



RESEARCH LETTER

10.1029/2020GL092360

Key Points:

- The predictive skill of winter seasonal hindcasts for teleconnection patterns in Europe varied over the twentieth century
- Seasonal hindcasts only partly reproduce the links between teleconnection patterns and winter temperature extremes
- The teleconnection/surface temperature relationships do not significantly change despite fluctuations in predictive skill

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

N. Schuhen,
nina.schuhen@cicero.oslo.no

Citation:

Schuhen, N., Schaller, N., Bloomfield, H. C., Brayshaw, D. J., Lledó, L., Cionni, I., & Sillmann, J. (2022). Predictive skill of teleconnection patterns in twentieth century seasonal hindcasts and their relationship to extreme winter temperatures in Europe. *Geophysical Research Letters*, 49, e2020GL092360. <https://doi.org/10.1029/2020GL092360>






Received 31 DEC 2020

Accepted 21 MAY 2022

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Predictive Skill of Teleconnection Patterns in Twentieth Century Seasonal Hindcasts and Their Relationship to Extreme Winter Temperatures in Europe

Nina Schuhen¹ , Nathalie Schaller¹ , Hannah C. Bloomfield^{2,3} , David J. Brayshaw^{2,4} , Llorenç Lledó⁵ , Irene Cionni⁶, and Jana Sillmann¹ 

¹Center for International Climate and Environmental Research (CICERO), Oslo, Norway, ²Department of Meteorology, University of Reading, Reading, UK, ³School of Geographical Sciences, University of Bristol, Bristol, UK, ⁴National Centre for Atmospheric Science, Reading, UK, ⁵Barcelona Supercomputing Center (BSC), Barcelona, Spain, ⁶Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (ENEA), Rome, Italy

Abstract European winter weather is dominated by several low-frequency teleconnection patterns, the main ones being the North Atlantic Oscillation, East Atlantic, East Atlantic/Western Russia, and Scandinavian patterns. We analyze the century-long ERA-20C reanalysis and ASF-20C seasonal hindcast data sets and find that these patterns are subject to decadal variability and fluctuations in predictive skill. Using indices for determining periods of extreme cold or warm temperatures, we establish that the teleconnection patterns are, for some regions, significantly correlated or anti-correlated to cold or heat waves. The seasonal hindcasts are however only partly able to capture these relationships. There do not seem to be significant changes to the observed links between large-scale circulation patterns and extreme temperatures between periods of higher and lower predictive skill.

Plain Language Summary Large-scale atmospheric patterns that influence European winter weather showed slowly evolving fluctuations over the course of the twentieth century. This affects how well the patterns can be predicted by seasonal forecasting models. However, the impact of the large-scale patterns on extreme winter temperatures mostly does not change over time, although forecasting models often struggle to correctly predict these impacts.

1. Introduction

Skillful seasonal forecasts provide valuable contextual information to various decision makers across a number of sectors, such as energy, public health, water management, disaster risk management, and agriculture (Bruno Soares & Dessai, 2016; White et al., 2021). In particular, the energy sector is strongly impacted by climate variability, as present-day power system demand is strongly correlated with near-surface temperatures (Bessec & Fouquau, 2008; Bloomfield et al., 2021b; Taylor & Buizza, 2003) and conventional (e.g., coal, gas, and nuclear power), as well as renewable energies (e.g., wind, solar, and hydro power) all have various degrees of sensitivity to weather and climate (Orlov et al., 2020). Some examples of challenges that may benefit from accurate seasonal forecasts of near-surface temperature include potential peak demand events, the ability to adjust the maintenance schedules of nuclear power plants and availability of cooling water for traditional energy generation during periods of extreme heat.

Recent studies have shown promising skill present in gridded sub-seasonal to seasonal forecasts of energy-industry-relevant variables (e.g., Beerli et al., 2017; De Felice et al., 2015; Dorrington et al., 2020; Gonzalez et al., 2021; Lledó et al., 2019; Torralba et al., 2017). Other authors have also demonstrated seasonal forecast skill in predicting large-scale atmospheric patterns, such as weather regimes, or atmospheric teleconnections known to have profound relevance for the energy industry (Scaife et al., 2014). The meteorological conditions relevant for energy system operations are related to these large-scale atmospheric patterns to various degrees (Bett et al., 2019; Bloomfield et al., 2020, 2021a; Brayshaw et al., 2011; Clark et al., 2017; Ely et al., 2013; Grams et al., 2017; Thornton et al., 2019; van der Wiel et al., 2019).

Clark et al. (2017) show that through predicting the winter North Atlantic Oscillation (NAO), skillful seasonal predictions of winter near-surface temperatures, and therefore electricity demand, can be made. Thornton



et al. (2019) find similar results for UK gas demand, with seasonal forecasts able to provide information on both the mean demand and the number of extreme gas demand days. Bett et al. (2019) also show promising results when using the NAO to predict wind and solar power generation across Europe. However, these studies of energy-system relevant skill all focus on a relatively recent forecast period, which does not take into account the potential of increased or decreased skill over the course of the century as highlighted by Shi et al. (2015) and Weisheimer et al. (2017). While several current operational seasonal forecast systems report skillful predictions of the NAO (e.g., Dunstone et al., 2018; Johnson et al., 2019; Scaife et al., 2014), the question remains whether they would also have skill during periods where the NAO is predominantly negative, and the intra-seasonal geopotential height variance is higher, as was the case in the middle of the twentieth century (Weisheimer et al., 2017, 2019), and might happen again in the future. In addition to the NAO, variability of European winter weather is related to several other teleconnections (e.g., Hall & Hanna, 2018; Josey et al., 2011; Lledó et al., 2020; Zubiate et al., 2017), whose long-term trends have not been investigated yet.

In this study, we are looking to answer three research questions related to the four main teleconnection patterns that dominate European winter weather:

1. Are other teleconnections besides the NAO subject to decadal variability and changing predictive skill in seasonal hindcast models?
2. Can we identify regions in Europe, where specific teleconnections modify the occurrence probability and duration of temperature extremes (represented by the heat and cold wave indices developed by Russo et al., 2015) and to what extent do forecasting systems reproduce these relationships?
3. Do these spatial patterns in the link between heat/cold wave indices and teleconnections vary over the course of the twentieth century?

The article is structured as follows: Section 2 is an overview over our reanalysis and seasonal hindcast data sets and the methods we used to determine the Euro-Atlantic teleconnection (EATC) patterns and extreme temperature indices. Results showing the time evolution of the teleconnection indices and their relationship to surface temperature extremes are presented in Section 3, followed by conclusions in Section 4.

2. Data and Methods

2.1. Data Sets

ERA-20C is a reanalysis based on the assimilation of surface observations only, covering the twentieth century—more specifically from 1900 to 2010—and is produced by the ECMWF using the atmosphere-only model version of their IFS model (Poli et al., 2013, 2015). The horizontal resolution over Europe is about 125 km, with 90 vertical model levels and a temporal frequency of 3 hr. It was created with the objective of providing a consistent reanalysis for the whole twentieth century using new assimilation methods. Contrary to the ERA-Interim reanalysis (Dee et al., 2011), it is not designed to be a “best-product” data set. ERA-Interim covers the years 1979–2019 and has a horizontal resolution of about 80 km. In this study, it is used to define the teleconnection patterns over the 1981–2016 time frame.

