Statistical Seasonal Prediction of European Summer Mean Temperature Using Observational, Reanalysis, and Satellite Data

MARIA PYRINA,^a MARCEL NONNENMACHER,^a SEBASTIAN WAGNER,^a AND EDUARDO ZORITA^a

^a *Helmholtz-Zentrum Hereon, Institute of Coastal Systems, Geesthacht, Germany*

(Manuscript received 18 December 2020, in final form 31 May 2021)

ABSTRACT: Statistical climate prediction has sometimes demonstrated higher accuracy than coupled dynamical forecast systems. This study tests the applicability of springtime soil moisture (SM) over Europe and sea surface temperatures (SSTs) of three North Atlantic (NA) regions as statistical predictors of European mean summer temperature (t2m). We set up two statistical-learning (SL) frameworks, based on methods commonly applied in climate research. The SL models are trained with gridded products derived from station, reanalysis, and satellite data (ERA-20C, ERA-Land, CERA, COBE2, CRU, and ESA-CCI). The predictive potential of SM anomalies in statistical forecasting had so far remained elusive. Our statistical models trained with SM achieve high summer t2m prediction skill in terms of Pearson correlation coefficient (r), with $r \geq 0.5$ over central and eastern Europe. Moreover, we find that the reanalysis and satellite SM data contain similar information that can be extracted by our methods and used in fitting the forecast models. Furthermore, the predictive potential of SSTs within different areas in the NA basin was tested. The predictive power of SSTs might increase, as in our case, when specific areas are selected. Forecasts based on extratropical SSTs achieve high prediction skill over south Europe. The combined prediction, using SM and SST predictor data, results in $r \geq 0.5$ over all European regions south of 50°N and east of 5°W. This is a better skill than the one achieved by other prediction schemes based on dynamical models. Our analysis highlights specific NA midlatitude regions that are more strongly connected to summer mean European temperature.

SIGNIFICANCE STATEMENT: The prediction of summer temperature is important for the estimation of weather related risks relevant for the society. Specifically for economy, agriculture, and implications on human health. Due to the chaotic system, and in the absence of only a single agent controlling the climate system, physics based dynamical models do not estimate correctly seasonal mean temperature. We tested whether a statistical model could achieve better predictions of summer temperature over Europe. Our data-driven model uses information from datasets of North Atlantic sea surface temperatures and European soil moisture, and achieves predictions with higher skill compared to physical numerical models. Moreover, we identified the regions in the North Atlantic Ocean that are more strongly related to the high prediction skill of European summer temperature. These findings suggest that statistical models should be used in operational seasonal forecasting.

KEYWORDS: Europe; Atmosphere-land interaction; Atmosphere-ocean interaction; Seasonal forecasting; Statistical forecasting; Seasonal variability; Machine learning

1. Introduction

Seasonal climate prediction focuses on producing forecasts at lead times on the order of a few months. Forecasting at these time scales can help sectors such as health, energy, and agriculture to anticipate changes in service requests, energy demand, and weather related risks (Curtis et al. 2017; Maracchi et al. 2005; Allegrini et al. 2012). Seasonal prediction can be accomplished by both dynamical and statistical models for several climatic variables, such as temperature (Kämäräinen et al. 2019; Manzanas et al. 2018), precipitation (Diez et al. 2011; Brands et al. 2012), and atmospheric pressure indices;

for instance, the North Atlantic Oscillation index (NAO) (Weisheimer et al. 2017; Wang et al. 2017). Sources of predictability at seasonal time scales for both dynamical and statistical models are the slowly varying components of the climate system. These components can act as boundary forcings for the troposphere and subsequently affect local weather and climate after some time lag. In this study, we design statistical models to forecast European summer temperatures one season in advance, based on the springtime predictors soil moisture and North Atlantic sea surface temperature.

At extratropical latitudes during summer and specifically over Europe, there is a link between summer atmospheric predictability and the Atlantic multidecadal variability (AMV). The AMV index summarizes the variability of the North Atlantic SSTs (Schlesinger and Ramankutty 1994) and influences multidecadal variations of European summer

Denotes content that is immediately available upon publication as open access.

Pyrina's current affiliation: ETH Zurich, Institute for Atmospheric and Climate Science, Zurich, Switzerland.

Corresponding author: Maria Pyrina, maria.pyrina@hereon.de

This article is licensed under a [Creative Commons Attribution 4.0 license](http://creativecommons.org/licenses/by/4.0/) (<http://creativecommons.org/licenses/by/4.0/>).

DOI: 10.1175/WAF-D-20-0235.1

temperatures (Ghosh et al. 2017; Sutton and Hodson 2005; Knight et al. 2006; Sutton and Dong 2012). Even though a link between European summer temperature and AMV has been established through the analysis of observations, this link is weaker in coupled dynamical model simulations (Qasmi et al. 2017). Therefore, the potential predictability of land summer temperatures is lower in simulations than in observations. One reason for the loss of prediction skill might be limitations of dynamical models, which mask the potential skill due to such boundary forcings (Cohen et al. 2019; Qasmi et al. 2017).

The seasonal forecasts of summer temperature based on dynamical models reveal limited skill over Europe, as determined either by their spatial anomaly correlation coefficient (ACC; Wilks 2011) or by their temporal correlation coefficient [r ; see Eq. (5)]. The ACC evaluates the spatial skill of the prediction and is used when the predictand variable is spatially averaged. Predictions based on dynamical models attain values of the order of $\text{ACC} \sim 0.3$, for forecasts with lead times longer than three weeks (Kirtman et al. 2014; Doblas-Reyes et al. 2013; Manzanos et al. 2018; Johnson et al. 2019). Higher seasonal forecasting skill of summer temperature ($r \sim 0.6$) is found by some dynamical model studies over specific European regions (Mishra et al. 2019; Bunzel et al. 2018). However, part of this skill may be an artifact, attributed to the common warming trend included in the respective temperature time series that artificially increases their temporal correlation (Doblas-Reyes et al. 2013; Doblas-Reyes et al. 2006). In view of this limited skill, there is an increasing interest in the usage of deep learning for seasonal forecasting (Scher and Messori 2019). However, these methods require large sample sizes for training. The limited availability of observational data for training complex deep learning models renders statistical modeling and simpler machine learning methods worth exploring as alternatives to dynamical seasonal forecasting (Wang et al. 2017; Cohen et al. 2019; Kämäräinen et al. 2019; Totz et al. 2017).

A few previous studies using machine-learning methods for seasonal forecasting have focused on the summer season (Min et al. 2019; Hartigan et al. 2020). Kämäräinen et al. (2019) predicted seasonal temperature averaged over five European domains. For the domains of Europe (EU), Scandinavia (SC), and western Europe (WE) the summer prediction skill in terms of ACC was found equal to 0.78, 0.48, and 0.61, respectively. For the regions of eastern Europe (EE) and Mediterranean (Med), the authors find that the skill is lower and sometimes considerably weaker than the skill achieved for the SC and WE regions. Aiming at increasing the summer prediction skill over the EE and Med regions and provide gridded predictions over those regions, this study tests the applicability of springtime predictors for mean summer temperature prediction over Europe, using reanalysis, observational and satellite gridded data products. Even though the study is focused on summer temperatures over the EE and Med regions, predictability beyond these regions is also assessed in order to determine the spatial extent of European predictability from these springtime predictor data.

The predictors are chosen based on their known influence as boundary forcings of European summer climate.

We implement two statistical learning models, which we set up in statistical-learning frameworks and take advantage of a ML algorithm for optimization of prediction skill. The seasonal prediction of European summer temperature is provided on a horizontal grid of $0.5^\circ \times 0.5^\circ$ ($\sim 50 \text{ km} \times 50 \text{ km}$). The paper is organized as follows. In section 2a, we present the predictor and predictand variables, followed by the description of the statistical methods in section 2b. In section 3, we evaluate the skill of the seasonal predictions. The discussion and conclusions are presented in sections 4 and 5, respectively. The statistical learning frameworks created by the current study are available in python code and are provided online at https://github.com/marpyr/forecast_predictability.

2. Data and methods

a. Data

1) SELECTION OF PREDICTOR PARAMETERS

Predictor data at different temporal lags are traditionally used for statistical seasonal forecasting, especially when the respective forecasts target all seasons. In this way, predictor–predictand lagged causal relationships are taken into account throughout the year. In our study, we use springtime predictor data (March–April–May) and test their predictive power for summer European temperature (mean over June–July–August). The predictive power of spring predictors is tested for the individual spring months, as well as for the whole-season mean and for the mean of late spring, i.e., April–May.

We use here for the first time soil moisture (SM) as statistical predictor of summer mean 2-m temperature (t_{2m}). The SM and t_{2m} variables are selected over the European region 10°W – 30°E , 34° – 70°N . Soil moisture affects the hydrological cycle, and in turn seasonal predictability, by influencing the surface energy fluxes, precipitation and evapotranspiration (van den Hurk et al. 2012; Thomas et al. 2016). Soil moisture anomalies have been extensively investigated and their potential for extratropical seasonal predictability highlighted by several studies (Hirschi et al. 2011; Ardilouze et al. 2017; Seneviratne et al. 2010; Seneviratne et al. 2013). Nevertheless, the effective contribution of soil moisture to summer temperature predictability, in terms of statistical seasonal forecasting, has not been evaluated yet. We take into account the volume of water in the soil layer from the surface down to 7-cm depth, as the topsoil layer is the reservoir for bare-soil evaporation (Stake and Hagemann 2016). Moreover, the topsoil layer is the part that receives precipitation and meltwater during spring.

