

Assessing the Skill of Precipitation and Temperature Seasonal Forecasts in Spain: Windows of Opportunity Related to ENSO Events

M. D. FRÍAS

Department of Applied Mathematics and Computer Science, University of Cantabria, Santander, Spain

S. HERRERA

Instituto de Física de Cantabria (CSIC-UC), Santander, Spain

A. S. COFIÑO

Department of Applied Mathematics and Computer Science, University of Cantabria, Santander, Spain

J. M. GUTIÉRREZ

Instituto de Física de Cantabria (CSIC-UC), Santander, Spain

(Manuscript received 3 September 2008, in final form 30 July 2009)

ABSTRACT

The skill of state-of-the-art operational seasonal forecast models in extratropical latitudes is assessed using a multimodel ensemble from the Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) project. In particular, probabilistic forecasts of surface precipitation and maximum temperature in Spain are analyzed using a high-resolution observation gridded dataset (Spain02). To this aim, a simple statistical test based on the observed and predicted tercile anomalies is used. First, the whole period 1960–2000 is considered and it is shown that the only significant skill is found for dry events in autumn. Then, the influence of ENSO events as a potential source of conditional predictability is studied and the validation to strong La Niña or El Niño periods is restricted. Skillful seasonal predictions are found in partial agreement with the observed teleconnections derived from the historical records. On the one hand, predictability is found in spring related to El Niño events for dry events over the south and the Mediterranean coast and for hot events in the southeast areas. In contrast, La Niña drives predictability in winter for dry events over the western part and for hot events in summer over the south and the Mediterranean coast.

This study considers both the direct model outputs and the postprocessed predictions obtained using a statistical downscaling method based on analogs. In general, the use of the downscaling method outperforms the direct output for precipitation, whereas in the case of the temperature no improvement is obtained.

1. Introduction

Seasonal forecasting is a promising research field with enormous potential influence in different socioeconomic sectors, such as agriculture (Challinor et al. 2005), health (Thompson et al. 2006), and energy (García-Morales and Dubus 2007). The goal is predicting climate seasonal anomalies a few months in advance, and its prospects have been discussed many times (Palmer and Anderson

1994; Trenberth 1997; Goddard et al. 2001; Troccoli et al. 2008). In the previous decades, a considerable effort has been made to improve the understanding of the physical phenomena responsible for the observed seasonal variability and to transfer the advances to the operational numerical forecasting systems, such the National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFS; Saha et al. 2006), the Predictive Ocean Atmosphere Model for Australia (POAMA; Wang et al. 2002), or the European Multimodel Seasonal to Interannual Prediction (EUROSIP; Palmer et al. 2004; Vitart et al. 2007). This transfer requires an appropriate assessment of the forecast skill achieved in different regions for different variables to evaluate

Corresponding author address: María Dolores Frías, Department of Applied Mathematics and Computer Science, University of Cantabria, Avd. de los Castros s/n, Santander 39005 Spain.
E-mail: friasmd@unican.es

future model improvements in terms of quality or value metrics [see Kirtman and Pirani (2008) for an informal review]. However, the continuous upgrading of operational systems is an obstacle for quality assessment because of the lack of a homogeneous representative prediction dataset.

The existing hindcasts (or retrospective forecasts) provide an opportunity for statistical quality assessment. For instance, the European Union (EU)-funded project Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER; the precursor of the operational EUROSIP model) considered seven state-of-the-art general circulation models developed in Europe and produced seasonal simulations covering a common period of 21 years (Palmer et al. 2004). The NCEP CFS has also developed a 30-yr hindcast to characterize the model statistics. These hindcast experiments have been used to globally analyze the predictability at seasonal time scales, especially in areas influenced by ENSO activity (Palmer et al. 2004; Gutiérrez et al. 2005). Moreover, although the potential predictability is lower at extratropical latitudes, recent studies have shown certain skill in those regions (e.g., in North America) associated with ENSO teleconnections (see, e.g., Quan et al. 2006), and also with other sources of seasonal predictability, such as the persistence of the North Pacific decadal oscillation (Gershunov and Cayan 2003) or the moisture content (Wang and Kumar 1998; Douville 2004). However, in most of the extratropics, the signals predicted by general circulation models are weak and do not add valuable information over a climatological forecast. In particular, no clear seasonal skill has as yet been found in Europe for surface variables, such as temperature or precipitation, although some preliminary studies have revealed some signals of skill for particular European regions, periods, and variables (Doblas-Reyes et al. 2000; Feddersen and Andersen 2005; Frías et al. 2005; Díez et al. 2005; Shongwe et al. 2007). Moreover, some traces of seasonal skill induced by ENSO events have been reported in Spain (Sordo et al. 2008).

The aim of the present study is to analyze in detail the skill of state-of-the-art seasonal forecasting systems in Spain for both maximum temperature and precipitation for the four seasons using a multimodel ensemble from the DEMETER project. Moreover, we consider both unconditional and conditional (restricted to strong ENSO events) skill, extending the ideas and methods introduced in Sordo et al. (2008). As a result, we report low unconditional skill (only dry episodes in autumn) but significant conditional skill related to ENSO events for certain seasons and variables (spring hot and/or dry events during El Niño and winter dry events or summer

hot events during La Niña), thus providing windows of opportunity for seasonal forecast in Spain.

Note that the Iberian Peninsula is an interesting area for this study because its climate regimes varies from maritime, influenced by the Atlantic (in the northwest), to Mediterranean (in the east). Moreover, the sharp orographic relief gives rise to different mechanisms involved in the observed climate variability, being particularly relevant for precipitation. Several studies have characterized the spatial and temporal variability of the precipitation and temperature over this area (Rodríguez-Puebla et al. 1998; Serrano et al. 1999; Sáenz et al. 2001; Cuadrat and Vide 2007). Furthermore, different rainfall and temperature spatial patterns have been identified in connection with several synoptic oscillation patterns, in particular with ENSO events (Rodó et al. 1997; Rocha 1999; Pozo-Vázquez et al. 2001, 2005).

The study is structured as follows. Section 2 describes the different datasets used in the study and the statistical downscaling approach applied to improve the skill at local scale. In section 3 we assess the skill of precipitation and maximum temperature seasonal predictions from the DEMETER project for the period 1960–2000. In section 4 the observed teleconnections between precipitation/temperature variability in Spain and El Niño/La Niña events are statistically assessed. The performance of the DEMETER seasonal forecasts conditioned to strong El Niño or La Niña events are evaluated in section 5; in this section we also compare the direct outputs from the multimodel ensemble system with the values obtained using the statistical downscaling method. Lastly, the main results and conclusions derived from this manuscript are summarized in section 6.

2. Data and methods

In this study we use a high-resolution $0.2^\circ \times 0.2^\circ$ daily gridded maximum temperature and precipitation dataset over Spain, covering the period 1950–2002 (see Fig. 1b). This grid was built from 2924 stations for precipitation and 865 for maximum temperature, obtained from the Spanish Meteorological Agency (AEMET). Those series were filtered by applying different quality and homogeneity tests to guarantee spatial and temporal homogeneity. [The grid Spain02 is described in detail in Herrera et al. (2009, manuscript submitted to *Int. J. Climatol.*)] We consider the Spain02 data both for validating the skill of seasonal forecasts and for evaluating the teleconnections of precipitation and maximum temperature with ENSO events.

