

## Performances of climatic indicators from seasonal forecasts for ecosystem management: The case of Central Europe and the Mediterranean

JM Costa-Saura <sup>a, b, \*</sup>, V Mereu <sup>b, a</sup>, M Santini <sup>c</sup>, A Trabucco <sup>b, a</sup>, D Spano <sup>a, b</sup>, V Bacciu <sup>b</sup>

<sup>a</sup> University of Sassari, Sassari 07100, Italy

<sup>b</sup> Foundation Euro-Mediterranean Center on Climate Change, Division Impacts on Agriculture, Forests and Ecosystem Services (IAFES), Sassari 07100, Italy

<sup>c</sup> Foundation Euro-Mediterranean Center on Climate Change, Division Impacts on Agriculture, Forests and Ecosystem Services (IAFES), Viterbo 01100, Italy



### ARTICLE INFO

#### Keywords:

Seasonal forecasts  
Climate services  
Bias correction  
Fire weather index  
Potential evapotranspiration  
Soil moisture deficit

### ABSTRACT

Climate change is affecting the interannual variability and the seasonal cycle for key meteorological variables, as well as the frequency, intensity, duration and timing of out-of-the-norm and extreme events. In such a context, seasonal forecasts (i.e., climate predictions from a few weeks to several months ahead) are compelling instruments able to anticipate upcoming climate risks and guide tactical decisions. In this study, we assess the performance of seasonal forecasts for the summer period across Central Europe and the Mediterranean region in predicting some key indicators serving agriculture and forestry sectors, i.e., Potential EvapoTranspiration (PET), Potential Soil Moisture Deficit (PSMD), and Fire Weather Index (FWI). We exploit authoritative climate data from Copernicus Climate Data Store, i.e., ERA5 reanalyses and hindcasts from CMCC SPSv3 and ECMWF SEAS5 seasonal prediction systems (SPSs), to evaluate, using both deterministic and probabilistic evaluation scores, differences in performance across the two SPSs, two start dates (March and May) and four correction techniques applied to overcome modelling bias, namely bias correction (BC), calibration (CAL), quantile mapping (QM) and detrending (DET). Results show that seasonal predictions of PET perform better in Western and Eastern Europe and some areas of North Africa. PSMD predictions follow a similar spatial pattern as PET, except that for some areas in Central (Eastern) Europe in which the performance increases (decreases). FWI predictions reveal better results in some areas of the Iberian Peninsula, North-Western Africa, Balkan Peninsula, and Ukraine. Results also suggest that QM might be the most suitable technique for bias correction. Furthermore, the start date of the forecast might imply varying correlation significance, with the start date closest to the forecasted period not always being the best. Overall, our study suggests the potential usefulness of seasonal forecasts for decision making under different geographical and environmental contexts, considering sensitivity inherent to different processing chains.

### 1. Introduction

Climate change is currently affecting the interannual variability and the seasonal cycle of key meteorological variables, as well as the frequency, magnitude, duration, and timing of extreme/out-of-the-norm events (Thornton et al., 2014). All these patterns are diverging from historical ones, which were traditionally considered by farmers and foresters for planning and management. Indeed, shifting seasonal patterns, associated with long-term changing climate conditions, is increasing uncertainties for effective decision-making. Thus, accurate seasonal climate predictions, clearly communicated to the stakeholders without mismatching concepts, might greatly benefit users, land

managers and/or decision-makers (Bruno Soares et al., 2018; Klemm and Mcpherson, 2017).

Seasonal forecasts, i.e., climate predictions made several months in advance, might be helpful to assist proactive decisions to better protect ecosystems and anthropogenic assets (Turco et al., 2019). Indeed, it is recognised an expansion of climate services applications as a tool for achieving broader societal objectives while managing climate risks (Jacobs and Street, 2020). Furthermore, thanks to improvements in numerical computations and climate modelling, seasonal forecasting has gained increasing interest (Capa-Morocho et al., 2016), with national and international agencies encouraging their use along with investments for large data facilities (e.g., the Copernicus programme from the

\* Corresponding author.

E-mail address: [jmcostasaura@uniss.it](mailto:jmcostasaura@uniss.it) (J. Costa-Saura).

European Union). Seasonal forecast are driven by initial boundary conditions, via data assimilation, than by changes in radiative forcings (e.g. greenhouse gas concentrations), and they are rather aiming at short-range predictions (Hurrell et al., 2009; Stockdale et al., 2010). Commonly, seasonal forecasts have focused mostly on general climate variables such as temperature and precipitation (Doblas-Reyes et al., 2013; Hemri et al., 2020; Manzanas et al., 2018a), however, the attention has recently move to their application over different sectors, e.g., water resources (Marcos et al., 2017), agricultural applications (Hayashi et al., 2018; Iizumi et al., 2018, 2013; Jha et al., 2019a, 2019b), wind power (Lledó et al., 2019), or fire risk (Bedia et al., 2018). However, most of these studies were based on previous versions of forecast systems (e.g., ECMWF Seasonal Forecast System 4), and found good performance only in specific regions and over short lead times, i.e., a short period from the start date of simulations/issue of the forecast. Indeed, using previous forecast systems, the accuracy of seasonal predictions of temperatures and precipitation was found moderate-high in tropical regions, but with limited reliability over temperate regions such as Europe (Doblas-Reyes et al., 2013). In fact using ECMWF System 4, Bedia et al. (2018) only found good accuracy for the Canadian Fire Weather Index (FWI) in a narrow area of eastern Europe, while Turco et al. (2018) did not find good correlations between forecasted precipitation indices and burned area for Euro-Mediterranean region.

Some sources of uncertainties in seasonal forecasting are induced by model approximations of physical processes and the imperfect knowledge of the system's initial conditions (Doblas-Reyes et al., 2013). In addition, in specific areas such as Europe the weak relationship between El Niño Southern Oscillation (ENSO) and climate variations makes predictions less accurate (Brands, 2017). Thus, several research programs such as SPECS, DEMETER, ENSEMBLES, and EUPHORIAS have been activated to overcome such modelling limitations and improve seasonal predictions. Indeed, these modelling developments are extremely dynamic and require continuous testing as a ground basis to guide further improvement for the next versions of seasonal systems. Recently, the ERA4cs MEDSCOPE project (<https://www.medscope-project.eu/>) aimed at understanding climate predictability sources across-Mediterranean area and at developing tools (i.e., the R package CSTools) to improve climate service production (e.g., Voces-Aboy et al., 2019). For instance, CSTools includes several functions to correct the systematic bias of climate data projections. Bias correction is of primary importance for decision making since it adjusts systematic modelling errors and allows stakeholders to directly use unbiased predicted values and thereafter correctly tune threshold values and run impact models (e.g., Brown et al., 2018; Ongutu et al., 2018).

Detailed performance assessments in seasonal forecasting, i.e., the influence of different start dates, lead times, climate models and post-processing techniques, have been mostly focused on main climate variables (e.g., temperature and precipitation, Manzanas et al., 2019, 2018a) generally lacking for complex climate indicators. Climate indicators, i.e., measures that operates further a single variable or integrates different variables, are recognized as valuable tools for rapid assessment of climate impacts helping to guide choices of best adaptation measures and robust decisions (de Rigo et al., 2017; Trnka et al., 2011). Single-variable indicators usually link to a few climate characteristics and related environmental processes, while a more articulated assessment of complex processes (e.g., ecosystem drought and fire risks) may address different physical components by including multiple variables into composite indicators (Talukder et al., 2017; Zscheischler et al., 2018). However, the rising complexity of indicators leads to a larger sensitivity to the different variables they depend on (Burgass et al., 2017). Thus, for a reliable picture supporting robust decision-making, it would be essential to assess multiple indicators considering the growing levels of complexity and sensitivity to errors inherent to the different input variables under different geographical and seasonal contexts.

Key Indicators for decision making in agriculture and forestry

systems include often evapotranspiration and water deficit processes, which might help farmers to select crop types and cultivars more adapted to dry/wet conditions and cold/warm seasons, reducing damages and increasing the yield (Olesen et al., 2011). In addition, they might help to optimize field operations for saving resources increasing farmers' incomes (Asseng et al., 2012; Ramírez-Rodrigues et al., 2016). Likewise, forest, land, and fire managers might also benefit from accurate seasonal predictions of fire risk indicators in different ways (Bedia et al., 2018). For instance, they might help at efficiently design the fire season campaign with more adequate assets and spatial distribution of fire suppression resources (e.g., truck water pumps or flame retardants).

Within the framework of the MEDSCOPE project, the present work aims at providing information and guidance to users of climate products serving the agriculture and forestry sectors, evaluating main features (e.g., available models, forecasts start dates), skills (e.g., deterministic and probabilistic scores) and different post-processing techniques for data correction. Using climatic indicators representing various landscape processes (in particular from evapotranspiration demand to water balance and fire risk), we assess two recently updated seasonal prediction systems provided by the Copernicus Climate Change Service (C3S), i.e., CMCC Seasonal Prediction System v3 (CMCC SPSv3) and ECMWF Seasonal Forecast system 5 (ECMWF SEAS 5), in Central Europe and the Mediterranean, for which previous climate predictions system showed limited accuracy (Doblas-Reyes et al., 2013). Furthermore, we assess how the spatial pattern and strength of results vary with increasing complexity of indicators, which is crucial for optimal decision making. In addition, the evaluation of climate prediction systems encompassed not only validation via correlation between observed and forecasted indicator values, but also assessed accuracy for out-of-the-norm and extreme years.

