Received: 5 March 2024 | Revised: 6 September 2024 | Accepted: 17 October 2024

DOI: 10.1002/qj.4900

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Hybrid statistical–dynamical seasonal prediction of summer extreme temperatures in Europe

Luca Famooss Paolini | **Paolo Ruggieri** | **Salvatore Pascale** | **Erika Brattich** | **Silvana Di Sabatino**

Department of Physics and Astronomy "Augusto Righi", University of Bologna, Bologna, Italy

Correspondence
Luca Famooss Paolini, Department of Physics and Astronomy "Augusto Righi", University of Bologna, Bologna, Italy.
Email: luca.famoosspaolini@unibo.it

Funding information
Horizon Europe Project "SoluTion foR mItiGatinG climate-induced hEalth thReats"–TRIGGER, Grant/Award Numbers: CUP J33C22002420005, (101057739)

Abstract

Although summer extreme temperatures over Europe are potentially predictable on seasonal time-scales, state-of-the-art dynamical seasonal prediction systems (SPSs) exhibit low skills in predicting such events in central and northern Europe. This limitation arises from the underestimation of predictable components of climate variability in the model ensemble. However, recent studies suggest that the skills in predicting extratropical climate can be largely improved through statistical postprocessing techniques, which increase the signal-to-noise ratio in the model ensemble. In this study, we evaluate the potential for improving the seasonal prediction skills of European summer extreme temperatures in a multimodel ensemble (MME) of SPSs by applying a teleconnection-based subsampling technique in the hindcast period 1993–2016. This technique is applied to the North Atlantic Oscillation (NAO) and East Atlantic (EA) modes, which are key drivers of summer extreme temperatures in Europe. Results show that the subsampling substantially improves the MME prediction skills of both the summer NAO and EA. Specifically, correlations between the observed and subsampled MME NAO indices improve from −0.38 to 0.77, and for the EA they improve from −0.11 to 0.84. Similarly, the root-mean-square error of the subsampled MME NAO (EA) index improves from 1.06 (1.02) to 0.65 (0.56). Moreover, retaining those ensemble members that accurately represent the NAO teleconnections enhances the MME prediction skills for the summer European climate, including the occurrence of summer extreme temperatures. This improvement is particularly pronounced in central and northern Europe; that is, the regions where current SPSs show the lowest skills in predicting European heat extremes. In contrast, selecting ensemble members that accurately represent the EA teleconnections does not improve the predictions of summer extreme temperatures. This is likely associated with the model deficiencies in realistically representing the spatial pattern of the summer EA and, thus, the physical processes driving summer extreme temperatures.

KEYWORDS

East Atlantic pattern, hybrid statistical–dynamical seasonal prediction, North Atlantic Oscillation, summer extreme temperatures, teleconnection-based subsampling

1 | INTRODUCTION

The occurrence of summer extreme temperatures in Europe has been increasing since the middle of the 20th century, and this is expected to further increase in the future because of anthropogenic global warming (Seneviratne *et al.*, 2021). Thus, predicting heat extremes and their statistics several months ahead is becoming increasingly crucial, particularly in light of their impacts on socio-economic and environmental systems and human health (e.gAlbergel *et al.*, 2019; Brás *et al.*, 2021; Ebi *et al.*, 2021). As an example, several studies show that long and frequent exposure to extreme temperatures can induce increased mortality from cardiorespiratory diseases, chronic kidney diseases, and mental health issues (e.g., Cheng *et al.*, 2019; Thompson *et al.*, 2018; Wong, 2023).

Summer extreme temperatures over Europe are potentially predictable on seasonal time-scales. Indeed, the summer European climate is significantly influenced by slowly evolving components of the climate system, such as sea-surface temperatures (SSTs) in the North Atlantic (Duchez *et al.*, 2016; Mecking *et al.*, 2019), Indian Ocean (Black & Sutton, 2007), and Mediterranean Sea (Black *et al.*, 2004), as well as European soil moisture (Hamilton *et al.*, 2012; Prodhomme *et al.*, 2016) and Arctic sea-ice (Screen, 2013). Moreover, the occurrence of summer European extreme temperatures has been associated with modes of low-frequency climate variability, such as the North Atlantic Oscillation (NAO; Folland *et al.*, 2009; Alvarez-Castro *et al.*, 2018), the East Atlantic (EA; Cassou *et al.*, 2005; Wulff *et al.*, 2017), and the El Niño–Southern Oscillation (ENSO; Luo & Lau, 2020; Martija-Díez *et al.*, 2021). Such influence of the low-frequency component of climate variability on the European summer offers a certain level of predictability for extreme temperatures, which can be exploited for seasonal forecast.

In this context, state-of-the-art dynamical seasonal prediction systems (SPSs) generally show good skills in predicting summer European climate and heat extremes in the Mediterranean and eastern Europe (Hamilton *et al.*, 2012; Mishra *et al.*, 2019). The skilful seasonal forecast for extreme temperatures in these regions is mainly explained by the preconditioning of soil moisture during the spring season and the strong atmosphere–land coupling (Miralles *et al.*, 2014; Prodhomme *et al.*, 2016). In particular, wet conditions and moist soils during the spring season lead to a reduced probability of hot days in the following summer, as the available surface energy tends to dissipate through latent heat fluxes, thus dampening the surface heating (Seneviratne *et al.*, 2010). Conversely, dry spring conditions are generally followed by a higher frequency of hot summer days, although their occurrence is particularly sensitive to the atmospheric circulation (Quesada *et al.*, 2012). In contrast, state-of-the-art SPSs show poor skills in predicting summer European climate and heat extremes in central and northern Europe (Prodhomme *et al.*, 2022). This may be due to model biases in representing persistent anticyclonic circulations that characterise summer heat extremes (Beverley *et al.*, 2019, 2021), especially in western Europe (Horton *et al.*, 2015; Rousi *et al.*, 2022). These model biases have also been shown to cause a mismatch between the observed warming trend of heat extremes in western Europe and the one predicted by CMIP6 models (Boe *et al.*, 2020; Vautard *et al.*, 2023). Furthermore, the low seasonal prediction skills of summer European climate may be induced by the underestimation of predictable components of climate variability in the model ensemble (Dunstone *et al.*, 2023; Eade *et al.*, 2014). The low signal-to-noise ratio in the North Atlantic region has been suggested to be caused by several factors, such as errors in the teleconnections affecting the North Atlantic variability, the lack of extratropical ocean–atmosphere coupling, errors in the parametrised processes, weak eddy feedback in current climate models, or weak response to boundary conditions and external radiative forcing (Gastineau *et al.*, 2013; Gray *et al.*, 2013; Minobe *et al.*, 2008; Scaife & Smith, 2018; Swingedouw *et al.*, 2017).

Recent studies have shown that the skills in predicting extratropical climate on different time-scales can be largely improved by refining the ensemble-based dynamical forecasts with statistical postprocessing techniques (Dobrynin *et al.*, 2022; Düsterhus, 2020; Neddermann *et al.*, 2019; Smith *et al.*, 2020). As an example, Dobrynin *et al.* (2018) have shown that the predictability of the winter NAO in a dynamical SPS is enhanced by subsampling the model ensemble; that is, by retaining only those ensemble members that better represent the link between the NAO and its autumn predictors. This subsampling technique also enhances the model prediction skills for atmospheric variables strongly influenced by the winter NAO, such as the surface temperature on a large portion of the Eurasian continent. The main advantages of statistical postprocessing techniques like subsampling are their ability to accurately capture the phase of observed variability and amplify the amplitude of predictable components of climate variability compared with the ensemble mean. This increases the signal-to-noise ratio in the model ensemble and, ultimately, its prediction skills. Remarkably, the recent study by Dunstone *et al.* (2023) has shown that the signal-to-noise ratio of the summer NAO in a state-of-the-art SPS is particularly low, mirroring results previously found for the winter NAO. This highlights the significant potential for implementing postprocessing techniques to enhance the signal-to-noise ratio during the summer season, thereby improving overall forecasting

accuracy. However, statistical postprocessing techniques for ensemble-based dynamical forecasts have not been employed in the context of summer European climate and temperature extremes.

This study assesses whether the seasonal prediction skill of summer extreme temperatures in Europe in state-of-the-art dynamical SPSs can be improved through subsampling. Specifically, we use the multimodel ensemble (MME) of SPSs contributing to the Copernicus Climate Change Service (C3S) and all the hindcast runs for the time period 1993–2016. A hybrid statistical–dynamical prediction system is presented, in which the dynamical model ensemble is subsampled by retaining only those ensemble members that predict the phase of NAO and EA, typically linked to summer extreme temperatures in Europe. This approach relies on spring predictors of the weather regimes and thus allows us to retain only those ensemble members with a reasonable representation of summer heat extreme teleconnections.

The article is structured as follows. In Section 2, the data and the teleconnection-based subsampling technique are described. In Section 3, the representation of summer European climate in the C3S MME and the effect of the subsampling on its prediction skills are presented. Finally, in Section 4, we discuss some open issues related to the results presented here, and in Section 5 we summarise the main findings of this work.

2 | DATA AND METHODOLOGICAL APPROACH

2.1 | Data

In this study, we analyse a MME consisting of seven dynamical SPSs (Table 1), all contributing to the C3S. In particular, we use the simulations initialised on May 1, for a total of 141 members. This allows us to assess the ability of the C3S MME to predict the summer European climate on a seasonal time-scale. Specifically, we analyse the hindcast time period 1993–2016, common to all dynamical SPSs analysed here. The predictions of the European climate in the C3S MME are compared with the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis v.5 (ERA5) atmospheric reanalysis (Hersbach *et al.*, 2020), used here as a surrogate for observations.

2.2 | Teleconnection-based subsampling

In order to improve the skill of the C3S MME in predicting summer extreme temperatures on a seasonal time-scale, a teleconnection-based subsampling is adopted here, similar to the approach suggested by Dobrynin *et al.* (2018, 2022). Generally speaking, this approach relies on subsampling a model ensemble by retaining only those ensemble members that better represent the link between a weather regime in a season and its predictors in the season before. Given that this methodology combines elements of both statistical and dynamical prediction, it is commonly referred to as a hybrid statistical–dynamical prediction approach.

In this study, we select the summer NAO and EA as weather regimes of interest. Indeed, these two weather regimes have been shown to play a key role in the development of summer extreme temperatures in Europe (Cassou *et al.*, 2005). The positive phase of the NAO has been related to a higher frequency of hot days, particularly in northern Europe (Bladé *et al.*, 2012; Drouard *et al.*, 2019; Folland *et al.*, 2009). Conversely, the positive phase of the EA has been associated with a higher occurrence of hot

TABLE 1 Dynamical seasonal prediction systems (SPSs) analysed in this study. All the SPSs contribute to the Copernicus Climate Change Service and are initialised on May 1 during 1993–2016. Columns detail the institution name, the system version, the system value adopted on the Climate Data Store, the ensemble size, and the reference.

Institution	System version	Climate Data Store system value	Ensemble size	Reference
Centro Euro-Mediterraneo sui Cambiamenti Climatici	CMCC-SPS3.5	35	40	Gualdi *et al.* (2020)
Deutscher Wetterdienst	GCFS 2.1	21	30	Fröhlich *et al.* (2021)
Environment and Climate Change Canada	GEM5-NEMO	3	10	Lin *et al.* (2021)
European Centre for Medium-Range Weather Forecasts	SEAS5	51	25	Johnson *et al.* (2019)
Météo-France	System 8	8	25	Battè *et al.* (2021)
National Centers for Environmental Prediction	CFSv2	2	4	Saha *et al.* (2014)
United Kingdom Met Office	GlobSea6	602	7	Williams *et al.* (2018)

days in central Europe (Cassou *et al.*, 2005), whereas its negative phase may play a role in the occurrence of hot days in eastern Europe and western Russia (Di Capua *et al.*, 2021). Northern and central Europe are the regions where the state-of-the-art dynamical SPSs feature low skills in predicting heat extremes on a seasonal time-scale (Prodhomme *et al.*, 2022). Thus, selecting only those C3S MME members that better represent the link between summer NAO and EA indices and their spring predictors is expected to enhance the seasonal prediction skill of the MME for the summer European climate, including the occurrence of extreme temperatures. In this context, in order to directly quantify the role played by the NAO and EA modes on the summer extreme temperatures in Europe, we project the summer extreme temperatures onto the two weather regimes. Refer to Section 2.5 for further details about this analysis.

