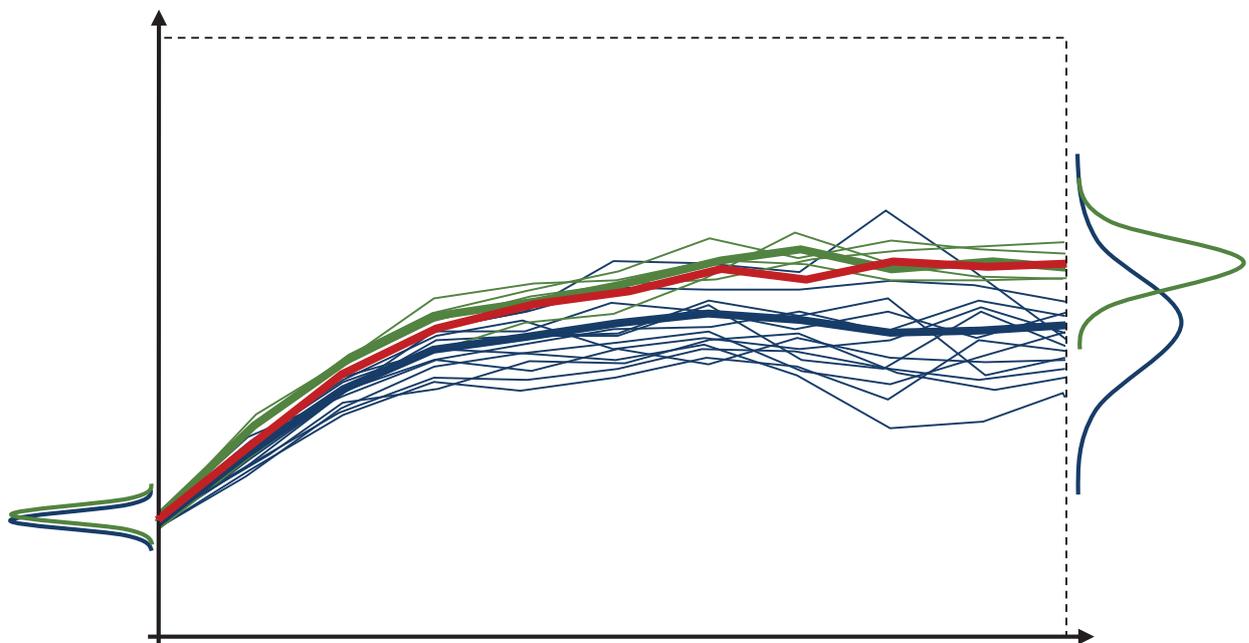




Seasonal prediction of European summer climate: A process-based approach



Nele-Charlotte Neddermann

Hamburg 2019

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Layout

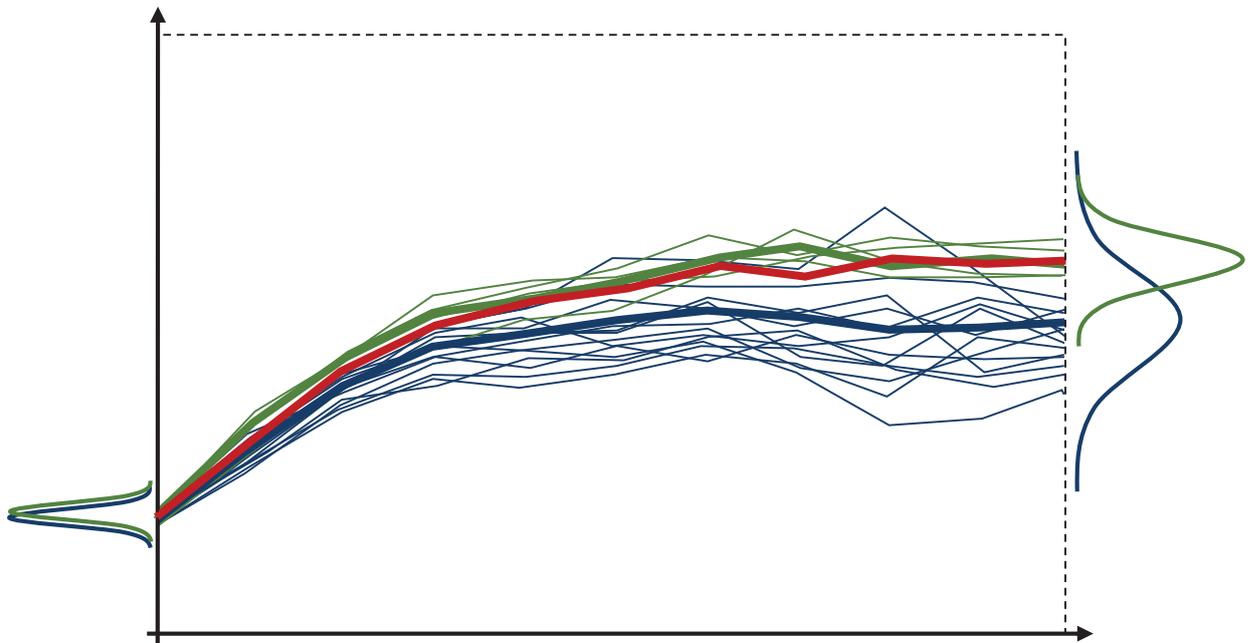
Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Photos below: ©MPI-M
Photos on the back from left to right:
Christian Klepp, Jochem Marotzke,
Christian Klepp, Clotilde Dubois,
Christian Klepp, Katsumasa Tanaka