## 2. Materials and methods

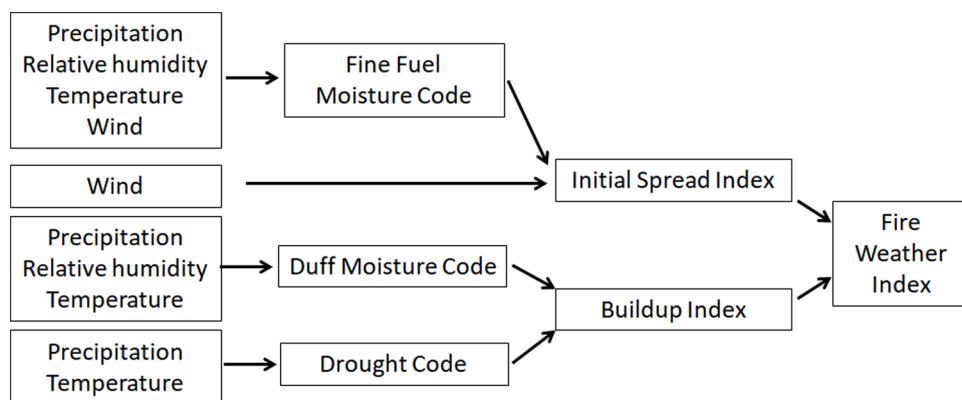
### 2.1. Study area

The study area covers Central Europe and the Mediterranean (see Fig. 2), for which previous climate predictions system showed limited accuracy (Doblas-Reyes et al., 2013) and thus was of interest for the MEDSCOPE project. This vast area allows considering different geographical subregions (see Fig. S4) that might potentially differ in climate predictability

### 2.2. Climate data

Time series of seasonal forecasts were downloaded from the Copernicus Climate Data Store (CDS, <https://cds.climate.copernicus.eu/>), wherein the C3S infrastructure provides high-quality climate information and tools to public and private institutions working on climate services as intermediate (more expert) or end (less expert) users. We used 6-hour retrospective seasonal forecasts, i.e., historical hindcasts, from the seasonal prediction systems (SPSs) CMCC SPSv3 and the ECMWF SEAS5 (also simply CMCC SPS and ECMWF SPS hereafter). Among available SPSs from CDS at the time of the analysis, the ones from CMCC and ECMWF were selected as they represent the best compromise among native horizontal resolution ( $1^{\circ}$  and  $\approx 0.32^{\circ}$  for CMCC and ECMWF, respectively) and the highest number of ensemble members (40 and 25 for CMCC and ECMWF, respectively), these representing different possible perturbation of initial conditions and thus accounting for uncertainties related to the imperfect representation of state variables which initialize the models. Indeed, the MeteoFrance and MetOffice SPSs have a good horizontal resolution ( $< 1^{\circ}$ ) but they have less members, while the DWD SPS has the same horizontal resolution than CMCC SPS but it includes 30 rather than 40 members.

In particular, the CMCC-SPSv3 improves its predecessor in terms of dynamical core and initialization data assimilation (Matera et al., 2017). In contrast, ECMWF SEAS5 implements new upgrades on ocean



**Fig. 1.** Variables required to produce the six subcomponents representing: fuel moisture (Fine Fuel Moisture Code, Duff Moisture Code, and Drought Code), rate of spread and fuel consumption (Initial Spread Index, Buildup Index), and the potential fire intensity (Fire Weather Index).

model, atmospheric resolution, and land surface initialization (Johnson et al., 2019). Because of data availability, we only considered in this study the hindcast period 1993–2015. For validation, we used the ERA5 climate reanalysis dataset also available from CDS. ERA5 is a climate reanalysis that combines past/current observations with climate models through data assimilation techniques to generate consistent retrospective time series at hourly time steps with a 0.25° resolution globally. The ERA5 climate dataset, referred to also as "observations" hereafter, was scaled to 1° resolution to compare against CMCC SPSv3 and ECMWF SEAS5 predictions. Available variables downloaded to calculate the indicators were dew point, maximum, minimum, mean temperatures, precipitation, and northward and eastward wind components. Dew point and mean temperatures were used to calculate relative humidity according to Lawrence (2005).

The SPSs contributing to C3S are coordinated to release forecasts at least every 1<sup>st</sup> of month and for at least 6 calendar months ahead. Here, analyses were focused on forecasts available at least one month before the season of interest (e.g., Hemri et al., 2020), i.e. initialized on the 1<sup>st</sup> of May, allowing time to formulate strategic decisions (i.e., lead time as the period between the issue of forecast and the occurrence of prediction). Although the inclusion and the performance assessment at one lead month forecast (e.g., for August by using predictions initialized on July) are feasible, seasonal forecast products/tools are commonly averaged over the whole season. In fact, it is expected and more feasible to make decisions considering climate for the complete season and not only for one specific month. For instance, firefighting agencies are interested to know, with due anticipation, what might happen during the whole dry season, not only in June, and thus providing information in July for August might be useless for them. Furthermore, to check the influence of longer lead times, the analysis using the 1<sup>st</sup> of March as start date for forecasting was also included, which is the first prediction release having the summer period JJA fully covered.

### 2.3. Climatic indicators

The climatic indicators as proxy for various landscape processes (in particular from evapotranspiration demand to water balance and fire risk) and serving decision and policy-makers in the agriculture and forestry sectors were specifically selected to compare indicators' performance across different level of combination complexity. The Potential Evapotranspiration (PET) was calculated monthly using the Hargreaves (1994) formula in equation 1 (as implemented in R package SPEI; <https://github.com/sbegueria/SPEI>), which is a simplified temperature-based indicator that relies on the mean, minimum and maximum temperatures ( $T$ , in °C) and extra-terrestrial solar radiation ( $R$ , in MJ m<sup>-2</sup> and converted in mm by 0.408 factor) based on the location's latitude and the calendar day following Allen et al. (1994).

$$PET = 0.0023(T_{mean} + 17.8)(T_{max} - T_{min})^{0.5} \quad (0.408 R) \quad (1)$$

The Potential Soil Moisture Deficit (PSMD, also called water balance, Beguería et al., 2014), as meteorological based indicator of soil water scarcity and stress for plants, was calculated as the cumulated deficit between PET (under the same formula above) and precipitation during the dry season at monthly time step (equation 2).

$$PSMD = PET - PRECIPITATION \quad (2)$$

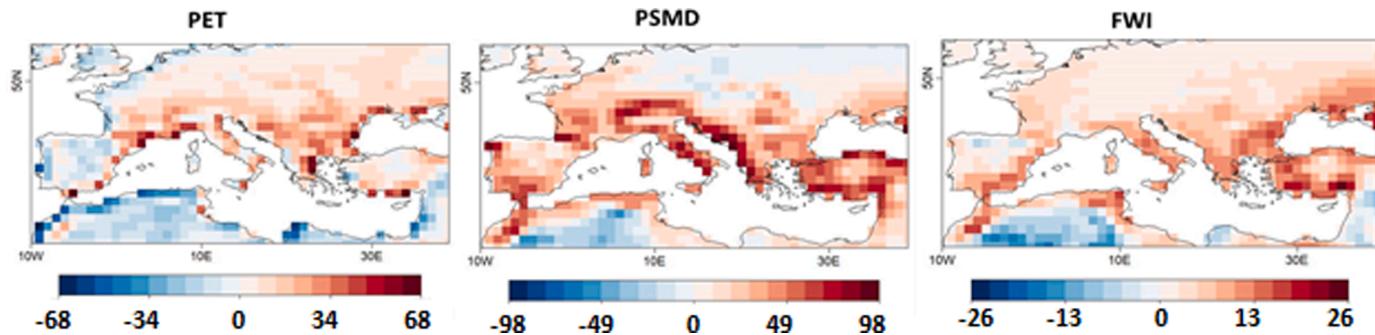
Finally, the Canadian Fire Weather Index (FWI), commonly used worldwide by national protection agencies for monitoring fire danger, is the most complex indicator based on temperature, precipitation, wind speed, and relative humidity. It is based on different subcomponents that account for the effects of both fuel moisture and wind on fire behaviour (Fig. 1, Van Wagner, 1987). The components are combined to produce the FWI to represent and rate the potential fire intensity, and also the difficulty of fire control, given the meteorological conditions. FWI was calculated using the R package *fireDanger* (<https://github.com/SantanderMetGroup/fireDanger>). After calculation, all the indicators were averaged for the summer season (i.e., June, July, and August; JJA).

### 2.4. Post-processing data

Different post-processing techniques implemented on the R package CSTools, developed in the MEDSCOPE project to improve forecast quality, were directly applied over the climatic indicators as suggested by Bedia et al. (2018). Most of the functions in this study are focused on correcting the systematic modelling biases of hindcasts against "observations". Each function alters the statistical properties of the input data differently. The CST\_BiasCorrection function (BC) adjusts the forecast to have the same mean and standard deviation as observations assuming both datasets having a Gaussian distribution (Torralba et al., 2017), whereas CST\_Calibration (CAL) ensures interannual variance equivalent to that of the observed dataset with no modification of the ensemble-mean correlation by minimizing a constrained mean-squared error (Doblas-Reyes et al., 2005). CST\_QuantileMapping (QM) is a more complex method since it affects higher-order moments of the distribution. QM aims to make the simulated distribution converge with the observed one. Furthermore, the effect of removing trends (Trend function from the S2dverification R package; <https://cran.r-project.org/web/packages/s2dverification/index.html>) was assessed since some studies suggest that it might affect some verification metrics (Bedia et al., 2018).