In this study, we refer to monthly, bimonthly, and seasonal means of SM. Therefore, in contrast to daily and weekly means of SM values, we do not expect noisy SM variability in the top soil level depths. Moreover, we use the top soil layer (0–7 cm), as we test the possibility of combining reanalysis data and satellite data in the same prediction scheme (see section 3d). The monthly climatology of CERA-20C SM products [see description in section 2a(2)] is given in the appendix (Fig. A12), for two European locations and for the four

available soil level depths (layer 1: 0–7 cm, layer 2: 7–28 cm, layer 3: 28–100 cm, layer 4: 100–289 cm). Moreover, in the [appendix](#) (Figs. A13, A14) we show the seasonal mean soil moisture values (mean over winter, spring, autumn, summer) for the period 1914–1970. For the time scales considered in this study, we do not expect that the selection of SM values from any of the layers 1, 2, or 3 would significantly affect the summer t2m predictability skill from springtime SM values.

The second predictor we use is springtime SSTs of the North Atlantic basin. Several studies have reported that the simultaneous SST variability of the NA basin ([Zampieri et al. 2017](#)), but also the SST variability over only its tropical ([Saeed et al. 2014](#)) or extratropical regions ([Ossó et al. 2018](#); [Ratcliffe and Murray 1970](#)) forces the European weather patterns in the one to two subsequent months. As we expect the variability of spring NA SSTs to be a precursor of EU summer temperature through the modulation of summer weather patterns, we test the SST predictive skill of three regions. The regions include what we define as the extratropical North Atlantic (ENA; 85°W–30°E, 34°–76°N), the tropical North Atlantic (TNA; 85°W–0°, 0°–33°N), and the North Atlantic (NA; 85°W–30°E, 0°–76°N).

2) GRIDDED PRODUCTS

Several gridded products are used in our analysis, including satellite products, reanalysis products, and observational datasets. The groups of predictor–predictand data that we used in our analysis are shown in [Table 1](#) (group 1-p1, group 1-p2, group 1-p3, group 2, group 3). The rationale in forming the respective groups relates to different combinations in the target and training dataset (i.e., group 1 forms the target dataset with CRU t2m data). For a better overview, we discuss in the main text only results for permutations within the group 1, unless stated differently. Results for the rest of the groups are presented in the [appendix](#). In the following, we describe in detail the individual data products.

(i) Reanalysis products

Climatic variables (SSTs, SM, t2m) are taken from the European twentieth century reanalysis ERA-20C, which is a product of the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA-20C provides global data for the period 1900–2010 on a horizontal resolution of approximately 125 km (1.25° longitude $\times$ 1.25° latitude grid). ERA-20C relies on a recent version of ECMWF's Integrated Forecast System (IFS) and it assimilates observations of only surface pressure and surface marine winds ([Poli et al. 2016](#)). Another twentieth century reanalysis product used is the CERA-20C, which is also an ECMWF product. It is based on the coupled data assimilation system for climate reanalysis CERA ([Lalouaux et al. 2016](#)), which assimilates surface pressure, marine wind observations, as well as ocean temperature and salinity profiles. CERA-20C provides with a 10-member ensemble of coupled climate reanalysis from 1901 to 2010 on a 125-km horizontal resolution, from which we used the SM data of one ensemble member. SM values from the ERA5-Land dataset are also used. The ERA5-Land is a reanalysis dataset that currently covers the period 1981–2020 (2–3 months before the present). The data are on a regular latitude-longitude grid of 0.1° $\times$ 0.1°, but interpolated for our analysis onto a 1.25° $\times$ 1.25° grid.

ERA5-Land will be extended back to 1950 and deliver timely updates, making it a promising dataset for seasonal statistical prediction. ERA5-Land is produced by a single simulation, without coupling to the atmospheric module of the ECMWF's Integrated Forecasting System (IFS) or to the ocean wave model of the IFS (Copernicus Climate Change Service (C3S) (2019): C3S ERA5-Land reanalysis. Copernicus Climate Change Service, September 2020. <https://cds.climate.copernicus.eu/cdsapp#!/home>).

The ECMWF soil moisture products present the advantage of being consistent over several decades and by design, they are more suitable than their operational counterparts for use in climate studies. The soil moisture product ERA-Interim (not used here) from ECMWF was compared to in situ soil moisture from 117 stations across the world, under different biome and climate conditions. The evaluation showed good skills regarding surface soil moisture variability, but an overestimation of soil moisture particularly for dry lands ([Albergel et al. 2012](#)). The ECMWF soil moisture products used in the current work are taken from the CERA-20C and ERA-20C datasets, which, compared to the scheme used in ERA-Interim, use upgraded versions of the IFS model that is the base of the ECMWF products data assimilation system ([Lalouaux et al. 2018](#)). The upgrades include an improved soil hydrology ([Balsamo et al. 2009](#)), a new snow scheme ([Dutra et al. 2010](#)), a multiyear satellite-based vegetation climatology ([Boussetta et al. 2013](#)) and a new soil moisture analysis scheme for the global land surface, based on a pointwise extended Kalman filter ([Drusch et al. 2009](#); [de Rosnay et al. 2013](#)). Those changes lead to the improvement of the surface soil moisture and the root-zone soil moisture analyses ([Albergel et al. 2012](#)). Moreover, compared to other widely used reanalysis data, [Li et al. \(2020\)](#) showed that the ERA5 reanalysis SM product shows higher skills and statistically significant correlations with SM observations.

(ii) Satellite data

We used soil moisture data generated in the framework of the Climate Change Initiative (CCI) project, which is part of the European Space Agency (ESA) Program on Global Monitoring of Essential Climate Variables (ECV) ([Dorigo et al. 2017](#); [Gruber et al. 2019](#)). For this dataset, soil moisture retrievals from space borne active and passive microwave instruments are merged into a single time series ([Gruber et al. 2017](#)). We use the SM data from the 4.7 version of the ESA-CCI that covers the period 1978–2019 with 25 km horizontal resolution.

(iii) Observational data

Observational t2m data are taken from the Climatic Research Unit gridded time series dataset (CRU TS). This observational dataset is derived by the interpolation of monthly climate anomalies from extensive networks of weather station observations, onto a 0.5° latitude $\times$ 0.5° longitude grid. Here we use the version CRU TS v4.04, which spans the period 1901–2019 ([Harris et al. 2020](#)). Observational monthly SST data are sampled from the version 2 of the Centennial in Situ Observation Based Estimates COBE2 ([Hirahara et al. 2014](#); [Ishii et al. 2005](#)). This dataset combines measurements from the ICOADS Rel.

2.5 (<http://icoads.noaa.gov/>), covers the period 1891–2019 and is interpolated to a $1^\circ \times 1^\circ$ latitude–longitude grid.

3) DATA PREPROCESSING

Due to the large uncertainty of the observations in the earlier years of the time series (Poli et al. 2016) we exclude from our analysis the period 1900–13, following the recommendation of Kämäräinen et al. 2019. To improve the stationarity of the data we removed the 1914–2010 centennial-scale linear trend, separately for each grid cell. The data were then divided in two periods, the training period (1914–70) and the test period being in the time range of seasonal hindcasts (1971–2010). Apart from linear detrending over the whole period, any other information from the test period was strictly barred from entering the fitting of the statistical models using data from the training period. This separation results in training sets of 57 values per grid point. For both periods and all variables, we calculate the anomalies with reference to the climatological mean of the period 1914–1970. Only the anomalies of ESA-CCI and ERA5-Land SM data are calculated with reference to the climatological mean of the available years. In the case that seasonal means are used, we calculate seasonal anomalies by subtracting the climatological seasonal mean. In the case of monthly values, the monthly anomalies are calculated by subtracting the climatological monthly means.

4) PREPROCESSING OF ESA-CCI DATA

The ESA-CCI SM time series are not complete for all grid points, exhibiting monthly and yearly diversity on the spatial location of the missing values. Considering that for our analysis we need as long time series as possible with no missing values, we filled the missing values of the ESA dataset using a bilinear interpolation of the four surrounding grid points. We applied the interpolation only to the months that had less than 6% of missing values. For this 6% limit, only the month May could construct a time series with more than one decade. The May values of 28 years were therefore used as the predictor variable of the ESA-CCI dataset. The 28-yr May period consists of the years 1983, 1985, 1986–87, 1992, 1993, 1995–97, 1999, 2000, 2002, and 2004–19. Because of the shortness of the 28-yr period for both statistical training and prediction, the SM data from the ESA-CCI dataset were used only for prediction, while for the training of the statistical model we used the CERA-20C SM data. To be able to use predictor data from different datasets their spatial dimensions have to match, we have therefore further interpolated the ESA-CCI SM data into $1.25^\circ \times 1.25^\circ$ grid, in order to match the CERA-20C grid resolution.

b. Methods

1) PCA AND GRADIENT DESCENT

To extract the relevant information from the datasets, the first step is to reduce the dimensionality of the data by decomposing the data into a new set of variables, the principal components (PCs), using principal component analysis (PCA) (Pearson 1901; Hotelling 1933). The maximum possible number of PCs depends on the sample size of the dataset. For

example, for a dataset with $x_1, \dots, x_m$ variables being sampled t times, the maximum possible number of PCs is equal to m . If m is a large number, with PCA we capture the essence of $x_1, \dots, x_m$ by a smaller set of variables $PC_1, \dots, PC_k$ having $k \ll m$. The PCs are temporally uncorrelated and successively maximize the described variance. In most studies, the data are decomposed and only the first few PCs are retained as predictors in a multilinear regression (MLR) model, usually the ones determined to have a physical interpretation (see example in Kämäräinen et al. 2019). However, the PCs that represent variance most efficiently do not necessarily have an underlying physical interpretation (Von Storch and Zwiers 2001).