The skill of seasonal forecasts over Spain is studied using the multimodel ensemble system provided by the European DEMETER project (Palmer et al. 2004). This

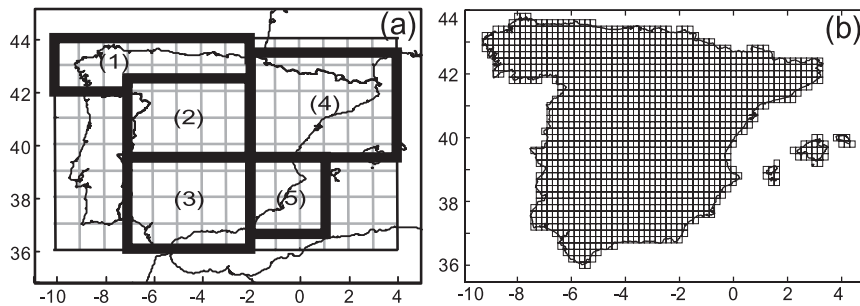


FIG. 1. (a) Area of study (Spain) displaying the uniform latitude–longitude 1° grid and the five regions used in the study: (1) north, (2) center, (3) South, (4) North Mediterranean, and (5) South Mediterranean. (b) High-resolution 0.2° grid of daily interpolated observation records.

project comprises seven global atmosphere–ocean coupled models, each running from an ensemble of nine initial conditions. In this work, simulations from the three models with the longest time series are considered, covering the period 1960–2000: the European Centre for Medium–Range Weather Forecast (ECMWF), the Centre National de Recherches Météorologiques (CNRM, Météo-France), and the Met office (UKMO). We want to remark that these models are the precursors of the operational EUROSIP multimodel seasonal forecasting system, which emerged as a follow up of the DEMETER project. Daily model outputs of precipitation and maximum temperature for the period 1960–2000 were interpolated to the 0.2° grid shown in Fig. 1b to compare with the observations. The DEMETER hindcasts were started from February, May, August, and November initial conditions, each running for six months. Here, the study is performed for winter [December–February (DJF)], spring [March–May (MAM)], summer [June–August (JJA)], and autumn [September–November (SON)]; in all cases we considered the forecasts initialized the month preceding the season (lead times of 1–3 months).

In addition to the direct model outputs, we also consider a statistical downscaling approach to test whether this methodology improves the skill of the predictions at the regional scale needed by end users; some previous results suggest an increment of the skill in some particular cases (Díez et al. 2005; Sordo et al. 2008), but further research is necessary to properly assess the added value of this methodology. Statistical downscaling procedures are particularly convenient in Spain, because the country shows a complex orography and a variety of climates in a relatively small area. We used the statistical downscaling method proposed by Gutiérrez et al. (2004), obtaining local temperature and precipitation forecasts from the synoptic circulation from the DEMETER models. Large-scale daily DEMETER fields for geopotential height Z , temperature T , U and V wind com-

ponents, and specific humidity H for 850 and 500 mb at 0000 UTC on the grids shown in Fig. 1a were used as predictors for the statistical downscaling method for each of the five regions shown in this figure (north, center, south, and North and South Mediterranean). Note that these regions comprise areas with different climatological characteristics (for more details, see Rodríguez-Puebla et al. 1998; Serrano et al. 1999; Muñoz-Díaz and Rodrigo 2004). The statistical model is a weather-type analog approach that applies offline a clustering technique to the daily fields to define the clusters or weather types characterizing the atmospheric circulation. In this work we considered 100 weather types for each region. [See Gutiérrez et al. (2004) for a detailed analysis of the optimal grid size and number of weather types to work with daily data.] These clusters were defined considering the 40-yr ECMWF Re-Analysis (ERA-40; Uppala et al. 2005) from 1957 to 2002 for the same predictors described above. The downscaling was executed independently in the five regions.

3. Assessing the skill of seasonal predictions

In this section we estimate the skill of DEMETER seasonal precipitation and maximum temperature forecasts over Spain for the period 1960–2000 using the probabilistic prediction given by the multimodel ensemble. We use the relative operating characteristic (ROC) skill area (RSA) score as a quality metric. The ROC curve shows the proportion of occurrences that were correctly forecast (the hit rate) versus the proportion of nonoccurrences that were incorrectly forecast [the false-alarm rate; see Jolliffe and Stephenson (2003) for an introduction to probabilistic forecast validation]. The area under the ROC curve (RA) is commonly used to evaluate the performance of probabilistic systems. The RSA is a skill measure, defined as $2RA - 1$, with values ranging from 1 (perfect forecasting system) to -1 (perfectly wrong forecasting system); a value of

zero indicates no skill, compared with a climatological prediction.

In this paper we consider a simple probabilistic prediction scheme based on tercile categories to explore whether the forecasting system captures correctly the observed dry/normal/wet or cold/normal/hot anomalies. The probabilities of these three categories represent a coarse-grained description of a probabilistic forecast and provide a reasonable first choice for seasonal forecast validation in the extratropics, where the signals/spreads are expected to be low/high. Thus, a significant departure from equal-category probabilities is an indication of predictability. Moreover, tercile-based predictions are common in the operational framework to provide seasonal forecast outlooks (see, e.g., Barnston et al. 2003).

The tercile probabilities are estimated from the ensemble prediction as the fraction of the members that fall in each category (counting method). Although more accurate parametric estimators could be applied (Tippett et al. 2007), we used this nonparametric approach to keep the consistency of the analysis for temperature and precipitation in the multimodel framework. Thus, this counting approach is applied to series of precipitation and maximum temperature from the DEMETER multimodel ensemble (3 models \times 9 members) to obtain a probabilistic forecast. For each model, the member and season of the daily predictions were averaged to obtain a single seasonal forecast. Then, the terciles for each member and grid point were computed independently for each season. Finally, for every year, the proportion of the individual ensemble members in each of the three categories was evaluated to estimate the RSA. We also performed the same analysis considering a single climatology per model (joining the ensemble members), obtaining similar results (not shown). Observations were also expressed in terciles using the Spain02 data for the same period. Thus, for each grid point, the 41 years of seasonal values are converted to a series of tercile categories by considering values mentioned earlier, between or below the terciles of the whole period. In this study we concentrate on the predictability of the interannual climate variability. We do not consider the predictability due to global warming, which in most extratropical regions nowadays gives most skill in temperature forecasts. The climate change signal and other low-frequency effects were approximated by linear trends and subtracted from both the predictions and observations before applying the analysis.

Figure 2 shows the spatial distribution of the RSA corresponding to each season (from top to bottom) for precipitation and Fig. 3 for maximum temperature. Only the lower (dry or cold) and upper (wet or hot) terciles,

corresponding to no-neutral anomalies, are displayed. The figures show only those values with a confidence higher than 95%, computed from 1000 bootstrapping runs (Mason and Graham 2002). We considered both the empirical distribution and a normal approximation, obtaining very similar results in both cases. The values shown in the figures correspond to the parametric option, because the Kolmogorov–Smirnov test revealed a good approximation to the normal distribution for almost every grid point. It is remarkable the low skill derived from the DEMETER models, except for precipitation in autumn, which shows a clear significant skill, especially for dry events in the western part of Spain; note that for the other cases, the predictability is quite limited. These results were compared with the RSA values derived from the downscaled precipitation and maximum temperature (not shown). As for the direct output, the skill is in general low, and the RSA spatial distribution is quite similar to the maps shown in Figs. 2 and 3. Thus, as could be expected, statistical downscaling methods do not provide extra information in the absence of signal.