### 2.5. Performance assessment

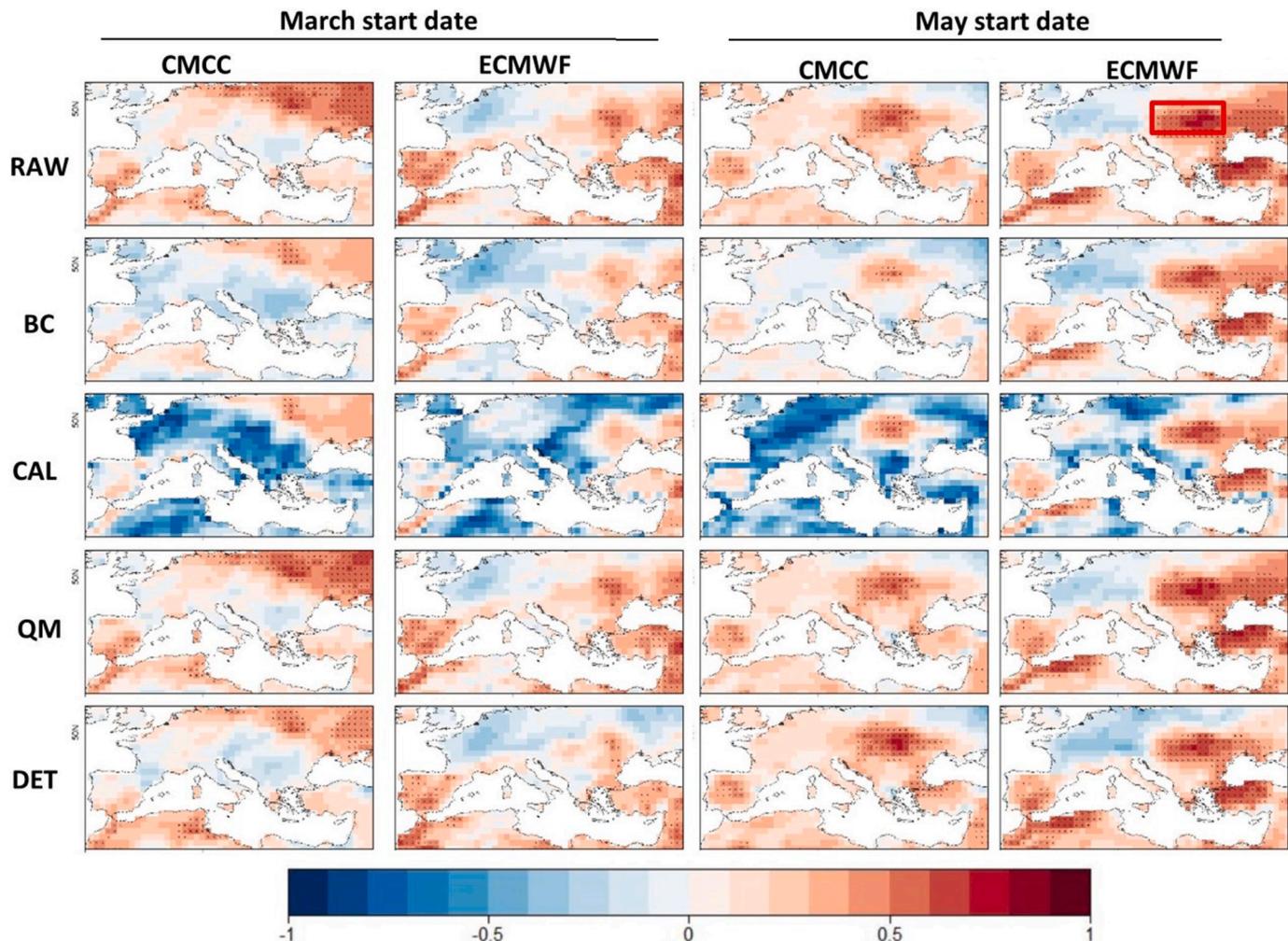
Predictions performance was assessed using deterministic and



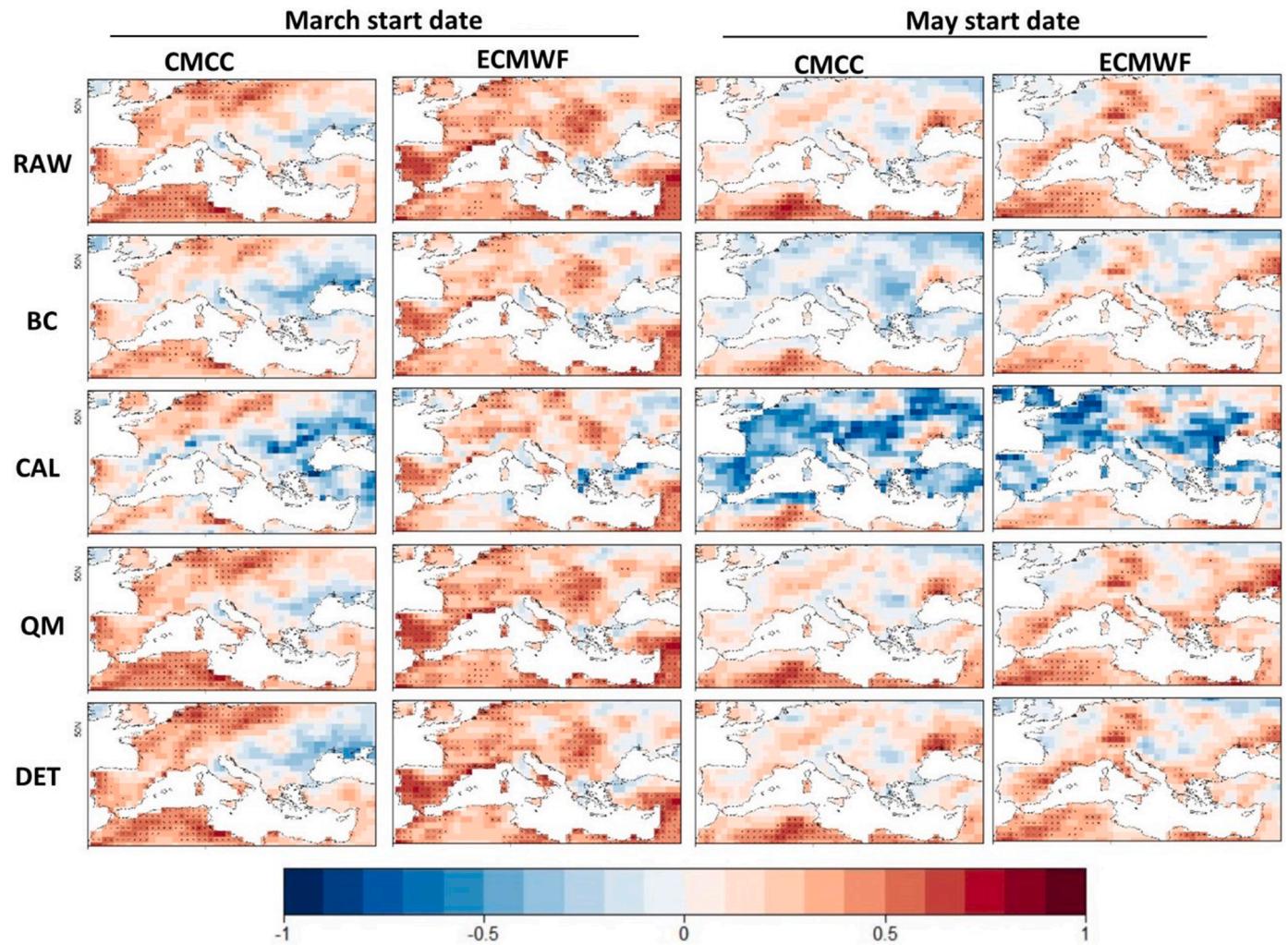
**Fig. 2.** Bias in climatology (June-July-August average 1993–2015) for each indicator (PET and PSMD in mm, FWI dimensionless) calculated between the hindcast ensemble mean (CMCC-SPSv3 start date May) and observations (ERA5 reanalysis). See supplementary materials for different SPSs and start dates and Figure S4 for referring to regional summary results.

probabilistic evaluation scores. As far as the former is concerned, we used the anomaly correlation coefficient (ACC), the most used metric in seasonal forecasting (Mishra et al., 2019). ACC allows evaluating the strength of the linear relationship between observed and forecasted datasets and thus if general trends can be potentially detected. If the ACC value equals 1 (-1), the datasets are positively (negatively) and perfectly correlated, whereas if the ACC is 0, there is no correlation between the datasets. Furthermore, we summarized the results across different regions (see supplemental materials), and calculated and mapped ACC

using the multi-model ensemble mean. We used the Brier Skill Score (BSS, Brier, 1950) to evaluate seasonal forecast skill on discriminating single events in a probabilistic manner (e.g., indicator values for a specific year being out-of-the-norm). BSS is used to assess the prediction's reliability and potential improvement against a reference dataset (i.e., climatology). For BSS, value 1 represents a perfect forecast, whereas negative values represent a forecast with lower skill than climatology (i.e., useless). The 66<sup>th</sup> and 95<sup>th</sup> percentiles have been used to assess forecast performance to discriminate out-of-the-norm and



**Fig. 3.** Time series correlations (1993–2015) between the hindcasts' ensemble mean and observations (ERA5 reanalysis) for potential evapotranspiration (PET) using ACC metric. Checked cells indicate a significant correlation ( $p$ -value  $< 0.05$ ).



**Fig. 4.** Time series correlations (1993–2015) between the hindcast ensemble mean and observations (ERA5 reanalysis) for Potential Soil Moisture Deficit (PSMD) using ACC metric. Checked cells indicate a significant correlation ( $p$ -value  $< 0.05$ ).

extreme events, respectively. As an exercise to provide further details of forecast skill over specific subareas (see red box in Fig. 3), we built tercile plots yearly representing observed versus predicted out of the norm events following Bedia et al. (2018) and using the visualize R package (<https://github.com/jwist/visualizeR> Frías et al., 2018).

### 3. Results

Results are organized to represent - for the different SPSs and forecast start dates - the absolute bias, against the observations, in the indicators for the whole climatological summer season, the year-to-year correlation and the ability to detect out-of-norm and extreme events.