The summer NAO has not received as much attention from the scientific community as its winter counterpart. Although numerous studies have demonstrated the connection between the autumn state of the ocean (Czaja & Frankignoul, 2002), sea-ice (Cohen *et al.*, 2014), snow cover (Gastineau *et al.*, 2017), stratosphere (Baldwin & Dunkerton, 1999), and the winter NAO, only a few studies have analysed the potential spring drivers of summer NAO variability (Baker *et al.*, 2019; Dong *et al.*, 2013; Hall *et al.*, 2017; Matsumura *et al.*, 2014; Screen, 2013; Wang & Ting, 2022; Zhao *et al.*, 2004). Nevertheless, these studies show that the anomalous state of these climatic components during the spring season can influence the summer atmospheric circulation over the North Atlantic sector, leading to NAO-like anomalies. Thus, following these results, here we adopt as predictors of the summer NAO the April state of the North Atlantic SST (90°W–10°E, 0°–80°N), the Arctic sea-ice concentration (SIC; north of 40°N), the Northern Hemisphere snow cover (SNOWC; north of 40°N), and the zonally averaged zonal wind in the lower stratosphere. In this context, although contrary to the common understanding of stratosphere–troposphere coupling (limited to the winter and spring seasons; Baldwin & Dunkerton, 1999; Ineson & Scaife, 2009), it has recently been shown that the influence of the stratospheric variability on the tropospheric circulation may extend into the summer season (Dunstone *et al.*, 2023; Wang & Ting, 2022). Specifically, the anomalous state of the lower stratospheric polar vortex during the spring season has been shown to propagate downward in a few weeks, inducing summer tropospheric circulation anomalies resembling the NAO-like circulation. For the link between the summer NAO and April stratospheric zonal wind, we consider the zonal wind zonally-averaged over the longitudinal range 75°W–15°E. In this way, we only focus on the stratosphere–troposphere coupling over the North Atlantic sector as in Wang and Ting (2022). Furthermore, the zonally averaged zonal wind is averaged in the lower stratosphere and over the North Atlantic latitudinal range (30–100 hPa; 35°N–70°N; green box in Supporting Information Figure S7), similar to Dunstone *et al.* (2023).

In contrast to the summer NAO, several studies have focused on the summer EA and its spring predictors. Specifically, the summer EA variability has been linked to the spring state of the North Atlantic (Cassou *et al.*, 2005; Gastineau & Frankignoul, 2015) and tropical Pacific Ocean (Wulff *et al.*, 2017), as well as the spring state of Arctic sea-ice (Gastineau & Frankignoul, 2015; Wu *et al.*, 2013). For this reason, in the context of the present work, we adopt the SST in the North Atlantic (90°W–10°E, 0°–80°N) and tropical Pacific (100°E–80°W, 25°S–25°N) and the SIC in the Arctic (north of 40°N) during April as predictors of the summer EA.

The link between the summer NAO and EA indices and their respective predictors in April is assessed through a leave-one-out cross-validated correlation analysis. Specifically, for every predictor and for every left-one-out year, we identify the regions where the correlation between a specific weather regime index and the predictor is significant. The correlation significance is assessed through a two-tailed Student's t test against the null hypothesis of coefficients equal to zero at the 90% confidence level. It is noted here that the locations and patterns of these regions can vary in time. As suggested in Dobrynin *et al.* (2018), we merge all the regions of significant positive correlation computed during each cross-validation step in one region and we calculate the predictor index as the detrended area-weighted anomalies over that region. For the link between the summer EA and April North Atlantic SST, we consider the regions of significant negative correlation, rather than positive ones, for the calculation of the predictor index. In this way, we give greater importance to the state of tropical North Atlantic, which has been shown to play a crucial role in the variability of summer EA (Cassou *et al.*, 2005). In order to further assess the robustness of the link between the summer NAO and EA and their respective April predictors, their correlation is also assessed by applying a leave-three-out cross-validation procedure. In the end, the predictor indices computed through the leave-one-out cross-validation procedure are adopted for the teleconnection-based subsampling of the MME, as described in Section 3.

Once the April predictor indices are defined, they provide the statistical predictions (or first guesses) of the summer NAO and EA, states. This means that, for each left-one-out year, we have a total of four and three first guesses (one per each April predictor) for the NAO and EA, respectively. The C3S MME is subsampled selecting only those ensemble members whose state of a specific

weather regime is closest to its respective first guesses. Specifically, 10 members are selected for every April predictor. Then, for the summer NAO (EA), the four (three) 10-member ensembles are merged into a single ensemble, ensuring that members selected by multiple predictors are not counted more than once. Consequently, the maximum size of a subsampled ensemble is 40 members for the summer NAO and 30 members for the summer EA (the 10 members selected by each predictor are all different), whereas the minimum size is 10 members for both the weather regimes (the 10 members selected by each predictor are all the same). The absence of duplications in the subsampled ensemble also leads to the number of members selected being able to vary from year to year. The teleconnection-based subsampling technique is applied to the C3S MME for both the NAO and EA, separately.

In order to assess whether the performance obtained by subsampling the full MME depends on individual models, the aforementioned teleconnection-based subsampling technique is also applied to the individual SPSs listed in Table 1, separately. In this case, 10, 7, 3, 6, 6, 1, and 5 members are selected for every April predictor for the Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC), Deutscher Wetterdienst (DWD), Environment and Climate Change Canada (ECCC), ECMWF, Météo-France (MF), National Centers for Environmental Prediction (NCEP), and United Kingdom Met Office (UKMO) SPSs, respectively.

To assess the influence of the ensemble size on the predictive skills of the subsampled C3S MME for the summer NAO and EA states, a sensitivity analysis is performed. Specifically, the teleconnection-based subsampling is repeated with a change in the number of members selected by each NAO and EA April predictor within the range 1–141. In this context, the sensitivity analysis is carried out using the predictors both collectively and individually, in order to assess their relative importance on the performance obtained through the subsampling.

2.3 | Weather regime indices

In ERA5, the NAO index is defined as the leading principal component of summer geopotential height at 500 hPa (Z500) over the North Atlantic sector (25°–80°N, 90°W–40°E). Concurrently, the EA index is defined as the second principal component of Z500 over the same geographical domain. Prior to the computation of these indices, we apply a scaling factor $\cos(\text{latitude})^{1/2}$ to Z500 data. This ensures that the contribution of various grid cells to the field variance is proportionate to the areas they represent.

In the C3S MME, we calculate the NAO and EA indices following a similar approach as for the observational dataset. Specifically, we perform the empirical orthogonal function (EOF) analysis for every SPS independently, merging all the respective members along the time dimension. This results in one common EOF pattern for every weather regime that represents the entire ensemble of each SPS. Subsequently, we decompose the principal components of summer Z500 back to the number of model ensemble members. This approach has been shown to be more reliable compared with performing EOF analysis on individual ensemble members separately (Dobrynin *et al.*, 2022). As for the observational dataset, the NAO index is defined as the leading principal component of the summer Z500 over the North Atlantic for all the SPSs. Concurrently, the EA index is defined as the second principal component of the summer Z500 only for the ECMWF, MF, NCEP, and UKMO SPSs. Differently, the EA index is defined as the third principal component of summer Z500 for the CMCC and DWD SPSs, and as fourth principal component of summer Z500 for the ECCC SPS. Indeed, the second most relevant weather regime in the CMCC, DWD, and ECCC SPSs is the Atlantic ridge, which is represented by the third principal component in the other SPSs (ECMWF, MF, NCEP, and UKMO SPSs) as well as in ERA5. For completeness, we show the Atlantic ridge pattern and index both for ERA5 and the MME in Supporting Information Figure S1. Finally, the third weather regime for the ECCC SPS is characterised by Z500 anomalies that cannot be identified as any known mode of the summer low-frequency atmospheric variability (not shown).

As detailed in the following section, the results from the EOF analysis indicate that the C3S MME exhibits deficiencies in accurately representing the summer low-frequency atmospheric variability, especially with regard to the EA. Given the significant role of the EA in the occurrence of summer extreme temperatures in Europe, this aspect should be carefully taken into account when interpreting the results described in the following section.

2.4 | Index for the summer extreme temperatures

Several indices have been proposed in the literature to characterise summer temperature extremes. These indices can range from relatively simple metrics, such as counting the number of days exceeding a certain temperature threshold (e.g., Moberg *et al.*, 2006), to more complex ones, such as those combining both intensity and duration of extreme temperature anomalies (e.g., Russo *et al.*, 2014). In our study we use the TX90p index, which counts the number of days when the maximum temperature at

2 m height (T2m) exceeds the 90th percentile of daily T2m during the summer season (June–August) (Lhotka & Kyselý, 2015; Sulikowska & Wypych, 2020). Although this index is simple compared with other indices that define heat extremes, it is a good diagnostic for assessing the tendency of a given summer to be extremely warmer than usual or not (Prodhomme *et al.*, 2022). However, the methodology described in the present work can be applied to other heat extreme indices as well.

In the C3S MME dataset, the TX90p index is computed for every member of each SPS. Specifically, the daily T2m climatology is computed taking into account all the members of a single SPS, thus obtaining one T2m threshold (the 90th percentile) for each day for every system. Then, the number of days when the T2m exceeds the respective threshold is computed for each member of the ensemble, separately.

2.5 | The NAO and EA role in explaining the summer extreme temperature variability in Europe

As already mentioned, in this study we select the summer NAO and EA as weather regimes of interest, because they have been shown to play a key role in the European summer extreme temperatures (e.g., Cassou *et al.*, 2005). In order to directly quantify their role in the summer extreme temperature variability in Europe, we perform a linear regression analysis of the summer TX90p index against the summer NAO and EA indices, separately. The statistical significance of linear regression coefficients is assessed through a two-tailed Student's t test against the null hypothesis of coefficients equal to zero at the 90% confidence level. Additionally, the R^2 coefficient is adopted here as a measure of the fraction of TX90p variance explained by the NAO and EA (DelSole & Tippett, 2022).

2.6 | Seasonal prediction skill and statistical test

The performance of the individual full and subsampled C3S SPS ensembles, as well as the full and subsampled C3S MME, in predicting the summer NAO and EA is evaluated here through correlation analysis and root-mean-square error (RMSE). The correlation significance is assessed through a two-tailed Student's t test, providing the confidence level for which the null hypothesis of correlation equal to zero is rejected. In order to further assess the robustness of these correlations and RMSE values, the link between the NAO and EA indices and the corresponding indices computed as ensemble mean of the full and subsampled C3S MME are also assessed applying a leave-one-out and leave-three-out cross-validation procedure. The results of these analyses are provided in Table 2, along with the results of the same analyses applied to the link between the summer NAO and EA and their respective April predictors.

Furthermore, the prediction skills of the individual full and subsampled C3S SPS ensembles, as well as the full and subsampled C3S MME, in predicting the summer

TABLE 2 Correlation (Corr.) and root-mean-square error (RMSE) between the North Atlantic Oscillation (NAO) and East Atlantic (EA) indices in European Centre for Medium-Range Weather Forecasts Reanalysis v.5 and the same indices computed as the ensemble mean of the full and subsampled multimodel ensemble (MME) as well as mean over all the respective April predictors. The teleconnection-based subsampling is performed by selecting 10 members for every April predictor. The minimum, mean, and maximum values of the correlations and RMSE calculated through a leave-one-out and leave-three-out cross-validation cross-validated correlation analysis are shown.

	Full MME				Predictors				Subsampled MME			
	NAO		EA		NAO		EA		NAO		EA	
	Corr.	RMSE	Corr.	RMSE	Corr.	RMSE	Corr.	RMSE	Corr.	RMSE	Corr.	RMSE
1-year out												
Min	−0.54***	0.97	−0.30	0.93	0.71***	0.57	0.80***	0.51	0.71***	0.61	0.80***	0.52
Mean	−0.38*	1.06	−0.11	1.02	0.78***	0.63	0.84***	0.55	0.77***	0.65	0.84***	0.56
Max	-0.30	1.08	0.02	1.04	0.82***	0.64	0.86***	0.56	0.81***	0.67	0.86***	0.57
3-year out												
Min	−0.52**	0.94	-0.36	0.9	0.67***	0.58	0.79***	0.5	0.65***	0.63	0.79***	0.50
Mean	−0.38*	1.06	-0.11	1.02	0.76***	0.65	0.84***	0.55	0.74***	0.68	0.84***	0.56
Max	−0.27	1.13	0.06	1.08	0.83***	0.69	0.87***	0.58	0.84***	0.72	0.88***	0.59

Note: Correlations that are statistically significant at the 90%, 95%, and 99% confidence levels are indicate with *, **, and ***, respectively.

