



# Seasonal predictability of European summer climate re-assessed

Nele-Charlotte Neddermann<sup>1,2</sup> · Wolfgang A. Müller<sup>3</sup> · Mikhail Dobrynin<sup>1</sup> · André Düsterhus<sup>1</sup> · Johanna Baehr<sup>1</sup>

Received: 23 July 2018 / Accepted: 9 February 2019 / Published online: 23 February 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

We improve seasonal hindcast skill of European summer climate in an ensemble based coupled seasonal prediction system by selecting individual ensemble members based on their respective consistent chain of processes that describe a physical mechanism. This mechanism is associated with the second mode of seasonal climate variability in the North-Atlantic-European sector and is contrary to the summer North Atlantic Oscillation. We initially analyse the mechanism in the ERA-Interim reanalysis and then test the influence of the mechanism on European hindcast skill in an initialised coupled seasonal climate model. We show that the mechanism originates in the tropical North Atlantic in spring, where either warm or cold sea surface temperature anomalies (SSTs) are connected with the European climate by an upper-level wave-train. This wave-train is accompanied by a zonal pressure gradient, that in turn influences the climate over central Europe in the following summer. We analyse the seasonal summer hindcast skill in a mixed resolution hindcast ensemble simulation generated by MPI-ESM, with 30 members starting every year in May. While the mean over the full ensemble shows no seasonal hindcast skill in summer, we achieve significant hindcast skill through forming a new mean over subselected ensemble members. For this selection, we test every ensemble member for the proposed consistent chain of connections between the wave-train, the zonal pressure gradient and their impact on European summer temperatures, and find that the processes that describe the mechanism are not represented in every ensemble member. Due to its influence on European summer climate, we use the condition of the persistent spring SSTs to anticipate the phase of the mechanism in each considered year. We thus use statistical relations to select ensemble members generated by a dynamical prediction system. With this approach, we significantly enhance the seasonal hindcast skill and the reliability of the hindcasts in the North-Atlantic-European sector, especially in the areas where the mechanism is showing a prominent signal. Since we only use knowledge that would be available in a real forecast set-up, this approach can potentially be applied in operational ensemble prediction systems.

## 1 Introduction

Current state-of-the-art prediction systems show seasonal predictability in various areas, including large parts of the North Atlantic, but their prediction skill for European climate is still very limited, particularly during the summer season (e.g., Arribas et al. 2011; Baehr et al. 2015). Seasonal predictions often lack an understanding of the physical processes (Doblas-Reyes et al. 2013) and while

seasonal European winter climate is mainly dependent on the phase of the North Atlantic Oscillation (NAO) (Hurrell 1996), various mechanisms are influencing the summer climate in Europe on seasonal time scales (e.g., Cassou et al. 2005). As shown in previous works (Domeisen et al. 2015; Dobrynin et al. 2018), improved seasonal hindcast skill can be achieved if driving mechanisms are included into the prediction through the selection of ensemble members via physical criteria. Here, we improve seasonal summer hindcast skill in an ensemble based seasonal prediction system, by selecting individual ensemble members for a mechanism that connects areas of high predictability in the tropical North Atlantic with the summer climate over Europe.

In the North-Atlantic-European sector, the tropical North Atlantic is a major source of low-frequency climate variability and has in turn a strong influence on seasonal variability in the tropics and mid-latitudes (Marshall et al. 2001). The high persistence of tropical sea surface temperatures (SSTs)

✉ Nele-Charlotte Neddermann  
nele.neddermann@uni-hamburg.de

<sup>1</sup> Institute for Oceanography, CEN, Universität Hamburg, Hamburg, Germany

<sup>2</sup> International Max Planck Research School on Earth System Modelling, Max Planck Institute for Meteorology, Hamburg, Germany

<sup>3</sup> Max Planck Institute for Meteorology, Hamburg, Germany

during spring and summer leads to high seasonal summer prediction skill in the tropical region. Due to the strong influence of tropical forcing on seasonal mid-latitude climate variability, seasonal predictability in the mid-latitudes then often originates from the seasonal predictability and persistence of tropical regions (e.g., Palmer and Anderson 1994).

In summer, warm SSTs in the tropical North Atlantic lead to extra convective heating in the atmosphere, which results in strong upper troposphere divergence in the tropics and convergence in the subtropics, acting as a Rossby wave source (e.g., Bjercknes 1966; Hoskins and Ambrizzi 1993; Gastineau and Frankignoul 2015). For such low-latitude sources, the resulting waves usually propagate polewards and eastwards (Hoskins and Karoly 1981). Such an evolving Rossby wave is known as the circumglobal teleconnection pattern (CGT, Branstator 2002), which is the leading mode of inter-annual variability of upper-tropospheric circulation and is associated with the subtropical jet stream wave guide (Ding and Wang 2005). The CGT traps signals along its path and, while linking regional climate variations, has a strong influence on the local climate systems over Europe and Asia. Branstator and Teng (2017) reveal a strong seasonality of the CGT and show that, opposed to its winter counterpart, the summertime CGT is not circumglobal, but rather splits into two separate waves. This is in accordance with the results of Saeed et al. (2014), who found a wavelike pattern similar to the CGT, but narrowed its domain to the North Atlantic and Eurasia. They suggest that SSTs in the Gulf of Mexico are the source of the CGT and show that the Eurasian CGT is related to a wavelike zonal pressure pattern over Europe.

This zonal pressure gradient is related to the second mode of summertime low-frequency variability in the North-Atlantic-European sector, which is known as the East Atlantic pattern (Wallace and Gutzler 1981; Barnston and Livezey 1987; Iglesias et al. 2014), the Atlantic Low (Cassou et al. 2005) or the summer East Atlantic mode (SEA, Wulff et al. 2017). For a positive (negative) phase of the CGT, the zonal pressure gradient is associated with anomalous high (low) pressure over the subtropical North Atlantic and low (high) pressure over central Europe. Over central Europe, a positive phase of the zonal pressure gradient is associated with low temperatures and enhanced precipitation (Wulff et al. 2017), while in its negative phase it is connected to European heat waves (Cassou et al. 2005; Duchez et al. 2016).

Wulff et al. (2017) confirm the connection between the extra-tropical Rossby wave-train and the zonal pressure gradient and suggest its source in the SST anomalies in the Caribbean and in the tropical North Pacific by showing that the tropical SSTs are leading the extra-tropical patterns by a few months. This temporal lag arises from the persistence of the SSTs in the tropical regions. In accordance with Wulff et al. (2017), we assume that a signal that originates in the tropical

North Atlantic in spring stays there until summer, which is why we focus on SSTs in spring, rather than in summer.

The summer North Atlantic Oscillation (SNAO, Folland et al. 2009), which is the leading mode of North-Atlantic-European atmospheric variability in summer, is not related to the zonal pressure gradient, since it has a more meridional pressure pattern and an influence on distinct different regions (e.g., Saeed et al. 2014). Li and Ruan (2018) find that the SNAO is also connected to a Rossby wave, but show that its pathway is clearly distinguishable from that of the CGT.

Several studies show potential predictability of the aforementioned zonal wind or pressure pattern in either prescribed or free model runs. Yasui and Watanabe (2010) find potential predictability of the CGT in a model run that is driven by prescribed diabatic heating, while Wulff et al. (2017) find that seasonal hindcast experiments forced with SSTs show skill in capturing the zonal pressure gradient. Saeed et al. (2014) test historical simulations of a global coupled climate model and show that the proposed wind and pressure patterns can be reproduced by a coupled climate model.

However, while parts of the mechanism have been captured by free and forced model runs, the tropical–extratropical teleconnection between spring SSTs in the tropical North Atlantic and summer wind, pressure and temperature over Europe have not yet been tested in a global coupled ensemble based seasonal prediction system. Since the mechanism connects areas of high prediction skill with seasonal climate over central Europe, it could lead to enhanced prediction skill of European summers if incorporated into a robust seasonal prediction system. Here, we test such a prediction system for the proposed connection and further also include this mechanism into the hindcast analysis with the aim to improve the seasonal hindcast skill of summers over central Europe.

As described, various mechanisms are influencing the summer climate in the North-Atlantic-European sector on seasonal time scales. Therefore, individual ensemble members in dynamical seasonal prediction systems are dominated by different physical processes. Hindcast analysis are usually conducted with the mean taken over all generated ensemble members, such that a mean is taken over the signal of various different physical mechanisms and the signal of individual mechanisms are then often averaged out. This usually results in an amplitude of the ensemble mean that is much lower than the amplitude of the observations, as for instance shown by Baker et al. (2018) for the NAO. Here, the goal is to amplify the signal of the proposed zonal mechanism by selecting only those ensemble members in which the described chain of physical processes associated with the proposed mechanism is represented.

Approaches in which a mechanism is incorporated into the prediction have already been tested by rejecting or

retaining individual ensemble members via physical or statistical criteria. Domeisen et al. (2015) select only ensemble members that contain sudden stratospheric warming events and further also years in which the El Niño–Southern Oscillation happened, and in turn improve seasonal winter hindcast skill over Europe with this approach. Dobrynin et al. (2018) base their ensemble selection on known physical links of the winter North Atlantic Oscillation (WNAO) with the autumn states of the ocean, sea ice, land surface and stratosphere and Düsterhus et al. (2017) use a similar methodology, but for the SNAO. This procedure results in enhanced hindcast skill in regions where the NAO has a strong influence on European climate, which is on northern and southern, but not on central Europe (Hurrell 1995).

Thus, we consider this different pattern of summer climate variability, which has an impact on central Europe and has not yet been tested in an approach that includes a mechanism into a seasonal prediction. In particular, we will include the zonal pressure gradient and its physical links into the prediction analysis by subsampling of ensemble members, which are generated by the initialised global seasonal prediction system MPI-ESM at mixed resolution and compare the results to ERA-Interim reanalysis data (Dee et al. 2011). While the studies by Domeisen et al. (2015) and Dobrynin et al. (2018) base their ensemble selection on individual initial conditions of the ocean or atmosphere, our selection is based on a chain of physical processes that are part of the analysed mechanism.

A brief description of the used data can be found in Sect. 2. Since a reliable representation of the mechanism is crucial for the applied method, the mechanism is first analysed in the ERA-Interim reanalysis data in Sect. 3, followed by a description of the subsampling method in Sect. 4. The effect of ensemble subsampling on seasonal hindcast skill is then accessed by the analysis of anomaly correlation and reliability diagrams in Sect. 5. Section 6 provides the discussion, followed by the summary and conclusions in Sect. 7.

## 2 Model and data

### 2.1 Reanalysis data

The analysis of the mechanism is carried out with the Interim European Centre for Medium-Range Weather Forecast (ECMWF) reanalysis (ERA-Interim, Dee et al. 2011) monthly-mean fields from 1982 to 2016. The results are used as a basis for and in comparison to the model output.