#### 3.1. Bias between observed and forecasted indicators

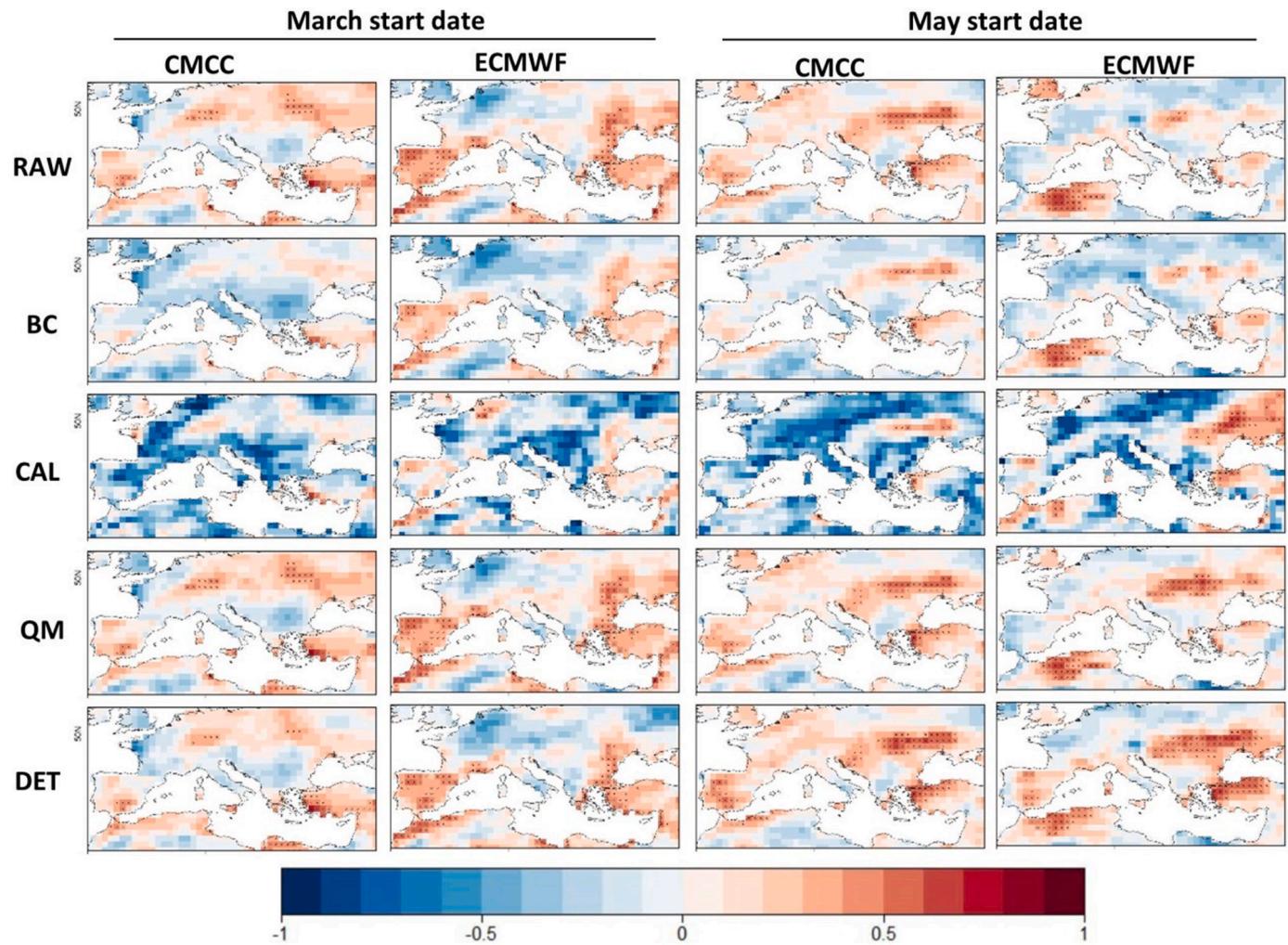
The bias for each indicator was evaluated as the climatological difference, for June-July-August (JJA), between the indicator's values derived from the ensemble mean of the retrospective forecast (hindcast) and the observations (i.e., the ERA5 reanalysis) taken as reference. Results show that the hindcasts predict higher PET, PSMD, and FWI values than the observations for most of the European countries surrounding the Mediterranean and for Turkey (Fig. 2 for CMCC SPS and May start date). Instead, a negative bias is observed for a large share of North Africa (except along the coasts for PSMD and FWI, both based on the combination of multiple variables), as well as across the Near East

excluding Turkey, along the Atlantic coasts and over the Iberian Peninsula for PET, whose calculation is based on temperature alone. For PSMD and FWI, respectively, a slight negative bias is also noticed over Eastern/North-Eastern Europe and in a few areas along the northern coast of Spain and over the southern part of the Near East. Overall, the largest absolute bias occurs along the Mediterranean coastal areas, while it tends to be weaker for eastern and central Europe towards the northernmost part of the domain. These patterns are similar across different start dates, and only slight differences are observed between climate models (see Supplementary Material).

#### 3.2. Correlations between observed and forecasted climatic indicators

Average anomaly correlation coefficient (ACC) values ( $\rho$  hereafter) per region are presented in Supplementary Table S1-S3 and Fig. S4, and below reported within the parenthesis when examined.

Significant correlations are found for PET (Fig. 3) with raw data over Eastern/North-Eastern Europe ( $\rho=0.32$ ), and secondarily over Central Europe ( $\rho=0.21$ ), North Africa ( $\rho=0.2$ ) and Iberian Peninsula (0.19) under the CMCC SPS using March as start date. Using QM and DET techniques for bias correction, the spatial pattern is conserved with respect to raw data, although a general worsening of correlation under DET is detected except for North Africa (here going from 0.2 to 0.25). With BC and CAL, the correlation is largely lower with good results remaining over Eastern/North-Eastern Europe ( $\rho=0.12$  for BC). When



**Fig. 5.** Time series correlations (1993–2015) between the hindcast ensemble mean and observations (ERA5 reanalysis) for Fire Weather Index (FWI) using ACC metric. Checked cells indicate a significant correlation ( $p$ -value  $< 0.05$ ).

taking May as start date lower correlation values are found for Eastern/North-Eastern Europe with both raw and corrected datasets, except over the Romania-Hungary-Ukraine regions, and for most of datasets for Central Europe and North Africa. Correlation is instead higher for the Central Mediterranean, Atlantic Europe and Near East (except under CAL), while for Iberian Peninsula and Central Europe it improves mostly under DET. Again, correlation spatial patterns from raw data are better maintained under QM and DET compared to BC and CAL.

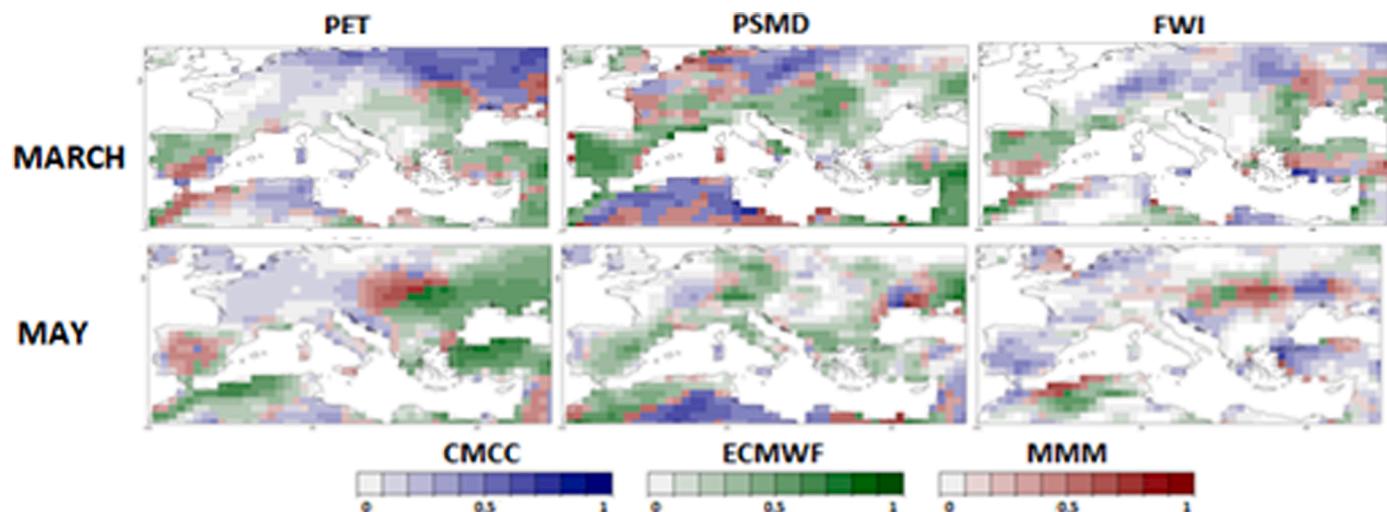
Using March as start date for the ECMWF SPS, the Near East shows the highest correlation ( $\rho=0.37$ ), and good values are also found for Iberian Peninsula ( $\rho=0.29$ ), as well as for the Eastern/North-Eastern Europe ( $\rho=0.23$ ). Satisfying results are also found over North Africa ( $\rho=0.18$ ) and Central Mediterranean ( $\rho=0.12$ ). Significance is concentrated around the Black Sea and the south-west of the domain. Switching to May as start date, higher and more significant correlations are found primarily in Eastern/North-Eastern Europe and over the Near East, with values  $\rho \geq 0.38$  under raw data and after QM, but with good results also found under other correction techniques ( $\rho \geq 0.17$ ). Noteworthy is also the correlation for the rest of the Mediterranean areas (North Africa, Central Mediterranean, Iberian Peninsula), with values  $\rho \geq 0.14$  except under the CAL technique. In contrast, the worst correlations are found for Central and Atlantic Europe ( $\rho \leq 0.03$ ).