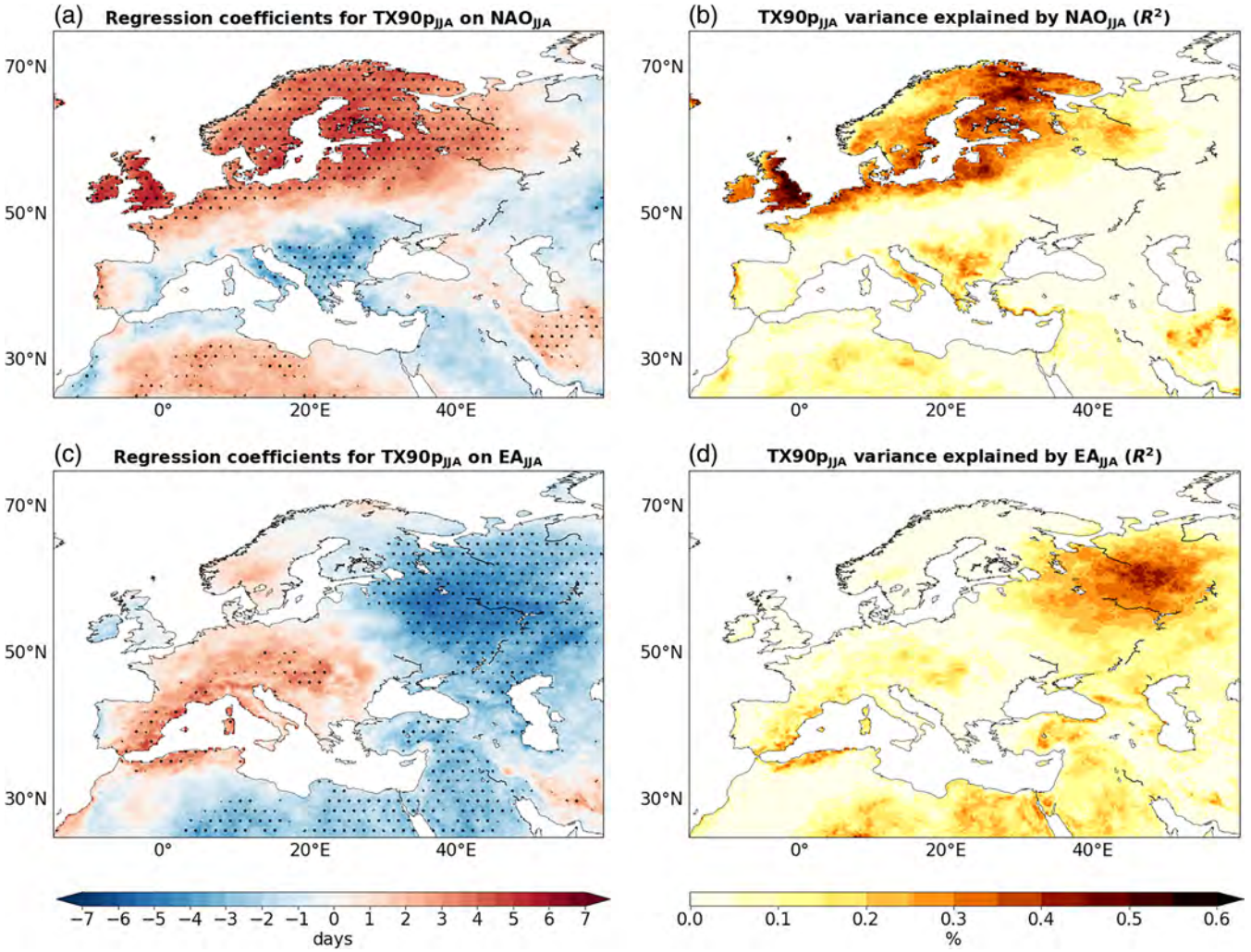

FIGURE 1 (a) Linear regression coefficients for summer TX90p (days) onto the summer North Atlantic Oscillation (NAO), using European Centre for Medium-Range Weather Forecasts Reanalysis v.5 data. Black dots denote linear regression coefficients that are statistically significant against the null hypothesis of coefficients equal to zero at the 90% confidence level. (b) R^2 coefficients (percentage) for the summer TX90p onto the summer NAO. The R^2 coefficient here is adopted as a measure of the fraction of TX90p variance explained by the NAO (DelSole & Tippett, 2022). (c, d) As (a) and (b) but for the East Atlantic (EA). TX90p: number of days when maximum temperature at 2 m height (T2m) exceeds the 90th percentile of daily T2m during June–August (JJA).

European climate are assessed by computing the anomaly correlation coefficients (ACCs) for Z500, T2m, and TX90p. The ACC statistical significance is assessed through a two-tailed Student's t test against the null hypothesis of ACC equal to zero at the 90% confidence level.

In order to evaluate the effect of subsampling on the model ensemble ability to predict summer European climate, we calculate the differences between the ACCs computed from the subsampled MME and the ones computed from the full MME. The ACC differences are computed for Z500, T2m, and TX90p. The statistical significance of the ACC differences is determined using a z test on Fisher z-transformed correlation coefficients (Fisher, 1920; Hinkle *et al.*, 2003), against the null hypothesis of zero differences at the 90% confidence level.

3 | RESULTS

Figure 1a shows that the positive phase of the NAO is associated with greater number of hot days in northern Europe, whereas its negative phase affects southern Europe, particularly Italy and the Balkan region. The impact of the negative NAO on southern Europe extreme temperatures aligns with the presence of positive Z500 anomalies characterising the summer NAO pattern over the Mediterranean region (Figure 2a). In this context, the NAO accounts for a significant portion of summer extreme temperature variance in northern Europe, with R^2 coefficients ranging from approximately 0.3 to 0.6 (Figure 1b). The NAO also explains a relevant portion of summer extreme temperature variance in Italy and the

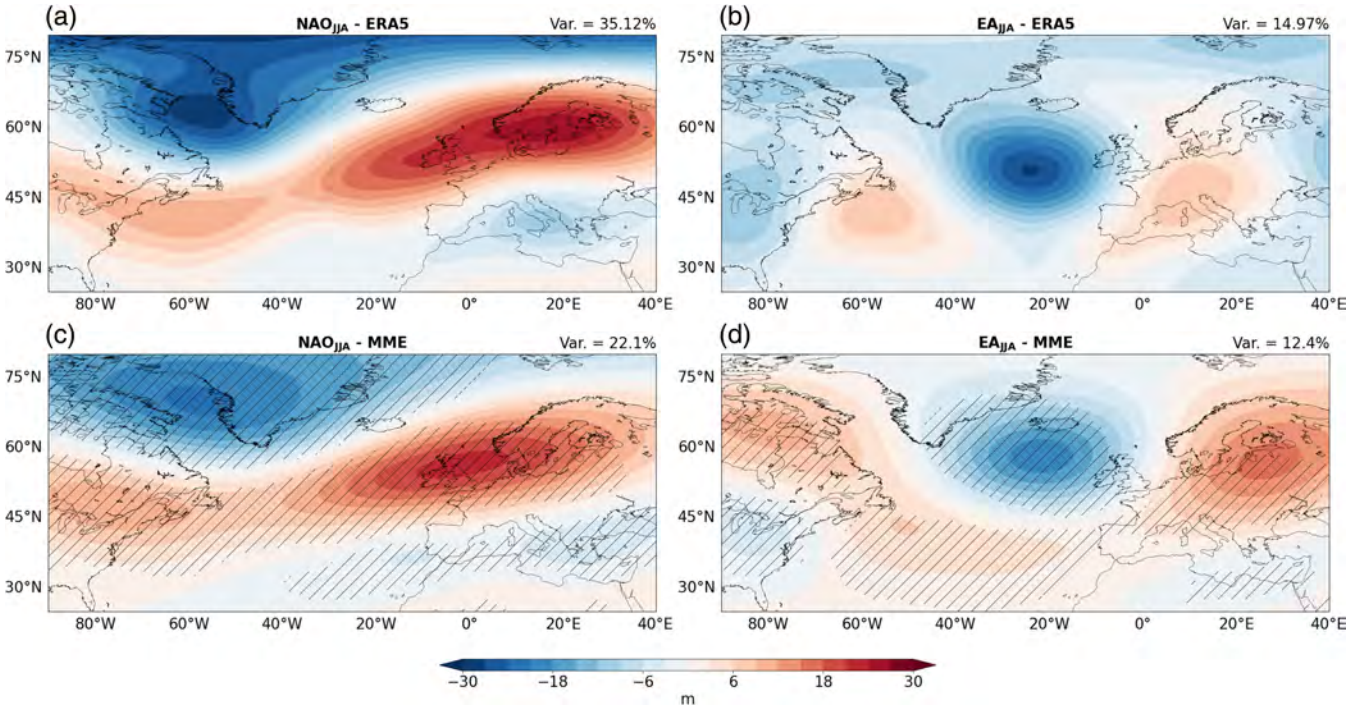

FIGURE 2 (a) Summer (June–August, JJA) North Atlantic Oscillation (NAO) pattern, calculated as the leading empirical orthogonal function (EOF) of summer Z500 over the North Atlantic sector (25°–80°N, 90°W–40°E) in European Centre for Medium-Range Weather Forecasts Reanalysis v.5 (ERA5). (b) Summer East Atlantic (EA) pattern, calculated as the second EOF of summer geopotential height at 500 hPa (Z500) over the North Atlantic sector in ERA5. (c, d) As (a) and (b) but in the Copernicus Climate Change Service (C3S) multimodel ensemble (MME). Refer to Section 2 for more details about the EOF analysis applied to the C3S MME. The variance of the summer low-frequency atmospheric variability explained by the NAO and EA in ERA5 and in the C3S MME is shown on the top right of each subplot. The results for the C3S MME are represented as MME mean. Hatched areas represent the regions with consistent sign in Z500 anomalies associated with the NAO and EA across all seasonal prediction systems.

Balkan region, although this influence is smaller than that in northern Europe.

Furthermore, Figure 1 shows that the summer extreme temperatures can be associated with the summer EA-like circulation. Specifically, the positive (negative) EA is associated with greater number of hot days in central Europe (eastern Europe and western Russia) (Figure 1c). The variance of summer extreme temperatures explained by the summer EA is particularly high in eastern Europe and western Russia, with the R^2 coefficient in the range 0.2–0.5. The EA also accounts for the summer extreme temperature variability in the western Mediterranean region and central Europe, though to a lesser degree than in eastern Europe and western Russia, with R^2 ranging between 0.1 and 0.3.

Consistent with previous studies (Di Capua *et al.*, 2021; Duchez *et al.*, 2016; Neddermann *et al.*, 2019), these results highlight the relevance of the NAO and EA modes in influencing summer extreme temperatures and explaining their variance in Europe, particularly in northern, central, and eastern regions. Therefore, incorporating these modes into the subsampling strategy may be beneficial to improve the seasonal prediction of summer extreme temperatures in state-of-the-art SPSs.

Figure 2a,c shows that the C3S MME successfully reproduces the key features of the summer NAO pattern in the observations, both in terms of spatial shape and amplitudes of Z500 anomalies. The NAO pattern is well reproduced by all single SPSs, as illustrated by the hatches in Figure 2c representing those regions with consistent sign in Z500 anomalies across all models. In terms of Z500 anomalies' amplitude, a small discrepancy emerges in the eastern portion of the Scandinavian region, where the positive Z500 anomalies are slightly lower in the MME than in ERA5. Despite the overall good representation of the NAO pattern, the variance of the summer low-frequency atmospheric variability explained by the NAO in the MME (22.1%) is lower than that in the observations (35.12%). In this context, the NAO-explained variance in the single SPSs ranges from a minimum of 17.91% (MF) to a maximum of 26.78% (ECCC).

Figure 2b,d also shows that the C3S MME is only partially able to capture the features of the summer EA pattern as observed in ERA5. In particular, the MME correctly

reproduces the position of the anomalous Z500 low over the subpolar gyre and the position of the anomalous Z500 high near the Gulf Stream region. However, the MME incorrectly represents the position of the positive Z500 anomalies over the European continent, which are shifted to the northeast. This is also true for the single SPSs, as shown by the hatched regions in Figure 2d. In this context, we specify that the absence of agreement among all models in northern Europe (as indicated by the lack of hatches) is only attributable to the UKMO SPS, which exhibits negative Z500 anomalies rather than positive ones (not shown). The position of the anomalous Z500 high over central (eastern) Europe during the positive (negative) EA phase plays a key role in the occurrence and variance of summer European extreme temperatures (Figure 1c,d). Specifically, the anomalous Z500 high can induce hot days through advection of warm air coming from lower latitudes, air subsidence, and sensible heat fluxes due to enhanced solar radiation (Domeisen *et al.*, 2023; Vautard *et al.*, 2023). Thus, these biases should be carefully taken into account when interpreting the MME skills in predicting the summer European climate and extreme temperatures, as well as assessing the effect induced by the subsampling on these skills, particularly in central and eastern Europe. Differently from the NAO, the variance of the summer low-frequency atmospheric variability explained by the EA in the MME (12.4%) is comparable to the that in the observations (14.97%). In this context, the EA-explained variance in the single SPSs ranges from a minimum of 9.04% (MF) to a maximum of 15.12% (UKMO).