4. Teleconnections with ENSO events

In this section we analyze ENSO events as a source of seasonal predictability, which may provide windows of opportunity for seasonal prediction in Spain, conditioned to the occurrence of the driving events. Although ENSO events affect the weather in many places around the world, the strength of the teleconnections depend strongly on the location and season (Halpert and Ropelewski 1992; Kiladis and Diaz 1989)—being stronger over tropical regions and over particular extratropical areas, such as North America, Australia, Japan, and India. In other regions far from the direct ENSO area of influence, such as Europe, the teleconnection is not so clear and sometimes it is hidden by other dynamical factors (Fraedrich 1990, 1994). The recent strong ENSO events have increased the interest about ENSO's effect on the European climate, especially related to precipitation variability [see Mariotti et al. (2005) for a detailed description of different studies]. For instance, a dipole of positive anomalies (wet episodes) in spring over a zonal belt from England and France to Ukraine, and negative anomalies (dry episodes) in Northern Africa and eastern Spain have been reported in different studies (Kiladis and Diaz 1989; Moron and Ward 1998; van Oldenborgh et al. 2000; Lloyd-Hughes and Saunders 2002; Knippertz et al. 2003; Pozo-Vázquez et al. 2005). The same results were identified in other studies focused on the Iberian Peninsula (Rodó et al. 1997; Rocha 1999). In this region, Pozo-Vázquez et al. (2005) also points out

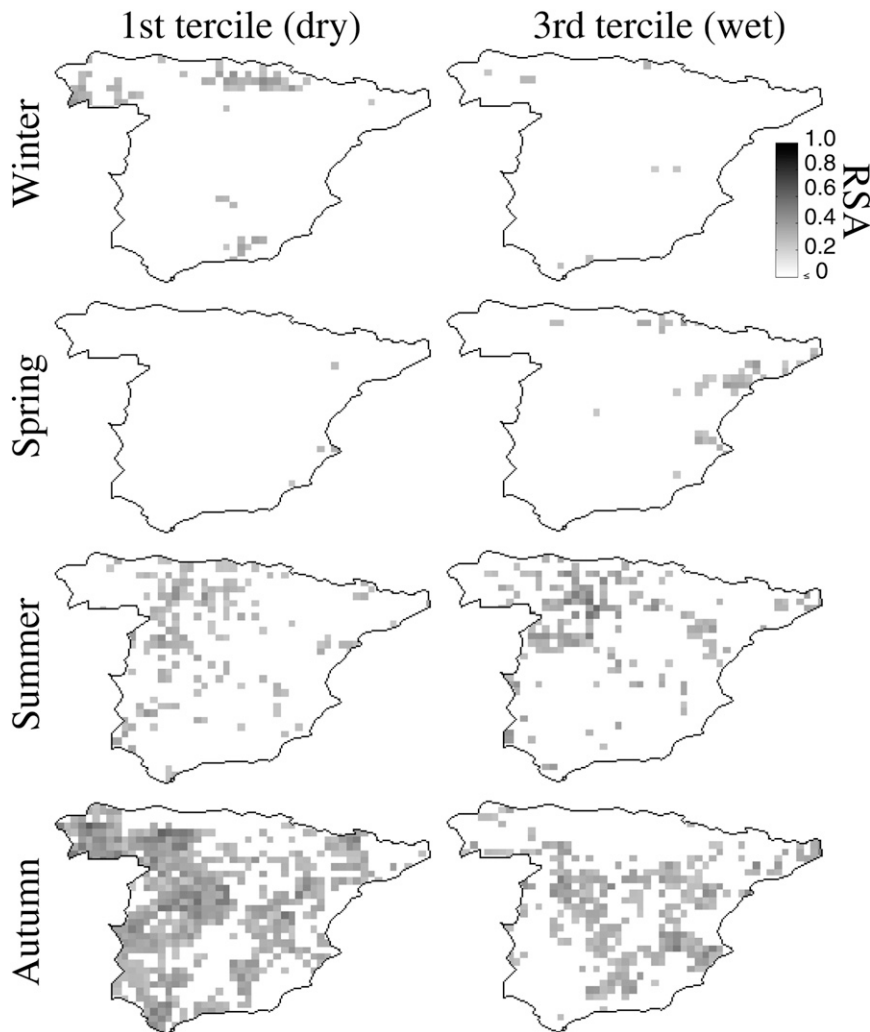


FIG. 2. RSA values for precipitation for the (top to bottom) different seasons using the direct output of the DEMETER models. Only lower and upper terciles are shown (in columns), and only the values significant at 95% confidence level using bootstrapping.

the influence of La Niña events on the precipitation using station data.

The availability of the high-resolution dataset (Spain02) considered in this study allows us to establish more confident relationships between ENSO events and temperature and precipitation over Spain for the four seasons. Thus, we can obtain an estimation of the potential predictability associated with ENSO events at seasonal time scale in Spain. To this end, the conditional climatology of years associated with the strongest La Niña and El Niño events are compared with the global 1950–2002 climatology to find statistically significant deviations. The ENSO years were selected from the Oceanic Niño index (ONI), provided by the Climate Prediction Center (available online at <http://www.cpc.noaa.gov>), which is defined from the winter (DJF) running mean

sea surface temperature (SST) anomalies in El Niño-3.4 region. The most extreme ENSO events for the period 1950 to 2002 were defined as those corresponding to the lower (La Niña) and upper (El Niño) quintile anomalies. In particular, the years 1950, 1956, 1965, 1971, 1974, 1976, 1985, 1989, 1999, and 2000 were obtained for La Niña and the years 1958, 1966, 1969, 1973, 1983, 1987, 1988, 1992, 1995, and 1998 for El Niño conditions. A similar selection of events can be found in Pozo-Vázquez et al. (2005), in which an amplitude-based threshold is applied to the SST normalized series over El Niño-3 region.

The connections with the ENSO events are analyzed for every season using a simple method based on tercile categories previously applied to winter season using station data in Sordo et al. (2008). For each grid point and season, a yearly series of tercile categories is generated