In summary, results for PET differ across SPSs, and QM and then DET present the highest correlations and the widest agreement with raw data in terms of spatial patterns. Such variability prevents extracting any

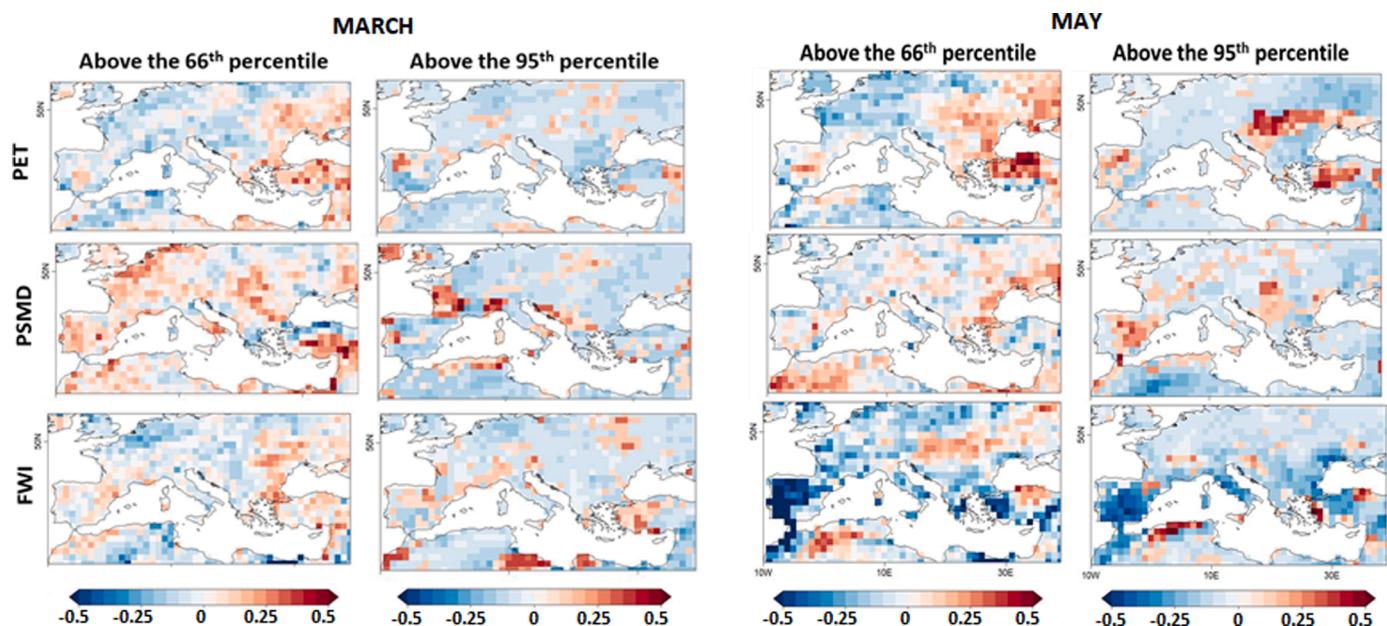
strong consensus between models and/or start dates, except better skills evidenced for the eastern half of the domain.

Hindcasts of PSMD (Fig. 4) under raw data show significant correlations over the western half of the domain under March start date for CMCC SPS, especially in North Africa ( $\rho=0.42$ ) and some areas over Central Europe and over the Iberian Peninsula ( $\rho=0.32$  either). A partial reduction of correlation is found when using BC and CAL methods, and also with DET over Eastern/North-Eastern Europe being the region with worst performances. When using May as start date, significant correlations in North Africa (0.36) and especially in its inland areas. Significant high correlation appear around the Black Sea's northern margins, more evident under DET, and over the Near East under both raw and QM datasets. In the rest of the domain correlation mostly decreases even becoming close to zero or negative under CAL. Hindcasts of PSMD from ECMWF SPS with March as start date have very good correlation in large parts of the domain, especially over Iberian Peninsula ( $\rho=0.46$ ) with lowest but however good values for Eastern/North-Eastern Europe ( $\rho=0.22$ ), except small areas in the northeast corner, over central Italy and surrounding of the Aegean Sea. May as start date causes overall drop in correlation: reduction is weaker in North Africa, with values switching from 0.32 to 0.28, and stronger over Atlantic Europe, where correlation declines from 0.26 to -0.14. Significant correlations remain over the Southern part of the Balkans, Central Europe, the Alps, the northern border of the Black Sea and the western part of the Mediterranean basin.

Hindcasts of FWI (Fig. 5) for raw data with CMCC SPS and March as



**Fig. 6.** Anomaly correlation coefficient (ACC) between SPSs and ERA5 for both forecast start dates after QM corrections. The figure shows the greatest ACC among models and the multi-model mean (MMM).

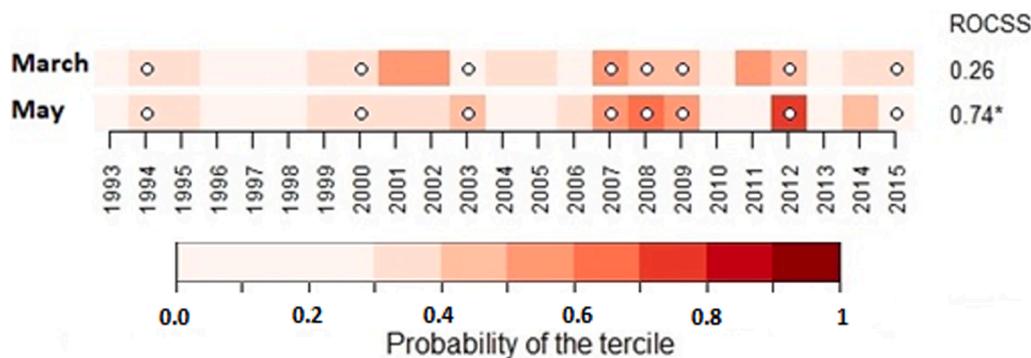


**Fig. 7.** Brier skill score (BSS) for out-of-the-norm and extreme events (i.e., above percentiles 66<sup>th</sup> and 95<sup>th</sup>, respectively) for multimodel hindcasts with respect to observations (ERA5 reanalysis) for both start dates after quantile mapping (QM) correction.

start date show a skilful correlation pattern in Turkey ( $\rho=0.18$  for Near East), Ukraine ( $\rho=0.15$  for Eastern/North-Eastern Europe), central Germany ( $\rho=0.14$  for Central Europe), Southern Spain ( $\rho=0.1$  for Iberian Peninsula), and northern Libya ( $\rho=0.06$  for North Africa). Results are similar after QM and DET, while when applying BC and CAL, many negative correlations emerge over most of the study domains. When considering May as start date, predictions improve significantly over the northern part of the Balkan peninsula, and partially for the Central Mediterranean, although still negative, while North-Eastern Europe witness lower correlation. Skills also increase over the Iberian Peninsula (correlation from -0.048 to 0.062), and slightly also around the southern borders of the North Sea, but the spatial average of correlation remains negative. In both forecast start dates, the QM and DET methods retain a spatial correlation as that of raw data. Considering the average among datasets, ECMWF SPS using March as start dates generally shows some positive correlation around the Black Sea ( $\rho \geq 0.014$ ), over the Iberian Peninsula ( $\rho=0.22$ ) and Morocco ( $\rho=0.03$ ), with similar results among

raw data and under QM and DET. When considering May as start date, skilful correlations appear in northern Algeria and, especially under QM and DET, over Romania-Hungary-Ukraine regions and north Turkey, leading there to correlation above 0.044. CAL procedure for this SPS confirms again to be the one with the worst performances. FWI predictions have lower skill than for other indicators, with wider negative correlation areas and more limited positive correlation and significance.

A multi-model combination of ACC results under the best performing QM technique (Fig. 6) defines the specific areas and spatial extent where either one or a combination of the two SPSs provides the best results for the three indicators. Results follow the spatial patterns of single SPSs, with better skills concentrated over the Eastern and Western half of the domain for PET and PSMD, respectively, and across less extended areas over Balkans and the South-Western corner of the domain for FWI.



**Fig. 8.** Forecast skill of PET (ECMWF SPS and QM) for a selected area in central-eastern Europe (see red rectangle in Fig. 3) using both start dates. Colours represent the probability of the upper tercile according to the hindcast, whereas white dots represent the observed tercile.

### 3.3. Accuracy predicting out-of-the-norm and extreme events

Predictions of out-of-the-norm events (i.e., those over the 66<sup>th</sup> percentile and assessed using the Brier Skill Score, BSS) follow similar patterns as correlations but with positive predictability constrained to smaller areas (results are focused on the multi-model mean after correction with QM, Fig. 7). Under both start dates, forecast of out-of-the-norm PET shows lower accuracy in the Iberian Peninsula and along the coastal areas in North Africa, maintaining good performances over eastern Europe and around the Black Sea instead, with highest score in Turkey for May start date. Out-of-the-norm PSMD forecasts also show lower accuracy in Spain and some central European areas under May start date, while the accuracy remains satisfying in Eastern Europe and northwest Africa. For FWI, the good accuracy over central-eastern Europe, northern Turkey and inland areas of North Africa is mostly conserved under May start date, while weaker under March start date. The accuracy assessment of detecting extreme events (i.e., out-of-normal over the 95<sup>th</sup> percentile) highlights that spatial patterns show a general decrease in skills except for some areas where they increase. Specifically for PET, the BSS is highest in eastern Europe and Turkey (May start date), and secondarily over Iberian Peninsula, also under March start date. For PSMD, the best BSS occurs over Southern France/Northwest Italy, limited areas in Portugal and across the Balkans' coastal areas under March start date, while when considering May as start date the BSS is lower over Balkans and Spain. For FWI, the best score is in northern Algeria and across both westernmost and northernmost Turkey under May start date. Good results are also found for the central part of North Africa and over Morocco under March start date.

Additionally, tercile plots are built for PET to better understand forecast quality over a small area that show the highest correlation skills for this indicator, under ECMWF SPS and QM, and encompassing central-eastern Europe (represented by the red box in Fig. 3). Above-normal seasonal PET conditions are computed yearly for the two start dates, and the observed terciles are also calculated (Fig. 8). The hindcast prediction attained a ROC of 0.26 and 0.74 for above-normal PET years under March and May start dates, respectively. Results show that hindcasts with March as start date tends to overestimate the number of the above-normal events, while when considering May as start date the hindcasts have the skill to discriminate the occurrence of the most severe years (such as 2003, 2007-2009 and 2012).