Seasonal prediction
of European summer climate:
A process-based approach



Nele-Charlotte Neddermann

Hamburg 2019

Nele-Charlotte Neddermann

from Bielefeld, Germany

Max-Planck-Institut für Meteorologie
The International Max Planck Research School on Earth System Modelling
(IMPRS-ESM)
Bundesstrasse 53
20146 Hamburg

Universität Hamburg
Geowissenschaften
Meteorologisches Institut
Bundesstr. 55
20146 Hamburg

Tag der Disputation: 29. Oktober 2019

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Johanna Baehr
Dr. Wolfgang A. Müller

Vorsitzender des Promotionsausschusses:

Prof. Dr. Dirk Gajewski

Dekan der MIN-Fakultät:

Prof. Dr. Heinrich Graener

ABSTRACT

Seasonal climate predictions show very limited skill over Europe, especially for the summer season. Those predictions are usually generated in ensembles and the skill is assessed as the mean over all ensemble members. Most scientific studies expect an increase in skill with an increase in ensemble size. However, the ensembles can spread out with increasing lead time, such that seasonal climate predictions over Europe show a high spread and an ensemble mean with low variability and values around the climatological mean.

Here, I show a way to refine an ensemble by grouping the members according to the physical process they represent. For this, I assess which processes dominate the climate of individual European summers and confirm that the dominant seasonal process can be explained by either a meridional or a zonal pressure gradient, in their positive or negative phase. The evaluated dynamical seasonal climate prediction model is able to represent the spatial pattern and overall frequency of occurrence of the assessed processes, but the individual members disagree on the process they predict for each summer. I thus show that the high spread of the ensemble results from the ensemble members predicting a variety of physical processes for European summers. A mean taken over all those members thus averages over different physical processes, which is not physically consistent. For a physically consistent prediction, I restrict the ensemble mean to those members, that predict the dominant physical process in each summer, which is obtained through observations. With such a refinement, significant hindcast skill can be achieved over many parts of Europe and the North Atlantic, showing that the model is capable of predicting European summers if the physical processes are considered.

In line with such a process-based approach I, instead of using observations to obtain the dominant physical process in each summer, show an alternative way in which I am able to predict the zonal pressure pattern and its teleconnections. I assess these connections in the ensemble through a chain of physical relations based on the process. I then refine the ensemble by choosing only those members that represent all proposed connections. A mean over only the chosen members then also leads to improved spread and significant skill over central Europe.

I show in this dissertation that maintaining the ensemble's physical consistency by focusing on physical processes leads to improved skill in the areas the processes are influencing. This process-based approach could be extended for further regions that are also influenced by several processes and for which the ensemble shows a large spread and allows for improved predictions in those regions.

ZUSAMMENFASSUNG

Saisonale Vorhersagen für Europa zeigen eine sehr geringe Vorhersagegüte, besonders für die Sommersaison. Die Vorhersagen werden üblicherweise in Ensembles erstellt und deren Vorhersagegüte dann als Mittel über alle Ensembleläufe evaluiert. Die meisten wissenschaftlichen Studien erwarten eine Verbesserung der Vorhersagegüte mit steigender Anzahl der Ensembleläufe. Die Streuung eines solchen Ensembles kann jedoch mit steigender Vorhersagezeit zunehmen. Saisonale Klimavorhersagen über Europa weisen deshalb eine hohe Streuung auf und resultieren somit in einem Ensemblemittel mit geringer Variabilität und Werten um das klimatologische Mittel.