Seasonal hindcasts are provided by the ASF-20C ensemble (Weisheimer et al., 2017), which is set up in such a way that its configuration matches ERA-20C. The ensemble has 51 members and is initialized four times per year in February, May, August, and November with forecast lead times of up to four months. As we are interested in winter extremes, we consider the November start date and analyze the predictions for December/January/February (DJF). All hindcasts and reanalyses were first regridded to a common $1^\circ \times 1^\circ$ grid over the area of interest.

2.2. Euro-Atlantic Teleconnection Patterns and Indices

We are using an approach that allows computing the four main EATC patterns, namely the NAO, East Atlantic (EA), East Atlantic/Western Russia (EA/WR), and Scandinavian (SCA) patterns, from a reanalysis and then obtain the associated teleconnection indices from both reanalysis and seasonal hindcast data sets, as described in Lledó et al. (2020).

The four observed EATC patterns and indices are calculated from ERA-Interim 500 hPa geopotential height fields by performing a Rotated Empirical Orthogonal Function (REOF) analysis over the Euro-Atlantic domain

[90°W–60°E; 20°N–80°N]: First, DJF anomalies for the 1981–2016 period are computed with respect to the same period mean. Then, the EOF analysis is carried out, and the first four variability modes are retained, before applying a varimax rotation. The four obtained REOF modes are reordered and their signs adjusted if needed so that the EATC patterns resemble as much as possible the positive phases of the NAO, EA, EA/WR, and SCA patterns as computed by Climate Prediction Center (2012).

To obtain the EATC indices for the whole twentieth century, the DJF mean anomalies (with respect to the 1900–2010 climatology) of the 500 hPa geopotential height fields in the ERA-20C reanalysis and each of the 50 ASF-20C ensemble members are projected onto the ERA-Interim patterns employing the scalar product. Anomalies for ERA-20C and ASF-20C are computed against the climatologies of the respective models. The fraction of explained variance of each EATC is computed, and the projected indices are scaled to have mean zero and variance one. For the purpose of this study, we assume that the pattern definitions remain stable over the time period covered.

2.3. Extreme Temperature Indices

In this study, we investigate links between EATCs and extreme winter temperatures using slightly adapted versions of the Heat Wave Magnitude Index daily (HWMId; Russo et al., 2016) and its equivalent for cold extremes, the Cold Wave Magnitude Index daily (CWMId; Brunner et al., 2018). Variations of these indices have, for instance, been used in studies by Lavaysse et al. (2018, 2019) to investigate temperature extremes with impact on human health. As they are calculated relative to the local climate conditions, the HWMId and CWMId are more suitable for analyzing and communicating pan-European seasonal temperature forecast than for example, the threshold-based heating and cooling degrees (e.g., Bloomfield et al., 2020), which are widely used in the energy sector.

Heat or cold waves are periods of at least three consecutive days where the maximum temperature exceeded the 90th percentile of the climatological reference (1981–2010) based on the ERA-20C and ASF-20C models, respectively. Russo et al. (2016) calculate the daily 90th percentile based on the daily maxima during a moving 31-day window. Because of the relatively narrow time windows in the ASF-20C hindcasts (i.e., one to 3 months), we use a 10 days window instead, with the first and last 5 days of the season having consequently marginally less robust statistics. At each grid point, the magnitude (as defined in Russo et al., 2015) for each day experiencing heat or cold wave is computed and finally, we sum these magnitudes over each winter to obtain the grid-point-wise HWMId and CWMId. We obtain 30 yearly values for each grid point and time period of interest (1951–1980 and 1981–2010).

3. Results and Discussion

3.1. Time Evolution of EATCs and Predictive Skill

Previous studies show that predictive skill levels for the winter NAO were subject to multi-decadal changes during the twentieth century (Weisheimer et al., 2017, 2019), with a low-skill phase during a period with predominantly negative NAO index in the middle of the century. We look at two periods of interest, an early period from 1951 to 1980 (starting with the winter season 1950/1951) and a late period from 1981 to 2010 (starting with the winter season 1980/1981). Figure 1 shows the time series of the four EATC indices (NAO, EA, EA/WR, and SCA) over 1901–2010, for both ERA-20C and the ASF-20C ensemble mean.

All four EATC indices show decadal fluctuation in the ERA-20C reanalysis, which the ASF-20C mean is generally able to reproduce. However, there are time periods with large errors, for example, between the 1950s and 1970s for NAO and from 1980 to 2010 for SCA. As shown in earlier studies, the NAO index is mostly negative during the early period and positive during the more recent period. Similar patterns can be seen for the EA, which switches from predominantly negative to positive, and the SCA, which is neutral during the first period and negative during the second. The two chosen periods do not coincide well with phase changes for the EA/WR, but this index also explains the smallest portion of the total variance and seems to be an outlier for most of the following analysis. For the later period, the ASF-20C mean and the ERA-20C time series are significantly correlated for NAO, EA, and SCA, which is not the case for the earlier period. This also points toward a difference in predictive skill for these three indices between the middle and the end of the century. Note that the statistical significance tests require independent and identically distributed data, which is likely not the case in this study. However,