## 4. Discussion

This work aims to support land planners and decision-making, by exploring effectiveness and reliability of seasonal forecast using agro-climatic indicators rather than on climate variables alone, as previous studies mostly did. Furthermore, assessing forecast performance of indicators across different levels of complexity defines various degrees of reliabilities for user's purposes compared to those of primary climate

variables. Indeed, users may integrate different variables when calculating such indicators representing/affecting specific ecosystem (services) of interest. In addition, assessing the performance of the latest seasonal prediction systems and post-processing chains is mandatory, since the development of forecasting systems and processing methodologies are continuously updating and require regular feedback for optimal development.

Our assessment of seasonal prediction systems and post-processing chains shows that, despite some differences across climate models, post-processing techniques, and start dates, agreement for some specific regions is found. In turn, this suggests the reliability of the potential use of SPSs for decision making under specific boundary conditions/assumptions. In the following, the main elements are discussed to support the community of potential end users, like land managers and decision-makers.

### 4.1. Bias across indicators, start dates, and Seasonal Prediction Systems

The bias between forecasts and observed data prevents, in some cases, the direct use of raw values for taking decisions on land management for supporting climate adaptation, e.g., by using a given threshold to implement decisions based on established hazard scales or by exploiting time series for running impact models. Results show that for both CMCC and ECMWF SPSs, forecasts for the summer period in the Euro-Mediterranean region are warmer and drier than observed data (i.e., higher PET, PSMD, and FWI), whereas in North Africa are wetter and colder. Previous studies with other climate systems, e.g., with the ECMWF System 4 or EC-EARTH, also showed a warmer/drier climate over the Euro-Mediterranean region (Bedia et al., 2018; Manzanas et al., 2018a), suggesting that, beyond own model's drift, these SPSs share commonalities when predicting climate over this region. Similar patterns of bias are observed across different start dates (see Supplementary Material), suggesting that systematic errors might be most likely linked to parameterization and model physics rather than initialization conditions. Indeed, it was shown that refining atmospheric parameters, e.g., through perturbed parameter ensemble and statistical emulations, the systematic surface temperature bias might substantially reduced (Li et al., 2019).

### 4.2. Forecast performance across indicators

PET and PSMD seasonal predictions show greater extent of areas with significant correlations than FWI. PET is based on temperatures only, whereas PSMD is based on precipitation and temperatures. FWI is the most complex metric, as it also includes wind speed and relative humidity. Thus, results suggest that increasing the indicator's complexity may decrease predictability since it relies on the combined forecast of different variables (Bedia et al., 2018). The best performing areas change across indicators, most likely due to the spatial

performances of the variables behind them. Indeed, rough comparisons with the accuracy of individual variables (see <https://www.ecmwf.int/en/forecasts/quality-our-forecasts> and <https://sps.cmcc.it/verifications/>) show that PET follows the spatial patterns of temperatures, whereas PSMD and FWI seem to be decoupled from temperatures and precipitations. Thus, results suggest that, when using any other complex indicators, their accuracy should be tested and understood (as in this study) to evaluate inheritance of inaccuracy from raw variables (i.e., temperatures or precipitation) that might ultimately lead to an unreliable decision-making process.

Results show some agreement and consolidation of reliability across indicators for specific areas in the Iberian Peninsula and Eastern Europe. PET performs better in western and eastern Europe and some areas of North Africa. PSMD follows similar spatial patterns, except for some areas of east and central Europe, where the performance decreases and increases, respectively. Previous studies also found similar skill patterns for temperatures associated with warming trends and weak inter-annual variability (Mishra et al., 2019). Precipitation skill has been found generally limited over Europe (Mishra et al., 2019), likely affecting the performance of indicators like PSMD and FWI. Good promising results are found for these indicators over some eastern, western (e.g., Spain), and northern-central areas in Europe and North Africa, which might be linked to the predictability of ENSO and sea surface temperature (Doblas-Reyes et al., 2013). Since agricultural water use constitutes a large share of total water withdrawal, especially in the Mediterranean region, more reliable seasonal predictions can anticipate/resolve conflicts for water use that involve other sectors that rely heavily on water during the summer period (domestic, tourism, ecosystems etc.). Furthermore, promising results for countries such as Ukraine, Romania, Poland that produce large amounts of cereals might prove useful to anticipate decisions for managing rainfed and irrigated crops, potentially improving production and lowering risks by changing crop calendars, selecting specific cultivars and scheduling proper agronomic practices. Moreover, predictions of PSMD, i.e., water availability for sustainable crop production (Wang et al., 2008), might inform and guide the decision-making process for effective management and more effective allocation of irrigation resources.

Prediction skills for FWI in addition to PET and PSMD over Mediterranean areas (e.g., Spain) might be of high interest for fire risk management. Indeed, as previously mentioned, FWI is one of the most used indices by regional and national agencies (and the European Union) to understand and manage fire risk (e.g., allocation of suppression resources and prevention actions). Our results contrast with those observed by Bedia et al. (2018), which found significant skill prediction over Greece and Bulgaria. Contrasting results suggest that using different climate models (they used the ECMWF system4) and slightly different time periods (they used an extended fire season of 4 months) might affect the results (see next paragraphs).

#### 4.3. Start date influence on forecasts performance

The results also show that the forecasts start date might alter prediction skills and its geographical distribution with significant correlations. For instance, as far as PET is concerned, a displacement of significant correlations is found in Eastern Europe from the north toward the Center Europe when the initialization time changes from March to May.. Results also show that in some areas, e.g., the Iberian Peninsula and North Africa, the good performance of indicators is mainly conserved across start dates, suggesting high confidence for predictions in regions with low rainfall. Previous studies indicated that the shorter the lead times (the distance of the considered period from the forecasts start date), the better the forecast performance (Doblas-Reyes et al., 2013; van den Hurk et al., 2012). However, our results also show that the skill might be higher with the most extended lead times (i.e., March start date) over specific areas, e.g., the northern north-eastern part of the domain, especially for PET and PSMD. Indeed, the spatial accuracy of

raw variables by the SPSs developers showed contrasting patterns across lead times (e.g., <https://sps.cmcc.it/verifications/>). We hypothesized that, potentially, in certain areas, some climate dynamics may be established by certain conditions from the end of winter whereas spring conditions could be potentially more variable and uncertain about building climate projections for the summer period. More detailed analyses linked with teleconnections or soil moisture initial conditions (Ceglar et al., 2018) might prove helpful in better understanding factors associated with changing predictability with start date and lead time. Thus, results suggest that, for instance, forecasts tailored for specific crops and management practices (e.g., crop calendars and field operation timing), and resource management (e.g. water allocation) should revise firstly the performance associated with different start dates and related lead times.

#### 4.4. Difference across climate models and multi-model performance

Although both CMCC and ECMWF SPSs share some common geographical areas with good predictability, some relevant differences are observed. Overall, the ECMWF SEAS5 seemed to perform better than the CMCC SPSv3, especially for PSMD, due to greater extent of areas with significant positive correlation. However, the worst correlation result is always observed for ECMWF SEAS5 under the CAL technique. It is important to notice that ECMWF SEAS5 is based on ERA-interim, which may favour its verification metrics (Hemri et al., 2020). Previous studies have already shown that different models at the basis of SPSs might provide different results because of different assumptions and simplifications (Manzanas et al., 2019; Mishra et al., 2019). However, combining multiple SPSs might give better results because of accounting for multiple plausible solutions (Hagedorn et al., 2005). Indeed, our results show that multi-SPS forecasts increased the extent of skilful predictions. Results suggest that including other SPSs, e.g., from the CDS archive, might improve seasonal forecasts' performance. Thus, managers are encouraged to ask their data and services purveyors to consider seasonal prediction multi-system for more reliable and robust predictions.

#### 4.5. The effect of correction techniques

Overall, post-processing techniques such as DET and QM may generally provide results with correlation levels in line with raw data. Indeed, Manzanas et al. (2019) showed that correction techniques, despite removing bias, which is of primary importance for users, slightly alter patterns compared to raw predictions. Previous studies (Bedia et al., 2018) suggested that detrending might remove spurious correlations, but we did not find a remarkable improvement in correlations for skillful areas. Although trends might affect some verification metrics, detrending is not a common practice in seasonal forecasting studies (Doblas-Reyes et al., 2013; Manzanas et al., 2019; Mishra et al., 2019), most likely due to its little potential impact on results (Pepler et al., 2015). Besides, potential benefits might be masked by possible distortions from data manipulation (e.g., not accomplishing method assumptions). Furthermore, it was observed that removing bias might eliminate the trends to a certain degree (Hempel et al., 2013).

Our results show that QM provides the best results while conserving the properties of the raw time series. QM is more complex than BC and CAL methods since adjusting higher-order moments of the distribution (not only the mean or the variance; Manzanas et al., 2018b; Torralba et al., 2017) and is the preferred method for some national and regional agencies (e.g., European Forest Fire Information System, EFFIS). BC and CAL show the worst results; the spatial extent of significant correlations highly decreased in all indicators, start date, and climate models, as well as the spread of negative correlations. More sophisticated recalibration methods have been tested in previous studies, e.g., Climate Conserving Recalibration, but only marginal benefits were found (Manzanas et al., 2019). Results suggest that data and services purveyors, supported by

**Table 1**

Climate data sources and main characteristics

| Climate dataset | <i>Native resolution</i> | <i>Number of members</i> |
|-----------------|--------------------------|--------------------------|
| CMCC SPSv3      | 1°                       | 40                       |
| ECMWF SEASS5    | 0.32°                    | 25                       |
| ERA5            | 0.28°                    | 10                       |

consolidated scientific findings, should evaluate if the decision-making process might require bias-corrected data since this might influence skillful areas, or if different and better performing bias correction techniques can be integrated.