Ich zeige hier, dass ein solches Ensemble präzisiert werden kann, indem ich die Läufe nach dem physikalischen Prozess gruppriere, den sie repräsentieren. Dazu analysiere ich, welche Prozesse einzelne europäische Sommer dominieren und belege, dass der dominante saisonale Prozess entweder durch einen meridionalen oder einen zonalen Druckgradienten, in jeweils ihrer positiven oder negativen Phase, erklärt werden kann. Das evaluierte dynamische saisonale Klimavorhersagemodell ist in der Lage, die räumlichen Strukturen und die generelle Häufigkeit des Auftretens der analysierten Prozesse darzustellen. Die einzelnen Läufe sagen jedoch unterschiedliche Prozesse als dominant für denselben Sommer voraus. Ich zeige damit, dass die hohe Streuung des Ensembles darauf zurückzuführen ist, dass die Ensembleläufe eine Vielzahl von physikalischen Prozessen für einen europäischen Sommer voraussagen. Ein Mittel über all diese Läufe bedeutet dementsprechend ein Mittel über verschiedene physikalische Prozesse zu nehmen und ist daher physikalisch nicht konsistent. Für eine physikalisch konsistente Vorhersage beschränke ich die Vorhersage auf die Läufe, die den dominanten physikalischen Prozess vorhersagen. Der dominante physikalische Prozess wird dabei den Beobachtungen entnommen. Mit einer solchen Präzision des Ensembles kann in weiten Teilen Europas und des Nordatlantiks eine signifikante Vorhersagegüte erreicht werden. Dies zeigt, dass das Modell in der Lage ist, europäische Sommer vorherzusagen, wenn die physikalischen Prozesse in der Vorhersage berücksichtigt werden.

Im Sinne eines solchen prozessbasierten Ansatzes zeige ich, anstatt Beobachtungen zu verwenden, um den dominanten physikalischen Prozess in jedem Sommer zu erhalten, einen alternativen Weg, bei dem ich in der Lage bin, das zonale Druckmuster und seine Telekonnektionen vorherzusagen. In den Ensembleläufen verwende ich dazu eine Kette von physikalischen Verbindungen, die auf dem Prozess basieren. Ich präzisiere das Ensemble, indem ich nur diejenigen Läufe

auswähle, die alle aufgezeigten Verbindungen aufweisen. Ein Mittel über nur diese ausgewählten Läufe führt ebenfalls zu einer verbesserten Streuung und signifikanter Vorhersagegüte über Mitteleuropa.

Ich zeige in dieser Dissertation, dass die Aufrechterhaltung der physikalischen Konsistenz eines Ensembles, durch Fokussieren auf physikalische Prozesse, zu einer verbesserten Vorhersagbarkeit in den Regionen führt, die von diesen Prozessen beeinflusst werden. Dieser prozessbasierte Ansatz könnte auf weitere Regionen ausgedehnt werden, die ebenfalls durch mehrere Prozesse beeinflusst werden und für die das Ensemble ebenfalls eine große Streuung aufweist und ermöglicht verbesserte Vorhersagen in diesen Regionen.

PARTS OF THIS DISSERTATION PRE-PUBLISHED OR
INTENDED FOR PUBLICATION

CHAPTER 2

Neddermann, N.-C., Müller, W.A., Düsterhus, A., Pohlmann, H., and Baehr, J., Dependence of seasonal hindcast skill on different mechanisms influencing European summers throughout the 20th century. *In Preparation*.

CHAPTER 3

Neddermann, N.-C., Müller, W.A., Dobrynin, M., Düsterhus, A., and Baehr, J. (2019), Seasonal predictability of European summer climate re-assessed. In: *Climate Dynamics*, 53.5, pp. 3039-3056.