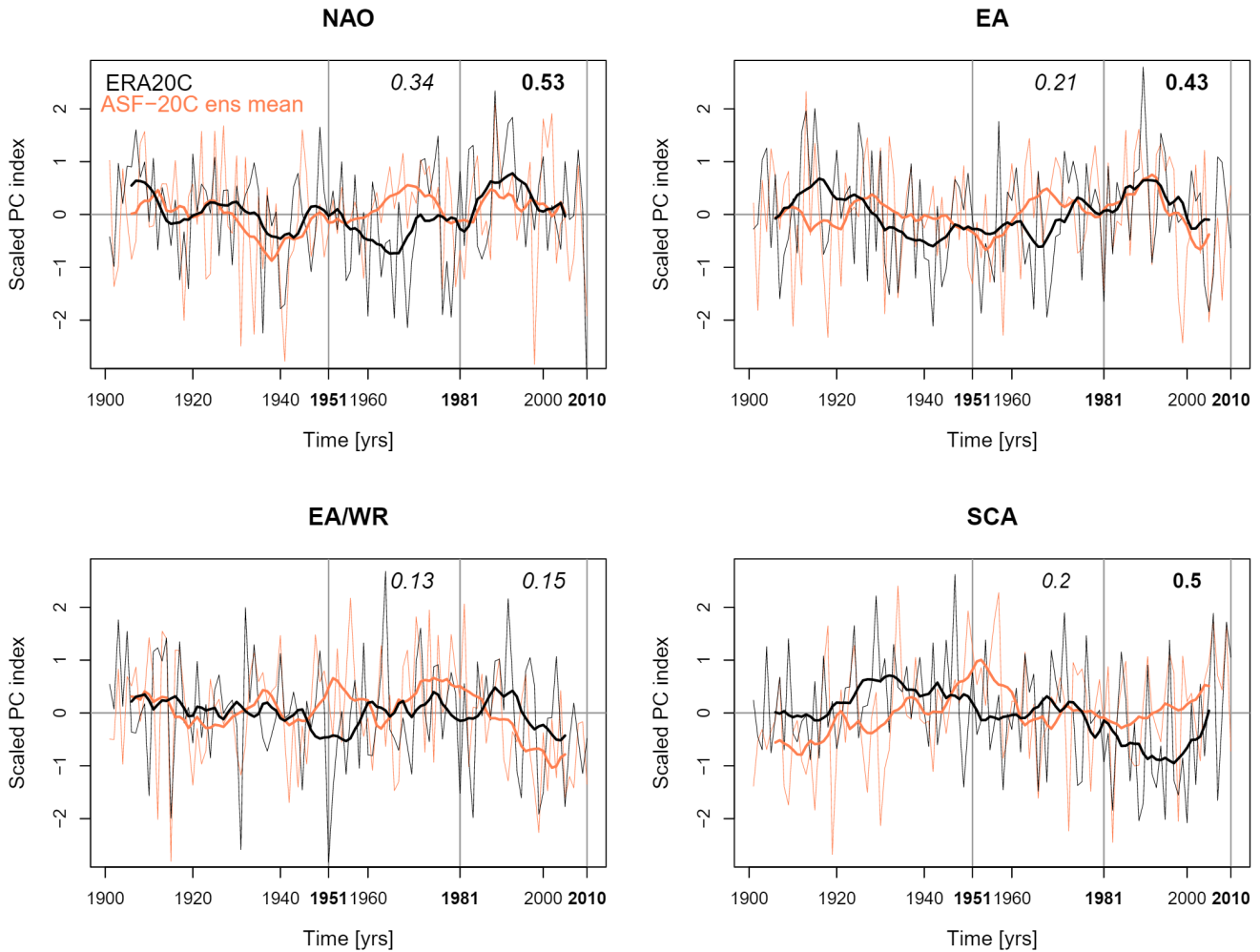


Figure 1. Time series of the four main Euro-Atlantic teleconnection (EATC) patterns during December/January/February (DJF) from 1901 to 2010, for ERA-20C and the ASF-20C ensemble mean. Thin lines are the scaled and centered yearly values and thick lines a 10 yr smoothing filter. The two periods of interest are marked with vertical lines, with the Pearson's correlation coefficients between the yearly ERA-20C and ASF-20C ensemble mean indices at the top. Numbers in bold are significant at the 5% level, based on a t-distributed test statistic.

we performed data analysis that shows that the results of the tests are not affected by relaxing this assumption. Nevertheless, significance results should be understood as guidance only.

To illustrate this point, we compute the 31 yr running anomaly correlation coefficient (ACC) between the ASF-20C mean and the ERA-20C reanalysis (Figure 2). Apart from EA/WR, all EATC patterns have a significant correlation—and therefore potentially increased predictive skill—during the 1981–2010 period. For SCA, there is a steady increase starting in the 1940s, while the correlation for EA/WR was initially high in the 1920s/1930s but has been oscillating around a non-significant level since about 1950. Further research into why the EA/WR often exhibits such different behavior compared to the other patterns is required; a potential reason could be mismatch between the EA/WR pattern definitions in ERA-Interim and ERA-20C/ASF-20C. We also note that the EA/WR resembles the Northern European blocking pattern (Scherrer et al., 2006), which has been shown to be difficult to predict by seasonal forecasting systems (e.g., Davini et al., 2020).

Also shown in Figure 2 are the smoothed squared error (SE) of the ensemble mean and the continuous ranked probability score (CRPS; e.g., Matheson & Winkler, 1976). Here, both scores are positively oriented and, for readability reasons, normalized by subtracting and then dividing by their long-term mean. NAO scores show a consistent decrease until about 1980 before rising again in the most recent decades. For the second half of the century, there is a slight, but not very clear similarity to ACC trends and changes in trend generally seem to lag

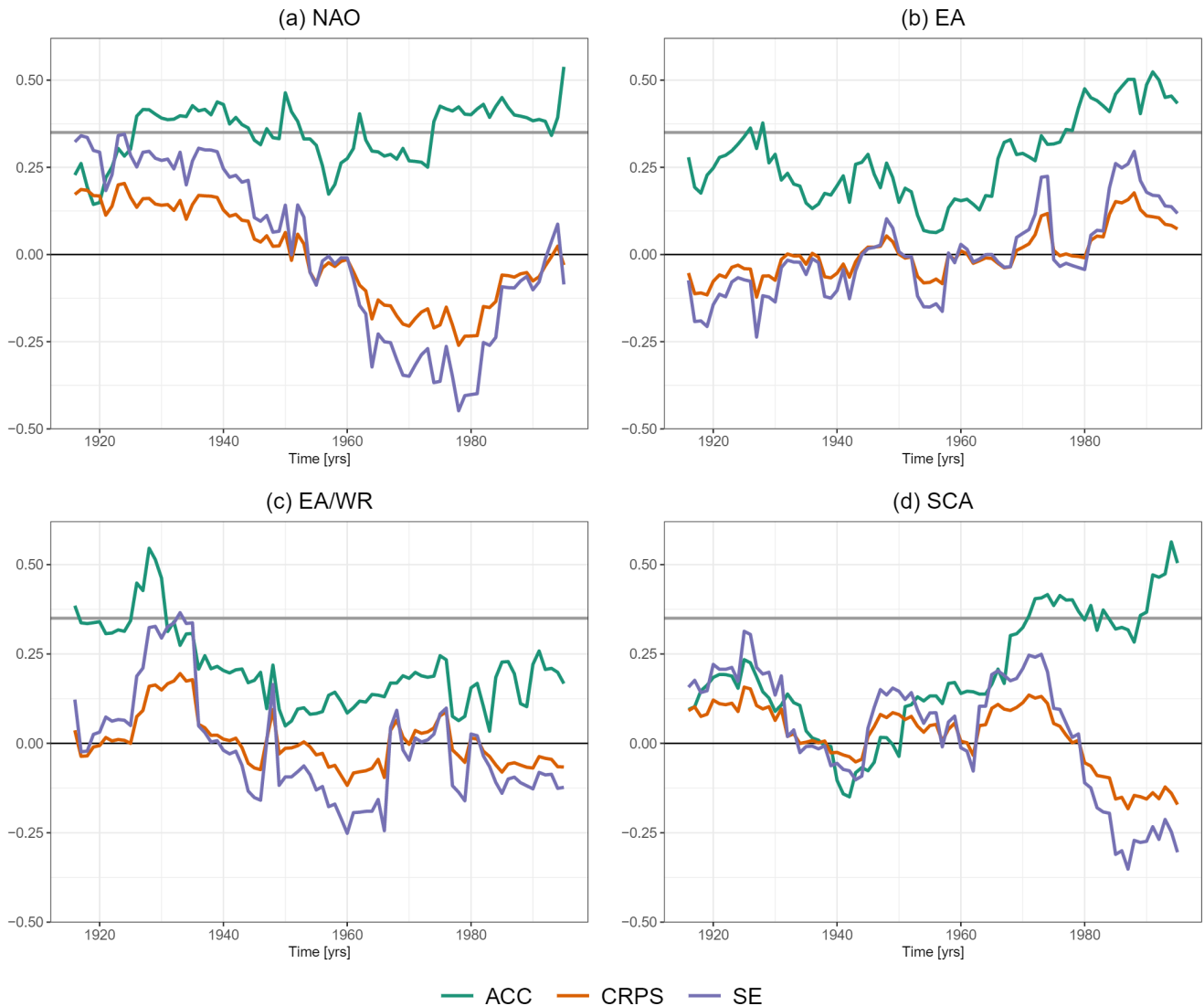


Figure 2. Time series of the Pearson's correlation coefficient between ERA-20C and the ASF-20C ensemble mean (ACC), the squared error (SE) of the ASF-20C ensemble mean against ERA-20C and the continuous ranked probability score (CRPS) of the ASF-20C ensemble against ERA-20C. The two scores are normalized by their long-term mean and smoothed over 31 yr windows. Correlations are computed over running windows of 31 yr and significant at the 5% level above the light gray line (based on a t test).