The primary analysed datasets are the monthly-mean skin temperature, sea level pressure, 500 hPa geopotential height and 200 hPa meridional wind. To eliminate long-term trends, we linearly detrend all fields. Only monthly data with respect to the climatological mean are considered.

In accordance with Folland et al. (2009), who show that the seasonal variability of summer climate in the North-Atlantic-European sector in June deviates from that in July and August, this study defines “summer” as the July–August (JA) mean. Nevertheless, the analysed patterns are similar in June–July–August (JJA), but the link between the individual parts of the mechanism is less pronounced. Consequently, we would have to adjust for these changes for an analyses in JJA.

### 2.2 Model setup

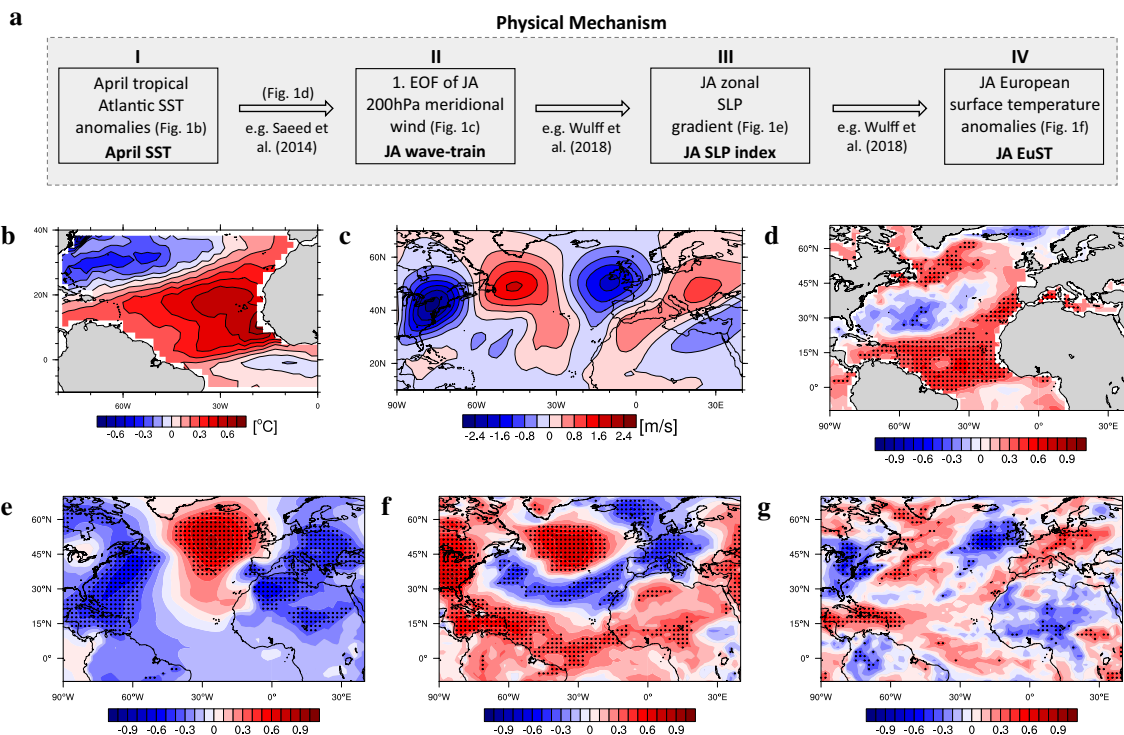
We use the dynamical seasonal prediction system (Dobrynin et al. 2018) based on the global Max Planck Institute Earth System Model at mixed resolution (MPI-ESM-MR) in the version as used for the CMIP5 simulations (Giorgetta et al. 2013). The model consists of the atmospheric component ECHAM6 (Stevens et al. 2013) with 200 km (1.875°) horizontal resolution and 95 vertical levels up to 0.01 hPa, coupled to the ocean component MPI-OM (Jungclaus et al. 2006) with a horizontal resolution of 40 km (0.4°) and 40 vertical layers.

To initialise the model, full-field nudging is used as an assimilation technique. The nudging is performed by Newtonian relaxation towards reanalysis data. In the atmosphere, vorticity, divergence, temperature and surface pressure are nudged towards ERA-Interim with a relaxation time scale of one day. In the ocean, the ECMWF ocean reanalysis system (ORAS4, Balmaseda et al. 2013) is used for nudging of temperature and salinity with a relaxation time scale of 10 days. Sea ice concentration is nudged towards the observational National Snow and Ice Data Center (NSIDC) sea ice concentration data (Fetterer et al. 2002) with an effective relaxation time of 20 days (Tietsche et al. 2013).

From the assimilation experiments, 30 ensemble members are initialised with slightly different initial conditions on the first of May each year from 1982 to 2016 (35 years). In the ocean, each ensemble member is perturbed using bred vectors with a vertically varying norm (Baehr and Piontek 2014). In the atmosphere, the diffusion coefficient in the uppermost layer is slightly disturbed.

### 2.3 Analysis

To identify the principle mode of variations in single fields, we calculate the empirical orthogonal functions (EOF) of the spatial variations from the fields by using their anomaly covariance matrix (North et al. 1982). In case coupled modes of variations between two fields are considered, we evaluate them with the singular value decomposition (SVD) of the covariance matrix of the two analysed fields (Bretherton et al. 1992). To derive sign definite regime patterns, we



**Fig. 1** a Schematic representation of the proposed physical mechanism analysed in the ERA-Interim reanalysis in 1982–2016. (I) The signal of the mechanism starts in the tropical North Atlantic in spring, where SST anomalies are the source of strong convection, depicted by **b** the first EOF of SST anomalies in April in the region [10°S–15°N, 80°W–20°W], which explains 41.2% of the total SST variance. (II) The strong convection induces a summertime wave-train, specified by **c** the first EOF of 200 hPa meridional wind (Vwind) in JA in the region [10°N–70°N, 90°W–40°E], explaining 22.4% of the

total variance. **d** The principal component associated with this first EOF of 200 hPa meridional wind in JA (PC1 Vwind), pointwise correlated with SSTs in April confirms this relation. (III) The wave-train is accompanied by a zonal SLP gradient, indicated by **e** the pointwise correlation between PC1 Vwind and the SLP in JA. (IV) This in turn has an influence on the summer climate over Europe as shown by the pointwise correlation of PC1 Vwind with **f** surface temperatures in JA and **g** total precipitation (TP) in JA. Dots represent significance at the 95% confidence level

further use the *k*-means cluster algorithm (Michelangeli et al. 1995) on the JA mean for the 35 analyzed summers.

Covariability between time series and a field is further derived through point-wise correlation. Significance of point-wise correlation is calculated via bootstrapping at the 95% confidence level using 500 samples.

The hindcast skill of the model output against the ERA-Interim data is assessed with the point-wise detrended anomaly correlation coefficient (ACC, Collins 2002). To account for the uncertainty of the ACC in the temporal dimension, we apply cross-validation by leaving out one year in the analysed period 1982–2016. The ACC is calculated for every cross-validated iteration between the reforecasted and reanalysed fields and shown as the mean over all cross validated iterations. Significance is derived for every iteration via bootstrapping at the 95% confidence level using 500 samples and depicted only for those regions that shown significance in every cross-validated iteration.

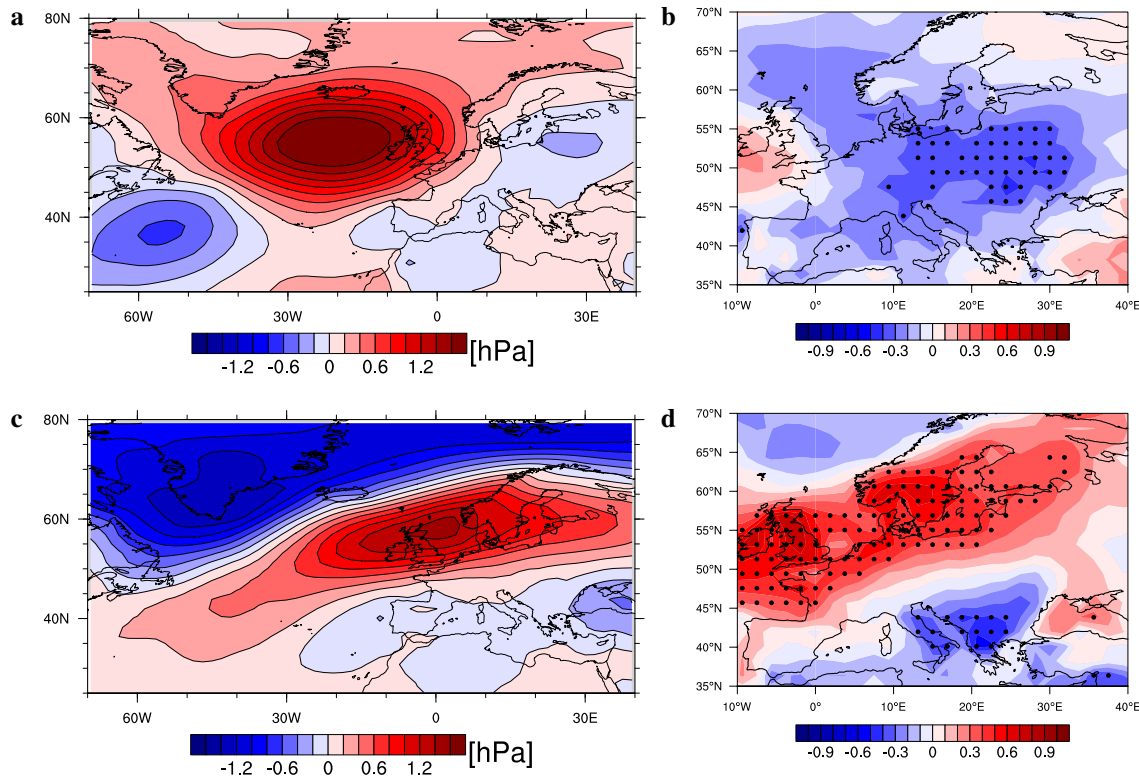
To further evaluate the hindcast skill, reliability diagrams (Wilks 2011) are used. Reliability diagrams are a tool to

quantify statistical reliability and show for a specified event the accordance between the observed relative frequency of the event and its forecasted probability. Here, we quantify events that lie above the climatology in the investigated region. For this, the analysed data are divided into ten different categories in dependence on the forecast probabilities of these events. Error bars are derived for every category via bootstrapping with 500 samples at the 95% confidence level.