#### 4.6. Predictions skill of out-of-the-norm and extreme events

Predictions of out-of-the-norm events (i.e., upper than the 66<sup>th</sup> percentile) show similar patterns as correlations but with relevant results limited to a more restricted spatial extent. Previous studies also showed similar patterns across accuracy measures (Pepler et al., 2015). Correlations allow assessing if trends of forecast values correspond to those of observed value, whereas probabilistic scores such as BSS allow evaluating the skill and the added value for forecasting specific events relative to climatology. Combined results suggest high confidence for specific areas and indicators and thus their potential use for decision making. When focusing on extreme events (above the 95<sup>th</sup> percentile), the extent of significant correlations decreased even further. However, in eastern central Europe, hindcasts exhibited a good forecast discrimination of above-normal events for PET, linked to drought events in 2003 (Rebetez et al., 2006), 2007, and 2012 (Stahl et al., 2016), which triggered severe impacts on a wide range of socio-economic sectors. It suggests that forecasts might be potentially useful for predicting catastrophic events for these areas. Indeed, in a global study, Pepler et al. (2015) found that the skill for extreme temperatures followed that of seasonal means but in some areas it was even higher and Prodhomme et al. (2021) showed significant prediction skills for detecting local warm extremes especially in Eastern Europe. It is noteworthy that BSS might be influenced by sample size with the rarer the event, the larger the number of additional samples needed to have reliable findings (Wheatcroft, 2019). Thus, the results for extreme events should be taken with caution by decision-makers and land managers if not supported by an extended period of time.

#### 5. Conclusions

Overall, our results showed the potential usefulness of seasonal forecasts for decision making in agriculture and forestry, especially over western Europe and some areas in the Iberian Peninsula. The study evidenced the sensitivity of results when using different seasonal prediction systems, start dates, and correction techniques. Notably, the effect of combining different climate variables when building indicators should be considered when using proxies for decision-making since performances might differ from general climate variables (e.g., temperature and precipitation). Increasing understandings on systematic errors across SPS forecasts and correction techniques might help to increase the robustness and consistency of predictions making final users more confident. Finally, we encourage clear communication with stakeholders without mismatching concepts to increase their comprehension and interest. Table 1

#### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

This work was developed under the MEDiterranean Services Chain based On climate Predictions (MESDCOPE) project (Grant Agreement N. 689029), funded under ERA4CS, the ERA-NET Consortium "European Research Area for Climate Services".

#### Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.agrformet.2022.108921](https://doi.org/10.1016/j.agrformet.2022.108921).

#### References

- Allen, R., Smith, M., Pereira, S., Perrier, A., 1994. An update for the calculation of reference evapotranspiration. *ICID Bull. Int. Comm. Irrig. Drain.* 35–92.
- Asseng, S., McIntosh, P.C., Wang, G., Khimashia, N., 2012. Optimal N fertiliser management based on a seasonal forecast. *Eur. J. Agron.* 38, 66–73. <https://doi.org/10.1016/j.eja.2011.12.005>.
- Bedia, J., Golding, N., Casanueva, A., Iturbide, M., Buontempo, C., Gutierrez, J.M., 2018. Seasonal predictions of Fire Weather Index: paving the way for their operational applicability in Mediterranean Europe. *Clim. Serv.* 9, 101–110. <https://doi.org/10.1016/j.ciser.2017.04.001>.
- Beguería, S., Vicente-Serrano, S., Reig, F., Latorre, B., 2014. Standardized Precipitation Evapotranspiration Index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *Int. J. Climatol.* 34, 3001–3023.
- Brands, S., 2017. Which ENSO teleconnections are robust to internal atmospheric variability? *Geophys. Res. Lett.* 44, 1483–1493. <https://doi.org/10.1002/2016GL071529>.
- Brier, G., 1950. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* 78, 1–3. [https://doi.org/10.1016/0016-0032\(94\)90228-3](https://doi.org/10.1016/0016-0032(94)90228-3).
- Brown, J.N., Hochman, Z., Holzworth, D., Horan, H., 2018. Seasonal climate forecasts provide more definitive and accurate crop yield predictions. *Agric. For. Meteorol.* 260–261, 247–254. <https://doi.org/10.1016/j.agrformet.2018.06.001>.
- Bruno Soares, M., Daly, M., Dessai, S., 2018. Assessing the value of seasonal climate forecasts for decision-making. *Wiley Interdiscip. Rev. Clim. Chang.* 9, 1–19. <https://doi.org/10.1002/wcc.523>.
- Burgass, M.J., Halpern, B.S., Nicholson, E., Milner-Gulland, E.J., 2017. Navigating uncertainty in environmental composite indicators. *Ecol. Indic.* 75, 268–278. <https://doi.org/10.1016/j.ecolind.2016.12.034>.
- Capa-Moroch, M., Ines, A.V.M., Baethgen, W.E., Rodríguez-fonseca, B., Han, E., Ruiz-Ramos, M., 2016. Crop yield outlooks in the Iberian Peninsula: Connecting seasonal climate forecasts with crop simulation models. *Agric. Syst.* 149, 75–87. <https://doi.org/10.1016/j.jagry.2016.08.008>.
- Ceglar, A., Toreti, A., Prodhomme, C., Zampieri, M., Turco, M., Doblas-Reyes, F.J., 2018. Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast. *Sci. Rep.* 8, 1–9. <https://doi.org/10.1038/s41598-018-19586-6>.
- de Rigo, G., Libertà, G., Durrant, T.H., Vivancos, T., San-Miguel-Ayanz, J., 2017. Forest fire danger extremes in Europe under climate change: variability and uncertainty. [doi:10.2760/13180](https://doi.org/10.2760/13180).
- Doblas-Reyes, F.J., García-Serrano, J., Lienert, F., Rodrigues, L.R.L., 2013. Seasonal climate predictability and forecasting: status and prospects. *WIREs Clim. Chang.* 4, 245–268. <https://doi.org/10.1002/wcc.217>.
- Doblas-Reyes, F.J., Hagedorn, R., Palmer, T.N., 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting - II. Calibration and combination. *Tellus, Ser. A Dyn. Meteorol. Oceanogr.* 57, 219–233. <https://doi.org/10.1111/j.1600-0870.2005.00103.x>.
- Frías, M.D., Iturbide, M., Manzanas, R., Bedia, J., Fernández, J., Herrera, S., Cofiño, A.S., Gutiérrez, J.M., 2018. An R package to visualize and communicate uncertainty in seasonal climate prediction. *Environ. Model. Softw.* 99, 101–110. <https://doi.org/10.1016/j.envsoft.2017.09.008>.
- Hagedorn, R., Doblas-Reyes, F.J., Palmer, T., 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting - I. Basic concept. *Tellus* 57, 219–233.
- Hargreaves, B.G.H., 1994. Defining and using reference evapotranspiration. *J. Irrig. Drain. Eng.* 120, 1132–1139.
- Hayashi, K., Llorca, L., Rustini, S., Setyanto, P., Zaini, Z., 2018. Reducing vulnerability of rainfed agriculture through seasonal climate predictions: A case study on the rainfed rice production in Southeast Asia. *Agric. Syst.* 162, 66–76. <https://doi.org/10.1016/j.jagry.2018.01.007>.
- Hempel, S., Frieler, K., Warszawski, L., Schewe, J., Piontek, F., 2013. A trend-preserving bias correction – the ISI-MIP approach. *Earth Syst. Dyn.* 4, 219–236. <https://doi.org/10.5194/esd-4-219-2013>.
- Henri, S., Bhend, J., Liniger, M.A., Manzanas, R., Siegert, S., Stephenson, D.B., Gutiérrez, J.M., Brookshaw, A., Doblas-Reyes, F.J., 2020. How to create an operational multi-model ensemble of seasonal forecasts? *Clim. Dyn.* 55, 1141–1157.
- Hurrell, J., Meehl, G.A., Bader, D., Delworth, T.L., Kirtman, B., Wielicki, B., 2009. A unified modeling approach to climate system prediction. *Bull. Am. Meteorol. Soc.* 90, 1819–1832. <https://doi.org/10.1175/2009BAMS2752.1>.
- Iizumi, T., Sakuma, H., Yokozawa, M., Luo, J.J., Challinor, A.J., Brown, M.E., Sakurai, G., Yamagata, T., 2013. Prediction of seasonal climate-induced variations in global food production. *Nat. Clim. Chang.* 3, 904–908. <https://doi.org/10.1038/nclimate1945>.