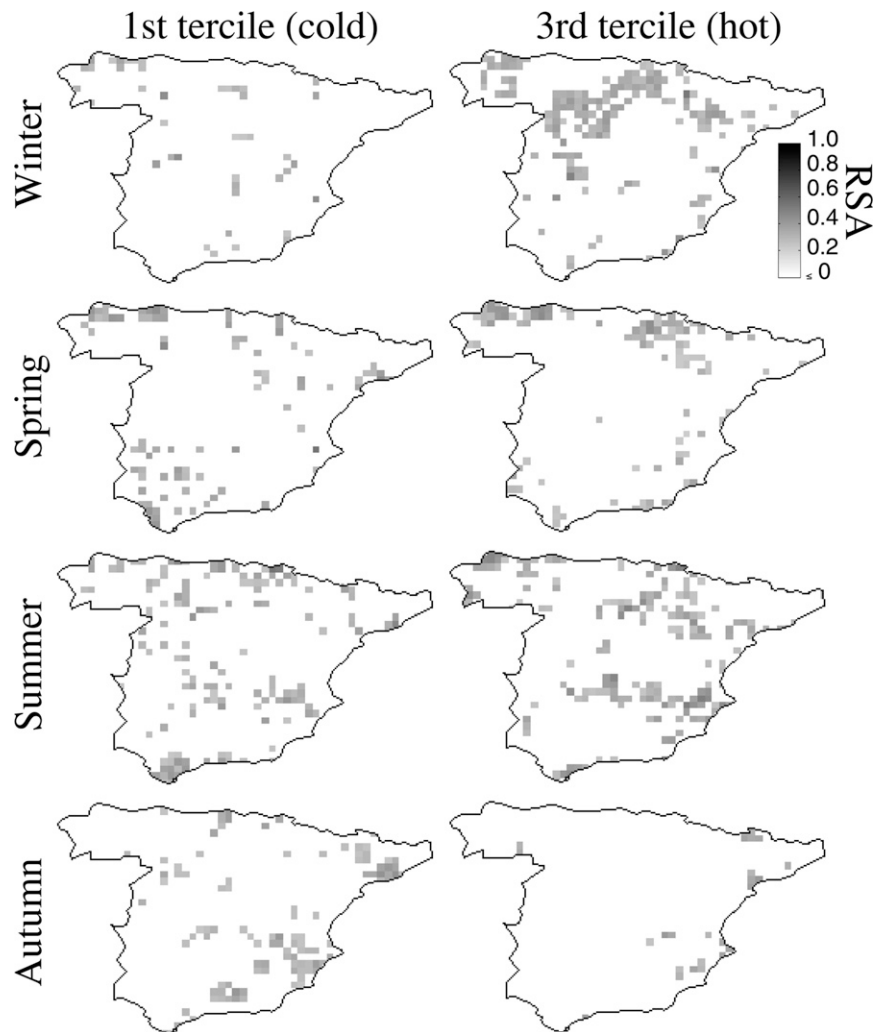


FIG. 3. As in Fig. 2 but for maximum temperature.

considering the tercile thresholds obtained from the whole period. Then, the frequency of each of the tercile categories conditioned to El Niño (La Niña) years is computed, considering only the El Niño (La Niña) years mentioned earlier. The resulting frequency for each category is then compared with the value $1/3$, corresponding to equal-category climatological probabilities. An estimation of the statistical significance of the departure from climatology is given by the confidence level corresponding to the null hypothesis of equal frequencies. This hypothesis was tested for each grid point using the standard test for differences of proportions (see, e.g., Sordo et al. 2008) and also a χ^2 -based test (Price et al. 1998), obtaining similar results. Next, only the significant teleconnections of ENSO events with precipitation and maximum temperature are described.

The influence of El Niño events on precipitation and temperature variability in Spain is shown in Fig. 4; in this

case, only significant teleconnections with spring climatology are found. To simplify the figures, we only show the results corresponding to the lower and upper terciles, because those teleconnections may give some seasonal predictability of unusual climate conditions (droughts, hot waves, among others), which are of great utility in different socioeconomic fields. Figures 4a and 4b represent the frequencies derived from El Niño years for spring precipitation for both terciles. Figures 4c and 4d show the confidence level corresponding to the null hypothesis of equal frequencies (or proportions); that is, $f = 1/3$, where f is the frequency for El Niño events. The values shown in the figures correspond to the χ^2 -based test mentioned earlier, and only values with confidence level larger than 95% are displayed. According to these results, years with an occurrence of El Niño events are related to dry events in spring over the south and the Mediterranean coast. In these regions the values are

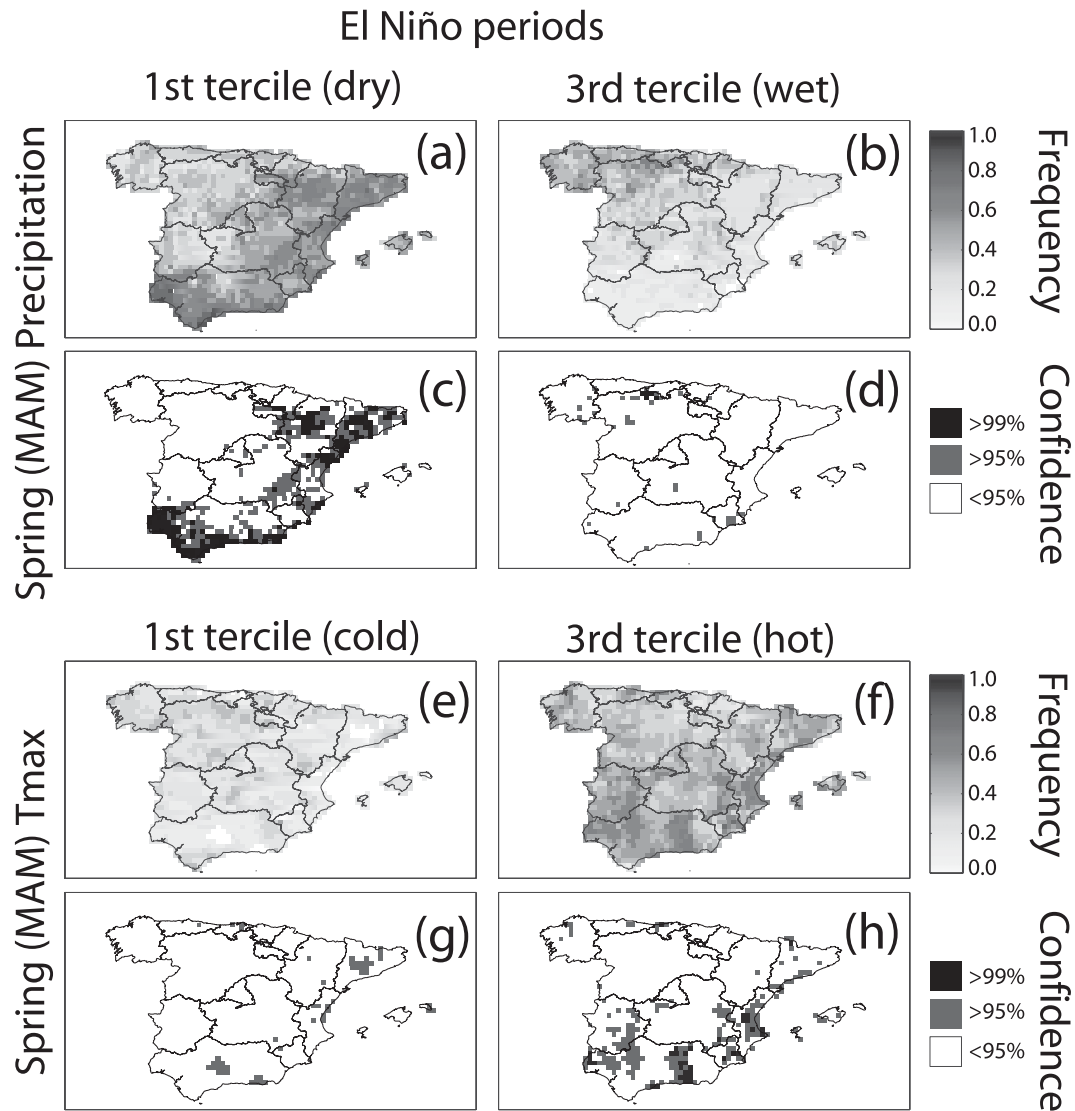


FIG. 4. Spatial distribution of the frequency of the lower and upper terciles during strong El Niño years for (a),(b) precipitation and (e),(f) maximum temperature in spring. (c),(d),(g),(h) The statistical significance is shown for each case.