CONTENTS

1	OVERCOMING CHALLENGES IN SEASONAL PREDICTION OF EUROPEAN SUMMER CLIMATE	1
1.1	Challenges in seasonal prediction of European summer climate	1
1.2	Physical processes driving European summer climate	6
1.3	Physical processes in an ensemble-based seasonal climate prediction system	7
1.4	Hindcast skill using a process-based approach	10
1.5	Process-based prediction in a real forecast setup	11
1.6	Process-based prediction in the future	13
2	THE INFLUENCE OF DIFFERENT PROCESSES ON SEASONAL PREDICTABILITY OF EUROPEAN SUMMER CLIMATE THROUGHOUT THE 20 TH CENTURY	15
2.1	Introduction	17
2.2	Model and Data	19
2.2.1	Reanalysis Data	19
2.2.2	Model Setup	19
2.2.3	Analysis	20
2.3	Spatio-temporal variability of the clusters	21
2.4	Seasonal Hindcast skill	25
2.4.1	Cluster dependency of the hindcast skill	27
2.4.2	Temporal variability of the hindcast skill	29
2.4.3	Grouping of clusters	31
2.5	Discussion	33
2.6	Summary and Conclusions	36
3	SEASONAL PREDICTABILITY OF EUROPEAN SUMMER CLIMATE REASSESSED	39
3.1	Introduction	41
3.2	Model and Data	44
3.2.1	Reanalysis Data	44
3.2.2	Model Setup	44
3.2.3	Analysis	45
3.3	Physical Mechanism	47
3.4	Ensemble Subsampling	49
3.5	Seasonal Hindcast Skill	57
3.6	Discussion	63
3.7	Summary and Conclusions	66
	BIBLIOGRAPHY	69

OVERCOMING CHALLENGES IN SEASONAL PREDICTION OF EUROPEAN SUMMER CLIMATE

*Even the longest journey begins
with a first step.*

— Lao Tzu

Decisions in climate-sensitive sectors in Europe, such as agriculture, finance, health, resource management, and energy rely on climate and its variations on seasonal time scales. The decisions-making process in those sectors would profit from accurate seasonal climate forecasts over Europe (e.g., Cantelaube and Terres, 2005; Soares and Dessai, 2015; Lowe et al., 2016; Lledó et al., 2019). However, the skill of seasonal forecasts for European climate compared to observations is still very low in most state-of-the-art prediction systems, especially for the summer season (e.g., Arribas et al., 2011; Kim et al., 2012; Mishra et al., 2018). This lack of skill in seasonal climate predictions hinders decision-makers to prepare for unexpected summer temperature, such as the many record-breaking European summers during the past two decades (e.g., Russo et al., 2014). Those events caused an extraordinary high number of heat-related deaths (e.g., Robine et al., 2008) and an increasing number of crop failure and forest fires, leading to severe financial losses and environmental damages (e.g., Zaneti et al., 2004). Better seasonal forecasts for European summers would be a valuable asset for the decision-making process and would lead to better-informed decisions (Soares and Dessai, 2015).

In this dissertation I address the question why we still lack skill in seasonal forecasts for European summer climate and investigate how better skill could be achieved.

1.1 CHALLENGES IN SEASONAL PREDICTION OF EUROPEAN SUMMER CLIMATE

Seasonal climate prediction refers to the prediction of climate on time scales between one month and one year. Forecasts on shorter time scales are investigated

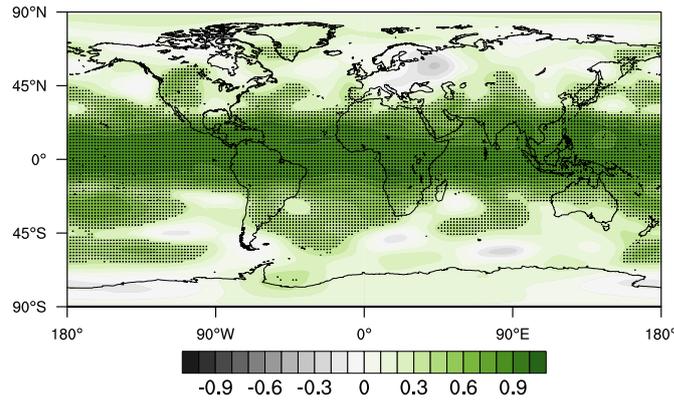


Figure 1.1: Seasonal climate prediction skill for 500 hPa geopotential height (Z_{500}) anomaly in summer (July-August mean) depicted as the Anomaly Correlation Coefficient (ACC) comparing the model predictions of the ensemble mean of the 30 ensemble member generated by MPI-ESM-MR to the ERA-Interim reanalysis in 1982-2016. Increasing positive correlation implies increasing seasonal climate prediction skill. Dots show significant correlation at the 95% confidence level derived via bootstrapping using 500 samples.

in weather and sub-seasonal prediction, longer ones in decadal prediction and climate projection (Doblas-Reyes et al., 2013). Unlike weather prediction, seasonal climate prediction does not attempt to forecast the actual day-to-day progression of the climate system, but instead the evolution of climate as a seasonal average (Council et al., 2010).