The PCA-based MLR regression model uses the normalized PCs of the predictor field as predictor variables and linearly predicts the PCs of the predictand field. We explore the effect that the number of retained PCs has on the prediction skill of the PCA-based MLR model for an equal number of predictor and predictand PCs $k \in [1, 50]$. The maximum number $k_{\max} = 50$ is chosen based on the cumulative variance captured by the PCs, which we set to be at least 99%. The optimal values of the linear regression parameters that use the PCA method for component selection are calculated with the mini-batch gradient descent algorithm (Baldi 1995).

Given a set of predictor–predictand variables (x_i, y_i) with $i \in [1, t]$ being the time step and $x_i \in R^k, y_i \in R$, our goal is to learn the linear function:

$$f(x) = \omega^T \mathbf{x} + b, \quad (1)$$

with model parameters $\omega \in R^k$ and intercept $b \in R$. To find the model parameters, we minimize the regularized training error given by

$$E(\omega, b) = \frac{1}{t} \sum_{i=1}^t L[y_i, f(x)_i] + \alpha R(\omega). \quad (2)$$

In Eq. (2), y_i represents each of the k retained PCs, $L[y_i, f(x)_i] = (1/2)[y_i - f(x)_i]^2$ is the least squares loss function measuring the models misfit, $R(\omega)$ is a regularization term that penalizes model complexity and $\alpha > 0$ is a nonnegative hyperparameter that controls the regularization strength. The higher the value α , the stronger the regularization. The value we fixed for $\alpha = 1e^{-3}$ and was set by trying out different α values on the training data. The regularization term is equal to $R(\omega) = (1/2) \sum_{j=1}^k \omega_j^2$. The algorithm shown by Eq. (2) iterates over the training data and for each iteration, the model parameters are updated. The number of iterations was set to 5000.

2) CCA

Canonical correlation analysis (CCA) is one of the ML linear methods that are commonly used in the environmental sciences (Hsieh 2009; Zorita et al. 1992). CCA displays some similarities to PCA. Instead of finding new variables that maximize the explained variance in one given dataset, CCA identifies pairs of new variables (the canonical pairs) in two datasets that are optimally correlated in time. The new canonical variables within each of the two datasets are

TABLE 1. Groups of target-predictor data used by our statistical models for training and forecasting, with the respective training and forecasting periods. The period of data availability is given in parentheses, for each of the datasets.

Data groups	Target data: Summer t2m (data availability)	Predictor data: Springtime SSTs (data availability)	Predictor data: Springtime SM (data availability)	Training and prediction periods
Group 1-p1	CRU (1900–2019)	COBE2-SST (1891–2019)	CERA (1901–2010)	Training: 1914–70 Prediction: 1971–2010
Group 1-p2	CRU (1900–2019)	–	CERA used for training (1901–2010) ESA-CCI used for prediction (1978–2019)	Training using CERA: 1914–82 Prediction using ESA-CCI: 1983–2019 ^a
Group 1-p3	CRU (1900–2019)	–	CERA used for training (1901–2010) ERA5 used for prediction (1981–2019)	Training using CERA: 1914–82 Prediction using ERA5-Land: 1983–2019
Group 2 ^b	ERA20c (1900–2010)	COBE2-SST (1900–2019)	CERA (1901–2010)	Training: 1914–70 Prediction: 1971–2010
Group 3 ^b	ERA20c (1900–2010)	ERA20c (1900–2010)	ERA20c (1900–2010)	Training: 1914–70 Prediction: 1971–2010

^a The ESA-CCI data used in this study include only 28 years during the period 1983–2019. For more details, see text.

^b The results given by these data combinations are shown in the [appendix](#).

uncorrelated. As in PCA, this optimality conditions lead to an eigenvalue problem. The resulting (high) correlated pairs of new variables can be used to set up a linear regression model.

We use CCA to select the variables included in the PCA-based MLR regression model. For a dataset with $x_1, \dots, x_m$ variables and another dataset with $y_1, \dots, y_l$ variables being sampled t times, the pairs of predictor–predictand variables

FIG. 1. Statistically significant prediction skill (r) calculated per grid point for the t2m prediction during 1971–2010, using the CCA simple model for $n = 1$ number of CCA pairs selected from $k = 2$ number of PC variables ($n = 1, k = 2$). The model was trained during 1914–1970 with data from group 1-p1. The results are given for the predictor variables (a) NA SST, (b) ENA SST, and (c) SM. Shown from left to right are the results calculated by predictor variables that include spring mean values, April–May mean values, March monthly values, April monthly values, and May monthly values. The dashed and solid black lines indicate the isopleth lines for $r = 0.3$ and $r = 0.5$, respectively. The correlation r , averaged for the European region (r_{aver}) is given on the top of each subpanel.

FIG. 2. As in Fig. 1, but for the best PCA-based MLR models given per predictor variable. The models were trained during 1914–1970 for $n = k$ number of PCs.

is chosen according to the criterion of maximum correlation (Hotelling 1992). Because of the large number of x_m , y_l variables it is common to use PCA to prefilter the data and reduce the dimensions of each dataset. Therefore, we decompose each dataset x_m , y_l with PCA into $k \in [1, 50]$ number of new variables PC_k^x and PC_k^y . Then the CCA method is applied to the k number of variables in order to find another set of variables, the canonical variables $u_n = \mathbf{f}_n^T PC^x$ and $v_n = \mathbf{g}_n^T PC^y$ with $n \in [1, k]$ and individual pairs that exhibit maximum correlation. For these canonical variables, the correlation between the pair u_l and v_l is maximized. Moreover,

the variables u_p , u_q and the variables v_p , v_q are uncorrelated for all values $p \neq q$.

Given the set of the predictor–predictand canonical pairs (u_n, v_n) with $\mathbf{u} \in R^n$, $\mathbf{v} \in R^n$, our goal is to learn the linear function:

$$\mathbf{f}(\mathbf{u}) = \mathbf{W}\mathbf{u} \quad (3)$$

for each of the n retained canonical pairs. To find the model parameters $\mathbf{W} \in R^{n \times n}$, we minimize the training error given by $L[v_n, f(u)_n] = \sum_{i=1}^n [v_n - f(u)_n]^2$, which is the least squares loss function. The optimal values of the linear models'

FIG. 3. (a) Statistically significant ensemble mean r values calculated for the predicted t2m during 1971–2010, using the CCA simple model ($n = 1$, $k = 2$) trained with May ENA SST data from the group 1-p1 during 1914–1970. (b) Statistically significant ensemble mean r values calculated for the predicted t2m using the persistence forecast. (c) Difference of CCA-based r values in (a) minus persistence r values in (b). The MSSS scores calculated for the reference forecasts of (d) persistence and (e) climatology. The solid black line indicates the isopleth line for $r = 0.5$.

FIG. 4. As in Fig. 3, but for the April–May mean SM predictor values.

regression parameters are calculated with the least squares method, which minimizes the sum of squared errors (SSE). We explore the effect that the number of retained canonical pairs has on the prediction skill of the CCA-based MLR model, by testing the prediction skill of $n \in [1, k_{\max}]$ retained pairs with $k_{\max} = 50$. Another possibility for constraining CCAs pairwise selection is to apply it to a smaller number k of retained PCs (i.e., $k_{\max} = 10$). We expect that the CCA based model with $n = 1$ will be different for i.e., $k_{\max} = 50$ and $k_{\max} = 10$. We have therefore additionally tested the performance of the simplest possible CCA-based model, being the model with $n = 1, k = 2$ i.e., just one canonical pair derived from two PCs of the predictor and two PCs of the predictand field. The model with $n = 1, k = 1$ is the simplest PCA-based regression.