clearly significant at 95% confidence level over a wide geographical area. This teleconnection has been already reported in several studies as the strongest in Europe (van Oldenborgh et al. 2000; Knippertz et al. 2003). Other studies focused in the Spanish area have also found the influence over this region using a long period of time (around 100 years)—for instance, Rodó et al. (1997) using 17 station data over Spain and Rocha (1999) considering nine grid points from a historical monthly precipitation grid. In contrast, results obtained for autumn using the high-resolution gridded dataset show no significant teleconnection of El Niño with wet events in Spain, as reported by Rocha (1999). The influence of El Niño is also detected on maximum tem-

perature in spring (Figs. 4e–4h). In this case, hot events over the southeastern part of Spain during spring are related to El Niño events (Fig. 4f). The frequencies in this case are also significantly different from $1/3$ at 95% confidence level over a large geographical area (Fig. 4h).

Figure 5 shows the results for La Niña events, with significant teleconnections with winter dry and summer hot events. The most remarkable influence of La Niña is on dry events in winter over the northwestern part of Spain (Fig. 5c). High significance levels are also found for the upper tercile (wet events) for low-frequency regions; thus, in this case, the predictability corresponds to the no occurrence of a wet event or, equivalently, the

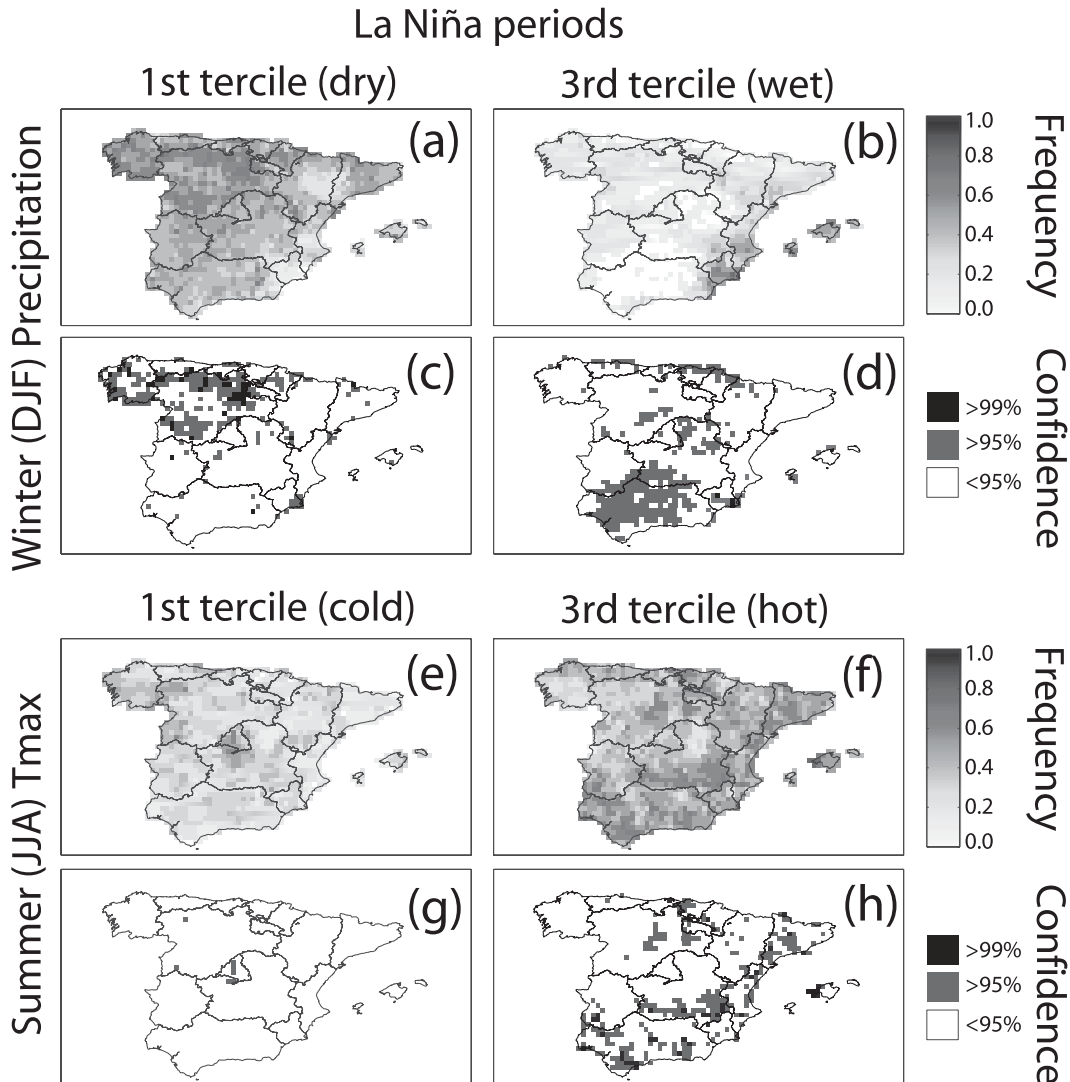


FIG. 5. As in Fig. 4 but for strong La Niña years for winter precipitation and summer maximum temperature.

occurrence of a normal or dry event. Similar results were found by Sordo et al. (2008) using station data from 123 gauges over Spain. Pozo-Vázquez et al. (2005) found a decrease of precipitation in winter over the southwestern and central parts during La Niña events using 48 monthly observed series during the period 1900–98. The domain of influence is located slightly to the north in our study, probably because of the higher resolution dataset used. An influence of La Niña phenomena is also found on maximum temperature. As shown in Fig. 5f, years with an occurrence of La Niña events are related to hot events in summer over the south and the Mediterranean coast.

Some other teleconnections have been also identified in Spain. For instance, higher temperatures in autumn in the Mediterranean coast and also wet events in summer

over the central part of Spain related to El Niño events, or higher than normal temperatures in Cantabrian coast in autumn for La Niña events. However, these teleconnections are weaker than those presented earlier and, as will be shown in the next section, there is not seasonal skill related to them. Therefore, the corresponding maps are not shown.