- Iizumi, T., Shin, Y., Kim, W., Kim, M., Choi, J., 2018. Global crop yield forecasting using seasonal climate information from a multi-model ensemble. *Clim. Serv.* 11, 13–23. <https://doi.org/10.1016/j.ciser.2018.06.003>.
- Jacobs, K.L., Street, R.B., 2020. The next generation of climate services. *Clim. Serv.* 20, 100199 <https://doi.org/10.1016/j.ciser.2020.100199>.
- Jha, P.K., Athanasiadis, P., Gualdi, S., Trabucco, A., Mereu, V., Shelia, V., Hoogenboom, G., 2019a. Evaluating the applicability of using daily forecasts from seasonal prediction systems (SPSs) for agriculture: a case study of Nepal's Terai with the NCEP CFSv2. *Theor. Appl. Climatol.* 135, 1143–1156. <https://doi.org/10.1007/s00704-018-2433-5>.
- Jha, P.K., Athanasiadis, P., Gualdi, S., Trabucco, A., Mereu, V., Shelia, V., Hoogenboom, G., 2019b. Using daily data from seasonal forecasts in dynamic crop models for yield prediction: A case study for rice in Nepal's Terai. *Agric. For. Meteorol.* 265, 349–358. <https://doi.org/10.1016/j.agrformet.2018.11.029>.
- Johnson, S.J., Stockdale, T.N., Ferranti, L., Balmaseda, M.A., Molteni, F., Magnusson, L., Tietze, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S.P.E., Mogensen, K., Zuo, H., Monge-sanz, B.M., 2019. SEAS5: the new ECMWF seasonal forecast system. *Geosci. Model Dev.* 1087–1117.
- Klemm, T., McPherson, R.A., 2017. The development of seasonal climate forecasting for agricultural producers. *Agric. For. Meteorol.* 232, 384–399. <https://doi.org/10.1016/j.agrformet.2016.09.005>.
- Lawrence, B.Y.M.G., 2005. The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air A Simple Conversion and Applications. *Am. Meteorol. Soc.* 225–233.
- Li, S., Rupp, D.E., Hawkins, L., Mote, P.W., McNeall, D., Sparrow, S.N., Wallom, D.C.H., Betts, R.A., Wettstein, J.J., 2019. Reducing climate model biases by exploring parameter space with large ensembles of climate model simulations and statistical emulation. *Geosci. Model Dev.* 12, 3017–3043. <https://doi.org/10.5194/gmd-12-3017-2019>.
- Lledó, L.I., Torralba, V., Soret, A., Ramon, J., Doblas-reyes, F.J., 2019. Seasonal forecasts of wind power generation. *Ren* 143, 91–100. <https://doi.org/10.1016/j.renene.2019.04.135>.
- Manzanas, R., Gutiérrez, J.M., Bhend, J., Hemri, S., Doblas-reyes, F.J., Torralba, V., Penabad, E., Brookshaw, A., 2019. Bias adjustment and ensemble recalibration methods for seasonal forecasting: a comprehensive intercomparison using the C3S dataset. *Clim. Dyn.* 53, 1287–1305.
- Manzanas, R., Gutiérrez, J.M., Fernández, J., Meigaard, E.Van, Calmant, S., Magariño, M.E., Cofiño, A., Herrera, S., 2018a. Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: added value for user applications. *Clim. Serv.* 9, 44–56. <https://doi.org/10.1016/j.ciser.2017.06.004>.
- Manzanas, R., Lucero, A., Weisheimer, A., Gutiérrez, J.M., 2018b. Can bias correction and statistical downscaling methods improve the skill of seasonal precipitation forecasts? *Clim. Dyn.* 1161–1176. <https://doi.org/10.1007/s00382-017-3668-z>.
- Marcos, R., Carmen, M., Quintana-seguí, P., Turco, M., 2017. Seasonal predictability of water resources in a Mediterranean freshwater reservoir and assessment of its utility for end-users. *Sci. Total Environ.* 575, 681–691. <https://doi.org/10.1016/j.scitotenv.2016.09.080>.
- Mishra, N., Prodhomme, C., Guemas, V., 2019. Multi - model skill assessment of seasonal temperature and precipitation forecasts over Europe. *Clim. Dyn.* 52, 4207–4225. <https://doi.org/10.1007/s00382-018-4404-z>.
- Oguttu, G.E.O., Franssen, W.H.P., Supit, I., Omondi, P., Hutjes, R.W.A., 2018. Probabilistic maize yield prediction over East Africa using dynamic ensemble seasonal climate forecasts. *Agric. For. Meteorol.* 250–251, 243–261. <https://doi.org/10.1016/j.agrformet.2017.12.256>.
- Olesen, J.E., Trnka, M., Kensebaum, K.C., Skjelvåg, A.O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozýra, J., Micale, F., 2011. Impacts and adaptation of European crop production systems to climate change. *Eur. J. Agron.* 34, 96–112. <https://doi.org/10.1016/j.eja.2010.11.003>.
- Pepler, A.S., Díaz, L.B., Prodhomme, C., Doblas-reyes, F.J., Kumar, A., 2015. The ability of a multi-model seasonal forecasting ensemble to forecast the frequency of warm, cold and wet extremes. *Weather Clim. Extrem.* 9, 68–77. <https://doi.org/10.1016/j.wace.2015.06.005>.
- Prodhomme, C., Materia, S., Ardilouze, C., White, R.H., Batté, L., Guemas, V., Frakoulidis, G., García-Serrano, J., 2021. Seasonal prediction of European summer heatwaves. *Clim. Dyn.* <https://doi.org/10.1007/s00382-021-05828-3>.
- Ramírez-Rodrigues, M.A., Alderman, P.D., Stefanova, L., Cossani, C.M., Flores, D., Asseng, S., 2016. The value of seasonal forecasts for irrigated, supplementary irrigated, and rainfed wheat cropping systems in northwest Mexico. *Agric. Syst.* 147, 76–86. <https://doi.org/10.1016/j.agry.2016.05.005>.
- Rebetez, M., Mayer, H., Dupont, O., Schindler, D., Kropp, P., Menzel, A., Rebetez, M., Mayer, H., Dupont, O., Schindler, D., Gartner, K., 2006. Heat and drought 2003 in Europe: a climate synthesis To cite this version: heat and drought 2003 in Europe: a climate synthesis. *Ann. For. Sci.* 63, 569–577.
- Stahl, K., Kohn, I., Blauhut, V., Urquijo, J., De Stefano, L., Acácio, V., Dias, S., Stagge, J. H., Tallaksen, L.M., Kampragou, E., Van Loon, A.F., Barker, L.J., Melsen, L.A., Bifulco, C., Musolino, D., De Carli, A., Massarutto, A., Assimacopoulos, D., Van Lanen, H.A.J., 2016. Impacts of European drought events: Insights from an international database of text-based reports. *Nat. Hazards Earth Syst. Sci.* 16, 801–819. <https://doi.org/10.5194/nhess-16-801-2016>.
- Stockdale, T.N., Alves, O., Boer, G., Deque, M., Ding, Y., Kumar, A., Kumar, K., Landman, W., Mason, S., Nobre, P., Scaife, A., Tomoaki, O., Yun, W.T., 2010. Understanding and predicting seasonal-to-interannual climate variability - The producer perspective. *Procedia Environ. Sci.* 1, 55–80. <https://doi.org/10.1016/j.proenv.2010.09.006>.
- Talukder, B., Hipel, K.W., VanLoon, G.W., 2017. Developing composite indicators for agricultural sustainability assessment: Effect of normalization and aggregation techniques. *Resources* 6, 1–27. <https://doi.org/10.3390/resources6040066>.
- Thornton, P.K., Erickson, P.J., Herrero, M., Challinor, A.J., 2014. Climate variability and vulnerability to climate change: A review. *Glob. Chang. Biol.* 20, 3313–3328. <https://doi.org/10.1111/gcb.12581>.
- Torralba, V., Doblas-reyes, F.J., Macleod, D., Christel, I., Davis, M., 2017. Seasonal climate prediction: a new source of information for the management of wind energy resources. *J. Appl. Meteorol. Climatol.* 56, 1231–1247. <https://doi.org/10.1175/JAMC-D-16-0204.1>.
- Trnka, M., Olesen, J.E., Kensebaum, K.C., Skjelvåg, A.O., Eitzinger, J., Seguin, B., Peltonen-Sainio, P., Rötter, R., Iglesias, A., Orlandini, S., Dubrovský, M., Hlavinka, P., Balek, J., Eckersten, H., Cloppet, E., Calanca, P., Gobin, A., Vučetić, V., Nejedlik, P., Kumar, S., Lalic, B., Mestre, A., Rossi, F., Kozyra, J., Alexandrov, V., Semerádová, D., Žalud, Z., 2011. Agroclimatic conditions in Europe under climate change. *Glob. Chang. Biol.* 17, 2298–2318. <https://doi.org/10.1111/j.1365-2486.2011.02396.x>.
- Turco, M., Jerez, S., Doblas-reyes, F.J., Aghakouchak, A., Llasat, M.C., Provenzale, A., 2018. Skilful forecasting of global fire activity using seasonal climate predictions. *Nat. Commun.* <https://doi.org/10.1038/s41467-018-05250-0>.
- Turco, M., Marcos-Matamoros, R., Castro, X., Canyameras, E., Llasat, M.C., 2019. Seasonal prediction of climate-driven fire risk for decision-making and operational applications in a Mediterranean region. *Sci. Total Environ.* 676, 577–583. <https://doi.org/10.1016/j.scitotenv.2019.04.296>.
- van den Hurk, B., Doblas-reyes, F.J., Balsamo, G., Koster, R., Seneviratne, S., Camargo, H., 2012. Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Clim. Dyn.* 349–362. <https://doi.org/10.1007/s00382-010-0956-2>.
- Van Wagner, C.E., 1987. Development and structure of the Canadian forest fire weather index system. *For. Techn. Rep.* 35.
- Voces-Aboy, J., Abia-Llera, I., Sánchez-García, E., Navascués, B., Rodríguez-Camino, E., Garrido-del-Pozo, M.N., García-Gómez, M.C., Álvarez-González, J.A., Pastor-Argüello, F., 2019. Web-based decision support toolbox for Spanish reservoirs. *Adv. Sci. Res.* 16, 157–163. <https://doi.org/10.5194/asr-16-157-2019>.
- Wang, E., Yu, Q., Wu, D., Xia, J., 2008. Climate, agricultural production and hydrological balance in the North of China Plain. *Int. J. Climatol.* 28, 1959–1970. <https://doi.org/10.1002/joc>.
- Wheatcroft, E., 2019. Interpreting the skill score form of forecast performance metrics. *Int. J. Forecast.* 35, 573–579. <https://doi.org/10.1016/j.ijforecast.2018.11.010>.
- Zscheischler, J., Westra, S., Van Den Hurk, B.J.M., Seneviratne, S.I., Ward, P.J., Pitman, A., Aghakouchak, A., Bresch, D.N., Leonard, M., Wahl, T., Zhang, X., 2018. Future climate risk from compound events. *Nat. Clim. Change* 8, 469–477. <https://doi.org/10.1038/s41558-018-0156-3>.