### 3 Physical mechanism

We describe a physical mechanism connecting SSTs in the tropical North Atlantic in spring and surface temperatures over Europe in summer (Fig. 1a). SSTs in the tropical North Atlantic show high persistence, such that a signal is apparent in the SSTs from spring until summer. In the tropical North Atlantic, the main SST variability in spring lies in the latitudinal band between 0°–20°N (Fig. 1b). Warm SSTs





**Fig. 2** **a** The second EOF of SLP in JA in the region [25°N–80°N, 70°W–40°E], explaining about 19.2% of the total variance. **b** The associated principal component of the second EOF of SLP in JA, pointwise correlated with surface temperatures in JA. **c** The first EOF of SLP in JA in the region [25°N–80°N, 70°W–40°E], which explains about

30.7% of the total variance. **d** The associated principal component of the the first EOF of SLP in JA, pointwise correlated with surface temperatures in JA. The patterns are computed using the ERA-Interim reanalysis in 1982–2016. Dots represent significance at the 95% confidence level

in this area are the source of strong convection in the tropical region, which act as a Rossby wave source that induces a tropical–extratropical teleconnection (e.g., Gastineau and Frankignoul 2015). Following Saeed et al. (2014), we associate the resulting Rossby wave with the first EOF over the North-Atlantic-European sector in July–August (JA). Here, we specify the EOF narrowed to the region of interest [10°–70°N, 90°W–40°E] (Fig. 1c). This results in a wavelike structure similar to the CGT (Branstator 2002; Ding and Wang 2005; Saeed et al. 2014), with four prominent alternating patterns extending from North America to eastern Europe at a latitudinal band corresponding to the latitudinal extend of the subtropical jet stream.

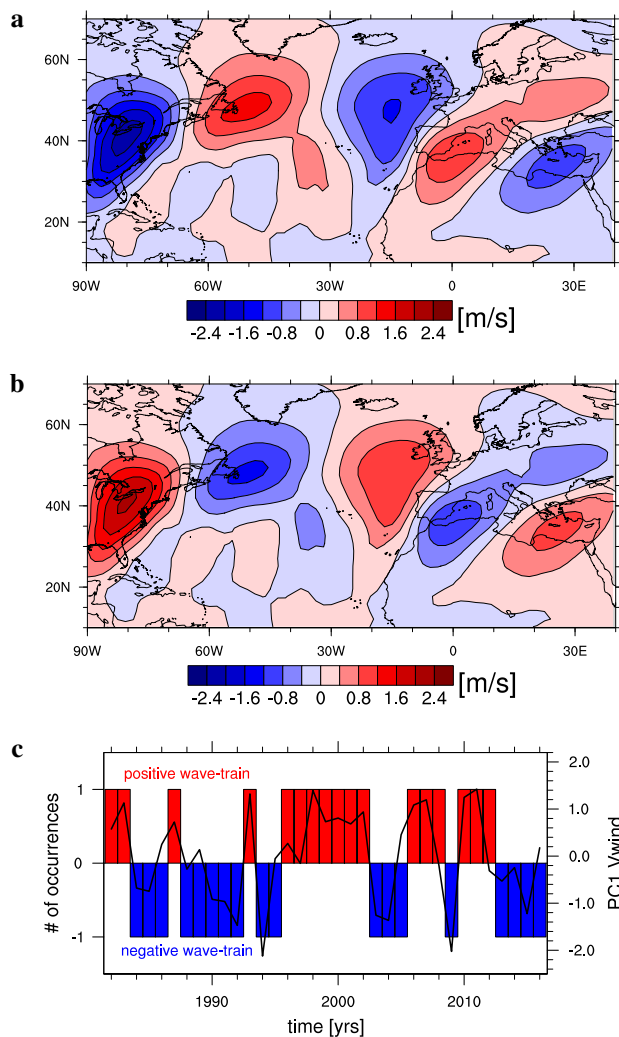
The forcing region of the wave-train is characterized in a pointwise correlation analysis between the temporal variability of the wave-train and SSTs in the North Atlantic (Fig. 1d), resulting in a significant correlation in the same region as the highest variability of the leading EOF of SSTs in the tropical North Atlantic in April (cf. Fig. 1b).

The wave-train has a strong influence on the summer climate in the North-Atlantic-European sector. It is accompanied by a zonal pressure gradient having a positive pressure

pattern over the northern North Atlantic and a negative one over eastern Europe (Fig. 1e), which is in accordance with previous findings (Saeed et al. 2014; Wulff et al. 2017). The zonal wind and pressure structures in turn influence the summer temperatures and precipitation over central Europe (Fig. 1f, g).

The SST signal, that is the source signal of this mechanism, is moving from the eastern and central tropical North Atlantic in spring to the western tropical North Atlantic and North America in summer (cf. Fig. 1d, f). The strong SST anomalies in the western tropical Pacific in JA are then accompanied by low pressure and high total precipitation anomalies in this region (cf. Fig. 1e, g), which is characteristic of a Gill-type response to the diabatic forcing within the Caribbean region (see e.g. Hodson et al. (2010) for details).

The zonal pressure gradient (Fig. 1e) is related to the second EOF of SLP in JA (Fig. 2a), explaining about 20% of the low-frequency pressure variability in summer. In its negative phase the patterns of the zonal pressure gradient are similar to the known East Atlantic pattern (Wallace and Gutzler 1981; Barnston and Livezey 1987; Iglesias et al. 2014), Atlantic Low pattern (Cassou et al. 2005) or summer East



**Fig. 3** **a** Positive and **b** negative wave-train cluster derived by the  $k$ -mean algorithms for ERA-Interim 200 hPa meridional wind in the area [10°N–70°N, 90°W–40°E] in 1982–2016. The pattern correlation with the wave-train EOF pattern (Fig. 1c) is  $\pm 87.1\%$ . **c** Frequency of the cluster over the whole period for the positive (red, frequency 48.6%) and the negative cluster (blue, frequency 51.4%), compared to the principal component (PC) associated with the wave-train (black line)

Atlantic mode (SEA; Wulff et al. 2017). The influence of the zonal EOF pattern on European summer temperatures is in strong agreement with the influence of the proposed wave-train mechanism (cf. Figs. 1f, 2b).

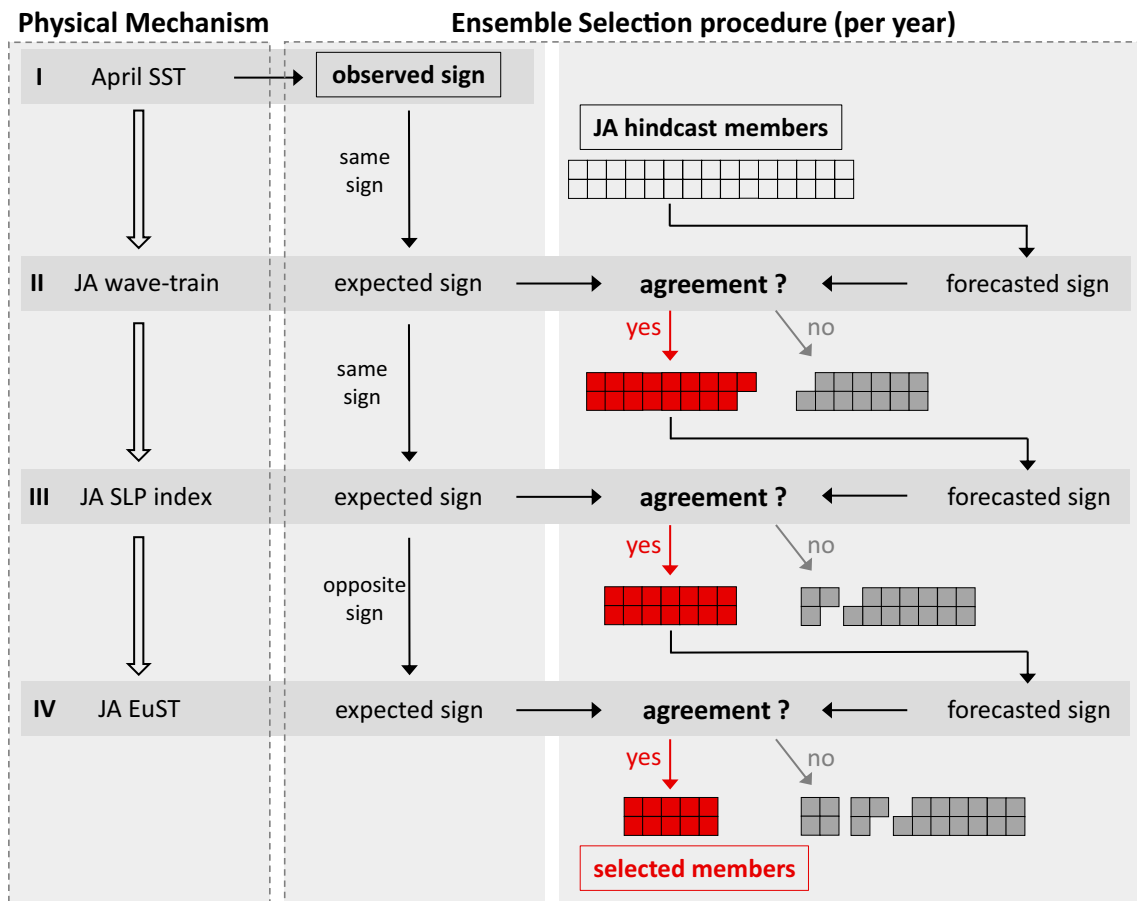
The zonal pressure gradient associated with the second EOF of SLP is contrary to the meridional pressure gradient of the first EOF, which explains about 30% of the low-frequency pressure variability in summer (Fig. 2c) and is associated with the SNAO (Folland et al. 2009). Its imprint on Europe is indicated by a north–south dipole with influence on northern and southern, but not on central Europe

(Fig. 2d), which resembles the findings of Folland et al. (2009) and Bladé et al. (2012).