3) VALIDATION METRICS

The metrics used for validating the quality of modeling against observed values were the temporal correlation coefficient (r) and the mean square skill score (MSSS). These metrics are calculated for each grid point during the test period. The r refers to the Pearson correlation coefficient calculated between the centered anomalies of the observed and predicted values and describes the model's ability to reproduce the observed temporal evolution of anomalies (independently of their amplitude). The MSSS informs about the model's relative accuracy compared to the given reference forecast and has values in the range $[-\infty, 1]$ with a score equal to one indicating a perfect accuracy. The MSSS and r for each grid cell are defined as follows:

$$\text{MSSS}_{\text{ref}} = 1 - \frac{\frac{1}{n} \sum_{t=1}^n (y_t^m - y_t^0)^2}{\frac{1}{n} \sum_{t=1}^n (y_t^{\text{ref}} - y_t^0)^2} = 1 - \frac{\text{MSE}_m}{\text{MSE}_{\text{ref}}}, \quad (4)$$

$$r = \frac{\sum_{t=1}^n (y_t^m - \bar{y}_m)(y_t^0 - \bar{y}_0)}{\sqrt{\sum_{t=1}^n (y_t^m - \bar{y}_m)^2} \sqrt{\sum_{t=1}^n (y_t^0 - \bar{y}_0)^2}}, \quad (5)$$

where for each time step t , y_t^m is the predicted value, y_t^0 is the target value for a given grid cell, and $\bar{y}_m, \bar{y}_0$ are their mean values over all time steps, respectively. The y_t^{ref} reference forecast refers to the persistence or to the climatology forecast at each time step.

Significance for the Pearson correlation values is derived for a given grid cell via bootstrapping at the 95% confidence level using 500 samples (Mudelsee 2014). Bootstrapping uses random sampling with replacement, meaning that a unit selected at random from the population is returned to the population and then a second element is selected at random. More specifically, after we predict t2m for the test period (1971–2010) we randomly shuffle the predicted t2m time series and calculate the correlation with the target t2m series. We repeat this procedure 500 times and build an empirical distribution from where the confidence intervals are derived.

3. Results

a. Predictors of summer temperature over Europe

The spatial extent of predictability of European summer mean t2m from springtime predictor data is assessed for

FIG. 5. As in Fig. 3, but for the combined SST and SM temperature prediction.

FIG. 6. (a),(b),(d),(e) Projected spatial patterns and (c),(f) their standardized temporal evolution for the $u-v$ pairs related to the SST-t2m prediction in (a)–(c) and the SM-t2m prediction in (d)–(f).

European SM and for the SSTs of three NA regions (NA, ENA, and TNA). The results regarding the predictions utilizing TNA SSTs are shown in the [appendix](#) (Fig. A10), as no statistically significant prediction skill is found for the datasets and model setups used. The predictive power of spring predictors is tested for the individual spring months, as well as for the whole-season-mean and for the mean of only April–May. To explore the effect of the number of retained predictors in the statistical models, we test two setups for the prediction models utilizing the CCA method and one setup for the models utilizing the PCA method, as described in [sections 2b\(1\)](#) and [2b\(2\)](#).

The gridded summer t2m prediction skill in terms of r values is given for the best temperature predictors in [Fig. 1](#) for the models based on the CCA method, and in [Fig. 2](#) for the models based on the PCA method. The best prediction skill for the CCA method, is achieved by the CCA simple model setup ($n = 1, k = 2$) with May ENA SSTs (fifth subpanel of [Fig. 1b](#)) and April–May mean SM (second subpanel of [Fig. 1c](#)). Similar skill is achieved by the PCA MLR model setup using April–May mean NA SSTs ($n = k = 4$) and April–May mean SM ($n = k = 2$), as shown in the second subpanels of [Figs. 2a](#) and [2c](#), respectively. The gridpoint correlation r is averaged over the European continent (r_{aver}) and is given on the top of each subpanel of [Figs. 1](#) and [2](#).

Even though similar skills are achieved by both methods, the CCA simple model setup is simpler to implement and provides

as skillful predictions as the PCA-based MLR model, with the inclusion of only one predictor variable in the regression model. The same applies for the other data groups, with the CCA simple model setup (see [Figs. A1–A3](#)) reaching similar prediction skill to the PCA MLR model (see [Figs. A8, A9, A11](#)). This implementation of the CCA simple model is different from the commonly used MLR setup of CCA, where CCA is set to identify optimal linear combinations of predictand–predictor PCs out of a large number of available PCs of each dataset. In the MRL case setup, CCA chooses optimal pairs in terms of temporal correlation. In the CCA-simple model case we restrain CCA to construct just one pair out of the first two leading predictand and two leading predictor PCs. By definition, the variance that relates to the original datasets is captured by the leading PCs and decreases as the PC rank increases. Therefore, restraining the CCA choice affects positively the results, as CCA is forced to select only one pair out of the two modes that explain most of the variability of the respective predictand and predictor variables.

Regarding the PCA-MLR model, we find that the optimal number of retained PCs depends on the predictor variable and on the dataset used, making it hard to define a dataset independent PC truncation (see [Figs. A5–A7](#)). For example, in the case of group 1-p1 and group 2, the best PCA SST predictor is April–May mean values from the NA basin for $k = 4$, while for

FIG. 7. (a) Statistically significant r values for the ESA-CCI prediction data (group 1-p2) calculated with the CCA simple model ($n = 1$, $k = 2$) using May monthly values of specific years. Statistically significant ensemble mean r values for the ERA5-Land prediction data (group 1-p3) calculated with the CCA simple model, using (b) May monthly values and (c) April–May mean values. The results shown in all subpanels, regard models trained with CERA-data during the period 1914–1982, and predict t2m during 1983–2019 (see Table 1). The dashed and solid black lines indicate the isopleth lines for $r = 0.3$ and $r = 0.5$, respectively. The correlation r , averaged for the European region (r_{aver}) is given on the top of each subpanel.

the group 3 the best prediction is achieved for $k = 2$ with the prediction skill degrading for a larger k .

b. Persistence, climatology, and combined forecasts

The two best performing CCA simple models were compared to the climatological and persistence forecasts. The results are shown in Figs. 3 and 4 for the model trained with SSTs and SM, respectively. Figure 3 displays the r values resulting from the correlation between the target summer and predicted summer t2m, using for prediction a) May ENA SSTs, b) the persistence forecast calculated from May t2m values, and c) the difference in r , CCA-predicted correlation minus persistence-predicted correlation. The r values might differ if we perform the calculation of the correlation within the same test period (1971–2010), but include slightly different years in the test set. Therefore, to estimate the spread of the test r during 1971–2010, we calculate ensembles of r values. This is done by calculating for each member of the ensemble, the correlation between the target and predicted t2m series by leaving one year out in every iteration, leading to a total number of 30 ensemble members. This method is commonly applied when evaluating the skill of seasonal predictions (Neddermann et al. 2019). For every iteration, the r was calculated and the mean r over all iterations is shown in Figs. 3a and 3b. Significance is derived for every iteration via bootstrapping at the 95% confidence level using 500 samples. The r ensemble mean is shown only over the regions where statistical significance is denoted in every iteration. The r ensemble mean (Fig. 3a) is found to be essentially the same to the r calculated for the full test period (Fig. 1b–May), denoting the robustness of the prediction within the test period. Our model's skill to reproduce the observed anomalies, in terms of r , is better than the persistence forecast over all regions east of 5°E and over the southern part of Spain (positive values displayed in Fig. 3c). The performance in terms of MSSS scores is shown in Figs. 3d and 3e for the reference forecasts of persistence and climatology, respectively. The relative accuracy of our prediction surpasses the

skill of both reference forecasts, with the skill of the climatological forecast being harder to improve. The respective analysis to the one shown in Fig. 3 is displayed in Fig. 4 for the CCA simple model that uses April–May mean SM values. In this case, the temperature persistence forecast was calculated using April–May mean temperature values as well. The results show that the CCA model built for soil moisture also surpasses the accuracy of both reference forecasts.

A generally better prediction than the two individual predictions, based on SM and on SST, can be achieved by combining both. The temperature predictions of the two best performing CCA simple models, trained with May ENA SSTs and April–May mean SM, were combined by taking the average of their individual t2m predictions. The correlation of the averaged prediction to the target summer t2m is shown in Fig. 5, together with the MSSS scores for the reference forecasts of persistence and climatology. The averaged prediction results in similar or higher skill over the European regions south of 50°N when compared to the two individual predictions. However, there is loss in skill for the regions north of 50°N ($r \approx 0.4$); over those regions, training with soil moisture alone reaches $r \geq 0.5$.

c. Spatial covariability patterns

To identify the regions that more strongly contribute to the canonical variables incorporated in the CCA simple models, we calculated the projection of the retained u and v variables on the field anomalies of the predictor and predictand data, respectively. The projected spatial pattern pairs are shown in Figs. 6a–c, for the u – v pair related to the best SST–t2m prediction, and in (Figs. 6d–f) for the u – v pair related to the best SM–t2m prediction. The rightmost panels of Fig. 6 show the standardized u – v variable pair as well as its Pearson correlation coefficient.

Our method relates a positive summer t2m pattern to a May SST pattern that exhibits its highest positive values south of Newfoundland as well as in the Baltic, North, and Norwegian Seas, and displays negative values on the southwest of Iceland and

in the Labrador Sea. A similar positive summer t2m pattern is also related by our method to a negative SM anomaly pattern (April–May SM). The SM anomaly pattern has its highest negative values over the eastern and western parts of the study region. The strength of the canonical correlation for the SST–t2m pair is equal to 0.4, whereas for the SM–t2m pair is 0.3. This analysis yields similar results when using the other data groups (not shown).