Note that upper-level circulation is directly influenced by large-scale ENSO teleconnections; however, surface variables such as precipitation or temperature are also influenced by local factors (e.g., orography), conditioning the response of these variables to circulation changes associated with teleconnections. This could explain the regional spatial scale of the teleconnections found in Figs. 4 and 5, corresponding to homogeneous climatological areas (the Mediterranean

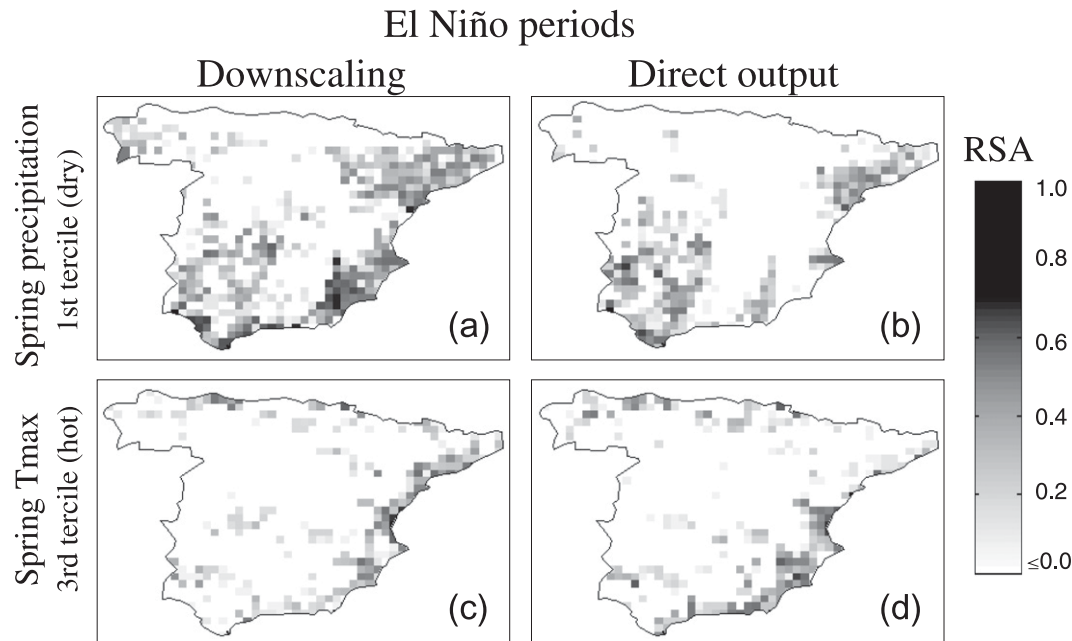


FIG. 6. RSA maps for events related to El Niño phenomena for the (right) direct model output and (left) downscaling method: (a),(b) spring dry events and (c),(d) spring hot events. The RSA values corresponding to the 90% confidence level obtained by bootstrapping with 1000 runs are 0.71, 0.65, 0.79, and 0.75 for (a)–(d), respectively.

coast or the northwestern part of Spain) with different large-scale driving mechanisms. Thus, the use of statistical downscaling methods could play a key role in projecting the large-scale predictability associated with different ENSO teleconnection into surface precipitation and temperature, by taking into account empirical relationships between large-scale and local variables. This possibility is analyzed in the next section.

Lastly, to estimate the statistical field significance of the previous results, we applied the methodologies described in Fraedrich et al. (1995) and Bretherton et al. (1999) to determine the effective spatial degrees of freedom of the grid points, and the method defined for Livezey and Chen (1983) to estimate the percentage of grid points necessary to assure a predefined significance (we considered a 95% confidence level). We analyzed the spatial significance of the results in the different zones (1–5) shown in Fig. 1a; the degrees of freedom for the different zones and seasons vary from 3 to 16. The results shown in Fig. 4c for El Niño periods are spatially significant for zones 3, 4, and 5, whereas no spatial significance was obtained for the rest of the cases (Figs. 4d, 4g, and 4h). For La Niña periods, the results for winter precipitation are spatially significant for zones 1 (Fig. 5c), 3, and 5 (Fig. 5d), whereas for temperature the results are spatially significant for zones 3 and 5 (Fig. 5h).

5. ENSO-conditioned seasonal predictability

The results shown in the previous section reveal the role of ENSO events as a driving factor of precipitation and temperature anomalies in Spain, thus providing a potential window of opportunity for seasonal forecast. In this section, we consider those forecasts corresponding to El Niño/La Niña years and assess the resulting conditional skill in Spain. As mentioned in the previous section, we analyze and compare both the direct model output from DEMETER and the downscaled values for precipitation and temperature. To simplify the results, we only present those corresponding to the variables and seasons that showed some predictability associated with ENSO events for the lower or upper terciles.

Figure 6 represents the RSA maps for those cases with El Niño teleconnections. The figures on the top correspond to the dry spring precipitation events for the direct output (right) and the downscaling model (left). The RSA values are positive over the south and the Mediterranean coast, reaching values around 0.8; in this case, the downscaling outputs improve the direct model predictions. Note that the spatial pattern of these plots is quite similar to the one presented in Fig. 4a, corresponding to the observed teleconnections, especially for the downscaling method, which exhibits a nearly perfect match. During El Niño events, the skill is also high for hot events in spring over the Mediterranean coast (Fig. 6,

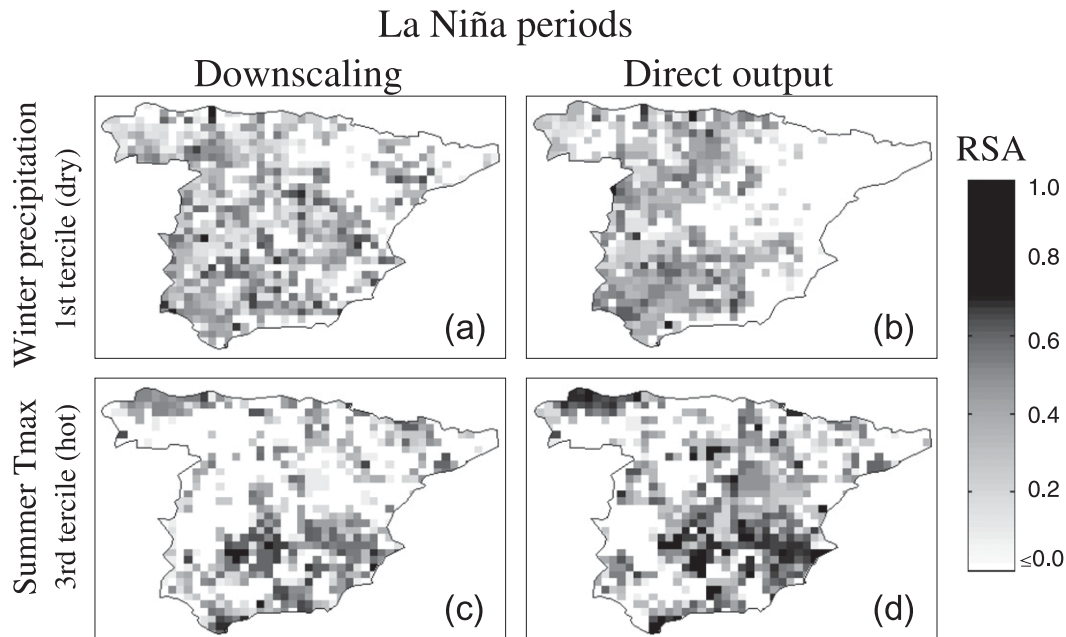


FIG. 7. RSA maps for events related to La Niña phenomena: (a),(b) winter dry events and (c),(d) summer hot events. The RSA values corresponding to the 90% confidence level obtained by bootstrapping with 1000 runs are 0.5, 0.45, 0.72, and 0.62 for (a)–(d), respectively.

bottom), which is also in agreement with the observed teleconnections (Fig. 4f), although neither the direct outputs nor the downscaled values capture the observed influence of El Niño events over the southwestern part of Spain. In this case, the statistical downscaling method does not present any improvement over the direct model outputs.