behind the ACC by 5–10 yr. The EA correlation initially decreases before rising from ca. 1950, whereas the corresponding scores follow an increasing trend over the entire time period. For EA/WR, the scores largely match the ACC evolution, with relatively low predictive skill for most of the twentieth century. Although the SCA pattern shows increasing and significant positive correlation over the last decades, the bias is also quite large at the same time. This coincides with the period where the SCA index is strongly negative. To investigate further, we also compute the ratio of the mean ensemble spread and the root-mean-square error, which is a measure of the confidence or calibration of the ensemble (e.g., Palmer et al., 2006), in Table S1 in Supporting Information S1 for the periods 1951–1980 and 1981–2010. Despite the larger SE values, the calibration of the SCA hindcasts improves from the early to the later period, which also holds true for the NAO and EA indices.

As in Weisheimer et al. (2017), we compare these results with the anomaly correlation of a persistence forecast based on the November ERA-20C values (Figure S1 in Supporting Information S1). There are short periods around the middle of the century, in case of NAO and EA, where the correlation of the persistence forecast is higher (however not or only barely significant) than that of the ASF-20C mean, coinciding with predominantly

negative indices. For both SCA and EA/WR, persistence has higher skill than the ensemble mean for somewhat longer periods starting in the 1920s. While SCA here shows a tendency toward positive values, there is no particular trend for EA/WR. In the latter case, the persistence correlation is however significant for several decades, which again hints at a potential issue with predicting this particular teleconnection.

Evidence has been found for trends in ERA-20C that can potentially be attributed to the lower availability and quality of observations in the early to mid-twentieth century (Bloomfield et al., 2018; Wohland et al., 2019). However, we find that the lack of significant correlation during the early period can likely not be explained by this, as there are clear periods with significant correlation during the first half of the century for the NAO and the EA/WR, as well as a reversal in trend for the EA and SCA correlations in the 1940s/1950s. A similar conclusion was made for the NAO in Weisheimer et al. (2017).

Our results thus indicate that there is substantial variability in the strength of the EATC pattern projection from the long-term data sets onto the modern-day atmospheric circulations. This not only confirms the result of Weisheimer et al. (2017) for the NAO, but now also points towards similar trends in the other three main EATCs. We have to consider that the patterns themselves might not be consistent over long time periods, although we see some ability in the ASF-20C hindcasts to reproduce the decadal-scale variability based on modern-day pattern definitions, especially for the NAO. The fluctuations point towards longer periods with higher or lower forecast skill—with recent decades demonstrating exceedingly good forecast quality—but it is unclear if and how much of this can be attributed to any evolution of the EATCs themselves. Therefore, we recommend that caution should be applied when assessing the long-term skill of REOF-based seasonal pattern predictions, as features and trends in forecast skill may not translate from relatively short recent periods (i.e., two to three decades) to longer time frames.

3.2. Relationship of EATCs and Surface Temperature Extremes

It has been established that teleconnection patterns and weather regimes influence weather and climate variability in Europe (e.g., van der Wiel et al., 2019). We now look to identify regions where the four teleconnections modify the occurrence and duration of surface temperature extremes. To this end, we first investigate the statistical link between the EATC time series and extreme temperature indices based on the reanalysis, and then analyze how well this link is reproduced in the hindcasts, independent of whether they correctly predict the teleconnection indices themselves. As we focus on winter temperatures and cold waves, the results for the HWMI_d can be found in the supporting material. For brevity, we also refer to the supplement for the results of the CWMI_d in the 1951–1980 period.

Figures 3 and 4 show the Spearman's rank correlation coefficient between the DJF values of the CWMI_d indices during 1981–2010 and the respective average EATC indices on a map of Europe. We again compare the reanalysis ERA-20C with the ASF-20C ensemble mean, with their difference depicted in the bottom left panel. Additionally, the plots in the bottom right panels show the number of ensemble members for each grid point that agree with the reanalysis on the significance and sign of the correlation. Note that we can not conclude from this statistic that there are individual ensemble members that reproduce the spatial correlation patterns in ERA-20C.

We see from Figure 3a that in ERA-20C, the NAO indices are correlated with fewer cold extremes in northern Europe and more cold extremes south of the Mediterranean. There is a small area with negative correlation over northern Europe for the ensemble mean, while only very few ensemble members (up to seven) match the correlations seen in the reanalysis. Positive values of EA indices are associated with warmer temperatures throughout Europe (Figure 3b), although there is barely any correlation at all in the hindcast mean. Again, there are only few members that agree with the reanalysis.

For the EA/WR pattern (Figure 4a), the reanalysis shows almost no link to cold extremes in Europe, only at some grid points with positive correlation in western Siberia. Thus it is surprising to find that the ASF-20C mean points toward a quite strong association with more cold extremes in eastern Europe and beyond. This adds to the previously seen conflicting results from this EATC. Figure 4b shows the correlations between SCA indices and the CWMI_d. Here, the patterns in ERA-20C and the ASF-20C mean are somewhat similar in that they link larger SCA indices to more cold extremes throughout Europe, even if they disagree on the location. This is slightly unexpected, as the SCA index was predominantly negative during the 1981–2010 period.