The spatial patterns of the predictand t2m patterns are very similar in either cases, using either the May SSTs or the April–May mean SM as predictor. In addition, the temporal evolution of the canonical variables (black solid line in Figs. 6c,f) is very similar. These results strongly suggest that the mechanisms involved in the t2m predictability based on SST and in the t2m predictability based on SM are likely related. A plausible interpretation is that the SST May pattern influences May precipitation over Europe, which in turn modifies soil moisture in May and, through persistence in summer. Summer soil moisture would then modulate summer t2m. However, to ascertain this suggestion, additional simulations with one or several Earth system models would be necessary, as we argue in the discussion section.

d. Potential of the ESA-CCI dataset for temperature prediction

A robust seasonal statistical forecast requires long, recent, and spatially as complete as possible gridded data for training and forecasting. The temporal coverage of the predictor data should be frequently updated with current values, in order for the statistical forecast to be operational. It is common that reanalysis datasets are either long, but not up to date, or they are short covering only the last two to three decades. However, long gridded datasets of t2m and SSTs exist from interpolated observations that are updated every year, but that is not the case for SM datasets. SM satellite data cover the recent decades, usually starting from the late 1970s; however, the sample size of these data is not large enough for training and testing the predictive power of a statistical model. Moreover, the unique ESA-CCI SM dataset has many missing values, therefore the effective amount of years that can be exploited is even less. In statistical prediction it is generally avoided, although not entirely uncommon (Kadow et al. 2020), to train and predict using the same climatic variable from different datasets. Nevertheless, in spite of the limitations of SM, we have tested this possibility for SM data.

The aim is to determine whether the ESA-CCI dataset can be utilized for t2m operational forecasting. We use the group 1-p2 datasets in order to train the statistical model with CERA SM reanalysis data from the period 1914–82 and forecast using certain years of ESA-CCI data during the period 1983–2019 [see section 2a(4)]. The predictive power, in terms of r , for the ESA-CCI dataset was only tested for the month May (Fig. 7a), as the majority of years for the rest of the spring months includes large regions over central Europe with missing values. The r values for the ESA-CCI prediction share a similar correlation pattern to the one shown by the data group 1-p1 regarding the prediction period 1971–2010 (May in Fig. 1c). The similarity of the correlation patterns shows that the different datasets contain similar information regarding the relationship of mean summer t2m and May SM values.

To be consistent with the test period 1983–2019 used for the ESA-CCI prediction, we have additionally tested whether the

technique of training and predicting using different datasets provides similar results for reanalysis datasets. Therefore, we train the statistical model using CERA SM data from the period 1914–1982 and forecast using ERA5-Land SM data of the full 1983–2019 period (group 1-p3). The predictive power of the ERA5-Land dataset was tested for May and April–May mean values. The correlation is shown in Figs. 7b and 7c, respectively. A correlation pattern similar to the one found for ESA-CCI May values, but with higher magnitude, is identified for the ERA5-Land prediction. Moreover, the ERA5-Land prediction that utilizes April–May mean predictor data shares a correlation pattern similar to the one shown by the respective prediction of group 1-p1 (Fig. 4a). Still, some differences exist, for example in the size of the area with statistically significant r values. The differences can be expected, as in the case of the ERA5-Land prediction the statistical model is trained with a different dataset.

Overall, the analysis showed that it is possible to use different datasets of soil moisture in the training and forecast modes of our operational statistical model, as the satellite and reanalysis data that we used display similar summer t2m predictability. The ESA-CCI dataset has potential for statistical prediction of European summer temperature; however, that will not be possible until missing grid cells for April over central Europe are filled in.

4. Discussion

a. Comparison to the prediction skill of dynamical and other statistical models

The prediction skill of our statistical models is comparable, or even higher, to that obtained from dynamical models. Dynamical model predictions are usually constructed from the simple or weighted mean of a multimodel or single model ensemble in order to improve the forecast skill. The model-spread is a first-order estimation of forecast uncertainty. The prediction skill of mean t2m in boreal summer over the period 1980–2005 for the ENSEMBLES multimodel forecast system (Weisheimer et al. 2009) is of the order of $r = 0.3$ (Doblas-Reyes et al. 2013). Over Europe, Doblas-Reyes et al. (2013) found r values mostly between 0.2 and 0.4, with few regions reaching r values of 0.4–0.6 (Fig. 5a in Doblas-Reyes et al. 2013). Similar t2m summer prediction skill ($r \approx 0.2$ –0.4) was estimated for the forecast system SEAS5 for the period 1981–2016 (Johnson et al. 2019). Lower prediction skill of summer mean t2m over EU was found for the grand NMME-1 multimodel ensemble ($r \approx 0$ –0.2) for the period 1982–2009 (Kirtman et al. 2014). However, the prediction skill of dynamical models depends on the hindcast length (Shi et al. 2015) and on the approach used for weighting the ensemble, which could lead to an improvement of the forecast skill (Slater et al. 2017).

The dynamical model output displays a warming temperature trend, which is pronounced since the 1980s. This trend is included when estimating the forecast quality and causes an overestimation of the forecast skill (Doblas-Reyes et al. 2006). The skill of predicting the variability around the warming trend is actually much lower (Weisheimer et al. 2011). Turco et al. (2017) evaluated the skill of the operational dynamical forecast system ECMWF System 4 using raw and detrended data. In their supplementary Fig. S3, Turco et al. (2017) show the r values for the target model output and for the detrended model output regarding monthly

temperature forecasts started in May, for the period of 1981–2015. The t2m skill for all forecasted months degrades when detrended data are used. Furthermore, even though the skill degrades as lead time increases, the ECMWF forecast during July and August (started in May) is more skillful over the south of Europe, and specifically over regions shown by our analysis as the most predictable regions by SST predictor data (Fig. 3a).

Our t2m prediction skill over eastern Europe and the Mediterranean region (Fig. 5a, $r \approx 0.5$ – 0.7) is higher than the skill reached by the statistical model of Kämäräinen et al. (2019) for the regional mean of those areas (ACC lower or considerably weaker than 0.48). Our study has also assessed t2m predictability beyond these regions. The spatial extent of t2m predictability from a combined SST and SM prediction is expected for all European regions south of 50°N and east of 5°W (Fig. 6a, $r \geq 0.5$). That was shown by both the PCA-MLR and CCA simple model, but for different number of optimal parameters.

We performed additional experiments with swapped prediction and training periods (not shown), to assess whether the prediction skill is maintained. Training during 1971–2010 and forecasting the full period 1914–70 results in considerably lower prediction skill ($r \approx 0.2$). However, forecasting during the more recent period (1941–70) leads to skill improvement ($r \approx 0.3$ – 0.4), suggesting that a reason for the low prediction skill for the swapped-period experiments is the data quality of the earlier years. Increased predictability of summer t2m during the recent period could mean that the statistical links established are stronger in recent years due to the changing climate (Hoffmann 2018; Quesada et al. 2012). Practically, that would not affect our framework for operational forecasting, as the model will be trained with all available data and rerun every year. Additional tests performed for dependent predictions, using predictors included in both the training and prediction periods, did not result in considerable changes in the prediction skill (see Figs. A17 and A18).

To assess the generalizability of our results we used out-of-sample (OOS) evaluation, where a section from the end of the time series is withheld for evaluation. This technique is suitable for real-world scenarios, when different sources of nonstationary variation are at play, in order to preserve the temporal order of observations (Cerqueira et al. 2020). Another widely used method to assess the generalizability of algorithms, is cross validation (CV). However, the serial correlation in the data, along with possible nonstationarities, make the use of CV problematic, as it does not account for these issues (Bergmeir and Benítez, 2012). We have additionally applied the CV technique and trained the CCA-based simple model during the period 1914–2010. In each training iteration, only one year of the period 1971–2010 is predicted. The predicted year is left out of the training period. The results indicate t2m predictability over south Europe for SSTs and over central and eastern Europe for SM, agreeing with the results shown in Figs. 3a and 4a (see Fig. A15).

b. Physical relationships regarding the skill of the predictions

The patterns of t2m predictability shown in Figs. 6b and 6e, and the very similar temporal evolution of the canonical variables shown Figs. 6c and 6f, suggest that the mechanisms that link May SST to summer t2m on the one hand, and SM to summer t2m on the other hand, may be related. It is plausible that the SST pattern

shown in Fig. 6a modulates May precipitation over Europe. May SM would then, through persistence, also influence summer SM and thus summer t2m. However, the proof for this causal chain of events is statistically challenging. The SST anomalies in the North Atlantic are not only the result of ocean circulation anomalies—an external driver in this context—but are strongly influenced by the atmospheric heat fluxes and the atmospheric circulation. The SST anomalies created by the atmospheric fluxes may also feedback onto the atmospheric circulation itself. Both variables are then coupled, and it is difficult to disentangle the cause-and-effect mechanisms. Disentangling where exactly reality lies requires dedicated simulations with global atmospheric models with prescribed SSTs. Our results provide a plausible set of hypotheses on the causality of t2m predictability that subsequent modeling studies may explore.