## 4 Ensemble subsampling

Based on the analysis of the proposed mechanism in the ERA-Interim reanalysis, we characterise the mechanism by four individual parts, namely the SSTs in the tropical North Atlantic in April and the wave-train, the zonal pressure gradient and the temperature anomalies over central Europe in JA (Fig. 1a), and are now looking for a consistent representation of all parts in individual ensemble members. To comparably identify the chain of physical relations in individual members, i.e. in 30 ensemble members and in each of the 35 analysed summers, we divide the mechanism into its individual parts, while we define each part in a way that it can be distinguished in single ensemble members at selected points in time:

- I. *April SST* The origin of the signal, which are either anomalous high or low SST anomalies in the tropical North Atlantic in April. Based on the correlation analysis between the wave-train and the SST anomalies in spring (Fig. 1d), the SSTs are averaged in the area [0°–15°N, 80°W–20°W].
- II. *JA wave-train* The wave-train in JA, which we so far defined by an EOF pattern (Fig. 1c). To identify the sign definite patterns of the wave-train, we perform a cluster analysis on the 200 hPa meridional wind JA means taken from ERA-Interim in the domain [10°N–70°N, 90°W–40°E] for the investigated 35 years. The analysis is conducted with  $k = 2$ , resulting in a positive and negative wave-train cluster (Fig. 3a, b, respectively), while the positive cluster occurred in 17 and the negative one in 18 years (Fig. 3c). There is a good agreement between the cluster and the EOF analysis, both for the patterns (cf. Figs. 1c, 3a, pattern correlation of  $\pm 87.1\%$ ) and occurrences (Fig. 3c), all in all confirming that the cluster analysis sufficiently represents the wave-train. To check if individual ensemble members represent either the positive or negative wave-train phase in the considered year, a pattern-matching algorithm in terms of the root-mean-square difference is used to assign each ensemble member to the closest cluster.
- III. *JA SLP index* The zonal pressure gradient in JA. Based on the zonal pressure pattern of the second EOF of SLP in JA (Fig. 2a) and on the correlation pattern of the wave-train with the SLP (Fig. 1e), the zonal pressure gradient is defined via an index as the normalized difference between the detrended summer SLP averaged in a region over the North Atlantic



**Fig. 4** Schematic representation of the subsampling method for 30 ensemble members selected for the proposed mechanism in 1982–2016 in one of the considered years. The proposed mechanism as depicted in Fig. 1a and described in Sect. 3, is based on observations in 1982–2016 and depicted here on the left. We divided the physical mechanism into four steps and use it as a guideline for the ensemble selection shown on the right. In every considered year we start with the JA output of the 30 ensemble members generated by MPI-ESM-MR in May and use the sign of ERA-Interim April SSTs (I) to anticipate the sign of the JA wave-train (II) and only select those ensemble members in which the sign predicted for the JA wave-train agrees with this anticipated sign. Those such selected ensemble mem-

bers are then individually tested if their predicted sign of the JA SLP index (III) agrees with the sign of the JA SLP index anticipated by the sign of ERA-Interim April SSTs. The remaining ensemble members are then in a last step checked for their sign of JA European surface temperatures (EuST, IV) and only those members are kept in which this predicted sign agrees with the anticipated sign. The mean over those remaining ensemble members is then taken as the selected ensemble mean in the considered year. The depicted selection procedure is repeated for every considered year, while the quantity of the remaining selected ensemble members varies and is listed in Table 1. The ensemble mean formed over the selected members in all years is termed “selected ensemble” (see e.g. Fig. 5)

[40°N–60°N, 40°W–10°W] and a region over Europe [40°N–60°N, 10°E–40°E]. The SLP index is in good agreement with the the second SLP EOF in ERA-Interim (correlation of 74.4%, not shown), approving that the SLP index can be used to represent the zonal pressure gradient.

IV. *JA EuST* The European surface temperature anomaly in JA. On the basis of the correlation analysis between the wave-train and temperature anomaly over Europe in summer (Fig. 1f), the temperature is averaged over a region in central Europe [35°N–55°N, 0°–20°E].

A schematic overview of this selection process can be found in Fig. 4. The summertime part of the mechanism, namely the wave-train cluster (II), the sign of the SLP index (III) and the sign of the averaged temperature anomaly over central Europe (IV), all in JA, can now be tested in the 30 hindcast ensemble members in every of the 35 analysed summers, with the aim to identify those members, that represent this entire chain of processes. Based on the correlation analysis in ERA-Interim (Fig. 1), the positive (negative) phase of the mechanism in summer corresponds to a positive (negative) wave-train cluster, a positive (negative) SLP index and a negative (positive) temperature anomaly over Europe. In accordance with these observations we thus

**Table 1** Overview over the phase of the mechanism in the analysed years (second column) anticipated by the sign of the April SST anomalies in the area [0°–15°N, 80°W–20°W] taken from ERA-Interim in 1982–2016

Year	Expected phase	# ensemble members	Wind cluster hit	SLP index hit	Hit with both
1982	–	10	o	o	o
1983	+	10	x	x	x
1984	–	10	x	o	o
1985	–	8	x	x	x
1986	–	10	x	o	o
1987	+	7	x	x	x
1988	+	6	o	o	o
1989	–	8	x	o	o
1990	+	4	o	o	o
1991	–	10	x	x	x
1992	–	4	x	x	x
1993	+	11	x	x	x
1994	–	7	x	x	x
1995	+	11	o	o	o
1996	+	7	o	x	o
1997	+	4	x	o	o
1998	+	3	x	x	x
1999	–	13	o	x	o
2000	+	10	x	x	x
2001	–	12	o	x	o
2002	–	13	o	o	o
2003	–	8	x	x	x
2004	+	11	o	o	o
2005	+	2	o	x	o
2006	+	4	x	x	x
2007	+	3	x	x	x
2008	–	10	x	x	x
2009	–	7	x	x	x
2010	+	7	x	x	x
2011	+	8	x	x	x
2012	–	10	x	x	x
2013	–	3	x	x	x
2014	–	4	x	o	o
2015	–	12	x	x	x
2016	+	0	o	o	o
# hits			24	23	19

“+” indicates an anticipated positive and “–” an anticipated negative phase. (third column) The number of selected ensemble members from MPI-ESM and (fourth column) if the chosen phase coincides with the ERA-Interim phase of the wind cluster, (fifth column) with the ERA-Interim phase of the SLP index or (last column) with both, the ERA-Interim phase of the wind cluster and of the SLP index. “x” represents a correctly determined phase, “o” an incorrectly determined one, while in the last row the actual number of hits are summed up

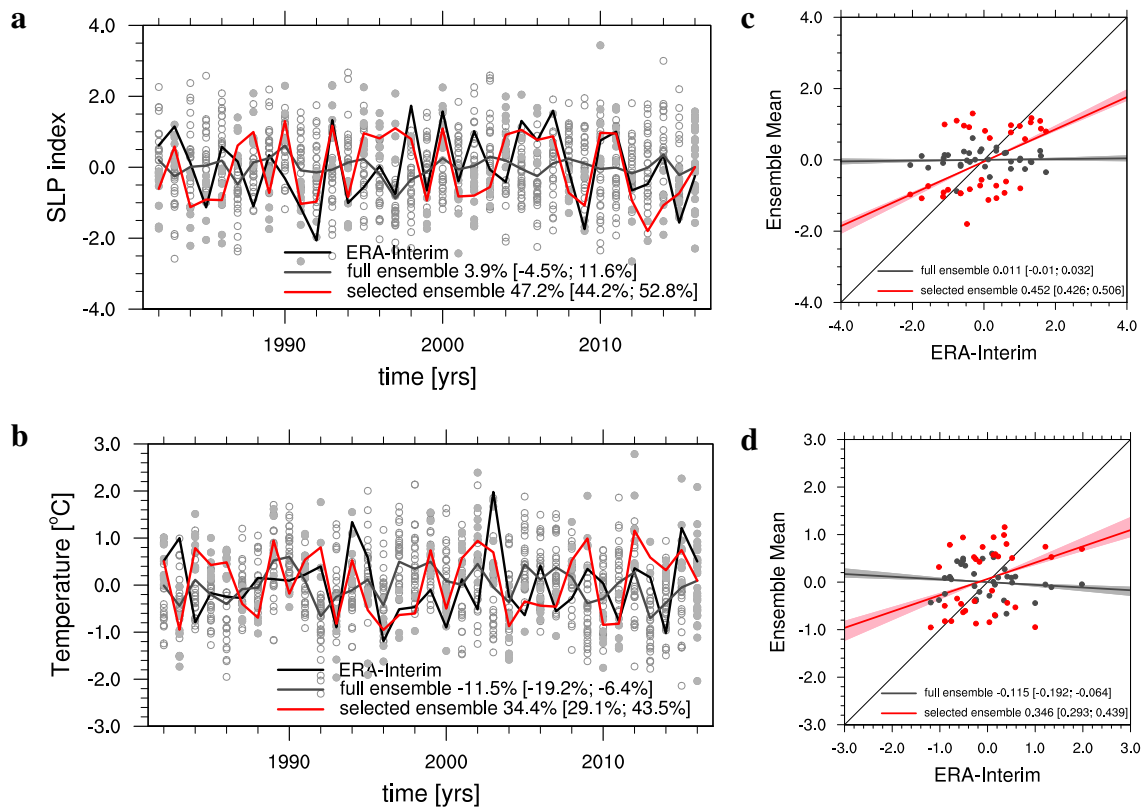
assume that the mechanism is only physically represented in those ensemble members, in which the signs of all three summer criteria are consistent to each other, either for the negative or for the positive phase. In the practical way we check the sign of each summertime criteria in every ensemble member in every considered year and verify the consistency of the signs of all three criteria. We can then categorise the individual ensemble members into the ones that predict a positive mechanism, the ones that predict a negative mechanism and the ones in which the signs of the three summer criteria are not consistent to each other and thus do not represent the mechanism. This analysis of consistency can be conducted within the hindcast setup for every individual summer without using observational information. However, while the consistency of the summertime mechanism can be tested within the hindcast setup, this analysis does not provide the information on the phase of the observed mechanism in each considered year.

Therefore, to determine if the mechanism in each considered year is in its positive or negative phase, we use the observed successive relation between the SSTs in the tropical North Atlantic in April and the mechanism in summer (Fig. 1d), which includes a lag of a few month. For this, we take the sign of the SST anomaly in April, which is before the initialisation of the ensemble prediction system, from observations and use it to anticipate the corresponding phase of the mechanism. According to the observed positive correlation between spring SSTs and the mechanism (cf. Fig. 1d), we assume that for a positive (negative) SST anomaly in April the mechanism in the considered year is positive (negative).

Following this procedure (see Fig. 4 for a schematic overview), each year is considered individually and an ensemble member is only retained if the sign of all three summertime criteria agree with the sign anticipated by the tropical SSTs that are observed in April in the examined year. This results in an ensemble size of 2 to 13 out of 30 ensemble members per year (Table 1), while in one year (2016) none of the ensemble members fulfil all criteria, where we use the mean of the full ensemble. In all other analysed years we then derive an ensemble mean by taking a mean over the selected rather than the full ensemble.