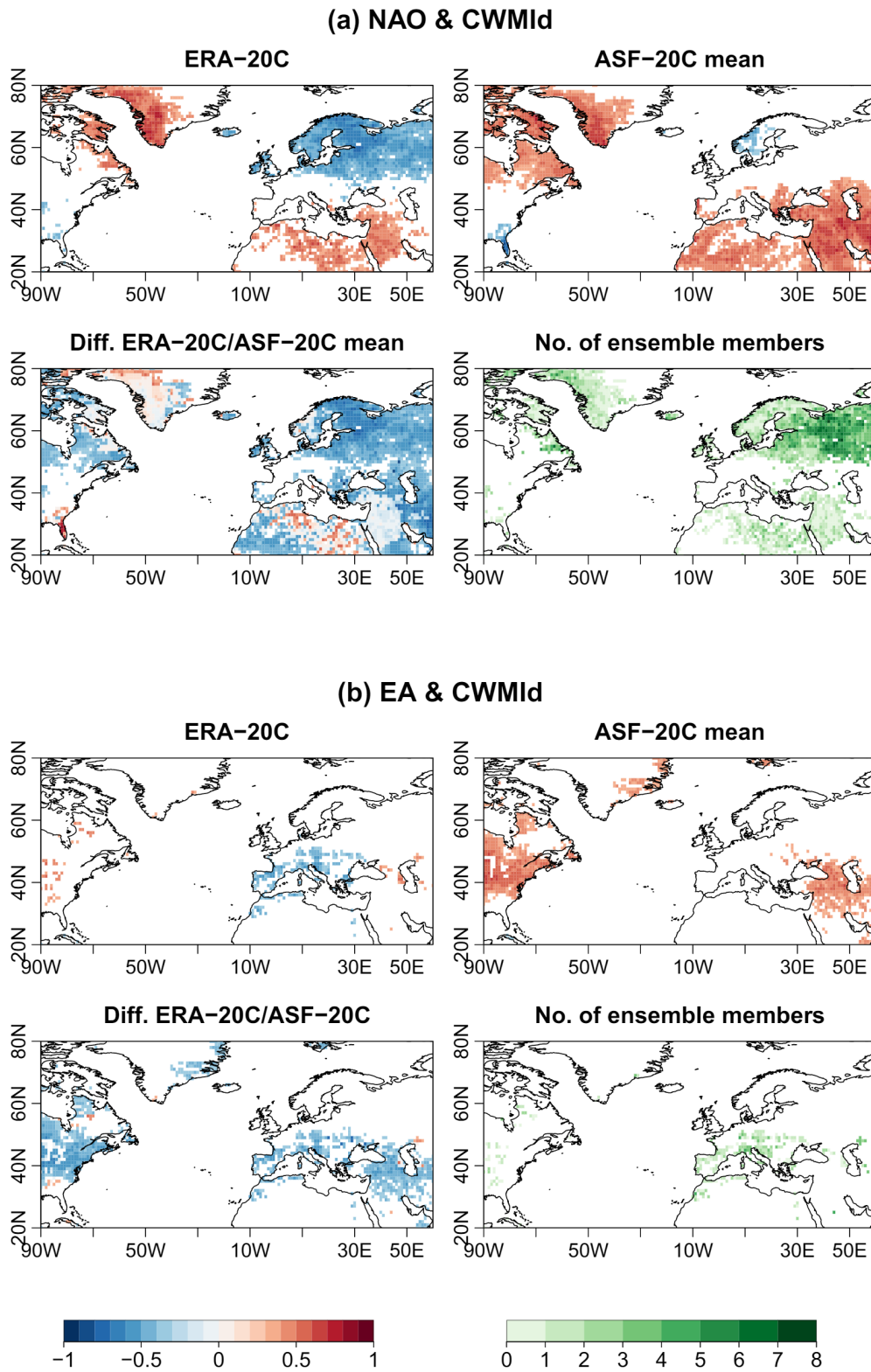


Figure 3.

We can compare these results with the earlier period from 1951 to 1980 (Figures S2 and S3 in Supporting Information S1). Overall, there are large similarities between the two time periods, apart from the SCA index, which has almost no relationship with cold waves in ERA-20C, but a substantial positive correlation in the ASF-20C mean. Also, the area of negative correlation over northern Europe for the NAO is much smaller in the earlier period. For a more quantitative result, we apply a Fisher transformation to the correlation at each grid point and test the differences between the early and the late period for significance. However, there are no areas in Europe with significantly different correlations. As the frequency of cold extremes is almost identical for the two time periods, any variability in correlation is likely not due to different sample sizes. The number of ensemble members that show grid-point-wise agreement with the reanalysis is much higher during the earlier period (in particular for the NAO), but as mentioned above, this does not necessarily mean that the spatial correlation patterns are reproduced in individual members.

The corresponding results for HWMId (Figures S4–S7 in Supporting Information S1) show very similar structures to the CWMId with reversed signs. The only outlier is the SCA, where a link to fewer warm extremes is present in the reanalysis and hindcast during the early period, as well as the reanalysis of the later period, but not in the modern hindcast mean.

4. Conclusions

Many applications across a multitude of sectors benefit from skillful sub-seasonal to seasonal climate prediction, for example, disaster preparedness for extreme events, long-term flood and drought risks or potential threats to food security (White et al., 2021). Here, we focused on large-scale circulation patterns and their impact on temperature extremes during winter, which are particularly relevant for the energy industry (Orlov et al., 2020). Due to the energy sector's increasing sensitivity to seasonal climate variability, advance knowledge about temperature extremes is crucial for mitigation and reduction of potential negative impacts. In this study, we propose using the heat and cold wave magnitude index daily (HWMId and CWMId) to identify temperature extremes, as they relate to percentiles of the local climatology. The HWMId can also indicate warm spells in winter, which might not be the case with other indicators, such as cooling degree days.

EATC patterns like the NAO are main drivers of winter temperatures in Europe and their predictions have shown considerable skill in the past (Lledó et al., 2020; Matsueda & Palmer, 2018). Weisheimer et al. (2017) established that the NAO experienced significant fluctuations in predictive skill during the twentieth century, with a period of particular high skill during the last 30 yr. We extend this earlier work by demonstrating that similar trends also exist for other EATCs, analyzed with help of the ERA-20C reanalysis. The EATC indices predicted by the ASF-20C ensemble also exhibit decadal variability, as illustrated by the 31 yr running window correlation between the reanalysis and the ensemble average, as well as the SE of the ensemble mean and the CRPS of the full ensemble. During the most recent three decades, we see a significant positive correlation between ERA-20C and the ASF-20C mean for the NAO, EA, and SCA indices, which also coincides with an improving trend in the SCA and CRPS in the case of NAO and EA. For SCA, however, both scores considerably deteriorate, which suggests that the ensemble mean is largely able to reproduce the most recent trend in SCA indices, but also suffers from a large bias. Overall, this hints at higher forecast skill in modern times compared to the middle of the century for some of the EATCs and, consequently, seasonal forecasting models may not be able to maintain current skill levels during future phases of for example, predominantly negative NAO indices. However, different metrics can show very different skill evolution over time (as seen for SCA), illustrating that multiple metrics should be used to assess skill and, in particular, that ACC results should be accompanied by a bias estimate (Thorarinsdottir & Schuhen, 2018).

Further research is needed as to the underlying causes of the EATC variability, but changes in interactions between tropical sea surface temperature anomalies and extratropical circulation are likely key (O'Reilly et al., 2017). While we assumed that the circulation pattern definitions did not change during the twentieth century, these might have actually evolved over time and thus contributed to the fluctuation of predictive skill. Moreover, it

Figure 3. Maps of the Spearman's rank correlation coefficient between the seasonal CWMId values over December/January/February (DJF) and (a) the North Atlantic Oscillation (NAO) index, (b) the East Atlantic (EA) index. Correlations are calculated from the ERA-20C reanalysis and the ASF-20C mean for the period 1981–2010. All correlations are significant at the 5% level based on a t-distributed test statistic. The bottom left panels show the difference between ERA-20C and the ASF-20C mean. The bottom right panels show the number of ensemble members at each grid point where both ERA-20C and the ensemble member show significant correlation and agree on the sign of correlation.