Temperature predictability over central and eastern EU owed to soil moisture, is found by our study over regimes characterized as transitional wet/dry regimes (Seneviratne et al. 2010). Transitional regimes can alter air temperature by up to 6–7 K, because over those regimes soil moisture exhibits large changes over its full range, while typical soil moisture variations can impact air temperature by up to 1.1–1.3 K (Schwingshackl et al. 2017). Especially over central and eastern European regions there is large climate model uncertainty in future temperature projections (Vogel et al. 2018; Seneviratne et al. 2013), pointing to the importance of an adequate representation of land–atmosphere feedbacks by climate models (Vogel et al. 2017; Guillod et al. 2014). Uncertainties in land–atmosphere feedbacks are also expected to affect the forecast skill of seasonal dynamical models, as these models are based on the same ESMs used for climate projections. Several studies have also indicated gains in seasonal forecast skill for a correct initialization of soil moisture (e.g., van den Hurk et al. 2012; Mahanama and Koster 2005). The results of those studies indicate the importance of using seasonal forecast systems for a better understanding of the mechanisms governing prediction on seasonal time scales, while our study demonstrates the decisive contribution of data driven statistical models for operational forecasting.

Apart from the framework on operational forecasting, our study provides a method that can possibly lead to skill improvement if applied in seasonal forecasts for ensemble member subselection. That can be accomplished by selecting only the ensemble members that adequately represent the statistical links related to summer t2m predictability. A further evaluation of the selected ensemble subset would be the representation of the SST and SM precursor patterns associated with summer t2m. Our framework's predictions regard summer and are thus highly related to agriculture. Therefore, our prediction of mean summer temperature could be a valuable input to other statistical models, such as of those predicting summer grain yield (Holzkämper et al. 2012; Robinson et al. 2016).

5. Conclusions

Observational benchmarks of climate predictability are essential in order to progress in model-based climate predictions. Here we have used reanalysis, observational and satellite data and established a robust statistical relation between springtime soil moisture, extratropical North Atlantic SSTs, and summer mean t2m.

Our framework for seasonal climate prediction is exemplified by gridded predictions of European summer climate. However, it must be noted that the statistical models built here regard only summer European mean temperature and may not be extensible to other variables, regions, and seasons. The merit of our CCA-based simple forecast model is that it offers the possibility of looking into the driving mechanisms of skillful predictions, as the predictor patterns and even their individual magnitude and importance are known for the predictand under investigation. Moreover, it is a very low-cost, but still effective predicting system. A similar setup, but possibly with other predictors could also be beneficial for other target regions and variables.

In summary, we can conclude that statistical schemes may provide useful seasonal predictions of European summer temperature with a skill higher to present model-based predictions. It is also likely that the main physical mechanisms responsible for summer predictability are related to summer soil moisture, which can be driven by the atmospheric circulation in the previous season.

Acknowledgments. We acknowledge the COBE-SST2 data provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, from their website at <https://psl.noaa.gov/>. We also acknowledge CRU TSv4 for providing the data via the Climatic Research Unit (CRU) website at <https://sites.uea.ac.uk/cru/data>. This work was funded by the project Reduced Complexity Models (REDMOD; <https://redmod-project.de/>), funded by the Helmholtz Association (Grant ZT-I-0010). The authors declare no conflict of interests.

Data availability statement. The ESA-CCI soil moisture data can be accessed at <https://www.esa-soilmoisture-cci.org/node/145>. The ERA20C, CERA, and ERA5-Land reanalysis data can be accessed at <https://www.ecmwf.int/en/forecasts/datasets>. For the rest of the datasets see the acknowledgements.

APPENDIX

Supplementary Results Showing the Prediction Skill of the Different Types of Statistical Models Tested, for all Groups of Predictor–Predictand Data Shown in Table 1

Figures A1–A3 show the performance of the CCA-simple model using data from the group 1-p1, group 2, and group 3. Figures A4–A7 show the effect of the number of retained predictors on the prediction skill of the CCA-based MRL model ($n \in [1, k_{\max} = 50]$ number of CCA pairs) and the PCA-based MLR model ($k \in [1, 50]$ number of PC variables). The prediction skill is averaged over Europe and is given for different data groups. The skill of the gridded predictions is given in Figs. A8–A11, but for fewer predictors (CCA: $n = 1-4$, $k = 50$ and PCA: $n = k = 1-4$). Figures A12–A14 show the monthly climatology and interannual seasonal variability of the four soil moisture layers for the CERA-20C data, for two European locations. Figures A15–A18 show whether the prediction skill of the best performing CCA-based models degrades when different training and prediction periods are used.

FIG. A1. As in Fig. 1, but no statistical testing is applied in the results shown in this figure.

FIG. A2. As in Fig. A1, but for group 2.

FIG. A3. As in Fig. A1, but for group 3.

FIG. A5. As in Fig. A4, but for the PCA-based MLR model, for $k \in [1, 50]$ number of PC variables.

FIG. A6. As in Fig. A5, but using data from group 2.

FIG. A7. As in Fig. A5, but for group 3.

FIG. A8. Gridpoint temporal correlation (r) calculated for the t2m prediction during 1971–2010, using MLR models trained during 1914–1970 with NA SST data from group 1. The results are given for the methods CCA ($n = 1-4, k = 50$) and PCA ($n = k = 1-4$), shown in the top and bottom rows, respectively; in (a) for the spring mean, (b) April–May mean, (c) April, and (d) May predictor data. Shown from left to right are the results for the predictor variables including $n = 1, 2, 3$, and 4 predictor variables, respectively. The thick black line indicates $ACC = 0.5$. No statistical testing is applied in the results of this figure.

FIG. A9. As in Fig. A8, but for ENA SST data from group 1.

FIG. A10. As in Fig. A8, but for TNA SST data from group 1.

FIG. A11. As in Fig. A8, but for SM data from group 1.

FIG. A12. Monthly climatology of the four soil moisture layers for the CERA-20C data (layer 1: 0–7 cm, layer 2: 7–28 cm, layer 3: 28–100 cm, and layer 4: 100–289 cm) and for two European locations. (left) The first location is in central Europe (50°N, 10°E), and (right) the second location is in Spain (40°N, 7°W).

FIG. A13. Interannual seasonal variability for the period 1914–70, for the location in central Europe (50°N, 10°E). The variability of the soil moisture layers (top) 1 and 2 and (bottom) 1 and 3.

FIG. A14. As in Fig. A13, but for the location in Spain (40°N, 7°W).

FIG. A15. Prediction skill (r) calculated per grid point for the t2m prediction during 1971–2010, using the CCA simple model for $n = 1$ number of CCA pairs selected from $k = 2$ number of PC variables ($n = 1, k = 2$). The model was trained during 1914–2010 using the cross-validation technique. The results are given for the predictor variables (a) May ENA SST and (b) April–May mean SM. Statistical significance is calculated via bootstrapping at the 95% confidence level using 500 samples.

FIG. A16. Prediction skill (r) calculated per grid point for the t2m prediction during 1971–2010, using the CCA simple model for $n = 1$ number of CCA pairs selected from $k = 2$ number of PC variables ($n = 1, k = 2$). The model was trained during 1946–70. The results are given for the predictor variables (a) May ENA SST and (b) April–May mean SM. Statistical significance is calculated via bootstrapping at the 95% confidence level using 500 samples.

FIG. A17. Prediction skill (r) calculated per grid point for the t2m prediction during 1951–2010, using the CCA simple model for $n = 1$ number of CCA pairs selected from $k = 2$ number of PC variables ($n = 1, k = 2$). The model was trained during 1914–70. The results are given for the predictor variables (a) May ENA SST and (b) April–May mean SM. Statistical significance is calculated via bootstrapping at the 95% confidence level using 500 samples.

FIG. A18. Prediction skill (r) calculated per grid point for the t2m prediction during 1961–2010, using the CCA simple model for $n = 1$ number of CCA pairs selected from $k = 2$ number of PC variables ($n = 1, k = 2$). The model was trained during 1914–70. The results are given for the predictor variables (a) May ENA SST and (b) April–May mean SM. Statistical significance is calculated via bootstrapping at the 95% confidence level using 500 samples.