Comparing the individual selection criteria in both the full and the selected ensemble (Fig. 5) demonstrates that before the selection, the individual ensemble members are spread over the whole range of values, resulting in a small temporal variability of the full ensemble mean (Fig. 5a, b). Through the selection process, the ensemble spread is reduced in every year, which then also results in higher temporal variability for the selected ensemble mean. Moreover, the values that are obtained in the selected compared to the full ensemble mean are in better agreement with the observed values (Fig. 5c, d). While the linear regression of the full ensemble results





**Fig. 5** **a, b** Comparison between ERA-Interim (black lines) and the ensemble mean over the full (grey lines) or the selected ensemble (red lines) from MPI-ESM in 1982–2016 including the mean and range (indicated in parentheses) leave-one-out cross-validated correlation values. Dots indicate the value of all ensemble members and full dots the ones of criterion selected ensemble members. **c, d** Scatterplots of ERA-Interim compared to the ensemble mean of MPI-ESM in 1982–2016 over the full (grey dots) or the selected ensemble (red dots) including linear regression. The mean and range (indicated

in parentheses and by shading) leave-one-out cross-validated linear regression slopes are derived for the full ensemble (grey line) and for the selected ensemble (red line). The black line indicates perfect linear regression. Plots depict **a, c** the SLP index defined as the difference in JA SLP between the North-Atlantic [40°N–60°N, 40°W–10°W] and Europe [40°N–60°N, 10°E–40°E] and **b, d** the temperature anomalies in Europe in JA averaged in the area [35°N–55°N, 0°–20°W]

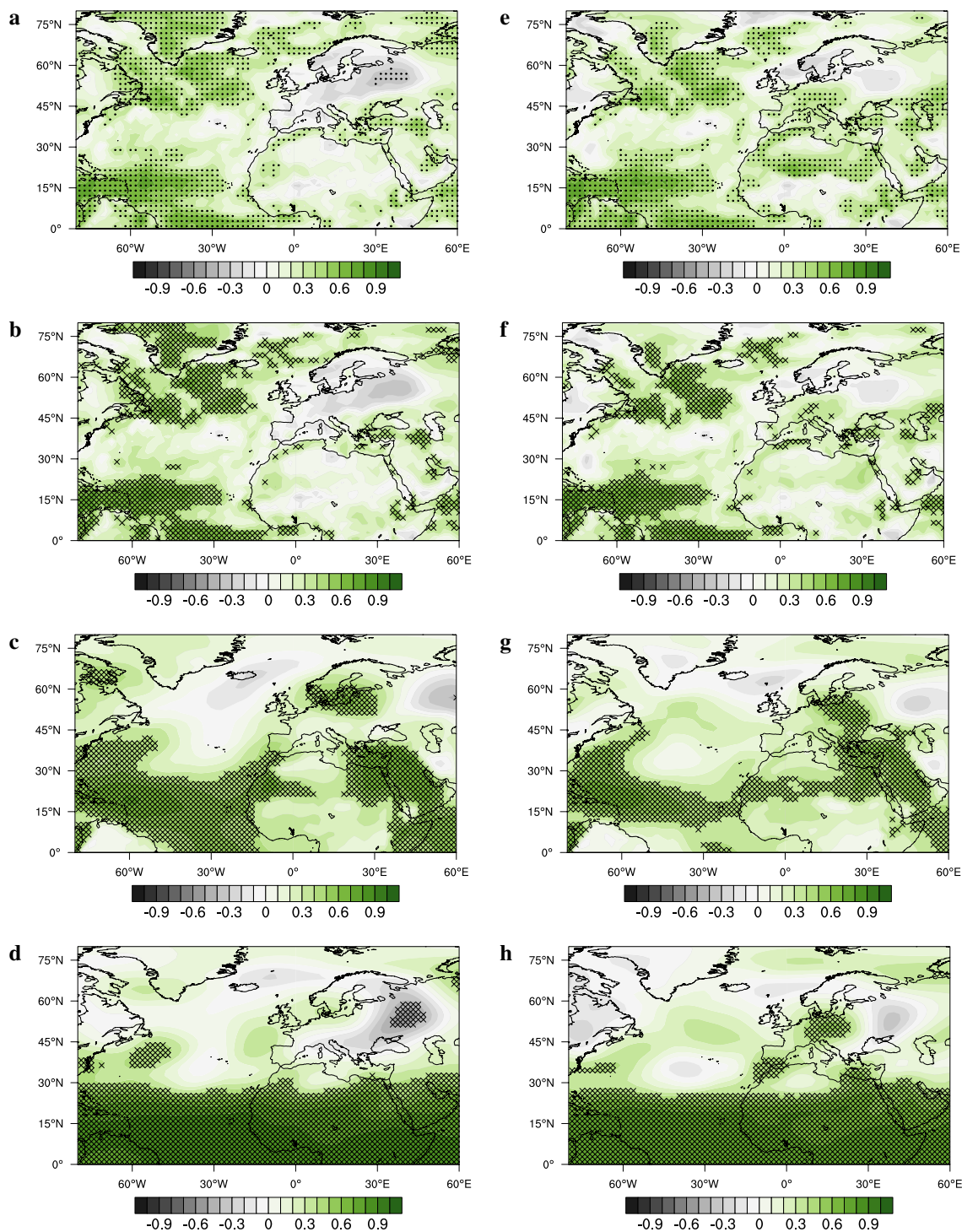
in a slope around zero and thus deviates strongly from ERA-Interim, the linear regression of the selected ensemble shows a positive slope of about 0.45 and 0.35, which is much closer to the line of perfect linear regression, that accords to equal values of ERA-Interim and the ensemble mean. In summary, the selection process yields a significant increase of correlation from no or negative correlation to about 47% for the SLP index and 34% for European surface temperatures.

As shown in Sect. 3, the proposed mechanism consists of a chain of processes including the wind, pressure and temperature systems over Europe and the North Atlantic. We assume here that the mechanism is only physically represented in those ensemble members that show all parts of this mechanism, which is practically determined through the listed criteria. Those ensemble members that do not fulfill the criteria, meaning that the sign of the three summertime criteria are not consistent to each other and do not agree with the sign of the observed spring SSTs, are thus rejected.

We assume that these rejected ensemble members do not represent the physical processes of the proposed mechanism, but rather of other mechanisms that influence European summers on seasonal time scales and could still be important for the representation of those mechanisms.

## 5 Seasonal hindcast skill

The subselection of the ensemble members allows to build a new ensemble mean for any of the simulated fields. We quantify the hindcast skill for surface temperature, SLP and 500 hPa geopotential height (Z500) in the North-Atlantic-European sector for the mean over the subselected ensemble and compare it to the hindcast skill for the full ensemble mean (Fig. 6). For surface temperatures, the full ensemble shows hindcast skill mainly over the North Atlantic and Greenland (Fig. 6a, b). The SLP for



**Fig. 6** Anomaly correlation (ACC) derived without cross-validation for (first row) surface temperatures and derived as the mean over all leave-one-out cross-validated correlation values for (second row) surface temperatures, (third row) SLP and (forth row) 500 hPa geopotential height (Z500) in summer (JA), comparing the model predictions

of MPI-ESM to ERA-Interim in 1982–2016. The ensemble mean is taken over **a–d** the full ensemble and **e–h** the selected ensemble. Dots show significance at the 95% confidence level, hatching represents areas in which significance is reached in every leave-one-out cross-validated iteration

the full ensemble shows significant skill in the tropical North Atlantic and over parts of northern Europe and Arabia (Fig. 6c), while for Z500 hindcast skill can only be achieved in the tropical belt (Fig. 6d). Neither of the fields achieve significant hindcast skill over central Europe.

With the ensemble selection on the other hand, surface temperature, SLP and Z500 show significant hindcast skill over central Europe (Fig. 6e–h), agreeing with the area where the proposed mechanism shows a significant imprint on seasonal European summer climate (cf. Figs. 1e–g, 2b). For SLP and Z500, improved skill can also be achieved over those parts of the North Atlantic where the zonal pressure gradient is located (cf. Fig. 1e). Other areas in the North-Atlantic-European sector stay at about the same hindcast skill. A slight decrease in ACC can only be found over Greenland and over parts of Scandinavia and Great Britain, which are the parts where the SNAO has its biggest influence (cf. Fig. 2d, Bladé et al. 2012).

To further analyse the robustness of the significant hindcast skill, we, aside from deriving the significant hindcast skill for all analysed years (Fig. 6a, e), also derive the hindcast skill that is significant in all cross-validated iterations (Fig. 6b, f). Due to the higher statistical robustness of the cross-validated hindcast skill, we restrict all further ACC plots to the ones including cross-validation.

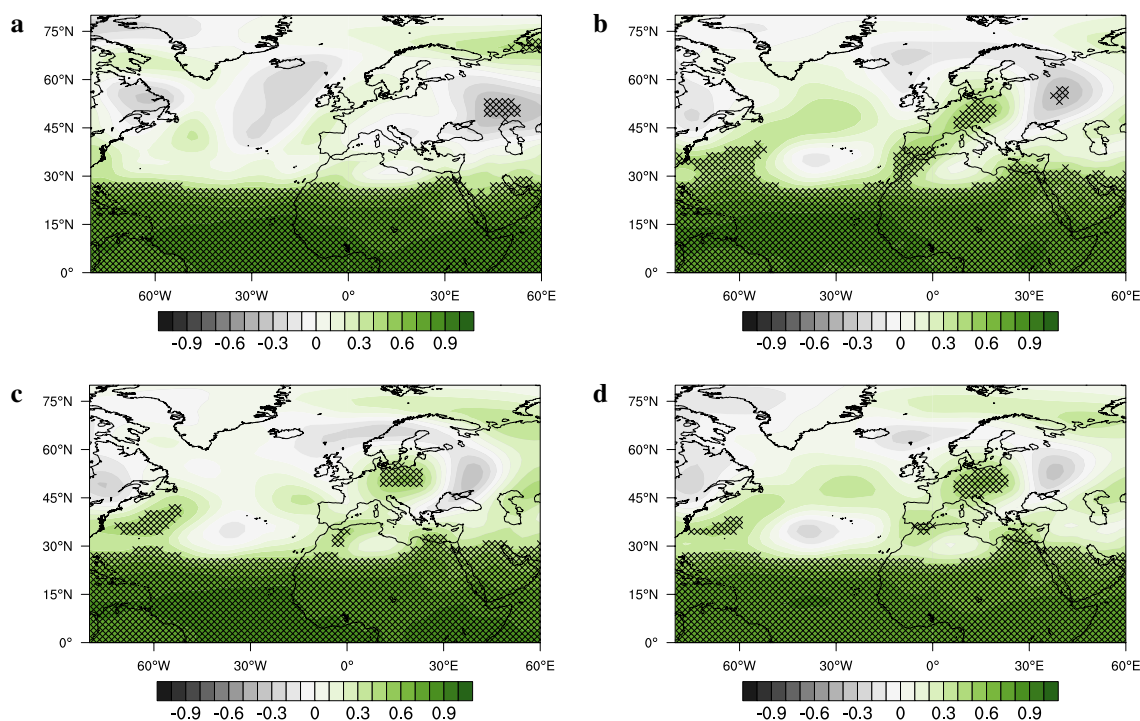
Additionally, we evaluate the influence of the chosen phase of the mechanism on the ACC (Fig. 7). So far, we determined the phase of the mechanism by the observed sign of spring SSTs in the tropics. In every analysed year we divide the 30 ensemble members into the ones that include a positive mechanism, the ones that include a negative mechanism and the ones do not contain the mechanism at all. Each of these three categories contains a different number of ensemble members every year. Instead of choosing the phase of the mechanism by the observed sign of spring SSTs in the tropics in every year, we could thus chose either the positive or negative phase, depending on which one contains more ensemble members in the respective year. However, the hindcast skill achieved with this procedure shows no improvement compared to the full ensemble (cf. Figs. 6d, 7a).