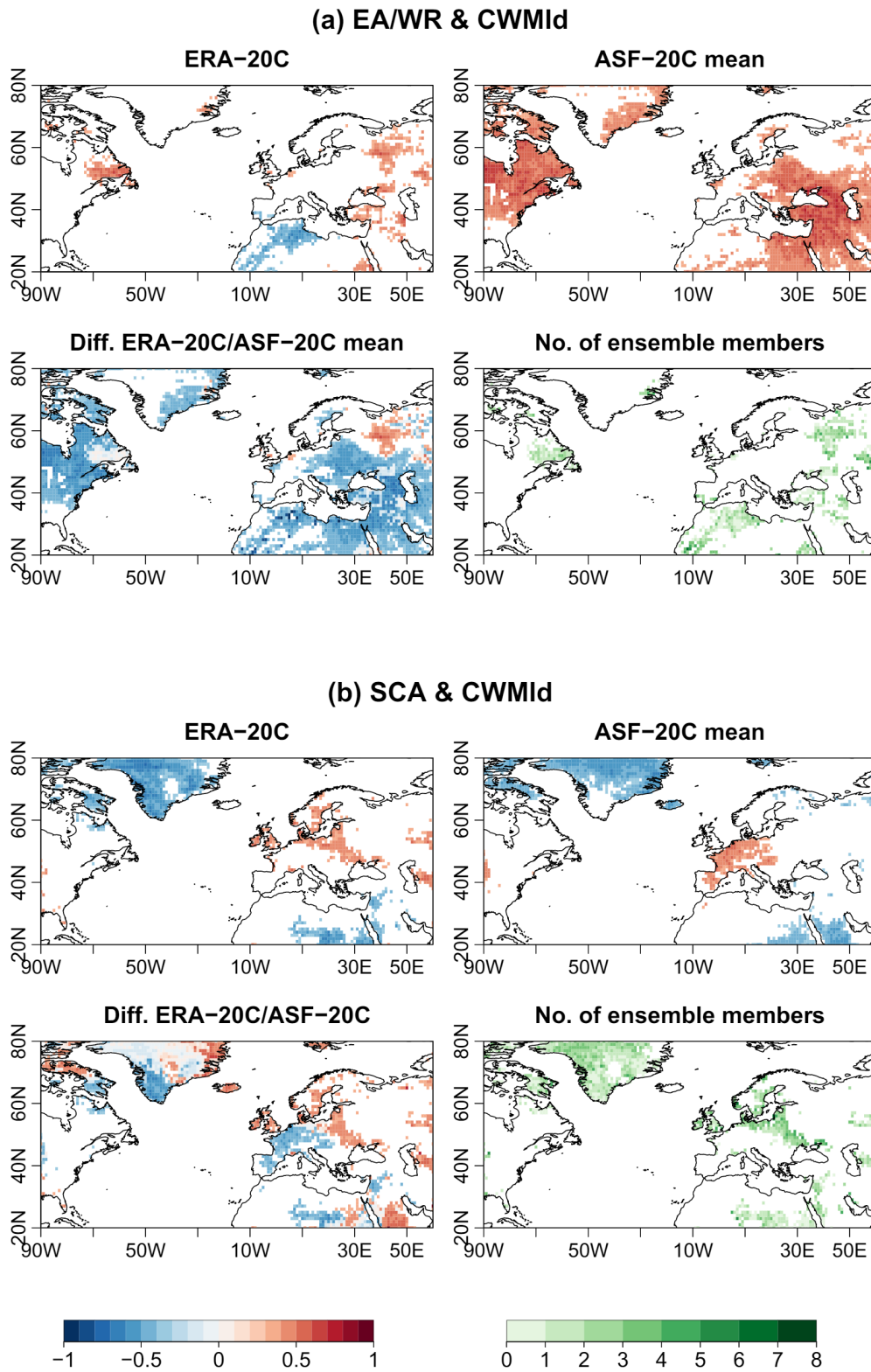


Figure 4.

unclear whether these fluctuations are driven by intrinsic predictability changes or limitations of the hindcast model. O'Reilly et al. (2017) find that hindcast skill for the Pacific/North American index (PNA) is also reduced during the middle of the twentieth century, and that there is a strong connection between PNA and NAO errors for this period. Decadal variability of NAO skill is also linked to skill variability of the Pacific Decadal Oscillation (PDO; Weisheimer et al., 2017), which is not surprising, as multidecadal PNA indices and PDO indices are highly correlated (see Figure S7 in Supporting Information S1 in O'Reilly et al., 2017).

In the final part of the study, we investigated the link between EATCs and the occurrence of winter surface temperature extremes. The ASF-20C ensemble mean only partially manages to reproduce the relationships seen in the reanalysis. For example, some of the negative correlation between the NAO index and cold extremes in northern Europe is not present in the hindcasts, while the EA/WR predictions show some suspiciously high correlation over most of Europe, which do not exist in the reanalysis. This means that even if teleconnection patterns would be highly predictable, the resulting impact on surface temperature extremes might not be accurate in some cases. An analysis of the number of ensemble members that agree with the reanalysis on the location and sign of significant correlation shows that, while the exact areas might not be reproduced, other ensemble statistics other than the mean might still give valuable insight. At the same time, we find that according to ERA-20C the link of EATCs to surface temperature extremes does not change significantly over the second half of the twentieth century. Similarly, Schaller et al. (2018) found that the relationship between atmospheric blocking and summer heat waves is not likely to change in the future.

As some recent studies use long-term hindcast ensembles to help quantify extreme events (e.g., Kay et al., 2020; Thompson et al., 2017), it would be worth exploring if the considerable difficulties in capturing the teleconnection-surface temperature relationships is due to ASF-20C being an atmosphere-only model and thus potentially of relatively poor quality. An extension of the ASF-20C hindcast system has recently been introduced (Weisheimer et al., 2020), for the first time providing seasonal hindcasts for the twentieth century from a coupled atmospheric-ocean-sea-ice model. Further analysis with the coupled model could help understand the role of a dynamic ocean in determining the long-term behavior of EATCs and their impact on extreme temperatures.

Concluding, we found that the four main teleconnection patterns exhibit various degrees of decadal variability and that their seasonal predictive skill in winter, based on the ASF-20C ensemble, fluctuated over the twentieth century, with three of the patterns experiencing higher skill over recent periods. The heat and cold wave indices often show areas with significant correlation to the teleconnection patterns. However, the representation of these regional correlation patterns in the ensemble mean forecasts was rather poor. Despite the long-term changes in forecast skill, there were no significant signs that the relation between teleconnections and surface temperature extremes are subject to the same variability.

Data Availability Statement

ERA-20C long-term reanalysis data can be obtained from ECMWF (2019) via <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-20c> under Creative Commons Attribution 4.0 International (CC BY 4.0). ASF-20C hindcast ensemble data are available from Weisheimer and O'Reilly (2019) via <https://catalogue.ceda.ac.uk/uuid/6e1c3df49f644a0f812818080bed5e45> under the Open Government License v3.0. ERA-Interim reanalysis data, which was used to define the teleconnection patterns, are published at ECMWF (2011) via <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim> under Creative Commons Attribution 4.0 International (CC BY 4.0).

Figure 4. Maps of the Spearman's rank correlation coefficient between the seasonal CWMId values over December/January/February (DJF) and (a) the East Atlantic/Western Russia (EA/WR) index, (b) the Scandinavian (SCA) index. Correlations are calculated from the ERA-20C reanalysis and the ASF-20C mean for the period 1981–2010. All correlations are significant at the 5% level based on a t-distributed test statistic. The bottom left panels show the difference between ERA-20C and the ASF-20C mean. The bottom right panels show the number of ensemble members at each grid point where both ERA-20C and the ensemble member show significant correlation and agree on the sign of correlation.

Acknowledgments

The authors have received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 776787 (S2S4E). The authors thank three anonymous reviewers for their comments, which helped us to improve and clarify this manuscript.