The fact that the C3S MME correctly reproduces the NAO pattern but fails to account for the variance explained by the NAO, and vice versa for the EA, indicates that the C3S MME has some deficiencies in representing the low-frequency summer atmospheric variability over the Euro-Atlantic sector. In addition, the C3S MME generally captures the wrong phase of the NAO and EA indices. In fact, the correlation between the NAO indices in the observations and in the MME mean is −0.38 (statistically significant at 93% confidence level; Figure 3a). This correlation is quite robust in time, as demonstrated by the ranges calculated through the leave-one-out (−0.54; −0.30) and leave-three-out (−0.52; −0.27) cross-validation procedure (Table 2). The negative correlation between the NAO indices in the observational and model datasets is also generally present in the individual SPSs, with correlations ranging from a minimum of −0.44 (CMCC) to a maximum of 0.05 (UKMO) (Supporting Information Figure S2). Furthermore, the C3S MME does not accurately capture the amplitude of the NAO index, as shown by the RMSE for the MME mean equal to 1.06 (Figure 3a) and that in the single SPSs ranging from a minimum of 1.03 (ECMWF and UKMO) to a maximum of 1.17 (ECCC and NCEP) (Supporting Information Figure S2). The RMSE values are consistent over time, with values ranging between 0.97 and 1.08 (0.94 and 1.13) for the leave-one-out (leave-three-out) cross-validation procedure (Table 2). Similar to the NAO, the correlation between the EA indices in the observations and in the MME mean is negative and equal to −0.11 (not statistically significant; Figure 3b). The fact that this value is smaller than that for the NAO index and not statistically significant is consistent with the fact that the correlation between the EA indices in the observational and model datasets is not robust in time, as shown by the ranges calculated through the leave-one-out (−0.30; 0.02) and leave-three-out (−0.36; 0.06) cross-validation procedure (Table 2). This is also in line with the correlations in the individual SPSs, ranging from a minimum of −0.38 (MF) to a maximum of 0.32 (DWD) (Supporting Information Figure S3). The wide range of the correlations in the C3S SPSs shows that these models have different skills in predicting the phase of the EA. The SPSs also show deficiencies in capturing the amplitude of the EA index, as shown by the RMSE for the MME mean equal to 1.02 (Figure 3b) and that in the single SPSs ranging from a minimum of 0.95 (DWD) to a maximum of 1.23 (UKMO) (Supporting Information Figure S3). Differently from the correlations, the RMSE values for the EA index are consistent over time, with values ranging between 0.93 and 1.04 (0.90 and 1.08) for the leave-one-out (leave-three-out) cross-validation procedure (Table 2). Finally, it is important to highlight that the predictable signal of the interannual variability in the C3S MME is much weaker than that in observations, as shown by the standard deviation of the MME mean equal to 0.13 for the NAO index and 0.10 for the EA index.

The low seasonal prediction skills of the C3S MME in predicting the NAO and EA are reflected in poor skills for Z500 over a large portion of the North Atlantic basin (Figure 6a). A similar lack of skill is observed for T2m and TX90p over central and northern Europe (Figure 6d,g). The NAO and EA represent the two most significant weather patterns characterising the low-frequency atmospheric variability in the North Atlantic region during the summer. Consequently, a suboptimal representation of their variability is expected to lead to diminished predictive performance for atmospheric variables strongly influenced by these patterns. Here, we point out that the C3S MME also has deficiencies in representing the third most important regime of low-frequency atmospheric variability over the Euro-Atlantic sector; that is, the Atlantic ridge (Supporting Information Figure S1). This is true both in terms of its pattern and variability and further explains the low skills of the C3S MME in predicting the summer European climate. In this context, it is highlighted that the low ACCs over the European region shown in Figure 6a,d,g, as well as in Figure 7a,d,g, are consistent

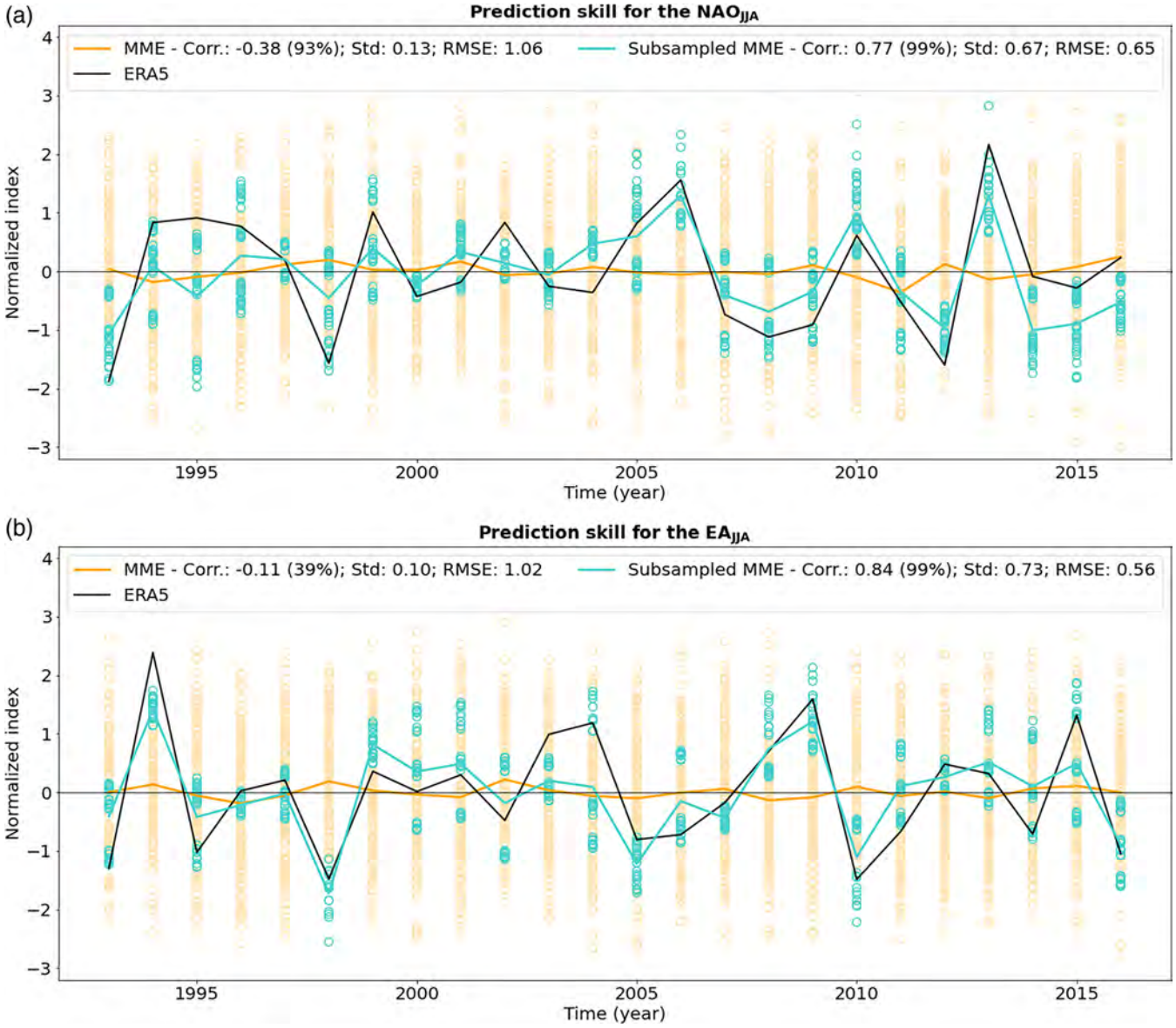

FIGURE 3 (a) The summer (June–August, JJA) North Atlantic Oscillation (NAO) index in European Centre for Medium-Range Weather Forecasts Reanalysis v.5 (ERA5; black), in the full Copernicus Climate Change Service (C3S) multimodel ensemble (MME; orange), and in the subsampled C3S MME (cyan). The orange (cyan) circles represent each member of the full (subsampled) MME. The orange (cyan) solid line represents the ensemble mean of the full (subsampled) MME. The subsampling of the MME is performed using all the April NAO predictors and selecting 10 members for every predictor. The correlation between the summer NAO index in ERA5 and the corresponding index as ensemble mean of the full and subsampled MME is presented in the legend, along with its confidence level in parentheses. The variability of the summer NAO index as ensemble mean of the full and subsampled MME is also shown in the legend as standard deviation (Std). (b) As (a) but for the summer East Atlantic (EA) index. RMSE: root-mean-square error.

with previous studies assessing the SPSs' ability in predicting the summer European climate and heat extremes (Hamilton *et al.*, 2012; Mishra *et al.*, 2019; Prodhomme *et al.*, 2022).

In order to enhance the seasonal prediction skills of the model ensembles in predicting the NAO and EA phases, we apply the teleconnection-based subsampling both to the C3S MME and individual SPSs, as described in Section 2. The first step of this methodology is to analyse the relationship between the summer NAO and EA and their April predictors, and thus to define the April predictor indices. The relationships between the summer NAO and EA and their April predictors can be found in Supporting Information Figures S4 to S10. Specifically, Supporting Information Figures S4 to S7 show that the mean (range) of the leave-one-out cross-validated correlations between the summer NAO and the North Atlantic SST, the Arctic SIC, the Northern Hemisphere SNOWC, and the zonally

averaged zonal wind in the lower stratosphere are 0.61 (0.54–0.67), 0.65 (0.59–0.81), 0.78 (0.72–0.80), and 0.56 (0.40–0.60), respectively. When we average all April NAO predictors together, the mean (range) of the correlation is 0.78 (0.71–0.82) (Table 2). On the other hand, Supporting Information Figures S8 to S10 show that the mean (range) of the leave-one-out cross-validated correlations between the summer EA and the North Atlantic SST, the tropical Pacific SST, and the Arctic SIC are 0.70 (0.65–0.76), 0.68 (0.66–0.74), and 0.68 (0.58–0.76), respectively. When the April EA predictors are combined, the mean (range) of the correlation is 0.84 (0.80–0.86). Similar values are obtained when applying the leave-three-out cross-validation procedure both for the NAO and EA. In general, these findings demonstrate the robustness of the relationships between the summer NAO and EA and the April predictors used in this analysis, with none of the years examined significantly influencing these relationships.

Once the April predictor indices are defined, they constitute the statistical predictions (or first guesses) of the summer NAO and EA states, through which the C3S MME and individual SPS ensembles can be subsampled. Figure 3 shows that the prediction skills of the C3S MME can be largely enhanced by applying the subsampling technique, using all the April predictors of the summer NAO and EA. In fact, the correlation between the NAO indices in ERA5 and in the subsampled C3S MME increases from −0.38 up to 0.77 (statistically significant at 99% confidence level), while the RMSE reduces from 1.06 to 0.65 in the subsampled MME. Similarly, the correlation for the EA index increases from −0.11 up to 0.84 (statistically significant at 99% confidence level) and the RMSE reduces from 1.02 to 0.56. Consistently, Supporting Information Figures S2 and S3 show that the subsampling also improves the prediction skills of the single SPSs, with correlations (RMSE) within the range 0.60–0.76 (0.70–0.80) and 0.60–0.85 (0.58–0.82) for the NAO and EA indices respectively. The correlations between the NAO and EA indices in ERA5 and in the subsampled SPSs are all statistically significant at 99% confidence level. The enhancement of the seasonal prediction skills of the subsampled MME is robust in time, as shown by the small ranges of the leave-one-out and leave-three-out cross-validated correlation for both the NAO (0.71–0.81; 0.65–0.84) and EA (0.80–0.86; 0.79–0.88) (Table 2). The same robustness applies to the RMSE values, with ranges of 0.61–0.67 (0.52–0.57) and 0.63–0.72 (0.50–0.59) for the leave-one-out and leave-three-out cross-validation procedure applied to the NAO (EA). Finally, we point out that the subsampling also enhances the interannual summer variability of the NAO and EA indices in the subsampled MME mean, reaching values of 0.67 and 0.73, respectively.

To assess the sensitivity of the results described thus far to the ensemble size, Figures 4 and 5 show the subsampled MME skills and variability of the NAO and EA indices with change in the number of members selected in the subsampling. In this context, the correlation, variability and RMSE are also computed using single April predictors, in order to assess their relative importance on the performance obtained. Specifically, Figure 4 shows that the MME skills in predicting the NAO and EA indices are pretty similar for a number of members selected within the range 1–100. After that, the correlation between the MME mean and observation indices decreases, reaching the minimum for an ensemble size equal to 141 (i.e., the full MME). Notably, this aspect shows that the MME mean indices computed with a great portion of the MME members (about 70%) generally capture the correct phases of the observed NAO and EA and that the overall MME correlation is strongly deteriorated by a very small portion of members (about 30%). Differently from the correlation, the interannual variability of the two indices regularly decreases with an increase of the ensemble size. On the other hand, the RMSE increases as the number of members selected increases (Figure 5). This is not entirely true for the single April predictors. Indeed, apart from using the SNOWC as April NAO predictor, the RMSE values for the NAO (EA) remain relatively stable for a number of members within the range 1–80 (1–50) and start to increase when approaching the full ensemble. These findings highlight the importance of a careful assessment of the number of members selected prior to implementing subsampling, in order to ensure a compromise between skilful predictions and high variability. In the context of the present work, the sensitivity analysis described herein confirms that using 10 members for every April predictor satisfies both these criteria.

Moreover, Figures 4 and 5 show that the improvement of the NAO prediction is mainly obtained by using the SNOWC in the Northern Hemisphere as April predictor (green line). Indeed, this April predictor allows one to obtain the highest correlation and the most significant reduction of the RMSE between the MME mean and observation NAO indices. At the same time, the North Atlantic SST, the Arctic SIC, and the wind in the lower stratosphere have relatively smaller impacts. In this context, the contributions of the North Atlantic SST and the Arctic SIC are comparable, whereas the stratospheric wind plays a minor role. In contrast, using all the April predictors results in higher prediction skills (higher correlation and lower RMSE) for the EA index compared with using the single April predictors. Notably, the SST in the North Atlantic and tropical Pacific exhibit slightly better skills in comparison with the Arctic SIC, especially with ensemble sizes smaller than about 60 members.