REFERENCES

- Albergel, C., P. de Rosnay, G. Balsamo, L. Isaksen, and J. Muñoz-Sabater, 2012: Soil moisture analyses at ECMWF: Evaluation using global ground-based in situ observations. *J. Hydrometeorol.*, **13**, 1442–1460, <https://doi.org/10.1175/JHM-D-11-0107.1>.
- Allegrini, J., V. Dorer, and J. Carmeliet, 2012: Influence of the urban microclimate in street canyons on the energy demand for space cooling and heating of buildings. *Energy Build.*, **55**, 823–832, <https://doi.org/10.1016/j.enbuild.2012.10.013>.
- Ardilouze, C., and Coauthors, 2017: Multi-model assessment of the impact of soil moisture initialization on mid-latitude summer predictability. *Climate Dyn.*, **49**, 3959–3974, <https://doi.org/10.1007/s00382-017-3555-7>.
- Baldi, P., 1995: Gradient descent learning algorithm overview: A general dynamical systems perspective. *IEEE Trans. Neural Network*, **6**, 182–195, <https://doi.org/10.1109/72.363438>.
- Balsamo, G., A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. K. Betts, 2009: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the integrated forecast system. *J. Hydrometeorol.*, **10**, 623–643, <https://doi.org/10.1175/2008JHM1068.1>.
- Bergmeir, C., and J. M. Benítez, 2012: On the use of cross-validation for time series predictor evaluation. *Info. Sci.*, **191**, 192–213, <https://doi.org/10.1016/j.ins.2011.12.028>.
- Boussetta, S., G. Balsamo, A. Beljaars, T. Kral, and L. Jarlan, 2013: Impact of a satellite-derived leaf area index monthly climatology in a global numerical weather prediction model. *Int. J. Remote Sens.*, **34**, 3520–3542, <https://doi.org/10.1080/01431161.2012.716543>.
- Brands, S., R. Manzanar, J. M. Gutiérrez, and J. Cohen, 2012: Seasonal predictability of wintertime precipitation in Europe using the snow advance index. *J. Climate*, **25**, 4023–4028, <https://doi.org/10.1175/JCLI-D-12-00083.1>.
- Bunzel, F., W. A. Müller, M. Dobrynin, K. Fröhlich, S. Hagemann, H. Pohlmann, T. Stacke, and J. Baehr, 2018: Improved seasonal prediction of European summer temperatures with new five-layer soil-hydrology scheme. *Geophys. Res. Lett.*, **45**, 346–353, <https://doi.org/10.1002/2017GL076204>.
- Cerqueira, V., L. Torgo, and I. Mozetič, 2020: Evaluating time series forecasting models: An empirical study on performance estimation methods. *Mach. Learn.*, **109**, 1997–2028, <https://doi.org/10.1007/s10994-020-05910-7>.
- Cohen, J., D. Coumou, J. Hwang, L. Mackey, P. Orenstein, S. Totz, and E. Tziperman, 2019: S2S reboot: An argument for greater inclusion of machine learning in subseasonal to seasonal forecasts. *Wiley Interdiscip. Rev.: Climate Change*, **10**, e00567, <https://doi.org/10.1002/wcc.567>.
- Curtis, S., A. Fair, J. Wistow, D. V. Val, and K. Oven, 2017: Impact of extreme weather events and climate change for health and social care systems. *Environ. Health*, **16**, 128, <https://doi.org/10.1186/s12940-017-0324-3>.
- de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen, 2013: A simplified extended Kalman filter for the global operational soil moisture analysis at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **139**, 1199–1213, <https://doi.org/10.1002/qj.2023>.
- Díez, E., B. Orfila, M. D. Frias, J. Fernández, A. S. Cofiño, and J. M. Gutiérrez, 2011: Downscaling ECMWF seasonal precipitation forecasts in Europe using the RCA model. *Dyn. Meteor. Oceanogr.*, **63**, 757–762, <https://doi.org/10.1111/j.1600-0870.2011.00523.x>.
- Doblas-Reyes, F., R. Hagedorn, T. Palmer, and J. J. Morcrette, 2006: Impact of increasing greenhouse gas concentrations in seasonal ensemble forecasts. *Geophys. Res. Lett.*, **33**, L07708, <https://doi.org/10.1029/2005GL025061>.
- , J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: Status and prospects. *Wiley Interdiscip. Rev.: Climate Change*, **4**, 245–268, <https://doi.org/10.1002/wcc.217>.
- Dorigo, W., and Coauthors, 2017: ESA CCI soil moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.*, **203**, 185–215, <https://doi.org/10.1016/j.rse.2017.07.001>.
- Drusch, M., K. Scipal, P. de Rosnay, G. Balsamo, E. Andersson, P. Bougeault, and P. Viterbo, 2009: Towards a Kalman filter based soil moisture analysis system for the operational ECMWF Integrated Forecast System. *Geophys. Res. Lett.*, **36**, L14041, <https://doi.org/10.1029/2009GL037716>.
- Dutra, E., G. Balsamo, P. Viterbo, P. M. A. Miranda, A. Beljaars, C. Schär, and K. Elder, 2010: An improved snow scheme for

- the ECMWF land surface model: Description and offline validation. *J. Hydrometeorol.*, **11**, 899–916, <https://doi.org/10.1175/2010JHM1249.1>.
- Ghosh, R., W. A. Müller, J. Baehr, and J. Bader, 2017: Impact of observed North Atlantic multidecadal variations to European summer climate: A linear baroclinic response to surface heating. *Climate Dyn.*, **48**, 3547–3563, <https://doi.org/10.1007/s00382-016-3283-4>.
- Gruber, A., W. A. Dorigo, W. Crow, and W. Wagner, 2017: Triple collocation-based merging of satellite soil moisture retrievals. *IEEE Trans. Geosci. Remote Sens.*, **55**, 6780–6792, <https://doi.org/10.1109/TGRS.2017.2734070>.
- , T. Scanlon, R. van der Schalie, W. Wagner, and W. Dorigo, 2019: Evolution of the ESA CCI soil moisture climate data records and their underlying merging methodology. *Earth Syst. Sci. Data*, **11**, 717–739, <https://doi.org/10.5194/essd-11-717-2019>.
- Guillod, B. P., and Coauthors, 2014: Land-surface controls on afternoon precipitation diagnosed from observational data: Uncertainties and confounding factors. *Atmos. Chem. Phys.*, **14**, 8343–8367, <https://doi.org/10.5194/acp-14-8343-2014>.
- Harris, I., T. J. Osborn, P. Jones, and D. Lister, 2020: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data*, **7**, 1–18, <https://doi.org/10.1038/s41597-020-0453-3>.
- Hartigan, J., S. MacNamara, and L. M. Leslie, 2020: Application of machine learning to attribution and prediction of seasonal precipitation and temperature trends in Canberra, Australia. *Climate*, **8**, 76, <https://doi.org/10.3390/cli8060076>.
- Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature analysis and its uncertainty. *J. Climate*, **27**, 57–75, <https://doi.org/10.1175/JCLI-D-12-00837.1>.
- Hirschi, M., and Coauthors, 2011: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe. *Nat. Geosci.*, **4**, 17–21, <https://doi.org/10.1038/ngeo1032>.
- Hoffmann, P., 2018: Enhanced seasonal predictability of the summer mean temperature in Central Europe favored by new dominant weather patterns. *Climate Dyn.*, **50**, 2799–2812, <https://doi.org/10.1007/s00382-017-3772-0>.
- Holzkämper, A., P. Calanca, and J. Fuhrer, 2012: Statistical crop models: Predicting the effects of temperature and precipitation changes. *Climate Res.*, **51**, 11–21, <https://doi.org/10.3354/cr01057>.
- Hotelling, H., 1933: Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, **24**, 417–441, <https://doi.org/10.1037/h0071325>.
- , 1992: Relations between two sets of variates. *Breakthroughs in Statistics*, S. Kotz and N. L. Johnson, Eds., Springer, 162–190, https://doi.org/10.1007/978-1-4612-4380-9_14.
- Hsieh, W. W., 2009: *Machine Learning Methods in the Environmental Sciences: Neural Networks and Kernels*. Cambridge University Press, 364 pp.
- Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe collection. *Int. J. Climatol.*, **25**, 865–879, <https://doi.org/10.1002/joc.1169>.
- Johnson, S. J., and Coauthors, 2019: SEAS5: The new ECMWF seasonal forecast system. *Geosci. Model Dev.*, **12**, 1087–1117, <https://doi.org/10.5194/gmd-12-1087-2019>.
- Kadow, C., D. M. Hall, and U. Ulbrich, 2020: Artificial intelligence reconstructs missing climate information. *Nat. Geosci.*, **13**, 408–413, <https://doi.org/10.1038/s41561-020-0582-5>.
- Kämäräinen, M., P. Uotila, A. Y. Karpechko, O. Hyvärinen, I. Lehtonen, and J. Räisänen, 2019: Statistical learning methods as a basis for skillful seasonal temperature forecasts in Europe. *J. Climate*, **32**, 5363–5379, <https://doi.org/10.1175/JCLI-D-18-0765.1>.
- Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, <https://doi.org/10.1175/BAMS-D-12-00050.1>.
- Knight, J. R., C. K. Folland, and A. A. Scaife, 2006: Climate impacts of the Atlantic Multidecadal Oscillation. *Geophys. Res. Lett.*, **33**, L17706, <https://doi.org/10.1029/2006GL026242>.
- Lalouay, P., M. Balmaseda, D. Dee, K. Mogensen, and P. Janssen, 2016: A coupled data assimilation system for climate reanalysis. *Quart. J. Roy. Meteor. Soc.*, **142**, 65–78, <https://doi.org/10.1002/qj.2629>.
- , and Coauthors, 2018: CERA-20C: A coupled reanalysis of the twentieth century. *J. Adv. Model. Earth Syst.*, **10**, 1172–1195, <https://doi.org/10.1029/2018MS001273>.
- Li, M., P. Wu, and Z. Ma, 2020: A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets. *Int. J. Climatol.*, **40**, 5744–5766, <https://doi.org/10.1002/joc.6549>.
- Mahanama, S. P. P., and R. D. Koster, 2005: AGCM biases in evaporation regime: Impacts on soil moisture memory and land–atmosphere feedback. *J. Hydrometeorol.*, **6**, 656–669, <https://doi.org/10.1175/JHM446.1>.
- Manzanas, R., J. Gutiérrez, J. Fernández, E. Van Meijgaard, S. Calmanti, M. Magariño, A. Cofiño, and S. Herrera, 2018: Dynamical and statistical downscaling of seasonal temperature forecasts in Europe: Added value for user applications. *Climate Serv.*, **9**, 44–56, <https://doi.org/10.1016/j.cliser.2017.06.004>.
- Maracchi, G., O. Sirotenko, and M. Bindi, 2005: Impacts of present and future climate variability on agriculture and forestry in the temperate regions: Europe. *Climatic Change*, **70**, 117–135, <https://doi.org/10.1007/s10584-005-5939-7>.
- Min, M., and Coauthors, 2019: Estimating summertime precipitation from Himawari-8 and global forecast system based on machine learning. *IEEE Trans. Geosci. Remote Sens.*, **57**, 2557–2570, <https://doi.org/10.1109/TGRS.2018.2874950>.
- Mishra, N., C. Prodhomme, and V. Guemas, 2019: Multi-model skill assessment of seasonal temperature and precipitation forecasts over Europe. *Climate Dyn.*, **52**, 4207–4225, <https://doi.org/10.1007/s00382-018-4404-z>.
- Mudelsee, M., 2014: *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*. 2nd ed. Springer, 486 pp.
- Neddermann, N.-C., W. A. Müller, M. Dobrynin, A. Düsterhus, and J. Baehr, 2019: Seasonal predictability of European summer climate re-assessed. *Climate Dyn.*, **53**, 3039–3056, <https://doi.org/10.1007/s00382-019-04678-4>.
- Ossó, A., R. Sutton, L. Shaffrey, and B. Dong, 2018: Observational evidence of European summer weather patterns predictable from spring. *Proc. Natl. Acad. Sci. USA*, **115**, 59–63, <https://doi.org/10.1073/pnas.1713146114>.
- Pearson, K., 1901: LIII. On lines and planes of closest fit to systems of points in space. *London Edinburgh Dublin Philos. Mag. J. Sci.*, **2**, 559–572, <https://doi.org/10.1080/14786440109462720>.
- Poli, P., and Coauthors, 2016: ERA-20C: An atmospheric reanalysis of the twentieth century. *J. Climate*, **29**, 4083–4097, <https://doi.org/10.1175/JCLI-D-15-0556.1>.
- Qasmi, S., C. Cassou, and J. Boé, 2017: Teleconnection between Atlantic multidecadal variability and European temperature: Diversity and evaluation of the Coupled Model Intercomparison Project phase 5 models. *Geophys. Res. Lett.*, **44**, 11 140–11 149, <https://doi.org/10.1002/2017GL074886>.