References

Beerli, R., Wernli, H., & Grams, C. M. (2017). Does the lower stratosphere provide predictability for month-ahead wind electricity generation in Europe? *Quarterly Journal of the Royal Meteorological Society*, *143*(709), 3025–3036. <https://doi.org/10.1002/qj.3158>

Bessec, M., & Fouquau, J. (2008). The nonlinear link between electricity consumption and temperature in Europe: A threshold panel approach. *Energy Economics*, *30*(5), 2705–2721. <https://doi.org/10.1016/j.eneco.2008.02.003>

Bett, P., Thornton, H. E., Troccoli, A., De Felice, M., Suckling, E., Dubus, L., et al. (2019). A simplified seasonal forecasting strategy, applied to wind and solar power in Europe. *EarthArXiv*. <https://doi.org/10.31223/osf.io/kzwqx>

Bloomfield, H. C., Brayshaw, D. J., & Charlton-Perez, A. J. (2020). Characterizing the winter meteorological drivers of the European electricity system using targeted circulation types. *Meteorological Applications*, *27*(1), e1858. <https://doi.org/10.1002/met.1858>

Bloomfield, H. C., Brayshaw, D. J., Gonzalez, P. L. M., & Charlton-Perez, A. (2021a). Pattern-based conditioning enhances sub-seasonal prediction skill of European national energy variables. *Meteorological Applications*, *24*(4), e2018. <https://doi.org/10.1002/met.2018>

Bloomfield, H. C., Brayshaw, D. J., Gonzalez, P. L. M., & Charlton-Perez, A. (2021b). Sub-seasonal forecasts of demand and wind power and solar power generation for 28 European countries. *Earth System Science Data*, *13*, 2259–2274. <https://doi.org/10.5194/essd-13-2259-2021>

Bloomfield, H. C., Shaffrey, L. C., Hodges, K. I., & Vidale, P. L. (2018). A critical assessment of the long-term changes in the wintertime surface Arctic Oscillation and Northern Hemisphere storminess in the ERA20C reanalysis. *Environmental Research Letters*, *13*, 094004. <https://doi.org/10.1088/1748-9326/aa5c55>

Brayshaw, D. J., Troccoli, A., Fordham, R., & Methven, J. (2011). The impact of large scale atmospheric circulation patterns on wind power generation and its potential predictability: A case study over the UK. *Renewable Energy*, *36*(8), 2087–2096. <https://doi.org/10.1016/j.renene.2011.01.025>

Brunner, L., Schaller, N., Anstey, J., Sillmann, J., & Steiner, A. K. (2018). Dependence of present and future European temperature extremes on the location of atmospheric blocking. *Geophysical Research Letters*, *45*(12), 6311–6320. <https://doi.org/10.1029/2018GL077837>

Bruno Soares, M., & Dessai, S. (2016). Barriers and enablers to the use of seasonal climate forecasts amongst organizations in Europe. *Climate Change*, *137*, 89–103. <https://doi.org/10.1007/s10584-016-1671-8>

Clark, R. T., Bett, P. E., Thornton, H. E., & Scaife, A. A. (2017). Skillful seasonal predictions for the European energy industry. *Environmental Research Letters*, *12*(2), 024002. <https://doi.org/10.1088/1748-9326/aa57ab>

Climate Prediction Center. (2012). *Northern Hemisphere teleconnection patterns*. Retrieved from <http://www.cpc.ncep.noaa.gov/data/teleconnections/telecontents.shtml>

Davini, P., Weisheimer, A., Balmaseda, M., Johnson, S. J., Molteni, F., Roberts, C. D., et al. (2020). The representation of winter Northern Hemisphere atmospheric blocking in ECMWF seasonal prediction systems. *Quarterly Journal of the Royal Meteorological Society*, *146*(733), 1344–1363. <https://doi.org/10.1002/qj.3974>

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. <https://doi.org/10.1002/qj.828>

De Felice, M., Alessandri, A., & Catalano, F. (2015). Seasonal climate forecasts for medium-term electricity demand forecasting. *Applied Energy*, *137*, 435–444. <https://doi.org/10.1016/j.apenergy.2014.10.030>

Dorrington, J., Finney, I., Palmer, T., & Weisheimer, A. (2020). Beyond skill scores: Exploring sub-seasonal forecast value through a case study of French month-ahead energy prediction. *Quarterly Journal of the Royal Meteorological Society*, *146*(733), 3623–3637. <https://doi.org/10.1002/qj.3863>

Dunstone, N., Scaife, A. A., MacLachlan, C., Knight, J., Ineson, S., Smith, D., et al. (2018). Predictability of European winter 2016/2017. *Atmospheric Science Letters*, *19*(12), e868. <https://doi.org/10.1002/asl.868>

ECMWF. (2011). The ERA-Interim reanalysis data set, Copernicus Climate Change Service (C3S). [Dataset]. ECMWF. Retrieved from <https://www.ecmwf.int/en/forecasts/datasets/archive-datasets/reanalysis-datasets/era-interim>

ECMWF. (2019). ERA-20C. [Dataset]. ECMWF. Retrieved from <https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-20c>

Ely, C. R., Brayshaw, D. J., Methven, J., Cox, J., & Pearce, O. (2013). Implications of the North Atlantic Oscillation for a UK—Norway renewable power system. *Energy Policy*, *62*, 1420–1427. <https://doi.org/10.1016/j.enpol.2013.06.037>

Gonzalez, P. M., Brayshaw, D. J., & Ziel, F. (2021). A new approach to extended-range multimodel forecasting: Sequential learning algorithms. *Quarterly Journal of the Royal Meteorological Society*, *147*(741), 4269–4282. <https://doi.org/10.1002/qj.4177>

Grams, C. M., Beerli, R., Pfenniger, S., Staffell, I., & Wernli, H. (2017). Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Climate Change*, *7*(8), 557–562. <https://doi.org/10.1038/nclimate3338>

Hall, R. J., & Hanna, E. (2018). North Atlantic circulation indices: Links with summer and winter UK temperature and precipitation and implications for seasonal forecasting. *International Journal of Climatology*, *38*, e660–e677. <https://doi.org/10.1002/joc.5398>

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*(3), 1087–1117. <https://doi.org/10.5194/gmd-12-1087-2019>

Josey, S. A., Somot, S., & Tsimplis, M. (2011). Impacts of atmospheric modes of variability on Mediterranean Sea surface heat exchange. *Journal of Geophysical Research*, *116*, C02032. <https://doi.org/10.1029/2010JC006685>

Kay, G., Dunstone, N., Smith, D., Dunbar, T., Eade, R., & Scaife, A. (2020). Current likelihood and dynamics of hot summers in the UK. *Environmental Research Letters*, *15*, 094099. <https://doi.org/10.1088/1748-9326/abab32>

Lavaysse, C., Cammalleri, C., Dosio, A., van der Schrier, G., Toreti, A., & Vogt, J. (2018). Towards a monitoring system of temperature extremes in Europe. *Natural Hazards and Earth System Sciences*, *18*, 91–104. <https://doi.org/10.5194/nhess-18-91-2018>

Lavaysse, C., Naumann, G., Alfieri, L., Salamon, P., & Vogt, J. (2019). Predictability of the European heat and cold waves. *Climate Dynamics*, *52*, 2481–2495. <https://doi.org/10.1007/s00382-018-4273-5>

Lledó, L., Cionni, I., Torralba, V., Bretonnière, P.-A., & Samsó, M. (2020). Seasonal prediction of Euro-Atlantic teleconnections from multiple systems. *Environmental Research Letters*, *15*, 074009. <https://doi.org/10.1088/1748-9326/ab87d2>