The most impact on the hindcast skill is thus made by the selected phase of the mechanism in the considered years. If chosen by the phase that contains the majority of ensemble members (as in Fig. 7a), the phase is only correctly determined in 11 out of 35 years for both the wave-train cluster and the SLP index, and in 15 or 17 years for either of the criteria (Table 2). This hit rate improves if the phase of the mechanism is anticipated by observed spring SSTs in the tropical North Atlantic, such that 19 out of 35 years for both and 24 or 23 years for either of the wave-train cluster or SLP index criteria are correctly determined (Table 1).

If the phase of the mechanism is determined by spring SSTs, we can also evaluate the influence of the different selection criteria, namely the wave-train cluster, the SLP index or the European temperature anomalies, on the ACC (Fig. 7b–d). Here, also those ensemble members are retained in which only one or two of the three summertime criteria agree with the anticipated sign. In this analysis, we find that improvements in the hindcast skill can be achieved with already one out of the three criteria. Including only the wave-train cluster into the prediction results in improved hindcast skill in the areas where the wave-train has its prominent patterns (Fig. 7b). If the temperature over Europe is the only included criterion, then the hindcast skill is stronger improved over central Europe (Fig. 7c), while a combination of the wave-train cluster and the European temperature criterion shows improved hindcast skill over larger areas of Europe (Fig. 7d). Including the SLP index as a third criterion then only slightly improves the hindcast skill over Europe (Fig. 6h). The small changes on ACC between the embedding of the different criteria is consistent with our analysis in Sect. 3, since we find that all three criteria are part of the same mechanism and should consequently show similar influence on the hindcast skill.

To further evaluate the improved reliability over central Europe, we examine reliability diagrams of temperature and Z500 over central Europe (Fig. 8). If reliability diagrams of the full ensemble are compared to the selected ensemble, improvements are achieved through subsampling by getting closer to the line of perfect reliability and thus resulting in more reliable hindcasts. This coincides with the findings in Fig. 5, in which we show that both, the variability and the values of the selected ensemble agree much better with ERA-Interim than the ones of the full ensemble. Further, the distribution, which expresses the frequency of each possible forecast probability, is more equally dispersed for the selected than for the ensemble mean. The frequency of the possible forecasts probability of the full ensemble deviates rarely from the average value. The selected ensemble mean on the other hand also shows frequency for extreme forecasts and thus results in more confident forecasts. This higher number of events for extreme forecast probabilities of the selected ensemble compared to the full ensemble also results in smaller errors.

According to our analysis in Fig. 7, the crucial step of our approach seems to be the anticipated phase of the mechanism. When estimated by the observed spring SST anomalies, the sign of the SLP index is only determined correctly in 23 out of 35 years (Table 1). If this estimation would be accurate for all evaluated years, a theoretically “perfect” ensemble selection could be achieved. In such a “perfect” analysis we use the same three criteria for the selection of the ensemble members, except that the phase of the mechanism in the individual years is not anticipated by



**Fig. 7** Anomaly correlation (ACC) for 500 hPa geopotential height (Z500) in summer (JA) comparing the model predictions of MPI-ESM to ERA-Interim in 1982–2016 derived as the mean over all leave-one-out cross-validated correlation values. **a** The ensemble members are selected by all three selection criteria (wave-train cluster, SLP index and averaged temperature anomaly), while the phase of the mechanism in each considered year is determined by the phase

that contains the majority of ensemble members. **b–d** The phase of the mechanism is anticipated by April SSTs in the tropical Atlantic, while the ensemble members are selected by a subset of criteria: **b** the wave-train, **c** the temperature or **d** both the wave-train and the temperature criteria. Hatching represents areas in which significance is reached in every leave-one-out cross-validated iteration

spring SSTs, but chosen by the sign of the SLP index in JA in ERA-Interim in the same year. Therefore, in this “perfect” analysis we are using information from the observations that occurred after the initialisation of the ensemble members, which is thus not feasible in a real forecast situation, but still of interest for reference.

The increase of the hindcast skill in this “perfect” analysis is in similar regions as for the ensemble selection (cf. Figs. 6f, h, 9a, b), but slightly more pronounced. The biggest improvement is found in the area in which the western part of the zonal pressure gradient is located (cf. Fig. 1e), where significant hindcast skill for Z500 is achieved in the “perfect” analysis. The “perfect” selection represents the hindcast skill that can be expected from an analysis in which the mechanism is perfectly predicted by the model and confirms our findings that the hindcast skill can be significantly improved in the areas where the mechanism is showing a prominent signal in the observations.

To ascertain the reliability that could maximally be achieved by including the proposed mechanism into the prediction, we examine the reliability diagrams of the “perfect” ensemble selection (Fig. 9c, d). For the temperature, the reliability diagram of the “perfect” ensemble is almost

identical to the one of the selected ensemble. For Z500, the reliability diagram of the “perfect” ensemble is closer to the line of perfect reliability, especially for the extreme forecast probabilities, while the distribution of the probabilities stays at about the same level of forecast confidence. These reliability diagrams demonstrate that the reliability that can maximally be achieved by including our mechanism into the prediction is limited.

Altogether we achieve better hindcast skill and reliability from an analysis that uses an area of high prediction skill and high persistence as a predictor for a mechanism that influences the European summer climate on seasonal time scales. The mechanism is included into the prediction by subselection of ensemble members through successive criteria based on the physical variables that define this mechanism.

## 6 Discussion

Our ensemble selection shows improved hindcast skill that is consistent throughout all analysed fields (Fig. 6e–h). Such a consistency in hindcast skill is not achieved in a hindcast analysis that uses the mean over all ensemble members



**Table 2** As in Table 1, but with the phase of the mechanism anticipated by the phase that contains the majority of ensemble members in MPI-ESM

Year	Expected phase	# ensemble members	Wind cluster hit	SLP index hit	Hit with both
1982	+	11	x	x	x
1983	+	10	x	x	x
1984	–	10	x	o	o
1985	+	10	o	o	o
1986	+	15	o	x	o
1987	–	8	o	o	o
1988	+	6	x	o	o
1989	–	8	x	o	o
1990	–	6	x	x	x
1991	–	10	x	x	x
1992	+	9	o	o	o
1993	+	11	x	x	x
1994	+	10	o	o	o
1995	+	11	o	o	o
1996	+	7	o	x	o
1997	–	18	o	x	o
1998	–	12	o	o	o
1999	–	13	o	x	o
2000	+	10	x	x	x
2001	–	12	o	x	o
2002	–	13	o	o	o
2003	+	11	o	o	o
2004	+	11	o	o	o
2005	–	12	x	o	o
2006	–	10	o	o	o
2007	–	8	o	o	o
2008	–	10	x	x	x
2009	+	11	o	o	o
2010	–	8	o	o	o
2011	+	8	x	x	x
2012	–	10	x	x	x
2013	+	12	o	o	o
2014	+	15	o	x	o
2015	–	12	x	x	x
2016	–	8	x	x	x
# hits			15	17	11

(Fig. 6a–d). We claim that this inconsistency occurs for the full ensemble mean, since a mean is taken over various different physical mechanisms, while in our analysis ensemble members are selected for just one mechanism. Since various mechanisms are influencing the summer climate in the North-Atlantic-European sector on seasonal time scales, the model only predicts the chain of processes of the proposed mechanism (cf. Fig. 1a) in certain ensemble members, which is why we select only those ensemble members that contain

the successive physical relations. We assume that all rejected ensemble members represent other mechanisms and could still be important for the analysis of those.

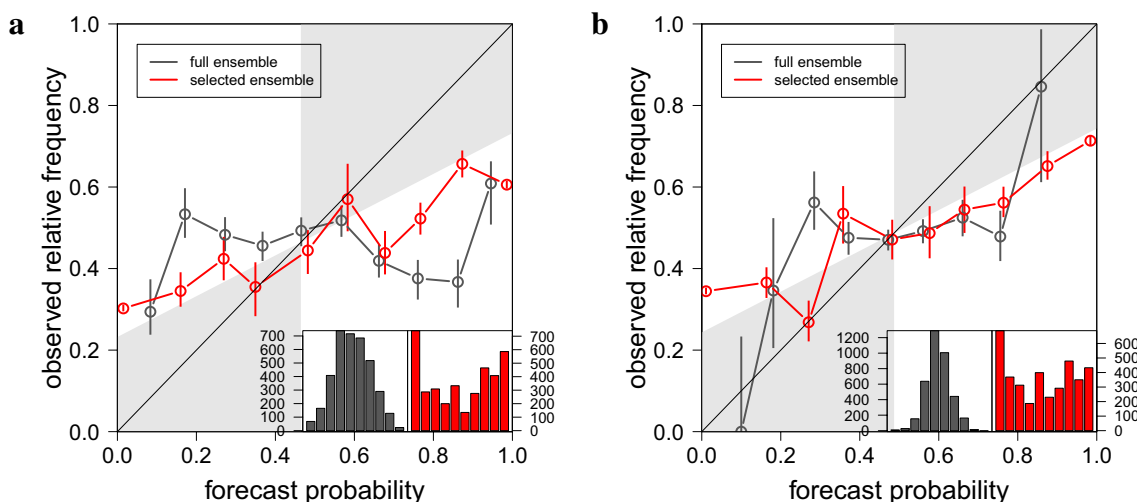
We further show that the spread of the full ensemble is too large and the mean over the full ensemble thus results in a variability much lower than the observed one (Fig. 5). This problem of a too large ensemble spread demonstrates that the signal of the ensemble mean is too low, while the noise is too high, and has been shown in previous studies for different ensemble prediction systems as well (e.g., Ho et al. 2013; Eade et al. 2014). Here we present an approach that reduces the ensemble spread through a selection of the ensemble members based on a prominent seasonal summer pattern and thus amplifies the signal, while suppressing the noise of seasonal summer climate over Europe and the North Atlantic.

However, the reliability diagrams in Fig. 8 demonstrate that only limited reliability can be achieved with our approach. Here, the limits of our method become apparent that probably arise from the mechanism explaining only a fraction of the variability of seasonal European summer climate and other mechanisms being prominent during European summers as well. Our method is based on only one of the mechanisms that are influencing European summer climate on seasonal time scales, such that the skill achieved by including our mechanism into the prediction has a natural limit. With our approach we further assume that the mechanism is present in every of the analysed years, which is not the case for all years, since various mechanisms are influencing the summer climate in Europe on seasonal time scales. Including additional mechanisms into the prediction analysis is thus a way to extend our method that could further affect the seasonal hindcast skill over Europe.

The most prominent mechanism in summer is the SNAO. We show that it differs from our proposed mechanism and that it also influences different areas over Europe (Fig. 2). While we find decreased hindcast skill over Greenland and parts of northern Europe, Düsterhus et al. (2017) show that, if the SNAO is included into the prediction, the hindcast skill can be improved in exactly those areas. Since Greenland and northern Europe are areas that are influenced by the SNAO, this underlines our finding that with our approach improved hindcast skill is found in areas where the mechanism, that is included into the prediction, shows pronounced influence.

To achieve improvements in the hindcast skill over areas that are related to either of the two mechanisms, an analysis could be tested that combines both the SNAO and the zonal mechanism.