To assess the significance of the results, we performed a bootstrapping test by randomly selecting from the whole period (1960–2000) 1000 samples with the same size as the number of ENSO events (El Niño or La Niña) considered in the study. The RSA values corresponding to a confidence level of 90% are shown in the caption of Fig. 6.

In the case of La Niña years, Fig. 7 shows the RSA values for dry events in winter from the downscaled precipitation (left) and the direct output (right). Both maps indicate better skill over the western part of Spain, in agreement with the observed teleconnections (Figs. 5c and 5d), which is focused on the northern area. There is also skill associated with La Niña episodes for hot events in summer (Fig. 7, bottom), in agreement with the observed teleconnections (Fig. 5f). Both downscaled maximum temperature and direct output show higher skill in the south and the Mediterranean coast. Values are particularly high in the southeastern part. Note that, again, in the case of temperature, the downscaling method does not improve the direct model predictions.

We have also compared the results obtained from the DEMETER multimodel with those obtained from per-

sistence, because some results indicate that some skillful predictions may be related to persistent patterns from the preceding season (see, e.g., Shongwe et al. 2007). Thus, we consider the persistence forecast as the tercile observed in the previous season (i.e., we assign one probability to this tercile category and zero to the others). We found no remarkable skill related to strong ENSO events, neither for precipitation nor for temperature, using this persistent prediction. Thus, the multimodel ensemble from DEMETER outperforms the persistence as seasonal forecast system over Spain.

6. Conclusions

This paper assesses the state of the seasonal forecasting in Spain using the multimodel ensemble from the DEMETER project (which is the precursor of the current EUROSIP operational forecasting system). Probabilistic predictions of surface precipitation and maximum temperature anomalies are first evaluated on the period 1960–2000. Significant skill is found only for dry events in autumn in the western part of Spain. To analyze in detail the predictability in this extratropical area, the study is then focused on one of the major source of predictability at seasonal time scales, the ENSO phenomenon, which may provide windows of opportunity for seasonal prediction.

Although different teleconnections with ENSO events have been reported in the literature for precipitation and

maximum temperature in the Iberian peninsula, we used a high-resolution gridded dataset over Spain to homogeneously check the existing teleconnections for these variables in all seasons. Significant influence is found in spring for dry events over the south and Mediterranean coast and for hot events over the southeastern part of Spain during El Niño years. The occurrence of La Niña events are related to dry events in winter over the northwestern part of Spain and to hot events in summer over the south and the Mediterranean coast. These results are in partial agreement with other studies, and they provide a global updated picture of the predictability due to forcing from the tropical Pacific at these latitudes.

The forecast skill of the multimodel ensemble is then evaluated for strong El Niño and La Niña years. Results show higher predictability for these events than for the whole period, with skill patterns resembling the observed teleconnections associated with maximum temperature and precipitation variability over Spain. Skillful seasonal predictions are found in spring for dry events over the south and the Mediterranean coast and for hot events in the southeast related to strong winter El Niño years. The connection with La Niña drives higher predictability for dry events over the western part in winter and for hot events in summer over the south and the Mediterranean coast. According to these results, the significant skill of DEMETER models predictions match the observed teleconnections with the ENSO events, thus the multimodel ensemble is able to perform skillful seasonal forecasts at extratropical latitudes.

We also compared the forecasts of the direct model outputs with those given by a statistical downscaling method based on analogs. The downscaling is shown to improve the precipitation skill, but no improvement is achieved for the temperature forecasts.

Some years ago, the predictability was placed on approximately 11–15 days. Nowadays, that limit can be pushed ahead as a result of the efforts from different centers around the world to develop climatic models potentially capable of making longer time-scale forecasts. The improvements in our knowledge about the aspects of the climate system that affect future climate has also played a main role in this advantage. This and other studies have shown that the predictability at seasonal time scales is not confined to tropical latitudes, but that the tropical Pacific Ocean also exerts some influence at higher latitudes. Skillful seasonal predictions have been detected in regions far away from the direct area of influence of the ENSO phenomenon, where, as shown in this study, at least for strong ENSO events there is some predictability. These results are of great utility in different socioeconomic fields. The ability to forecast unusual climate conditions, such as droughts or

hot seasons, a few months in advance is particularly interesting in agriculture, energy requirements, and other human affairs to avoid severe consequences. Nevertheless, efforts should continue to improve those forecasts, taking into account other atmospheric circulation patterns that can also be important at extratropical latitudes.

Acknowledgments. The authors are grateful to the 6th FP EU projects ENSEMBLES (Grant GOCE-CT-2003-505539) and to the CICYT projects (Grants CGL-2007-64387/CLI and CGL2005-06966-C07-02/CLI) for partial support of this work. We also want to acknowledge DEMETER partners for freely providing the data resulting from the project and the two reviewers for their constructive comments.

REFERENCES

- Barnston, A. G., S. J. Mason, L. Goddard, D. G. Dewitt, and S. E. Zebiak, 2003: Multimodel ensembling in seasonal climate forecasting at IRI. *Bull. Amer. Meteor. Soc.*, **84**, 1783–1796.
- Bretherton, C., M. Widmann, V. Dymnikov, J. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, **12**, 1990–2009.
- Challinor, A. J., J. M. Slingo, T. R. Wheeler, and F. J. Doblas-Reyes, 2005: Probabilistic simulations of crop yield over western India using the DEMETER seasonal hindcast ensembles. *Tellus*, **57A**, 498–512.
- Cuadrat, J. M., and J. M. Vide, Eds., 2007: *Spanish Climatology. Past, Present and Future*. Prensa Universitaria de Zaragoza, 575 pp.
- Díez, E., C. Primo, J. A. García-Moya, J. M. Gutiérrez, and B. Orfila, 2005: Statistical and dynamical downscaling of precipitation over Spain from DEMETER seasonal forecasts. *Tellus*, **57A**, 409–423.
- Doblas-Reyes, F. J., M. Déqué, and J. P. Piedelievre, 2000: Multimodel spread and probabilistic seasonal forecast in PROVOST. *Quart. J. Roy. Meteor. Soc.*, **126**, 2069–2087.
- Douville, H., 2004: Relevance of soil moisture for seasonal atmospheric predictions: Is it an initial value problem? *Climate Dyn.*, **22**, 429–446.
- Fedderson, H., and U. Andersen, 2005: A method for statistical downscaling of seasonal ensemble predictions. *Tellus*, **57A**, 398–408.
- Fraedrich, K., 1990: European grosswetter during the warm and cold extremes of the El Niño/Southern Oscillation. *Int. J. Climatol.*, **10**, 21–31.
- , 1994: An ENSO impact on Europe? *Tellus*, **46A**, 541–552.
- , C. Ziehmann, and F. Sielmann, 1995: Estimates of spatial degrees of freedom. *J. Climate*, **8**, 361–369.
- Frías, M. D., J. Fernández, J. Sáenz, and C. Rodríguez-Puebla, 2005: Operational predictability of monthly average maximum temperature over the Iberian Peninsula using DEMETER simulations and downscaling. *Tellus*, **57A**, 448–463.
- García-Morales, M., and L. Dubus, 2007: Forecasting precipitation for hydroelectric power management: How to exploit GCM's seasonal ensemble forecasts. *Int. J. Climatol.*, **27**, 1691–1705.
- Gershunov, A., and D. R. Cayan, 2003: Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. *J. Climate*, **16**, 2752–2765.