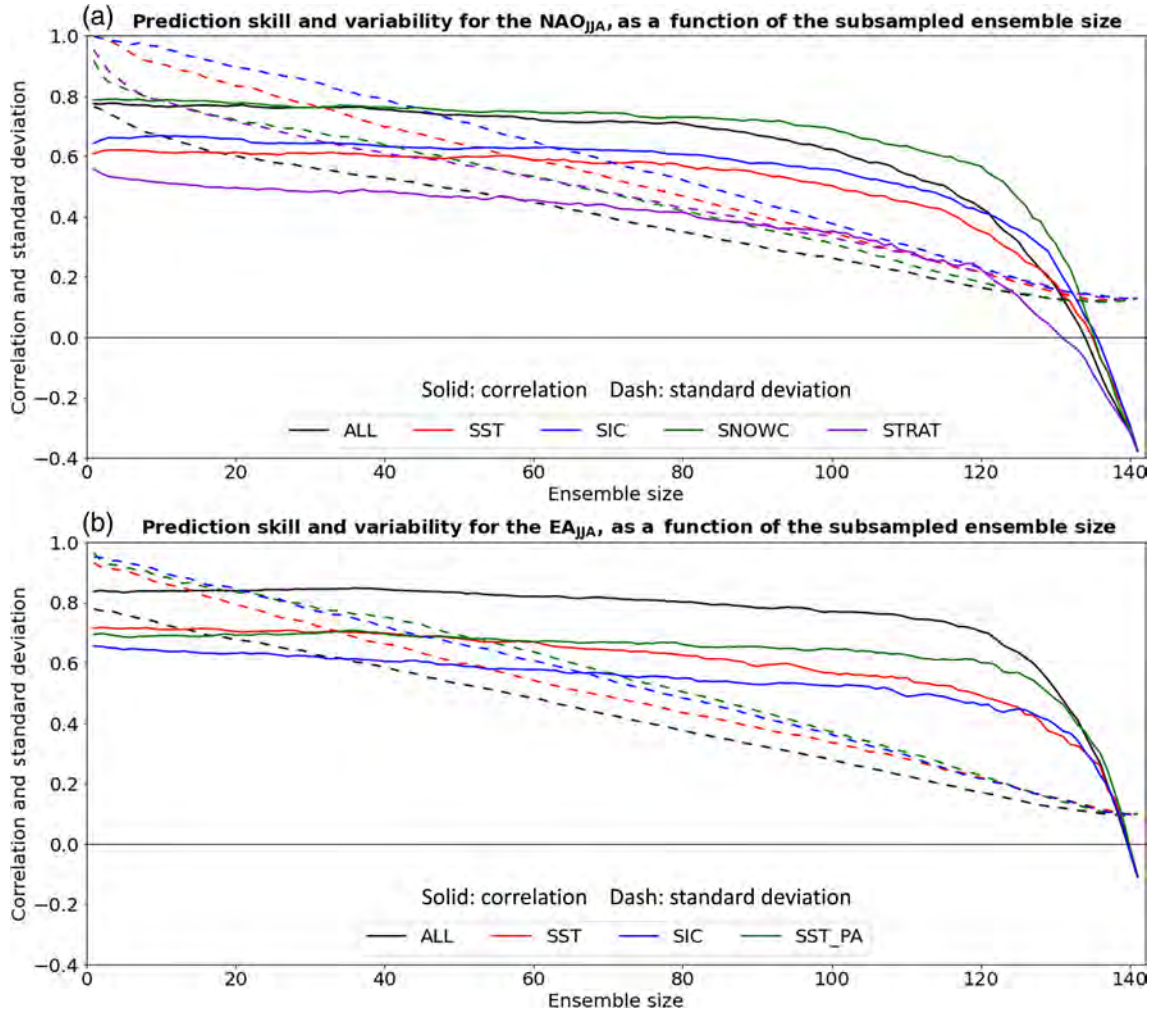

FIGURE 4 (a) The prediction skill (correlation: solid line) and variability (standard deviation: dashed line) for the summer (June–August, JJA) North Atlantic Oscillation (NAO) in the subsampled Copernicus Climate Change Service (C3S) multimodel ensemble (MME). The subsampling of the C3S MME is performed using all the April NAO predictors (black) and using only the North Atlantic sea-surface temperature (SST; red), the Arctic sea-ice concentration (SIC; blue), the Northern Hemisphere snow cover (SNOWC; green), and the zonally averaged zonal wind in the lower stratosphere (STRAT; purple) in April. Furthermore, the subsampling is performed by changing the number of members selected within the range 1–141 for every predictor. The prediction skill is computed with respect to the summer NAO index in European Centre for Medium-Range Weather Forecasts Reanalysis v.5. (b) As (a) but for the summer East Atlantic (EA) index. In this case the green lines refer to the subsampling of the C3S MME performed using the tropical Pacific SST (SST_PA) in April.

Finally, Figure 4 shows that the variability of the NAO and EA indices obtained by using the single April predictors is always higher than that obtained by using all the predictors together. This is consistent with the fact that, when all the predictors are considered, the subsampled MME mean indices are typically computed over a larger ensemble size, which tends to dampen the variability of the indices themselves.

Based on the aforementioned results, we assess the impact of the subsampling on the MME seasonal prediction skills for Z500, T2m, and TX90p, with the latter here adopted as the index to describe the summer extreme temperatures (Lhotka & Kyselỳ, 2015; Sulikowska & Wypych, 2020). Figure 6c shows that retaining only those ensemble members with a good representation of NAO teleconnections improves the MME seasonal prediction skills for the Z500 over a large portion of the North Atlantic region. This is especially true over the central and eastern North Atlantic, where the subsampled MME exhibits positive ACCs for Z500 (Figure 6b), whereas the full MME exhibits negative ACCs (Figure 6a). This is reflected in higher ACCs for the T2m and, importantly, TX90p over the land of northern Europe (Figure 6f,i). Notably, the ACC differences are particularly pronounced and statistically significant over regions such as the United Kingdom, parts of the Scandinavian Peninsula and western Russia; that

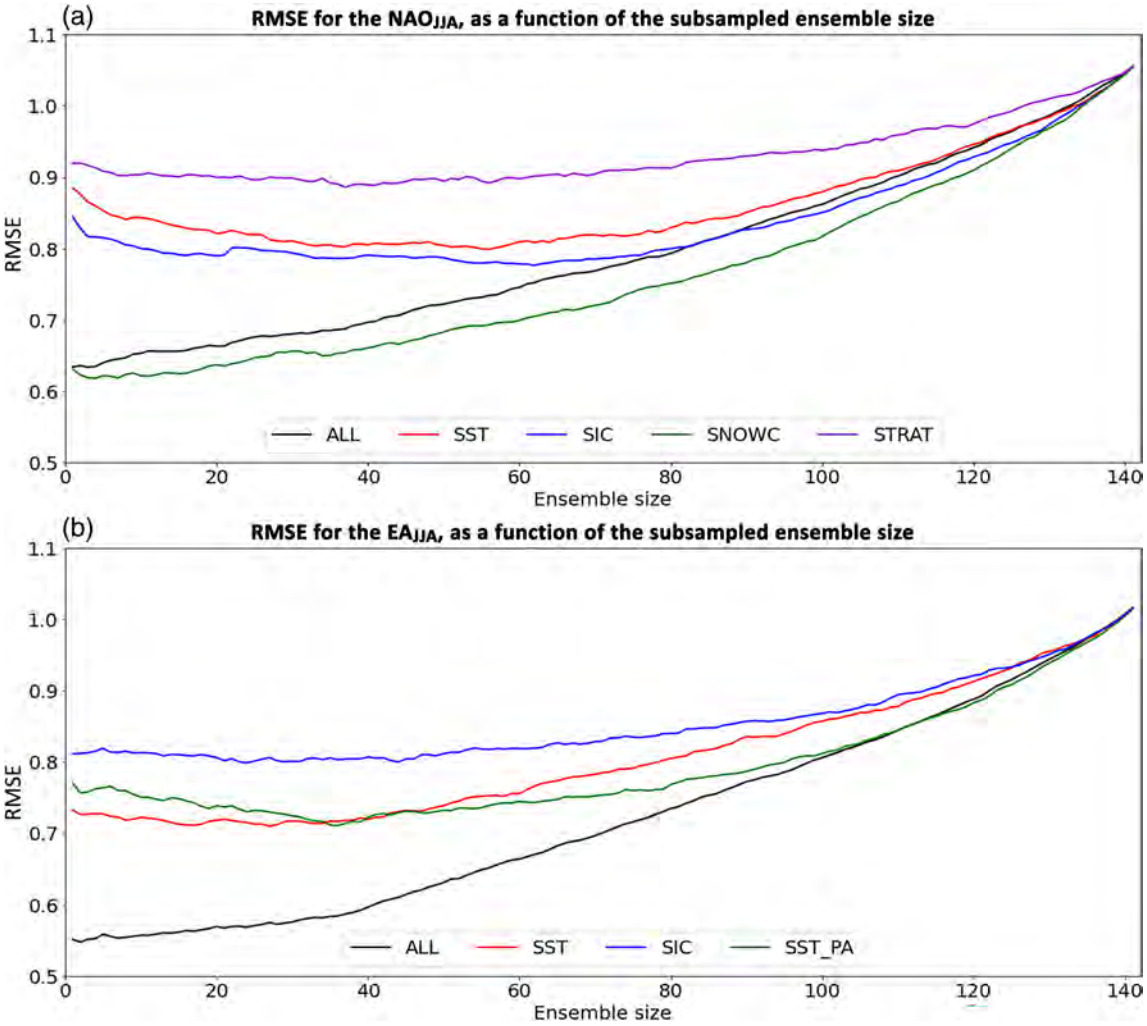

FIGURE 5 (a) The root-mean-square error (RMSE) for the summer (June–August, JJA) North Atlantic Oscillation (NAO) in the subsampled Copernicus Climate Change Service (C3S) multimodel ensemble (MME). The subsampling of the C3S MME is performed using all the April NAO predictors (black) and using only the North Atlantic sea-surface temperature (SST; red), the Arctic sea-ice concentration (SIC; blue), the Northern Hemisphere snow cover (SNOWC; green), and the zonally averaged zonal wind in the lower stratosphere (STRAT; purple) in April. The subsampling is performed by changing the number of members selected within the range 1—141. The RMSE is computed with respect to the summer NAO index in European Centre for Medium-Range Weather Forecasts Reanalysis v.5. (b) As (a) but for the summer East Atlantic (EA) index. In this case the green line refers to the subsampling of the C3S MME performed using the tropical Pacific SST (SST_PA) in April.

is, those regions where the state-of-the-art SPSs have been shown to be particularly deficient in predicting the summer European climate (e.g., Prodhomme *et al.*, 2022). However, it should be noted that the subsampled MME exhibits positive and often statistically significant skills for T2m and TX90p over a much broader area when compared with the full MME, covering western and central Europe, the Balkan region, and the entire Scandinavian Peninsula (Figure 6e,h). The fact that the improvement in T2m and TX90p ACCs, as shown by ACCs differences, is not always significant in certain regions may be due to the already high predictive skills of the full MME in those areas. Nevertheless, this should not detract from the higher capacity of the subsampled MME to predict the summer European climate compared with the full MME, thanks to the correct representation of the NAO phase. The enhanced prediction skills of the subsampled MME are also evident in its improved ability to capture the anomalous states of climatic variables during specific summers. For instance, Supporting Information Figure S11 shows that the subsampled MME more accurately represents the Z500, T2m, and TX90p anomalies observed in ERA5 during summer 2006, here selected as an example owing to its intense positive NAO phase. The improvement is particularly evident in the North Atlantic sector for Z500 and in northern and southern Europe for T2m and TX90p, regions where the