Further, it could be tested how big the influence of strong SNAO years are on the hindcast skill achieved here, meaning to analyse in which years the SNAO is prominent and how big the effect of those years are on the hindcast skill. Such



**Fig. 8** Reliability diagrams comparing the mean over the full (gray) and the selected (red) ensemble of MPI-ESM to ERA-Interim in 1982–2016 for **a** temperature and **b** Z500 in the area [35°N–55°N, 0°–20°W]. The diagonal line indicates perfect reliability, meaning that the observed relative frequency of the considered event accords perfectly with its forecasted probabilities. The shaded grey box is set by

the vertical line, that marks the climatological probability of the event in the forecasts and observations, and by the “no-skill” line. Points that lie inside the grey box contribute positively to the forecast skill, based on the climatological reference. Vertical lines show the error bootstrapped at the 95% confidence level. The histograms depict the distribution of the data

a yearly selection has already been done by Domeisen et al. (2015) and could, apart from excluding strong SNAO years, also be applied for only those years in which our mechanism is the dominant one.

The East Atlantic pattern has in its positive phase a negative pressure anomaly over the subtropical North Atlantic and a positive anomaly over Europe and depicts a similar structure than the SLP difference shown here, but in the opposite phase. However, the index of the East Atlantic pattern as defined by the National Weather Service Climate Prediction Center via Rotated Principal Component Analysis of 500 hPa geopotential height in the northern hemisphere, if averaged for July and August, is not related to the here defined SLP index (correlation of about  $-0.18$ ).

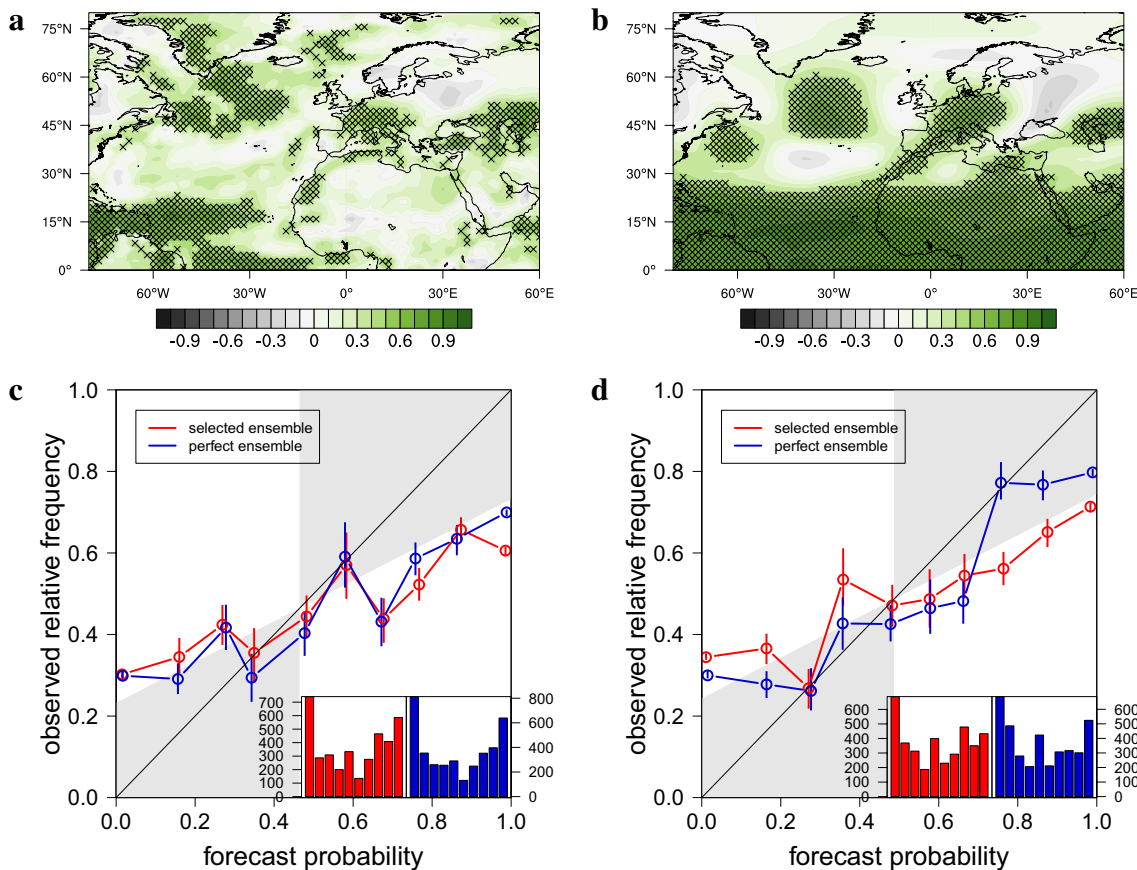
Wulff et al. (2017) suggest a relation between their zonal mechanism and the El Niño–Southern Oscillation (ENSO), while Ding and Wang (2005) claim that the CGT is independent of ENSO. In this work, the connection of the mechanism to ENSO is not further analysed, since we restrict our analysis to the North-Atlantic-European sector. In general, we do not rule out a connection to ENSO. According to Lau and Nath (2001) and Alexander et al. (2002), SSTs in different ocean basins are linked via an “atmospheric bridge”, such that SST anomalies in the North Atlantic in spring and summer are lead by SST anomalies in the tropical Pacific. This assumption does not disagree with our findings. On the contrary, since ENSO is leading the spring SSTs in the North Atlantic, which are here used as a predictor, it even provides the opportunity to use ENSO as a further predictor.

Recent studies further show that summer variability is better represented in model runs with higher resolution (e.g., Müller et al. 2018). However, our analysis focuses on large-scale teleconnections and patterns. Therefore, we expect the analysed patterns and results to be similar in higher resolved model runs.

We only consider mean temperatures over the whole summer season, such that no attention is paid to extreme events. Cassou et al. (2005) and Duchez et al. (2016) show that the zonal pressure gradient in its negative phase can lead to European heat waves. This is in agreement with our findings, since our correlation analysis confirms that the zonal mechanism in its negative phase is accompanied by warm temperatures over Europe (cf. Figs. 1f, 2b). Additionally, we show that with our ensemble selection the generated predictions are more reliable for extreme values than they are in the full ensemble (Fig. 8). Thus, including our mechanism into the prediction of extreme events should potentially lead to useful prediction skill.

Wu et al. (2016) find that the CGT also exists on inter-decadal time scales and associate it with the Atlantic multi-decadal oscillation (AMO). This is in agreement with the findings of Gastineau and Frankignoul (2015) who show that the SSTs, for which we reveal a connection to the zonal pressure gradient, are influenced by the AMO. To investigate the influence of the AMO on the mechanism and the hindcast skill analysed in this work, a seasonal hindcast run longer than the available 35 years would be needed.

The “perfect” ensemble selection shows a hindcast analysis in which the phase of the mechanism would be known each



**Fig. 9** Hindcast skill for the “perfect” ensemble selection derived with the known state of the SLP index. **a, b** Anomaly correlation (ACC) for **a** surface temperatures and **b** 500 hPa geopotential height (Z500) in summer (JA), comparing the model predictions of MPI-ESM to ERA-Interim in 1982–2016 derived as the mean over all leave-one-out cross-validated correlation values. Hatching represents areas in which significance is reached in every leave-one-out cross-

validated iteration. **c, d** Reliability diagrams comparing the mean over the selected (red) and the “perfect” (blue) ensemble of MPI-ESM to ERA-Interim in 1982–2016 for **c** temperature and **d** Z500 in the area [35°N–55°N, 0°–20°W]. Vertical lines show the error bootstrapped at the 95% confidence level. The histograms depict the distribution of the data

year in advance, such that it reveals the hindcast skill that can be expected if the mechanism would be perfectly predicted by the model (Fig. 9). This “perfect” analysis results in improved skill in those areas, that are expected to be influenced by the proposed mechanism (cf. Fig. 1). The hindcast skill achieved with the anticipated phase of the mechanism is in good agreement with the “perfect” prediction and shows improved skill in the same areas, just less pronounced, so that a consistent improvement in hindcast skill is found throughout our analysis.

### 7 Summary and conclusions

We assess the summer seasonal hindcast skill of the MPI-ESM-MR ensemble based seasonal prediction system over the North-Atlantic-European sector with regard to a mechanism that is influencing this region on seasonal time scales. The proposed mechanism, here analysed in

the ERA-Interim reanalysis has its origin in the tropical North Atlantic in spring, where persistent SST anomalies are the source of a Rossby wave-train that is accompanied by a zonal pressure gradient and has in turn an influence on European summer climate (Saeed et al. 2014; Wulff et al. 2017). We show the statistical relation between the different parts of the mechanism and include the mechanism into the hindcast analysis by selecting only those ensemble members in which its entire chain of processes is represented. The selection is thus built on three successive criteria that are based on the physics that define the mechanism. The starting signal of the mechanism is in the tropical SSTs in April, which is before the initialisation of the prediction system and can therefore be used to anticipate the phase of the mechanism in individual years. We thus use statistical relations to select ensemble members generated by a dynamical prediction system. Since the indicated statistical relations of the physical processes

are limited, the success of this method is restricted to the degree of this statistical relations.

From this analysis we conclude the following:

- Seasonal European summer climate variability is in many summers significantly influenced by a mechanism, whose signal originates in the tropical North Atlantic in spring and is transported to central Europe via zonal wind and pressure systems (Fig. 1).
- This mechanism is the second leading mode of seasonal climate variability in the North-Atlantic-European sector in summer and shows distinct different characteristics than the SNAO, which are a zonal, instead of a meridional pressure gradient and an influence on central, instead of on northern and southern Europe (Fig. 2).
- The individual parts of the summertime mechanism can be characterized by the wave-train cluster, the SLP index and the averaged European summer temperatures. Accounting for the proposed mechanism in the hindcast analysis by selecting only those ensemble members in which this entire chain of physical processes is represented, results in the reduction of the ensemble spread and a better representation of the variability of the proposed mechanism in the model (Fig. 5).
- The crucial step in the hindcast analysis is to anticipate the phase of the proposed mechanism, which is in most cases falsely anticipated if chosen by the phase that contains the majority of ensemble members in the respective year (Table 2).
- In a “perfect” prediction, in which the phase of the mechanism would be known in every analysed year, we demonstrate the hindcast skill that could be achieved if the mechanism is properly represented in the dynamical seasonal prediction system (Fig. 9). The areas of improved hindcast skill coincide with the areas in which the mechanism is influencing the North Atlantic-European climate on seasonal time scales (Fig. 1).
- Due to the demonstrated influence of the tropical spring SSTs on European summers (Fig. 1), we make use of the high persistence and predictability of tropical regions and use the observed SSTs in April to anticipate the phase of the mechanism in the ensemble system in each year. With this approach we achieve significantly improved hindcast skill over Europe and parts of the North Atlantic for surface temperature, SLP and Z500 (Fig. 6).

