRICH, 2018: Parametric Decadal Climate Forecast Recalibration (DeFoReSt 1.0). – Geosci. Model Dev. **11**, 351–368. DOI:10.5194/gmd-11-351-2018.

PAXIAN, A., D. SEIN, H.-J. PANITZ, M. WARSCHER, M. BREIL, T. ENGEL, J. TÖDTER, A. KRAUSE, W.D. CABOS NARVAEZ, A.H. FINK, B. AHRENS, H. KUNSTMANN, D. JACOB, H. PAETH, 2016: Bias reduction in decadal predictions of West African monsoon rainfall using regional climate models. – J. Geophys. Res. Atmos. **121**, 1715–1735. DOI:10.1002/2015JD024143.

PETERSON, T.C., R.S. VOSE, 1997: An overview of the Global Historical Climatology Network Temperature Database. – Bull. Amer. Meteor. Soc. **78**, 2837–2849. DOI:10.1175/1520-0477(1997)078<2837:AOOTGH>2.0.CO;2.

PIETZSCH, S., P. BISSOLLI, 2011: A modified drought index for WMO RA VI. – Adv. Sci. Res. **6**, 275–279. DOI:10.5194/asr-6-275-2011.

POHLMANN, H., W.A. MÜLLER, K. KULKARNI, M. KAMESWARRAO, D. MATEI, F.S.E. VAMBORG, C. KADOW, S. ILLING, J. MAROTZKE, 2013: Improved forecast skill in the tropics in the new MiKlip decadal climate predictions. – Geophys. Res. Lett. **40**, 5798–5802. DOI:10.1002/2013GL058051.

RÄISÄNEN, J., J.S. YLHÄISI, 2011: How much should climate model output be smoothed in space? – J. Climate **24**, 867–880. DOI:10.1175/2010JCLI3872.1.

RAMESH, N., M.A. CANE, R. SEAGER, D.E. LEE, 2017: Predictability and prediction of persistent cool states of the Tropical Pacific Ocean. – Climate Dyn. **49**, 2291–2307. DOI: 10.1007/s00382-016-3446-3.

REASON, C.J., W. LANDMAN, W. TENNANT, 2006: Seasonal to Decadal Prediction of Southern African Climate and Its Links

with Variability of the Atlantic Ocean. – Bull. Amer. Meteor. Soc. **87**, 941–955, DOI:10.1175/BAMS-87-7-941.

Rockel, B., A. Will, A. Hense, 2008: The regional climate model COSMO-CLM(CCLM). – Meteorol. Z. **17**, 347–348. DOI:10.1127/0941-2948/2008/0309.

Ropelewski, C.F., J.E. Janowiak, M.S. Halpert, 1984: The Climate Anomaly Monitoring System (CAMS). – Climate Analysis Center, NWS, NOAA, Washington DC, 39 pp.

Salvi, K., G. Villarini, G.A. Vecchi, 2017: High resolution decadal precipitation predictions over the continental United States for impacts assessment. – J. Hydrol. **553**, 559–573. DOI:10.1016/j.jhydrol.2017.07.043.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, M. Ziese, 2015: GPCC Full Data Reanalysis Version 7.0 at 1.0°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. – DWD GPCC, DOI:10.5676/DWD_GPCC/FD_M_V7_100.

Schubert, S.D., M.J. Suarez, P.J. Pegion, R.D. Koster, J.T. Bacmeister, 2004: On the Cause of the 1930s Dust Bowl. – Science **19**, 1855–1859, DOI:10.1126/science.1095048.

Sheen, K.L., D.M. Smith, N.J. Dunstone, R., Eade; D.P. Rowell, M. Vellinga, 2017: Skilful prediction of Sahel summer rainfall on inter-annual and multi-year timescales. – Nat. Comm. **8**, 14966. DOI:10.1038/ncomms14966.

Stevens, B., M. Giorgetta, M. Esch, T. Mauritsen, T. Crueger, S. Rast, M. Salzmann, H. Schmidt, J. Bader, K. Block, R. Brokopf, I. Fast, S. Kinne, L. Kornblueh, U. Lohmann, R. Pincus, T. Reichler, E. Roeckner, 2013: Atmospheric component of the MPI-M Earth system model: ECHAM6. – J. Adv. Model. Earth Sys. **5**, 146–172. DOI: 10.1002/jame.20015.

Taylor, K.E., R.J. Stouffer, G.A. Meehl, 2012: An Overview of CMIP5 And The Experiment Design. – Bull. Amer. Meteor. Soc. 93, 485–498. DOI:10.1175/BAMS-D-11-00094.1.

Thornthwaite, C., 1948: An approach towards a rational classification of climate. – Geogr. Rev. **38**, 55–94. DOI:10.2307/210739.

Tsakiris, G., H. Vangelis, 2005: Establishing a Drought Index Incorporating Evapotranspiration. – European Water 9/10, 3–11. http://ewra.net/ew/pdf/EW_2005_9-10_01.pdf (accessed at 15.12.2017).

Uppala, S., P.W. Kallberg, A.J. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J.K. Gibson, J. Haseler, A. Hernandez, G.A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R.P. Allan, E. Andersson, K. Arpe, M.A. Balmaseda, A.C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B.J. Hoskins, L. Isaksen, P.A.E.M. Janssen, R. Jenne, A.P. Mcnally, J.-F. Mahfouf, J.-J. Morcrette, N.A. Rayner, R.W. Saunders, P. Simon, A. Sterl, K.E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, J. Woollen, 2005: The ERA-40 re-analysis. – Quart. J. Roy. Meteor. Soc. **131**, 2961–3012. DOI:10.1256/qj.04.176.

Van Oldenborgh, G.J., F.J. Doblas-Reyes, B. Wouters, W. Hazeleger, 2012: Decadal prediction skill in a multi-model ensemble. – Climate Dyn. **38**, 1263–1280. DOI: 10.1007/s00382-012-1313-4.

Vicente-Serrano, S.M., S.A.S. Begueria, , J.I. Lopez-Moreno, 2010: A multiscalar Drought Index sensitive to global warming: the standardized precipitation evapotranspiration index. – J. Climate **23**, 1696–1718. DOI: 10.1175/2009JCLI2909.1.

Wilks, D., 2011: Statistical methods in the atmospheric sciences. – Elsevier Academic Press, Oxford, Amsterdam, 676 pp.

WMO, 2009: Lincoln Declaration on Drought Indices. – World Meteorological Organization, Lincoln, 2 pp. http://www.wmo.int/pages/prog/wcp/agm/meetings/wies09/documents/Lincoln_Declaration_Drought_Indices.pdf (accessed at 15.12.2017).

Wu, H., M.D. Svoboda, M.J. Hayes, D.A. Wilhite, F. Wen, 2007: Appropriate application of the standardized precipitation index in arid locations and dry seasons. – Int. J. Climatol. **27**, 65–79. DOI:10.1002/joc.1371., 2007.

Ziese, M., U. Schneider, A. Meyer-Christoffer, K. Schamm, J. Vido, P. Finger, P. Bissolli, S. Pietzsch, A. Becker, 2014: The GPCC Drought Index – a new, combined and gridded global drought index. – Earth Syst. Sci. Data **6**, 285–295. DOI:10.5194/essd-6-285-2014.

The pdf version (Adobe Java Script must be enabled) of this paper includes an electronic supplement:
Table of content – Electronic Supplementary Material (ESM)

Figures: S1, S2, S3, S4, S5, S6, S7
Please download the electronic supplement and rename the file extension to .zip (For security reasons Adobe does not allow to embed .zip, .rar, .exe etc. files)