Lledó, L., Torralba, V., Soret, A., Ramon, J., & Doblas-Reyes, F. J. (2019). Seasonal forecasts of wind power generation. *Renewable Energy*, *140*, 91–100. <https://doi.org/10.1016/j.renene.2019.04.135>

Matheson, J. E., & Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Management Science*, *22*(10), 1087–1096. <https://doi.org/10.1287/mnsc.22.10.1087>

Matsueda, M., & Palmer, T. N. (2018). Estimates of flow-dependent predictability of wintertime Euro-Atlantic weather regimes in medium-range forecasts. *Quarterly Journal of the Royal Meteorological Society*, *144*(713), 1012–1027. <https://doi.org/10.1002/qj.3265>

19448007, 2022, 11, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2020GL092360 by NASA Shared Services Center (NSSC), Wiley Online Library on [25/05/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

- O'Reilly, C. H., Heatley, J., MacLeod, D., Weisheimer, A., Palmer, T. N., Schaller, N., & Woollings, T. (2017). Variability in seasonal forecast skill of Northern Hemisphere winters over the twentieth century. *Geophysical Research Letters*, *44*, 5729–5738. <https://doi.org/10.1002/2017GL073736>
- Orlov, A., Sillmann, J., & Vigo, I. (2020). Better seasonal forecasts for the renewable energy industry. *Nature Energy*, *5*, 108–110. <https://doi.org/10.1038/s41560-020-0561-5>
- Palmer, T., Buizza, R., Hagedorn, R., Lawrence, A., Leutbecher, M., & Smith, L. (2006). *Ensemble prediction: A pedagogical perspective* (ECMWF Newsletter No. 106—Winter 2005/2006). European Centre for Medium-Range Weather Forecasts.
- Poli, P., Hersbach, H., Berrisford, P., Dee, D., Simmons, A., & Lalayaux, P. (2015). *ERA-20C deterministic* (ECMWF ERA Report Series: No. 20). Retrieved from <https://www.ecmwf.int/node/11700>
- Poli, P., Hersbach, H., Tan, D. G. H., Dee, D. P., Thépaut, J.-N., Simmons, A., et al. (2013). *The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C)* (ECMWF ERA Report Series: No. 14). Retrieved from <https://www.ecmwf.int/node/11699>
- Russo, S., Marchese, A. F., Sillmann, J., & Immé, G. (2016). When will unusual heat waves become normal in a warming Africa? *Environmental Research Letters*, *11*(5), 054016. <https://doi.org/10.1088/1748-9326/11/5/054016>
- Russo, S., Sillmann, J., & Fischer, E. M. (2015). Top ten European heat waves since 1950 and their occurrence in the coming decades. *Environmental Research Letters*, *10*(12), 124003. <https://doi.org/10.1088/1748-9326/10/12/124003>
- Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., et al. (2014). Skillful long-range prediction of European and North American winters. *Geophysical Research Letters*, *41*(7), 2514–2519. <https://doi.org/10.1002/2014GL059637>
- Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M., & Russo, S. (2018). Influence of blocking on Northern European and Western Russian heat waves in large climate model ensembles. *Environmental Research Letters*, *13*(5), 054015. <https://doi.org/10.1088/1748-9326/13/5/054015>
- Scherrer, S. C., Croci-Maspoli, M., Schwierz, C., & Appenzeller, C. (2006). Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *International Journal of Climatology*, *26*, 233–249. <https://doi.org/10.1002/joc.1250>
- Shi, W., Schaller, N., MacLeod, D., Palmer, T. N., & Weisheimer, A. (2015). Impact of hindcast length on estimates of seasonal climate predictability. *Geophysical Research Letters*, *42*(5), 1554–1559. <https://doi.org/10.1002/2014GL062829>
- Taylor, J. W., & Buizza, R. (2003). Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting*, *19*(1), 57–70. [https://doi.org/10.1016/S0169-2070\(01\)00123-6](https://doi.org/10.1016/S0169-2070(01)00123-6)
- Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., & Belcher, S. E. (2017). High risk of unprecedented UK rainfall in the current climate. *Nature Communications*, *8*, 107. <https://doi.org/10.1038/s41467-017-00275-3>
- Thorarindottir, T. L., & Schuhen, N. (2018). Verification: Assessment of calibration and accuracy. In S. Vannitsem, D. S. Wilks, & J. W. Messner (Eds.), *Statistical postprocessing of ensemble forecasts* (pp. 155–186). Elsevier. <https://doi.org/10.1016/B978-0-12-812372-0.00006-6>
- Thornton, H. E., Scaife, A. A., Hoskins, B. J., Brayshaw, D. J., Smith, D. M., Dunstone, N., et al. (2019). Skillful seasonal prediction of winter gas demand. *Environmental Research Letters*, *14*(2), 024009. <https://doi.org/10.1088/1748-9326/aaf338>
- 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. *Journal of Applied Meteorology and Climatology*, *56*(5), 1231–1247. <https://doi.org/10.1175/JAMC-D-16-0204.1>
- van der Wiel, K., Bloomfield, H. C., Lee, R. W., Stoop, L. P., Blackport, R., Screen, J. A., & Selten, F. M. (2019). The influence of weather regimes on European renewable energy production and demand. *Environmental Research Letters*, *14*(9), 094010. <https://doi.org/10.1088/1748-9326/14/9/094010>
- Weisheimer, A., Befort, D. J., MacLeod, D., Palmer, T., O'Reilly, C., & Strømmen, K. (2020). Seasonal forecasts of the twentieth century. *Bulletin of the American Meteorological Society*, *101*(8), E1413–E1426. <https://doi.org/10.1175/BAMS-D-19-0019.1>
- Weisheimer, A., Decremier, D., MacLeod, D., O'Reilly, C., Stockdale, T. N., Johnson, S., & Palmer, T. N. (2019). How confident are predictive estimates of the winter North Atlantic Oscillation? *Quarterly Journal of the Royal Meteorological Society*, *145*(1), 140–159. <https://doi.org/10.1002/qj.3446>
- Weisheimer, A., & O'Reilly, C. (2019). *Initialized seasonal forecast of the 20th century (accessed August 2019)*. Available from Centre for Environmental Data Analysis. Retrieved from <https://catalogue.ceda.ac.uk/uuid/6e1c3df49f644a0f812818080bed5e45>
- Weisheimer, A., Schaller, N., O'Reilly, C., MacLeod, D. A., & Palmer, T. (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. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 917–926. <https://doi.org/10.1002/qj.2976>
- White, C. J., Domeisen, D. I. V., Acharya, N., Adefisan, E. A., Anderson, M. L., Aura, S., et al. (2021). Advances in the application and utility of subseasonal-to-seasonal predictions. *Bulletin of the American Meteorological Society*, 1–57. In Press. <https://doi.org/10.1175/BAMS-D-20-0224.1>
- Wohland, J., Omrani, N.-E., Withaut, D., & Keenlyside, N. S. (2019). Inconsistent wind speed trends in current twentieth century reanalysis. *Journal of Geophysical Research: Atmospheres*, *124*, 1931–1940. <https://doi.org/10.1029/2018JD030083>
- Zubieta, L., McDermott, F., Sweeney, C., & O'Malley, M. (2017). Spatial variability in winter NAO-wind speed relationships in Western Europe linked to concomitant states of the East Atlantic and Scandinavian patterns. *Quarterly Journal of the Royal Meteorological Society*, *143*, 552–562. <https://doi.org/10.1002/qj.2943>