- Goddard, L., S. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152.
- Gutiérrez, J. M., A. S. Cofiño, R. Cano, and M. A. Rodríguez, 2004: Clustering methods for statistical downscaling in short-range weather forecasts. *Mon. Wea. Rev.*, **132**, 2169–2183.
- , R. Cano, A. S. Cofiño, and C. Sordo, 2005: Analysis and downscaling multi-model seasonal forecasts in Peru using self-organizing maps. *Tellus*, **57A**, 435–447.
- Halpert, M. S., and C. F. Ropelewski, 1992: Surface temperature patterns associated with the Southern Oscillation. *J. Climate*, **5**, 577–593.
- Jolliffe, I. T., and D. B. Stephenson, Eds., 2003: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Wiley, 254 pp.
- Kiladis, G. N., and H. F. Diaz, 1989: Global climatic anomalies associated with extremes in the Southern Oscillation. *J. Climate*, **2**, 1069–1089.
- Kirtman, B., and A. Pirani, 2008: WCRP position paper on seasonal prediction. Tech. Rep., WCRP Informal Rep. 3/2008, ICPO Publication 127, 25 pp.
- Knippertz, P., U. Ulbrich, F. Marques, and J. Corte-Real, 2003: Decadal changes in the link between El Niño and springtime North Atlantic oscillation and European-North African rainfall. *Int. J. Climatol.*, **23**, 1293–1311.
- Livezey, R., and W. Chen, 1983: Statistical field significance and its determination by monte carlo techniques. *Mon. Wea. Rev.*, **111**, 46–59.
- Lloyd-Hughes, B., and M. A. Saunders, 2002: Seasonal prediction of European spring precipitation from El Niño-Southern Oscillation and local sea-surface temperatures. *Int. J. Climatol.*, **22**, 1–14.
- Mariotti, A., J. Ballabrera-Poy, and N. Zeng, 2005: Tropical influence on Euro-Asian autumn rainfall variability. *Climate Dyn.*, **24**, 511–521.
- Mason, S., and N. Graham, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quart. J. Roy. Meteor. Soc.*, **128**, 2145–2166.
- Moron, V., and M. Ward, 1998: ENSO teleconnections with climate variability in the European and African sector. *Weather*, **53**, 287–295.
- Muñoz-Díaz, D., and F. Rodrigo, 2004: Spatio-temporal patterns of seasonal rainfall in Spain (1912–2000) using cluster and principal component analysis: Comparison. *Ann. Geophys.*, **22**, 1435–1448.
- Palmer, T. N., and D. L. T. Anderson, 1994: The prospects for seasonal forecasting - A review paper. *Quart. J. Roy. Meteor. Soc.*, **120**, 755–793.
- , and Coauthors, 2004: Development of a European multi-model ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85**, 853–872.
- Pozo-Vázquez, D., M. J. Esteban-Parra, F. S. Rodrigo, and Y. Castro-Díez, 2001: The association between ENSO and winter atmospheric circulation and temperature in the North Atlantic region. *J. Climate*, **14**, 3408–3420.
- , S. R. Gámiz-Fortis, J. Tovar-Pescador, M. J. Esteban-Parra, and Y. Castro-Díez, 2005: El Niño-Southern oscillation events and associated European winter precipitation anomalies. *Int. J. Climatol.*, **25**, 17–31.
- Price, C., L. Stone, A. Huppert, B. Rajagopalan, and P. Alpert, 1998: A possible link between El Niño and precipitation in Israel. *Geophys. Res. Lett.*, **25**, 3963–3966.
- Quan, X. W., M. P. Hoerling, J. S. Whitaker, G. T. Bates, and T. Y. Xu, 2006: Diagnosing sources of U.S. seasonal forecast skill. *J. Climate*, **19**, 3279–3293.
- Rocha, A., 1999: Low-frequency variability of seasonal rainfall over the Iberian Peninsula and ENSO. *Int. J. Climatol.*, **19**, 889–901.
- Rodó, X., E. Baert, and F. A. Comin, 1997: Variations in seasonal rainfall in Southern Europe during the present century: Relationships with the North Atlantic Oscillation and the El Niño-Southern Oscillation. *Climate Dyn.*, **13**, 275–284.
- Rodríguez-Puebla, C., A. H. Encinas, S. Nieto, and J. Garmendia, 1998: Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula. *Int. J. Climatol.*, **18**, 299–316.
- Sáenz, J., C. Rodríguez-Puebla, J. Fernández, and J. Zubillaga, 2001: Interpretation of interannual winter temperature variations over southwestern Europe. *J. Geophys. Res.*, **106**, 20 641–20 652.
- Saha, S., and Coauthors, 2006: The NCEP climate forecast system. *J. Climate*, **19**, 3483–3517.
- Serrano, A., J. E. García, V. L. Mateos, M. L. Cancillo, and J. Garrido, 1999: Monthly modes of variation of precipitation over the Iberian Peninsula. *J. Climate*, **12**, 2894–2919.
- Shongwe, M., C. Ferro, C. Coelho, and G. van Oldenborgh, 2007: Predictability of cold spring seasons in Europe. *Mon. Wea. Rev.*, **135**, 4185–4201.
- Sordo, C. M., M. D. Frías, S. Herrera, A. S. Cofiño, and J. M. Gutiérrez, 2008: Interval-based statistical validation of operational seasonal forecasts in Spain conditioned to El Niño-Southern Oscillation events. *J. Geophys. Res.*, **113**, D17121, doi:10.1029/2007JD009536.
- Thompson, M. C., F. J. Doblas-Reyes, S. J. Mason, R. Hagedorn, S. J. Connor, T. Phindela, A. P. Morse, and T. N. Palmer, 2006: Malaria early warnings based on seasonal climate forecasts from multi-model ensembles. *Nature*, **439**, 576–579.
- Tippett, M. K., A. G. Barnston, and A. W. Robertson, 2007: Estimation of seasonal precipitation tercile-based categorical probabilities from ensembles. *J. Climate*, **20**, 2210–2228.
- Trenberth, K. E., 1997: Short-term climate variations: recent accomplishments and issues for future progress. *Bull. Amer. Meteor. Soc.*, **78**, 1081–1096.
- Troccoli, A., M. Harrison, D. L. T. Anderson, and S. J. Mason, 2008: *Seasonal Climate: Forecasting and Managing Risk*. NATO Science Series IV: Earth and Environmental Sciences, Vol. 82, Springer, 467 pp.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012.
- van Oldenborgh, G. J., G. Burgers, and A. K. Tank, 2000: On the El Niño teleconnection to spring precipitation in Europe. *Int. J. Climatol.*, **20**, 565–574.
- Vitart, F., and Coauthors, 2007: Dynamically-based seasonal forecasts of Atlantic tropical storm activity issued in June by EURO-SIP. *Geophys. Res. Lett.*, **34**, L16815, doi:10.1029/2007GL030740.
- Wang, G., R. Kleeman, N. Smith, and F. Tseitkin, 2002: The BMRC coupled general circulation model ENSO forecast system. *Mon. Wea. Rev.*, **130**, 975–991.
- Wang, W., and A. Kumar, 1998: A GCM assessment of atmospheric seasonal predictability associated with soil moisture anomalies over North America. *J. Geophys. Res.*, **103**, 28 637–28 646.