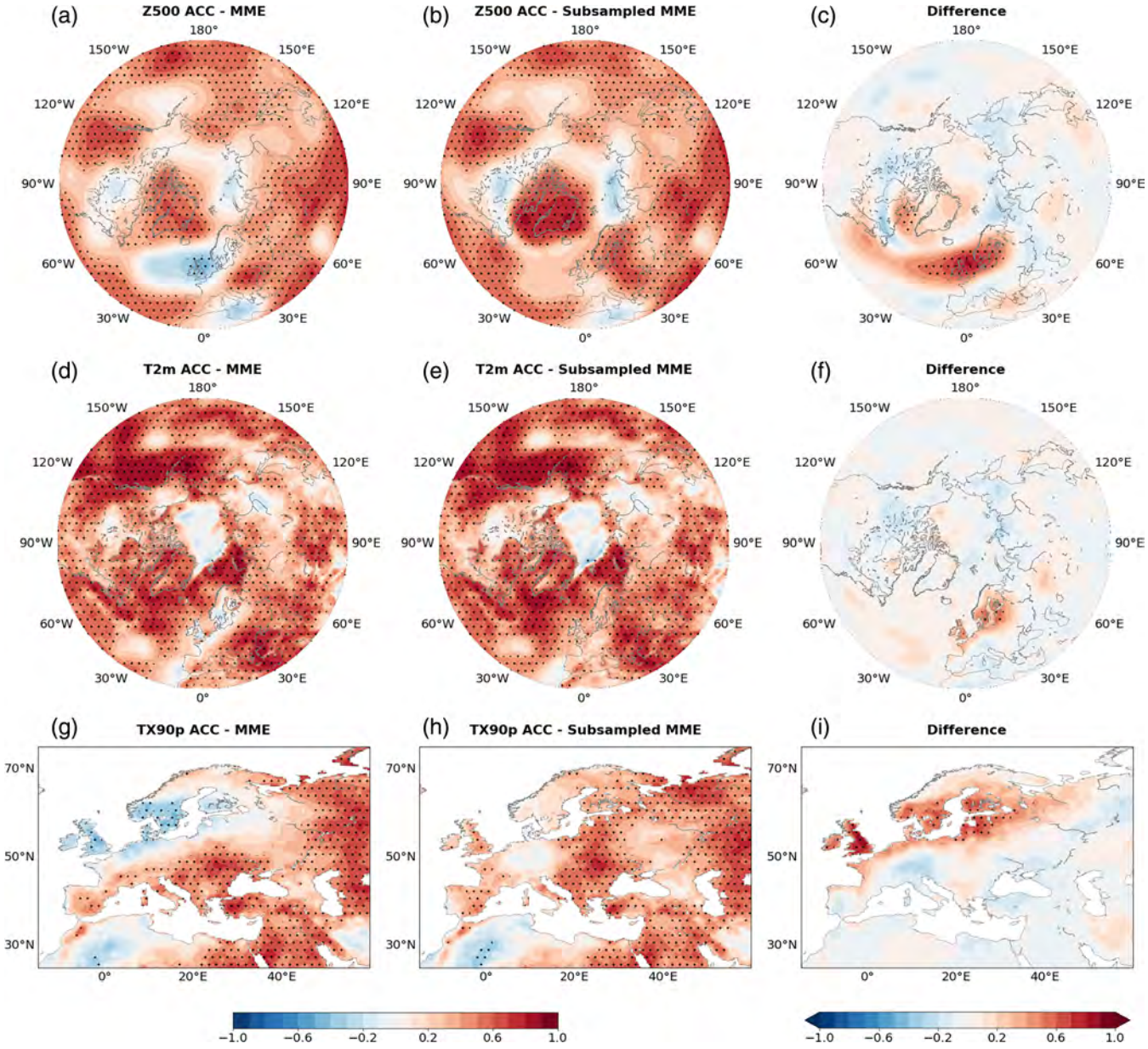

FIGURE 6 (a) Anomaly correlation coefficient (ACC) computed for the full Copernicus Climate Change Service (C3S) multimodel ensemble (MME) mean with respect to European Centre for Medium-Range Weather Forecasts Reanalysis v.5 for geopotential height at 500 hPa (Z500). (b) As (a) but for the subsampled C3S MME mean. The subsampling of the MME is performed using all the April NAO predictors and selecting 10 members for every predictor. (c) Difference between the ACC computed for the subsampled MME mean and the ACC computed for the full MME mean for Z500. Black dots denote ACC and differences that are statistically significant at the 90% confidence level. (d–f) As (a)–(c) but for temperature at 2 m height (T2m). (g–i) As (a)–(c) but for number of days when maximum temperature at T2m exceeds the 90th percentile of daily T2m during June–August (TX90p).

NAO significantly influences summer extreme temperature variability (Figure 1). In contrast, the subsampled MME does not capture the highest occurrence of hot days (positive TX90p anomalies) during summer 2006 in parts of France, Germany, Poland, and northern Italy. However, it is important to highlight that the summer extreme temperature variability in these regions is not strongly influenced by the NAO (Figure 1), so an improvement in TX90p prediction with subsampling is not expected there.

On the other hand, Figure 7 reveals that retaining only those ensemble members with a good representation of EA teleconnections does not generally improve the MME seasonal prediction skills for the summer European climate. In fact, even if the subsampling induces an increase of the ACCs for Z500 in the middle of the North Atlantic, it also deteriorates the ACCs over a large portion of eastern Europe (Figure 7c). This is associated with a reduction of the ACCs for T2m and TX90p in the same region, with the

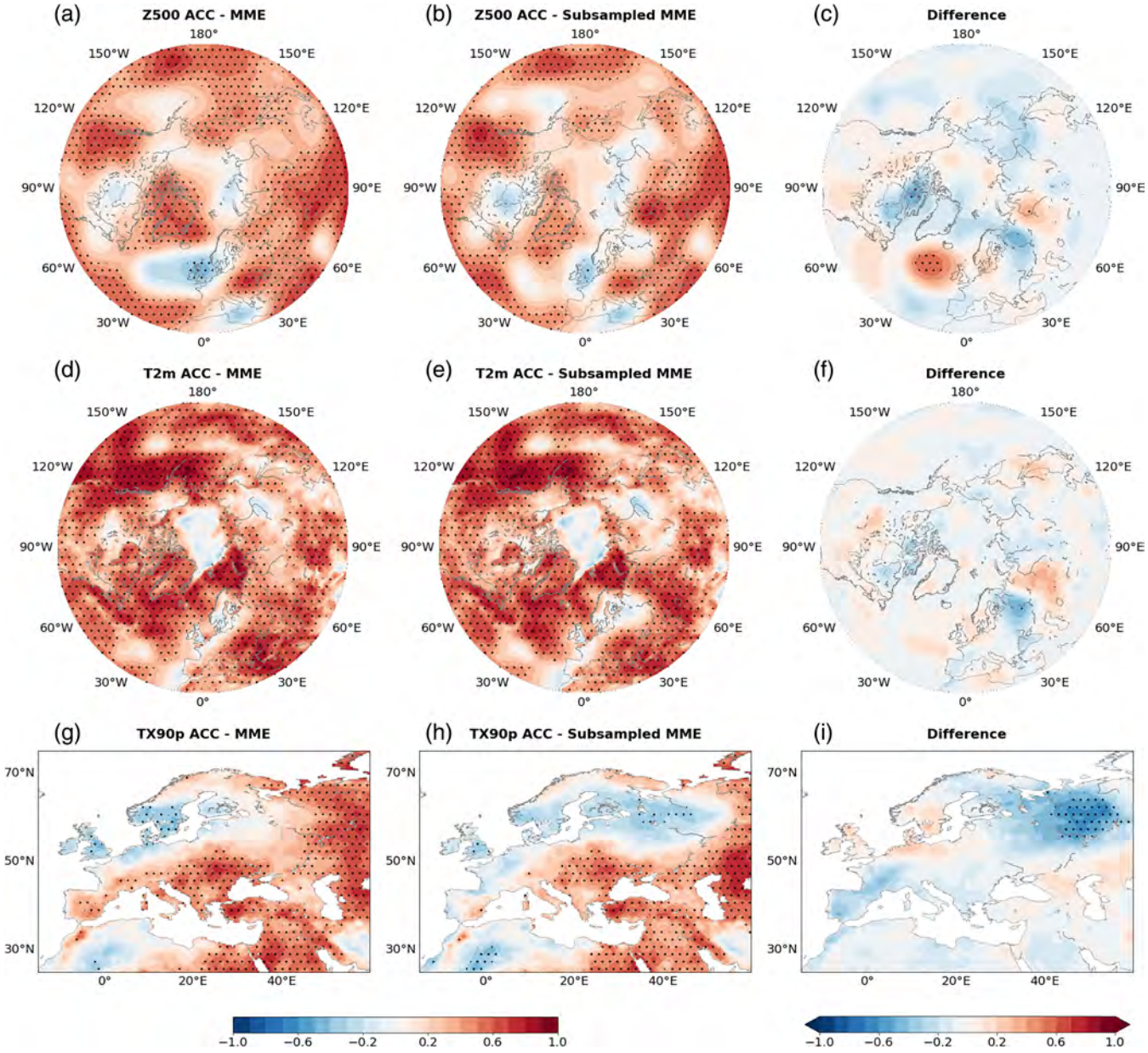

FIGURE 7 Same as Figure 6 but performing the subsampling of the Copernicus Climate Change Service multimodel ensemble (MME) using all the April East Atlantic predictors.

latter statistically significant at the 90% confidence level (Figure 7h,i). The ACCs' differences for T2m and TX90p are also negative in southern and Mediterranean Europe, even if not statistically significant. In line with the general decrease in the ACCs, Figure 7b,e,h reveals that the seasonal prediction skills of the subsampled MME are considerably lower and statistically significant within a more limited geographical area than with the full MME. For instance, the seasonal prediction skills of the subsampled MME for TX90p remain notably high and statistically significant in the Balkan region, but they drop to nearly zero in the rest of Europe. In this context, it is important to highlight that the regions experiencing the most significant reduction in ACCs, notably eastern Europe, coincide with the areas where the MME exhibits the least accurate representation of the EA pattern (Figure 2d). This means that the overall poor representation of the EA pattern within the MME limits the potential for skill improvement through the subsampling, particularly in those regions where the atmospheric circulation significantly influences T2m variability and, consequently, summer extreme temperatures (as shown in Figure 1c,d). In contrast, the skills remain high in those areas, such as the Balkan Peninsula, where the extreme temperatures are mainly determined by soil moisture preconditioning during the spring season and strong atmosphere–land coupling (Miralles *et al.*, 2014; Prodhomme *et al.*, 2016), whereas the atmospheric circulation plays a less relevant role.

Before concluding this section, we specify that the subsampling performed with predictors defined through leave-three-out cross-validation procedures gives similar seasonal prediction skills for Z500, T2m, and TX90p, as shown in Figures 6 and 7 (not shown). This is consistent with the fact that the hybrid statistical–dynamical predictions of the NAO and EA conducted with leave-one-out and leave-three-out procedures show comparable performance, as illustrated in Table 2.

4 | DISCUSSION

The results described in the previous section demonstrate that the skills of the C3S MME in predicting the NAO and EA phases on a seasonal time-scale can be largely enhanced through the teleconnection-based subsampling. This important aspect is summarised in Table 2, which presents the leave-one-out and leave-three-out cross-validated correlations and RMSE between the NAO and EA indices in ERA5 and the corresponding indices calculated as full and subsampled MME mean. The subsampled MME mean correlations are notably higher than the full MME mean correlations. At the same time, the RMSE for the subsampled MME is notably lower than the one computed for the full MME. However, Table 2 also shows that the skills achieved with the hybrid statistical–dynamical prediction systems are comparable to the ones obtained with the statistical prediction systems relying only on the April predictors of the NAO and EA. This could raise doubts about the actual utility of implementing a subsampling methodology when similar results can be achieved with much simpler statistical forecasts. Nevertheless, it should be pointed out that the hybrid statistical–dynamical approach provides much more information about the climate system. This is possible by generating subsampled ensemble means for any field of interest from the ensemble members with a reasonable representation of the teleconnections associated with the NAO and EA. In contrast, once the weather regime phase is predicted with a statistical forecast, another statistical model has to be defined to extract the link between the weather regime phase and any field of interest, which increases the overall complexity of the forecast process and may cause the loss of important dynamical information.

Regardless of whether a statistical or a hybrid statistical–dynamical system is used to predict the summer NAO and EA phases, the choice of appropriate April predictors represents a key stage in providing reliable seasonal predictions. In the present work, we adopt the North Atlantic SST, the Arctic SIC, the Northern Hemisphere SNOWC, and the zonally averaged zonal wind in the lower stratosphere as summer NAO predictors, and the SST in the North Atlantic and tropical Pacific and the Arctic SIC as summer EA predictors. Results in Supporting Information Figures S4 to S10 show that these variables correlate well with the summer NAO and EA phase. This is consistent with previous studies showing the role played by the spring state of the North Atlantic and tropical Pacific Ocean, Arctic sea-ice, Northern Hemisphere SNOWC, and stratospheric circulation on the summer North Atlantic atmospheric variability (e.g., Dunstone *et al.*, 2023; Gastineau & Frankignoul, 2015; Hall *et al.*, 2017; Wulff *et al.*, 2017). Nevertheless, further investigation is warranted to consolidate the predictors of the summer North Atlantic variability previously suggested in literature and to explore additional or more effective ones. This is especially true for the summer NAO, which has so far received limited attention from the scientific community. As an example, Dunstone *et al.* (2023) have recently shown that the variability of the polar vortex during the spring season is driven not only by the internal variability but also by sudden stratospheric warming events occurring in late winter. Indeed, the easterly stratospheric wind anomalies associated with these events can propagate in the troposphere in late March–early April, reducing the upward wave propagation and dynamical heating of the stratosphere. Consequently, the stratospheric cooling results in an anomalously strong polar vortex by the end of April, with westerly wind anomalies that propagate downward and induce positive NAO-like tropospheric circulation that can persist throughout the summer. Thus, the sudden stratospheric warming events that occur in late winter may represent windows of opportunity to further improve the prediction of the summer NAO and extend its range into the winter season. Research efforts in this direction are particularly relevant in order to extend the methodology described in this work from the retrospective forecast (hindcast) environment to the operational seasonal prediction systems.

