Physics and Chemistry of the Earth 94 (2016) 155–166

Contents lists available at ScienceDirect

Physics and Chemistry of the Earth

journal homepage: www.elsevier.com/locate/pce

Seasonal drought predictability in Portugal using statistical–dynamical techniques

A.F.S. Ribeiro, C.A.L. Pires*

Instituto Dom Luiz (IDL), Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal

ARTICLE INFO

Article history:
Received 10 December 2014
Received in revised form
18 March 2015
Accepted 23 April 2015
Available online 16 May 2015

Keywords:
Long-range forecasts
Predictability
Statistical prediction
Drought index SPI
Precipitation forecasting

ABSTRACT

Atmospheric forecasting and predictability are important to promote adaption and mitigation measures in order to minimize drought impacts. This study estimates hybrid (statistical–dynamical) long-range forecasts of the regional drought index SPI (3-months) over homogeneous regions from mainland Portugal, based on forecasts from the UKMO operational forecasting system, with lead-times up to 6 months. ERA-Interim reanalysis data is used for the purpose of building a set of SPI predictors integrating recent past information prior to the forecast launching. Then, the advantage of combining predictors with both dynamical and statistical background in the prediction of drought conditions at different lags is evaluated. A two-step hybridization procedure is performed, in which both forecasted and observed 500 hPa geopotential height fields are subjected to a PCA in order to use forecasted PCs and persistent PCs as predictors. A second hybridization step consists on a statistical/hybrid downscaling to the regional SPI, based on regression techniques, after the pre-selection of the statistically significant predictors. The SPI forecasts and the added value of combining dynamical and statistical methods are evaluated in cross-validation mode, using the R^2 and binary event scores. Results are obtained for the four seasons and it was found that winter is the most predictable season, and that most of the predictive power is on the large-scale fields from past observations. The hybridization improves the downscaling based on the forecasted PCs, since they provide complementary information (though modest) beyond that of persistent PCs. These findings provide clues about the predictability of the SPI, particularly in Portugal, and may contribute to the predictability of crops yields and to some guidance on users (such as farmers) decision making process.

1. Introduction

Drought episodes in the Iberian Peninsula (IP) have become more frequent and severe (Vicente-Serrano et al., 2014; Sousa et al., 2011), causing several consequences on the socioeconomic and ecological sectors, and leading to major impacts on vegetation (Vicente-Serrano et al., 2013; Gouveia et al., 2012, 2009; Vicente-Serrano, 2007). Extreme episodes, such as the recent 2011–2012, 2004–2005 droughts at Iberia (Trigo et al., 2013; García-Herrera et al., 2007) and the heatwave of 2003 in Europe (Trigo et al., 2005) can seriously affect the agricultural and the hydrologic activities, conditioning the crops and, consequently, requiring adaptation and mitigation measures often entailing high costs. A tendency towards a drier Mediterranean during the 21st century (Giorgi and Lionello, 2008) and the strong variability of the precipitation regime in the IP (Esteban-Parra et al., 1998) enhance the occurrence of droughts and promote the need for drought forecast (either seasonal and climatic), in order to prevent the mentioned impacts.

The predictability of the atmospheric seasonal variability (e.g. Brankovic et al., 1994; e.g. Van den Dool, 1994) has many potential applications, in particular those related to risk management due to drought conditions. Predictions on seasonal time-scales rely on the low-frequency variability of the atmosphere, since the major part of the atmospheric variability show variations on long time scales (seasons and years), which can be predictable to some extent. One of the most important of these low-frequency modes in Europe is the North Atlantic Oscillation (NAO), the dominant mode of winter climate variability over the North Atlantic-European (NAE) region (e.g. Hurrel, 1995), which plays the most important role in precipitation variability over Iberia (Trigo et al., 2004; Rodríguez-Puebla et al., 2001; Pires and Perdigão, 2007). Nevertheless, despite the

* Corresponding author.
E-mail address: clpires@fc.ul.pt (C.A.L. Pires).

http://dx.doi.org/10.1016/j.pce.2015.04.003

much lower forecast skill over extratropics than in tropical regions (e.g. Rowell, 1998; Zhu et al., 2014) it is not negligible, and large-scale features has long been regarded as more predictable than small scales (Van den Dool and Saha, 1990). In a simple way, low-frequency modes can be interpreted as sources of predictability or "windows of opportunity" of potential predictability: the ability to predict well the low-frequency modes, may lead to a good prediction of the regional features, such as droughts.

Long-range forecasts in Europe have been carried out by many operational meteorological centers, such as the European Center for Medium-Range Weather Forecast (ECMWF), the UK Met Office (UKMO) and the Météo-France. Dynamical predictions systems are widely used for long-range forecasts (e.g. Doblas-Reyes et al., 2009; Vitart et al., 2007), as well as statistical approaches (e.g. Folland et al., 2012; Morid et al., 2007). Given the availability of producing forecasts based on both dynamical and statistical systems, the hybrid combination of statistical and dynamical methods emerged as a commonly and useful approach for long-range forecasts in order to take advantage of both approaches using, for instance: probabilistic forecasting based on Gaussian distribution and deterministic forecasting (Déqué and Stroe, 1994), Bayesian joint probability (BJP) (Peng et al., 2014), space—time principal components (ST-PCs) (Vautard et al., 1996; Sarda et al., 1996) — obtained from the Multi-Singular Spectrum Analysis (MSSA) — regression-based models (Kim and Webster, 2010), and Maximum Covariance Analysis (MCA) (Coelho et al., 2006).