We show an alternative approach, in which the ensemble size, instead of being further increased, is decreased through ensemble selection on the basis of a chain of known physical relations. We make use of the influence of the high persistent and predictable spring SSTs on European summer climate

which we expose in a chain of physical processes to enhance seasonal hindcast skill over central Europe. Here, we focus on one mechanism that influences European summer climate on seasonal time scales and demonstrate that including a mechanism into the prediction through subsampling of the ensemble members by using successive physical relations is an effective method to achieve significant seasonal prediction skill. This approach could be extended with further mechanisms or also be applied for other mechanism and allows for improved predictions in other regions. Since we only use observations in April every year, which is before the initialisation of the model and knowledge of observations is not required after the initialisation, this approach can be applied to operational ensemble prediction systems.

**Acknowledgements** The authors would like to thank the two anonymous reviewers for their helpful remarks. Many thanks go also to the *Climate Modelling* group at the University Hamburg for the discussions and their feedback on the findings of this paper. This work was funded by the German Federal Ministry for Education and Research (BMBF) through the second Regional Atlantic Circulation and Global Change Project (RACE II; NCN, JB) and through the MiKlip project FLEXFORDEC (grant number 01LP1519A; WM), by the Copernicus Climate Change Service (contract number C3S 433 DWD; MD, JB). It was further funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the University Hamburg’s Cluster of Excellence Integrated Climate System Analysis and Prediction (CliSAP; AD, JB), and under Germany’s Excellence Strategy – EXC 2037 ‘Climate, Climatic Change, and Society’ (CliCCS) – Project Number: 390683824, as contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg (AD, JB). The model simulations were performed using the high-performance computer at the German Climate Computing Center (DKRZ).

## References

- Alexander MA, Bladé I, Newman M, Lanzante JR, Lau NC, Scott JD (2002) The atmospheric bridge: the influence of ENSO teleconnections on air-sea interaction over the global oceans. *J Clim* 15(16):2205–2231
- Arribas A, Glover M, Maidens A, Peterson K, Gordon M, MacLachlan C, Graham R, Fereday D, Camp J, Scaife A et al (2011) The GloSea4 ensemble prediction system for seasonal forecasting. *Mon Weather Rev* 139(6):1891–1910
- Baehr J, Piontek R (2014) Ensemble initialization of the oceanic component of a coupled model through bred vectors at seasonal-to-interannual timescales. *Geosci Model Dev* 7(1):453–461
- Baehr J, Fröhlich K, Botzet M, Domeisen DI, Kornbluh L, Notz D, Piontek R, Pohlmann H, Tietsche S, Mueller WA (2015) The prediction of surface temperature in the new seasonal prediction system based on the MPI-ESM coupled climate model. *Clim Dyn* 44(9–10):2723–2735
- Baker L, Shaffrey L, Sutton R, Weisheimer A, Scaife A (2018) An intercomparison of skill and overconfidence/underconfidence of the wintertime North Atlantic Oscillation in multimodel seasonal forecasts. *Geophys Res Lett* 45:7808–7817
- Balmaseda MA, Mogensen K, Weaver AT (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q J R Meteorol Soc* 139(674):1132–1161



- Barnston AG, Livezey RE (1987) Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon Weather Rev* 115(6):1083–1126
- Bjerknes J (1966) A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature. *Tellus* 18(4):820–829
- Bladé I, Liebmann B, Fortuny D, van Oldenborgh GJ (2012) Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region. *Clim Dyn* 39(3–4):709–727
- Branstator G (2002) Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic Oscillation. *J Clim* 15(14):1893–1910
- Branstator G, Teng H (2017) Tropospheric waveguide teleconnections and their seasonality. *J Atmos Sci* 74(5):1513–1532
- Bretherton CS, Smith C, Wallace JM (1992) An intercomparison of methods for finding coupled patterns in climate data. *J Clim* 5(6):541–560
- Cassou C, Terray L, Phillips AS (2005) Tropical Atlantic influence on European heat waves. *J Clim* 18(15):2805–2811
- Collins M (2002) Climate predictability on interannual to decadal time scales: the initial value problem. *Clim Dyn* 19(8):671–692
- Dee DP, Uppala S, Simmons A, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137(656):553–597
- Ding Q, Wang B (2005) Circumglobal teleconnection in the Northern Hemisphere summer. *J Clim* 18(17):3483–3505
- Doblas-Reyes FJ, García-Serrano J, Lienert F, Biescas AP, Rodrigues LR (2013) Seasonal climate predictability and forecasting: status and prospects. *Wiley Interdiscip Rev Clim Change* 4(4):245–268
- Dobrynin M, Domeisen DI, Müller WA, Bell L, Brune S, Bunzel F, Düsterhus A, Fröhlich K, Pohlmann H, Baehr J (2018) Improved teleconnection-based dynamical seasonal predictions of boreal winter. *Geophys Res Lett* 48:3605–3614
- Domeisen DI, Butler AH, Fröhlich K, Bittner M, Müller WA, Baehr J (2015) Seasonal predictability over Europe arising from El Niño and stratospheric variability in the MPI-ESM seasonal prediction system. *J Clim* 28(1):256–271
- Duchez A, Frajka-Williams E, Josey SA, Evans DG, Grist JP, Marsh R, McCarthy GD, Sinha B, Berry DI, Hirschi JJ (2016) Drivers of exceptionally cold North Atlantic Ocean temperatures and their link to the 2015 European heat wave. *Environ Res Lett* 11(7):074004
- Düsterhus A, Dobrynin M, Domeisen DI, Pohlmann H, Baehr J (2017) A statistical-dynamical seasonal prediction of the summer North Atlantic Oscillation. In: *Proceedings of the 19th EGU General Assembly, EGU2017*. Vienna, pp 7153
- Eade R, Smith D, Scaife A, Wallace E, Dunstone N, Hermanson L, Robinson N (2014) Do seasonal-to-decadal climate predictions underestimate the predictability of the real world? *Geophys Res Lett* 41(15):5620–5628
- Fetterer F, Knowles K, Meier W, Savoie M (2002) Sea ice index. National Snow and Ice Data Center, Boulder
- Folland CK, Knight J, Linderholm HW, Fereday D, Ineson S, Hurrell JW (2009) The summer North Atlantic Oscillation: past, present, and future. *J Clim* 22(5):1082–1103
- Gastineau G, Frankignoul C (2015) Influence of the North Atlantic SST variability on the atmospheric circulation during the twentieth century. *J Clim* 28(4):1396–1416
- Giorgetta MA, Jungclaus J, Reick CH, Legutke S, Bader J, Böttinger M, Brovkin V, Crueger T, Esch M, Fieg K et al (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *J Adv Model Earth Syst* 5(3):572–597
- Ho CK, Hawkins E, Shaffrey L, Bröcker J, Hermanson L, Murphy JM, Smith DM, Eade R (2013) Examining reliability of seasonal to decadal sea surface temperature forecasts: the role of ensemble dispersion. *Geophys Res Lett* 40(21):5770–5775
- Hodson DL, Sutton RT, Cassou C, Keenlyside N, Okumura Y, Zhou T (2010) Climate impacts of recent multidecadal changes in Atlantic Ocean sea surface temperature: a multimodel comparison. *Clim Dyn* 34(7–8):1041–1058
- Hoskins BJ, Ambrizzi T (1993) Rossby wave propagation on a realistic longitudinally varying flow. *J Atmos Sci* 50(12):1661–1671
- Hoskins BJ, Karoly DJ (1981) The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J Atmos Sci* 38(6):1179–1196
- Hurrell JW (1995) Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science* 269(5224):676–679
- Hurrell JW (1996) Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature. *Geophys Res Lett* 23(6):665–668
- Iglesias I, Lorenzo MN, Taboada JJ (2014) Seasonal predictability of the East Atlantic pattern from sea surface temperatures. *PLoS One* 9(1):86439
- Jungclaus J, Keenlyside N, Botzet M, Haak H, Luo JJ, Latif M, Marotzke J, Mikolajewicz U, Roeckner E (2006) Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM. *J Clim* 19(16):3952–3972
- Lau NC, Nath MJ (2001) Impact of ENSO on SST variability in the North Pacific and North Atlantic: seasonal dependence and role of extratropical sea-air coupling. *J Clim* 14(13):2846–2866
- Li J, Ruan C (2018) The North Atlantic-Eurasian teleconnection in summer and its effects on Eurasian climates. *Environ Res Lett* 13(2):024007
- Marshall J, Kushnir Y, Battisti D, Chang P, Czaja A, Dickson R, Hurrell J, McCARTNEY M, Saravanan R, Visbeck M (2001) North Atlantic climate variability: phenomena, impacts and mechanisms. *Int J Climatol* 21(15):1863–1898
- Michelangeli PA, Vautard R, Legras B (1995) Weather regimes: recurrence and quasi stationarity. *J Atmos Sci* 52(8):1237–1256
- Müller W, Jungclaus J, Mauritsen T, Baehr J, Bittner M, Budich R, Bunzel F, Esch M, Ghosh R, Haak H et al (2018) A higher-resolution version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *J Adv Model Earth Syst* 10(7):1383–1413
- North GR, Bell TL, Cahalan RF, Moeng FJ (1982) Sampling errors in the estimation of Empirical Orthogonal Functions. *Mon Weather Rev* 110(7):699–706
- Palmer TN, Anderson DL (1994) The prospects for seasonal forecasting—a review paper. *Q J R Meteorol Soc* 120(518):755–793
- Saeed S, Van Lipzig N, Müller WA, Saeed F, Zanchettin D (2014) Influence of the circumglobal wave-train on European summer precipitation. *Clim Dyn* 43(1–2):503–515
- Stevens B, Giorgetta M, Esch M, Mauritsen T, Crueger T, Rast S, Salzmann M, Schmidt H, Bader J, Block K et al (2013) Atmospheric component of the MPI-M Earth System Model: ECHAM6. *J Adv Model Earth Syst* 5(2):146–172
- Tietsche S, Notz D, Jungclaus J, Marotzke J (2013) Assimilation of sea-ice concentration in a global climate model—physical and statistical aspects. *Ocean Sci* 9(1):19
- Wallace JM, Gutzler DS (1981) Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon Weather Rev* 109(4):784–812
- Wilks DS (2011) *Statistical methods in the atmospheric sciences*, vol 100. Academic, New York

Wu B, Lin J, Zhou T (2016) Interdecadal circumglobal teleconnection pattern during boreal summer. *Atmos Sci Lett* 17(8):446–452

Wulff CO, Greatbatch RJ, Domeisen DI, Gollan G, Hansen F (2017) Tropical forcing of the Summer East Atlantic pattern. *Geophys Res Lett* 44:21

Yasui S, Watanabe M (2010) Forcing processes of the summertime circumglobal teleconnection pattern in a dry AGCM. *J Clim* 23(8):2093–2114

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.