Finally, it should be taken into account that whereas the NAO and EA pattern explain around 50% of the summer low-frequency summer (Figure 2), they are not the only large-scale atmospheric circulation patterns that influence extreme summer temperatures across Europe. Multiple studies have shown that European summer extreme temperatures may also be linked with large-scale wavelike structures propagating along the summer upper tropospheric polar and subtropical jet stream (e.g., Kornhuber *et al.*, 2019; Xu *et al.*, 2020; Xu *et al.*, 2021). In this context, as noted by previous articles (Neddermann *et al.*, 2019; Wulff *et al.*, 2017), it is worth mentioning that the EA pattern can be seen as a local manifestation of circumglobal wave-train patterns, with intense geopotential height anomalies also outside the North Atlantic sector. Therefore, incorporating the EA pattern into the teleconnection-based subsampling

approach described here partly accounts for the influence of large-scale wavelike structures previously identified as significant for European extreme temperatures. Nevertheless, integrating additional large-scale atmospheric circulation patterns besides the NAO and EA into the hybrid statistical–dynamical approach could potentially enhance the seasonal forecasting capabilities of the subsampled MME. Obviously, this approach would rely on the availability and understanding of seasonal predictors for these patterns.

5 | CONCLUSION

The results presented in this study show that the teleconnection-based subsampling technique significantly improves the seasonal prediction of the summer NAO and EA phase in the C3S SPS MME (Figure 3; Supporting Information Figures S2 and S3). Specifically, the correlations between the observed and subsampled MME mean NAO indices improve from −0.38 to 0.77, whereas for the EA they improve from −0.11 to 0.84. Furthermore, the RMSE for the NAO reduces from 1.06 in the full MME ensemble to 0.65 in the subsampled MME ensemble, whereas for the EA it reduces from 1.02 to 0.56. The enhancement of the seasonal prediction skills is robust over time, as evidenced by the narrow ranges of leave-one-out and leave-three-out cross-validated correlations and RMSE for both the NAO and EA (Table 2). It is important to highlight that the correlation improvement is rather insensitive to the number of members selected used for each April predictor ranging from 1 to about 100 (Figure 4). This is not completely true for the NAO and EA variability and RMSE. The variability decreases almost linearly with an increase of the ensemble size. The RMSE generally increases as the number of members selected increases when all April predictors are used, whereas it remains relatively stable when adopting individual April predictors for a number of members below approximately 50–80. In this context, using all the April NAO predictors or only the Northern Hemisphere SNOWC provide comparable improvements of the NAO prediction skills within the C3S MME (Figures 4a and 5a). In contrast, all the April EA predictors contribute to the improvement of the EA prediction skills within the C3S MME (Figures 4b and 5b).

The results also show that retaining only those ensemble members that accurately represent the NAO teleconnections enhances the MME seasonal prediction skills for the summer European climate (Z500 and T2m), including the occurrence of summer extreme temperatures as described by the TX90p index (Figure 6). This improvement is particularly pronounced in central and northern Europe, where both individual full C3S SPSs and MME exhibit limitations in predicting summer European climate and heat extremes.

Conversely, selecting only those ensemble members that accurately represent the EA teleconnections does not improve either the predictions of summer European climate or the predictions of summer extreme temperatures (Figure 7). These results can be explained by the fact that the C3S SPSs exhibits deficiencies in accurately representing the summer low-frequency atmospheric variability. In particular, the C3S SPSs exhibit biases in the summer EA pattern. Notably, the positive Z500 anomalies associated with this weather regime, which are typically found over central Europe, are displaced towards the northeastern region of the European continent (Figure 2). This aspect can lead to an inaccurate representation of the factors that drive summer extreme temperatures (such as air subsidence and advection of warm air from lower latitudes; Domeisen *et al.*, 2023) in the model environment, and thus lower skills in predicting such events, particularly in regions where the summer extreme temperatures are strongly influenced by EA variability (Figure 1).

In conclusion, the current study highlights the potential for enhancing seasonal predictions of summer European climate by employing a hybrid statistical–dynamical approach that involves the subsampling of dynamical prediction systems. In this context, the seasonal predictions of summer extreme temperatures are particularly improved over northern and central Europe, where the skills of state-of-the-art SPSs in predicting their occurrence have been shown to be particularly low. Future endeavours will focus on developing and adapting the methodology presented in this study to operational seasonal prediction environments, addressing the prediction of summer extreme temperatures and, eventually, other climate extremes across various seasons. Moreover, additional research should focus on understanding the sources of model biases in representing the patterns of summer low-frequency atmospheric variability, as these biases limit the benefits provided by postprocessing techniques. For instance, rectifying or minimising these biases could make it reasonable to employ the subsampling technique detailed in this study by selecting those ensemble members that better represent both the NAO and EA phase simultaneously.

ACKNOWLEDGEMENTS

We sincerely thank the anonymous reviewers for their careful reading of the manuscript and their valuable feedback, which has helped to improve the quality of our work. LFP, PR, SP, EB, and SDS were supported by Horizon Europe Project "Solution for mitigating climate-induced health threats"–TRIGGER (CUP J33C22002420005) (101057739).

DATA AVAILABILITY STATEMENT

ERA5 data are available from the Copernicus Climate Data Store (CDS) at https://cds.climate.copernicus.eu/#!/search?text=ERA5. The MME of seasonal prediction systems contributing to the C3S are available from the CDS at https://cds.climate.copernicus.eu/#!/search?text=seasonal%20prediction%20system.

ORCID

Luca Famooss Paolini https://orcid.org/0000-0002-0394-7713

REFERENCES

Albergel, C., Dutra, E., Bonan, B., Zheng, Y., Munier, S., Balsamo, G., de Rosnay, P., Muñoz-Sabater, J., & Calvet, J.-C. (2019) Monitoring and Forecasting the Impact of the 2018 Summer Heatwave on Vegetation. *Remote Sensing*, 11(5), 520. Available from: https://doi.org/10.3390/rs11050520

Alvarez-Castro, M.C., Faranda, D. & Yiou, P. (2018) Atmospheric Dynamics Leading to West European Summer Hot Temperatures Since 1851. *Complexity*, 2018, 2494509. Available from: https://doi.org/10.1155/2018/2494509

Baker, H.S., Woollings, T., Forest, C.E. & Allen, M.R. (2019) The linear sensitivity of the North Atlantic Oscillation and eddy-driven jet to SSTs. *Journal of Climate*, 32(19), 6491–6511.

Baldwin, M.P. & Dunkerton, T.J. (1999) Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research: Atmospheres*, 104(D24), 30937–30946. Available from: https://doi.org/10.1029/1999JD900445

Battè, L., L. Dorel, C. Ardilouze, and J.-F. Guérémy, 2021: Documentation of the METEO-FRANCE seasonal forecasting system 8. https://www.umr-cnrm.fr/IMG/pdf/system8-technical.pdf.

Beverley, J.D., Woolnough, S.J., Baker, L.H., Johnson, S.J. & Weisheimer, A. (2019) The northern hemisphere circumglobal teleconnection in a seasonal forecast model and its relationship to European summer forecast skill. *Climate Dynamics*, 52(5), 3759–3771. Available from: https://doi.org/10.1007/s00382-018-4371-4

Beverley, J.D., Woolnough, S.J., Baker, L.H., Johnson, S.J., Weisheimer, A. & O'Reilly, C.H. (2021) Dynamical mechanisms linking Indian monsoon precipitation and the circumglobal teleconnection. *Climate Dynamics*, 57(9), 2615–2636. Available from: https://doi.org/10.1007/s00382-021-05825-6

Black, E., Blackburn, M., Harrison, G., Hoskins, B. & Methven, J. (2004) Factors contributing to the summer 2003 European heatwave. *Weather*, 59(8), 217–223. Available from: https://doi.org/10.1256/wea.74.04

Black, E. & Sutton, R. (2007) The influence of oceanic conditions on the hot European summer of 2003. *Climate Dynamics*, 28(1), 53–66. Available from: https://doi.org/10.1007/s00382-006-0179-8

Bladé, I., Liebmann, B., Fortuny, D. & van Oldenborgh, G.J. (2012) Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region. *Climate dynamics*, 39, 709–727.

Boe, J., Terray, L., Moine, M.-P., Valcke, S., Bellucci, A., Drijfhout, S., Haarsma, R., Lohmann, K., Putrasahan, D.A., Roberts, C., Roberts, M., Scoccimarro, E., Seddon, J., Senan, R., & Wyser, K. (2020) Past long-term summer warming over western Europe in new generation climate models: role of large-scale atmospheric circulation. *Environmental Research Letters*, 15(8), 084038.

Brás, T.A., Seixas, J., Carvalhais, N. & Jägermeyr, J. (2021) Severity of drought and heatwave crop losses tripled over the last five decades in europe. *Environmental Research Letters*, 16(6), 065012. Available from: https://doi.org/10.1088/1748-9326/abf004

Cassou, C., Terray, L. & Phillips, A.S. (2005) Tropical Atlantic Influence on European Heat Waves. *Journal of Climate*, 18(15), 2805–2811. Available from: https://doi.org/10.1175/JCLI3506.1

Cheng, J., Cheng, J., Xu, Z., Bambrick, H., Prescott, V., Wang, N., Zhang, Y., Su, H., Tong, S., & Hu, W. (2019) Cardiorespiratory effects of heatwaves: A systematic review and meta-analysis of global epidemiological evidence. *Environmental research*, 177(108), 610.

Cohen, J., Screen, J.A., Furtado, J.C., Barlow, M., Whittleston, D., Coumou, D., Francis, J., Dethloff, K., Entekhabi, D., Overland, J. et al. (2014) Recent Arctic amplification and extreme mid-latitude weather. *Nature Geoscience*, 7(9), 627–637. Available from: https://doi.org/10.1038/ngeo2234

Czaja, A. & Frankignoul, C. (2002) Observed Impact of Atlantic SST Anomalies on the North Atlantic Oscillation. *Journal of Climate*, 15(6), 606–623. Available from: https://doi.org/10.1175/1520-0442(2002)015¡0606:OIOASA¿2.0.CO;2

DelSole, T. & Tippett, M. (2022) *Statistical methods for climate scientists*. Cambridge University Press.

Di Capua, G., Sparrow, S., Kornhuber, K., Rousi, E., Osprey, S., Wallom, D. et al. (2021) Drivers behind the summer 2010 wave train leading to russian heatwave and pakistan flooding. *npj Climate and Atmospheric Science*, 4(1), 55.

Dobrynin, M., Düsterhus, A., Fröhlich, K., Athanasiadis, P., Ruggieri, P., Müller, W.A. et al. (2022) Hidden Potential in Predicting Wintertime Temperature Anomalies in the Northern Hemisphere. *Geophysical Research Letters*, 49(20), e2021GL095063. Available from: https://doi.org/10.1029/2021GL095063

Dobrynin, M., Domeisen, D.I.V., Müller, W.A., Bell, L., Brune, S., Bunzel, F. et al. (2018) Improved Teleconnection-Based Dynamical Seasonal Predictions of Boreal Winter. *Geophysical Research Letters*, 45(8), 3605–3614. Available from: https://doi.org/10.1002/2018GL077209

Domeisen, D.I.V., Eltahir, E.A.B., Fischer, E.M., Knutti, R., Perkins-Kirkpatrick, S.E., Schär, C., Seneviratne, S.I., Weisheimer, A., & Wernli, H. (2023) Prediction and projection of heatwaves. *Nature Reviews Earth & Environment*, 4(1), 36–50. Available from: https://doi.org/10.1038/s43017-022-00371-z

Dong, B., Sutton, R.T., Woollings, T. & Hodges, K. (2013) Variability of the North Atlantic summer storm track: mechanisms and impacts on European climate. *Environmental Research Letters*, 8(3), 034037.

Drouard, M., Kornhuber, K. & Woollings, T. (2019) Disentangling dynamic contributions to summer 2018 anomalous weather over Europe. *Geophysical Research Letters*, 46(21), 12537–12546.

Duchez, A., Frajka-Williams, E., Josey, S.A., Evans, D.G., Grist, J.P., Marsh, R., McCarthy, G.D., Sinha, B., Berry, D.I., & Hirschi, J.J.M. (2016) Drivers of exceptionally cold North Atlantic Ocean temperatures and their link to the 2015 European heat wave. *Environmental Research Letters*, 11(7), 074004. Available from: https://doi.org/10.1088/1748-9326/11/7/074004

Dunstone, N., Smith, D.M., Hardiman, S.C., Hermanson, L., Ineson, S., Kay, G., Li, C., Lockwood, J.F., Scaife, A.A., Thornton, H. et al. (2023) Skilful predictions of the Summer North Atlantic Oscillation. *Communications Earth & Environment*, 4(1), 409.