- Quesada, B., R. Vautard, P. Yiou, M. Hirschi, and S. I. Seneviratne, 2012: Asymmetric European summer heat predictability from wet and dry southern winters and springs. *Nat. Climate Change*, **2**, 736–741, <https://doi.org/10.1038/nclimate1536>.
- Ratcliffe, R., and R. Murray, 1970: New lag associations between North Atlantic sea temperature and European pressure applied to long-range weather forecasting. *Quart. J. Roy. Meteor. Soc.*, **96**, 226–246, <https://doi.org/10.1002/qj.49709640806>.
- Robinson, H. A., A. Pinharanda, and D. Bensasson, 2016: Summer temperature can predict the distribution of wild yeast populations. *Ecol. Evol.*, **6**, 1236–1250, <https://doi.org/10.1002/ece3.1919>.
- Saeed, S., N. Van Lipzig, W. A. Müller, F. Saeed, and D. Zanchettin, 2014: Influence of the circumglobal wave-train on European summer precipitation. *Climate Dyn.*, **43**, 503–515, <https://doi.org/10.1007/s00382-013-1871-0>.
- Scher, S., and G. Messori, 2019: Weather and climate forecasting with neural networks: Using general circulation models (GCMs) with different complexity as a study ground. *Geosci. Model Dev.*, **12**, 2797–2809, <https://doi.org/10.5194/gmd-12-2797-2019>.
- Schlesinger, M. E., and N. Ramankutty, 1994: An oscillation in the global climate system of period 65–70 years. *Nature*, **367**, 723–726, <https://doi.org/10.1038/367723a0>.
- Schwingshackl, C., M. Hirschi, and S. I. Seneviratne, 2017: Quantifying spatiotemporal variations of soil moisture control on surface energy balance and near-surface air temperature. *J. Climate*, **30**, 7105–7124, <https://doi.org/10.1175/JCLI-D-16-0727.1>.
- Seneviratne, S. I., and Coauthors, 2013: Impact of soil moisture–climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment. *Geophys. Res. Lett.*, **40**, 5212–5217, <https://doi.org/10.1002/grl.50956>.
- , T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, <https://doi.org/10.1016/j.earscirev.2010.02.004>.
- Shi, W., N. Schaller, D. MacLeod, T. Palmer, and A. Weisheimer, 2015: Impact of hindcast length on estimates of seasonal climate predictability. *Geophys. Res. Lett.*, **42**, 1554–1559, <https://doi.org/10.1002/2014GL062829>.
- Slater, L. J., G. Villarini, and A. A. Bradley, 2017: Weighting of NMME temperature and precipitation forecasts across Europe. *J. Hydrol.*, **552**, 646–659, <https://doi.org/10.1016/j.jhydrol.2017.07.029>.
- Stacke, T., and S. Hagemann, 2016: Life time of soil moisture perturbations in a coupled land-atmosphere simulation. *Earth Syst. Dyn.*, **7**, 1–19, <https://doi.org/10.5194/esd-7-1-2016>.
- Sutton, R. T., and B. Dong, 2012: Atlantic Ocean influence on a shift in European climate in the 1990s. *Nat. Geosci.*, **5**, 788–792, <https://doi.org/10.1038/ngeo1595>.
- , and D. L. R. Hodson, 2005: Atlantic Ocean forcing of North American and European summer climate. *Science*, **309**, 115–118, <https://doi.org/10.1126/science.1109496>.
- Thomas, J. A., A. A. Berg, and W. J. Merryfield, 2016: Influence of snow and soil moisture initialization on sub-seasonal predictability and forecast skill in boreal spring. *Climate Dyn.*, **47**, 49–65, <https://doi.org/10.1007/s00382-015-2821-9>.
- Totz, S., E. Tziperman, D. Coumou, K. Pfeiffer, and J. Cohen, 2017: Winter precipitation forecast in the European and Mediterranean regions using cluster analysis. *Geophys. Res. Lett.*, **44**, 12 418–12 426, <https://doi.org/10.1002/2017GL075674>.
- Turco, M., A. Ceglar, C. Prodhomme, A. Soret, A. Toreti, and J. D.-R. Francisco, 2017: Summer drought predictability over Europe: Empirical versus dynamical forecasts. *Environ. Res. Lett.*, **12**, 084006, <https://doi.org/10.1088/1748-9326/aa7859>.
- van den Hurk, B., F. Doblas-Reyes, G. Balsamo, R. D. Koster, S. I. Seneviratne, and H. Camargo, 2012: Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dyn.*, **38**, 349–362, <https://doi.org/10.1007/s00382-010-0956-2>.
- Vogel, M. M., R. Orth, F. Cheruy, S. Hagemann, R. Lorenz, B. J. J. M. van den Hurk, and S. I. Seneviratne, 2017: Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks. *Geophys. Res. Lett.*, **44**, 1511–1519, <https://doi.org/10.1002/2016GL071235>.
- , J. Zscheischler, and S. I. Seneviratne, 2018: Varying soil moisture–atmosphere feedbacks explain divergent temperature extremes and precipitation projections in central Europe. *Earth Syst. Dyn.*, **9**, 1107–1125, <https://doi.org/10.5194/esd-9-1107-2018>.
- Von Storch, H., and F. W. Zwiers, 2001: *Statistical Analysis in Climate Research*. Cambridge University Press, 496 pp.
- Wang, L., M. Ting, and P. Kushner, 2017: A robust empirical seasonal prediction of winter NAO and surface climate. *Sci. Rep.*, **7**, 279, <https://doi.org/10.1038/s41598-017-00353-y>.
- Weisheimer, A., and Coauthors, 2009: ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions—Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.*, **36**, L21711, <https://doi.org/10.1029/2009GL040896>.
- , F. J. Doblas-Reyes, T. Jung, and T. Palmer, 2011: On the predictability of the extreme summer 2003 over Europe. *Geophys. Res. Lett.*, **38**, L05704, <https://doi.org/10.1029/2010GL046455>.
- , N. Schaller, C. O'Reilly, D. A. MacLeod, and T. Palmer, 2017: Atmospheric seasonal forecasts of the twentieth century: Multi-decadal variability in predictive skill of the winter North Atlantic Oscillation (NAO) and their potential value for extreme event attribution. *Quart. J. Roy. Meteor. Soc.*, **143**, 917–926, <https://doi.org/10.1002/qj.2976>.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.
- Zampieri, M., A. Toreti, A. Schindler, E. Scoccimarro, and S. Gualdi, 2017: Atlantic multi-decadal oscillation influence on weather regimes over Europe and the Mediterranean in spring and summer. *Global Planet. Change*, **151**, 92–100, <https://doi.org/10.1016/j.gloplacha.2016.08.014>.
- Zorita, E., V. Kharin, and H. von Storch, 1992: The atmospheric circulation and sea surface temperature in the North Atlantic area in winter: Their interaction and relevance for Iberian precipitation. *J. Climate*, **5**, 1097–1108, [https://doi.org/10.1175/1520-0442\(1992\)005<1097:TACASS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<1097:TACASS>2.0.CO;2).