A hybrid approach that has been followed relies in a two-step procedure, in which the predictors are first numerically forecasted and then statistically processed, and finally used to empirically estimate the predict and (e.g. Pires, 1996; Sarda et al., 1996). Recently, using the most recent seasonal Integrated Forecast Systems (IFS) from the EMCWF, Vitart (2014) have shown that the skill of the NAO monthly forecasts have been improving, which is particularly interesting given the importance of this index for the precipitation variability over the NAE region. One of the goals of this work is to take advantage of the skill of the dynamical models in forecasting the low-frequency modes (Peng et al., 2014; Kim et al., 2012), using the first leading Principal Components (PCs) of dynamical seasonal forecasts of the geopotential at 500 hPa (first-step of the hybrid procedure), and then perform a statistical downscaling of those predictors to the regional scale, based on regression techniques (second-step of the hybrid procedure). The first-step of the hybrid procedure also includes the PCs of prior (forecasting initialization date) observations of geopotential, gathering a pool of predictors from both dynamical and statistical backgrounds. The influence of each predictor is here assessed, and the added value of hybridizing dynamical and statistical based predictors is evaluated.

Besides the NAO, evidence is found that other main low-frequency modes are linked to the precipitation regime over IP, which is marked by a strong variability enhance the occurrence of droughts (Esteban-Parra et al., 1998), for instance: the Scandinavian Pattern (SCAND) and the East-Atlantic pattern (EA) (e.g. Rodriguez-Puebla et al., 1998); the East-Atlantic/West Russian Pattern (EA/WR) (Vicente-Serrano and López-Moreno, 2006); which suggests possible drought predictive skills of these teleconnection patterns. Another source of predictability is the influence, though weaker, of the El Niño—Southern Oscillation phenomenon (ENSO) on precipitation (e.g. deCastro et al., 2006; Vicente-Serrano, 2005; Pozo-Vázquez et al., 2001; Rodriguez-Puebla et al., 1998).

An important tool when analyzing drought conditions are the drought indices, which allows for the assignment of different degrees of intensity, duration and spatial extent of droughts. Vicente-Serrano (2005) has discussed the influence of El Niño and La Niña events on droughts in the IP using the Standardized Precipitation Index (SPI), developed by McKee et al. (1993), which is based on a probabilistic approach of the precipitation. Another widely employed drought index is the Palmer Drought Severity Index (PDSI), developed by Palmer (1965) which, besides precipitation, is based on the soil water budget. More recently, Vicente-Serrano et al. (2010) has formulated an index that seeks the combined influence of precipitation and evapotranspiration: the Standardized Precipitation Evaporation Index (SPEI).

The present work aims to contribute to the assessment and improvement of the seasonal predictability of the SPI at a time scale of 3-months (predictand) in the mainland Portugal (western Iberia), with lead-times up to 6 months, based on hybridization techniques. Coelho et al. (2006) has developed a single integrated forecast of seasonal of rainfall that gathers prediction information from three dynamical systems and an empirical model. More recently, Coelho and Costa (2010) discusses the challenges of integrating seasonal forecasts to help the end user managing process. Similarly, the present study also intends to contribute to an improved understanding of the relevance of combining dynamical and statistical techniques in the prediction of droughts over the Portuguese territory, in order to provide some guidance on users (such as farmers) decision making process.

In Section 2 the data and the applied methodology is described, focusing on the two-step hybridization procedure, the selection of the best predictors and the formulation of the statistical downscaling. In Section 3 the results of the forecasts (statistical and hybrid) of the drought index are evaluated according to the lead-time, season and region. A discussion and conclusions follow in Section 4.

2. Data and methods description

2.1. Regional drought index

In this study the Standard Precipitation Index (SPI) is computed based on monthly precipitation time series, providing a measure of anomalies in precipitation (McKee et al., 1993). In order to consider homogeneous climate regions for the computation of regional SPI time-series, the recent high-resolution ($0.2^\circ \times 0.2^\circ$) gridded dataset PT02 developed by Belo-Pereira et al. (2011) is considered. This new dataset based on a dense network of observed data over Portugal (several hundred stations) is a rather good representation of the precipitation regime over Portugal, resulting from the interpolation of observations using the ordinary kriging method, after quality control and homogenization steps.

In this study a Principal Component Analysis (PCA) (Hannachi et al., 2007 and references therein) and a subsequent varimax rotation is applied to the PT02 dataset to search for independent climate sub-regions (Fig. 1). The eigenvalues and eigenvectors are computed from the correlation matrix of the monthly precipitation anomalies with respect to annual cycle. The spatial orthogonal patterns (Empirical Orthogonal Functions (EOFs)) given by the eigenvectors, are proportional to the correlation between the monthly precipitation anomalies and the corresponding principal component time series. The first three modes explain much of the variance fraction (83%) of the precipitation anomalies field, which are retained under orthogonal rotation of the standardized PCs. The varimax orthogonal rotation of the three leading standardized PCs gives three independent sub-regions according to the zero loading lines of the second and third rotated EOFs (REOFs) (Figs. 1 and 2). The zero loading lines from the three REOFs (including the first REOF) would give approximately four regions, however the possible north-eastern region (defined by the first REOF) would be rather small than the other regions and therefore we decided to include it into the north-western region.