Düsterhus, A. (2020) Seasonal statistical–dynamical prediction of the North Atlantic Oscillation by probabilistic post-processing and its evaluation. *Nonlinear Processes in Geophysics*, 27(1), 121–131. Available from: https://doi.org/10.5194/npg-27-121-2020

Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L. et al. (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophysical Research Letters*, 41(15), 5620–5628. Available from: https://doi.org/10.1002/2014GL061146

Ebi, K.L., Capon, A., Berry, P., Broderick, C., de Dear, R., Havenith, G., Honda, Y., Kovats, R.S., Ma, W., Malik, A. et al. (2021) Hot weather and heat extremes: health risks. *The Lancet*, 398(10301), 698–708. Available from: https://doi.org/10.1016/S0140-6736(21)01208-3

Fisher, R. (1920) On the "probable error" of a coefficient of correlation. *Metron*, 1, 3–32.

Folland, C.K., Knight, J., Linderholm, H.W., Fereday, D., Ineson, S. & Hurrell, J.W. (2009) The Summer North Atlantic Oscillation: Past, Present, and Future. *Journal of Climate*, 22(5), 1082–1103. Available from: https://doi.org/10.1175/2008JCLI2459.1

Fröhlich, K., Dobrynin, M., Isensee, K., Gessner, C., Paxian, A., Pohlmann, H., Haak, H., Brune, S., Früh, B. & Baehr, J. (2021) The German climate forecast system: GCFS. *Journal of Advances in Modeling Earth Systems*, 13(2), e2020MS002 101.

Gastineau, G., D'Andrea, F. & Frankignoul, C. (2013) Atmospheric response to the north atlantic ocean variability on seasonal to decadal time scales. *Climate dynamics*, 40, 2311–2330.

Gastineau, G. & Frankignoul, C. (2015) Influence of the North Atlantic SST Variability on the Atmospheric Circulation during the Twentieth Century. *Journal of Climate*, 28(4), 1396–1416. Available from: https://doi.org/10.1175/JCLI-D-14-00424.1

Gastineau, G., García-Serrano, J. & Frankignoul, C. (2017) The Influence of Autumnal Eurasian Snow Cover on Climate and Its Link with Arctic Sea Ice Cover. *Journal of Climate*, 30(19), 7599–7619. Available from: https://doi.org/10.1175/JCLI-D-16-0623.1

Gray, L.J., Scaife, A.A., Mitchell, D.M., Osprey, S., Ineson, S., Hardiman, S., Butchart, N., Knight, J., Sutton, R. & Kodera, K. (2013) A lagged response to the 11 year solar cycle in observed winter atlantic/european weather patterns. *Journal of Geophysical Research: Atmospheres*, 118(24), 13–405.

Gualdi, S., Borrelli, A., Cantelli, A., Davoli, G., del Mar Chavesmontero, M., Masina, S., Navarra, A., Sanna, A. & Tibaldi, S. (2020) The new CMCC Operational Seasonal Prediction System. https://doi.org/10.25424/CMCC/SPS3.5 CMCC Technical Note RP0288.

Hall, R.J., Jones, J.M., Hanna, E., Scaife, A.A. & Erdélyi, R. (2017) Drivers and potential predictability of summer time North Atlantic polar front jet variability. *Climate Dynamics*, 48, 3869–3887.

Hamilton, E., Eade, R., Graham, R.J., Scaife, A.A., Smith, D.M., Maidens, A. et al. (2012) Forecasting the number of extreme daily events on seasonal timescales. *Journal of Geophysical Research: Atmospheres*, 117(D3), D03114. Available from: https://doi.org/10.1029/2011JD016541

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D. et al. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. Available from: https://doi.org/10.1002/qj.3803

Hinkle, D.E., Wiersma, W. & Jurs, S.G. (2003) *Applied Statistics for the Behavioral Sciences*. Boston, USA: Houghton Mifflin.

Horton, D.E., Johnson, N.C., Singh, D., Swain, D.L., Rajaratnam, B. & Diffenbaugh, N.S. (2015) Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature*, 522(7557), 465–469.

Ineson, S. & Scaife, A.A. (2009) The role of the stratosphere in the European climate response to El Niño. *Nature Geoscience*, 2(1), 32–36. Available from: https://doi.org/10.1038/ngeo381

Johnson, S.J., Stockdale, T.N., Ferranti, L., Balmaseda, M.A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G. et al. (2019) SEAS5: the new ECMWF seasonal forecast system. *Geoscientific Model Development*, 12(3), 1087–1117. Available from: https://doi.org/10.5194/gmd-12-1087-2019

Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S. et al. (2019) Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environmental Research Letters*, 14(5), 054002.

Lhotka, O. & Kyselỳ, J. (2015) Hot Central-European summer of 2013 in a long-term context. *International Journal of Climatology*, 35(14), 4399–4407.

Lin, H., Merryfield, W.J., Muncaster, R., Smith, G.C., Markovic, M., Dupont, F., Roy, F., Lemieux, J.-F., Dirkson, A., Kharin, V.V. et al. (2021) The Canadian Seasonal to Interannual Prediction System version 2.1 (CanSIPSv2.1). https://collaboration.cmc.ec.gc.ca/cmc/cmoi/product·guide/docs/tech·notes/technote·cansips-210·e.pdf

Luo, M. & Lau, N.-C. (2020) Summer heat extremes in northern continents linked to developing ENSO events. *Environmental Research Letters*, 15(7), 074042. Available from: https://doi.org/10.1088/1748-9326/ab7d07

Martija-Díez, M., Rodríguez-Fonseca, B. & López-Parages, J. (2021) ENSO Influence on Western European Summer and Fall Temperatures. *Journal of Climate*, 34(19), 8013–8031. Available from: https://doi.org/10.1175/JCLI-D-20-0808.1

Matsumura, S., Zhang, X. & Yamazaki, K. (2014) Summer Arctic atmospheric circulation response to spring Eurasian snow cover and its possible linkage to accelerated sea ice decrease. *Journal of Climate*, 27(17), 6551–6558.

Mecking, J.V., Drijfhout, S.S., Hirschi, J.J.-M. & Blaker, A.T. (2019) Ocean and atmosphere influence on the 2015 European heatwave. *Environmental Research Letters*, 14(11), 114035. Available from: https://doi.org/10.1088/1748-9326/ab4d33

Minobe, S., Kuwano-Yoshida, A., Komori, N., Xie, S.-P. & Small, R.J. (2008) Influence of the gulf stream on the troposphere. *Nature*, 452(7184), 206–209.

Miralles, D.G., Teuling, A.J., van Heerwaarden, C.C. & Vilà-Guerau de Arellano, J. (2014) Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature Geoscience*, 7(5), 345–349. Available from: https://doi.org/10.1038/ngeo2141

Mishra, N., Prodhomme, C. & Guemas, V. (2019) Multi-model skill assessment of seasonal temperature and precipitation forecasts over Europe. *Climate Dynamics*, 52(7), 4207–4225. Available from: https://doi.org/10.1007/s00382-018-4404-z

Moberg, A., Jones, P.D., Lister, D., Walther, A., Brunet, M., Jacobeit, J., Alexander, L.V., Della-Marta, P.M., Luterbacher, J., Yiou, P.

et al. (2006) Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000. *Journal of Geophysical Research: Atmospheres*, 111(D22), D22106.

Neddermann, N.-C., Müller, W.A., Dobrynin, M., Düsterhus, A. & Baehr, J. (2019) Seasonal predictability of European summer climate re-assessed. *Climate Dynamics*, 53, 3039–3056.

Prodhomme, C., Doblas-Reyes, F., Bellprat, O. & Dutra, E. (2016) Impact of land-surface initialization on sub-seasonal to seasonal forecasts over Europe. *Climate Dynamics*, 47(3), 919–935. Available from: https://doi.org/10.1007/s00382-015-2879-4

Prodhomme, C., Materia, S., Ardilouze, C., White, R.H., Batté, L., Guemas, V. et al. (2022) Seasonal prediction of European summer heatwaves. *Climate Dynamics*, 58(7), 2149–2166. Available from: https://doi.org/10.1007/s00382-021-05828-3

Quesada, B., Vautard, R., Yiou, P., Hirschi, M. & Seneviratne, S.I. (2012) Asymmetric European summer heat predictability from wet and dry southern winters and springs. *Nature Climate Change*, 2(10), 736–741.

Rousi, E., Kornhuber, K., Beobide-Arsuaga, G., Luo, F. & Coumou, D. (2022) Accelerated western European heatwave trends linked to more-persistent double jets over Eurasia. *Nature communications*, 13(1), 3851.

Russo, S., Dosio, A., Graversen, R.G., Sillmann, J., Carrao, H., Dunbar, M.B., Singleton, A., Montagna, P., Barbola, P. & Vogt, J.V. (2014) Magnitude of extreme heat waves in present climate and their projection in a warming world. *Journal of Geophysical Research: Atmospheres*, 119(22), 12–500.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., van den Dool, H., Zhang, Q., Wang, W., Chen, M. & Becker, E. (2014) The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208. Available from: https://doi.org/10.1175/JCLI-D-12-00823.1

Scaife, A.A. & Smith, D. (2018) A signal-to-noise paradox in climate science. npj Climate and Atmospheric. *Science*, 1 (1), 28. Available from: https://doi.org/10.1038/s41612-018-0038-4

Screen, J.A. (2013) Influence of Arctic sea ice on European summer precipitation. *Environmental Research Letters*, 8(4), 044015. Available from: https://doi.org/10.1088/1748-9326/8/4/044015

Seneviratne, S., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S. et al. (2021) *Weather and Climate Extreme Events in a Changing Climate, chap*, Vol. 11. Cambridge University Press, pp. 1513–1766. Available from: https://doi.org/10.1017/9781009157896.013

Seneviratne, S.I., Corti, T., Davin, E.L., Hirschi, M., Jaeger, E.B., Lehner, I. et al. (2010) Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3-4), 125–161.

Smith, D.M., Scaife, A.A., Eade, R., Athanasiadis, P., Bellucci, A., Bethke, I., Bilbao, R., Borchert, L.F., Caron, L.-P., Counillon, F. et al. (2020) North Atlantic climate far more predictable than models imply. *Nature*, 583(7818), 796–800. Available from: https://doi.org/10.1038/s41586-020-2525-0

Sulikowska, A. & Wypych, A. (2020) Summer temperature extremes in Europe: how does the definition affect the results? *Theoretical and Applied Climatology*, 141, 19–30.

Swingedouw, D., Mignot, J., Ortega, P., Khodri, M., Menegoz, M., Cassou, C. et al. (2017) Impact of explosive volcanic eruptions on the main climate variability modes. *Global and Planetary Change*, 150, 24–45.

Thompson, R., Hornigold, R., Page, L. & Waite, T. (2018) Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. *Public health*, 161, 171–191.

Vautard, R., Cattiaux, J., Happé, T., Singh, J., Bonnet, Rémy, Cassou, C., Coumou, D., D'andrea, F., Faranda, D., Fischer, E. et al. (2023) Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends. *Nature Communications*, 14(1), 6803. Available from: https://doi.org/10.1038/s41467-023-42143-3

Wang, L. & Ting, M. (2022) Stratosphere-Troposphere Coupling Leading to Extended Seasonal Predictability of Summer North Atlantic Oscillation and Boreal Climate. *Geophysical Research Letters*, 49(2), e2021GL096 362.

Williams, K.D., Copsey, D., Blockley, E.W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H.T., Hill, R. et al. (2018) The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, 10(2), 357–380. Available from: 10.1002/2017MS001115

Wong, C. (2023) Extreme heat harms health-what is the human body's limit? *Nature*. Available from: https://doi.org/10.1038/d41586-023-02482-z

Wu, B., Zhang, R., D'Arrigo, R. & Su, J. (2013) On the Relationship between Winter Sea Ice and Summer Atmospheric Circulation over Eurasia. *Journal of Climate*, 26(15), 5523–5536.

Wulff, C.O., Greatbatch, R.J., Domeisen, D.I.V., Gollan, G. & Hansen, F. (2017) Tropical Forcing of the Summer East Atlantic Pattern. *Geophysical Research Letters*, 44(21), 11166–11173. Available from: https://doi.org/10.1002/2017GL075493

Xu, P., Wang, L., Liu, Y., Chen, W. & Huang, P. (2020) The record-breaking heat wave of June 2019 in Central Europe. *Atmospheric Science Letters*, 21(4), e964.

Xu, P., Wang, L., Vallis, G.K., Geen, R., Screen, J.A., Wu, P., Ding, S., Huang, P. & Chen, W. (2021) Amplified waveguide teleconnections along the polar front jet favor summer temperature extremes over northern Eurasia. *Geophysical Research Letters*, 48(13), e2021GL093 735.

Zhao, P., Zhang, X., Zhou, X., Ikeda, M. & Yin, Y. (2004) The sea ice extent anomaly in the North Pacific and its impact on the East Asian summer monsoon rainfall. *Journal of Climate*, 17(17), 3434–3447.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Famooss Paolini, L., Ruggieri, P., Pascale, S., Brattich, E. & Di Sabatino, S. (2024) Hybrid statistical–dynamical seasonal prediction of summer extreme temperatures in Europe. *Quarterly Journal of the Royal Meteorological Society*, 1–20. Available from: https://doi.org/10.1002/qj.4900