Seasonal climate prediction considers the various components of the climate system - the atmosphere, ocean, land and cryosphere (e.g., Shukla and Kinter III, 2006). These components vary on different time scales and react to perturbations with different response times. The different response time scales can be seen as the memory of climate processes. While the atmosphere has memory of only a few days, the memory of sea ice, land and ocean can last up to several years (e.g., Boer et al., 2016).

Skill in seasonal climate prediction generally arises from the interaction of the atmosphere with the components of the climate system that have a longer memory and are thus predictable, such as the ocean, land and cryosphere. Those components of the climate system store heat and moisture, and the response of the atmosphere that interacts with them through coupling is a premise for seasonal climate predictability (Palmer and Anderson, 1994).

Even given the memory of the climate system, the quality of forecasts vary substantially. It is therefore necessary to measure the skill of a forecast. The skill of seasonal climate prediction is assessed relative to some reference, preferably observations. Since we do not know how the future evolves, predictions are performed for the past in so-called hindcasts, such that they can be compared to

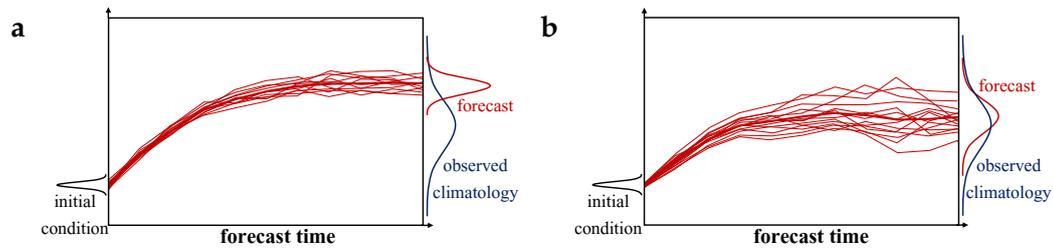


Figure 1.2: A schematic illustrating an ensemble forecast. An ensemble of forecasts (thin red lines) for any variable starts from slightly different initial condition of this variable from where it naturally spreads out with increasing forecast time. Ensemble forecasts are usually assessed by the ensemble mean (thick red lines), which is the mean over all ensemble members. **a** Optimally the spread of the ensemble results in a probability density function (PDF) that is clearly distinguishable from the PDF of the observed climatological values. **b** In some forecasts the ensemble spread results in a PDF close to that of the observed climatology.

existing records of observed past data. The similarity between past observations and the hindcast can then be evaluated as so-called hindcast skill.

Seasonal hindcasts performed with the Max Planck Institute Earth System Model (MPI-ESM) are skilful mainly in the tropical and ocean areas, but less skilful in the extra-tropical and land regions (Fig. 1.1). Due to the persistence of the climate in the tropics, tropical regions are predictable for a longer lead time (i.e. for more time in advance) than the extra-tropical regions (Palmer and Anderson, 1994). Predictability in the extra-tropical regions is most often associated with connections to the predictable tropics through teleconnections, that connect variability in the climate on seasonal timescales (e.g., Shukla et al., 2000). Teleconnections are physical processes that must be represented by the climate model, so that it is able to generate predictability. Seasonal climate prediction skill thus depends on how well the relevant physical processes are represented in a prediction model.

To account for the different physical pathways that a climate simulation could take, a prediction is usually assessed in an ensemble. The climate system is chaotic and thus very sensitive to small perturbations in the initial state of the system. Ensemble forecasting takes into account the sensitivity of the model to initial conditions by running a large number of realisations of the model started from slightly different initial conditions (Fig. 1.2). The resulting different forecasts show the many different possible physical pathways of the climate system. Seasonal climate prediction then results in a range of possible forecasts, the spread of the ensemble.

Seasonal climate prediction aims at forecasting an anomaly from the observed climatological mean of the assessed season (Council et al., 2010). The spread of the ensemble is seen as the uncertainty of the prediction (e.g., Ho et al., 2013). In an optimal ensemble prediction, the individual ensemble members agree on the

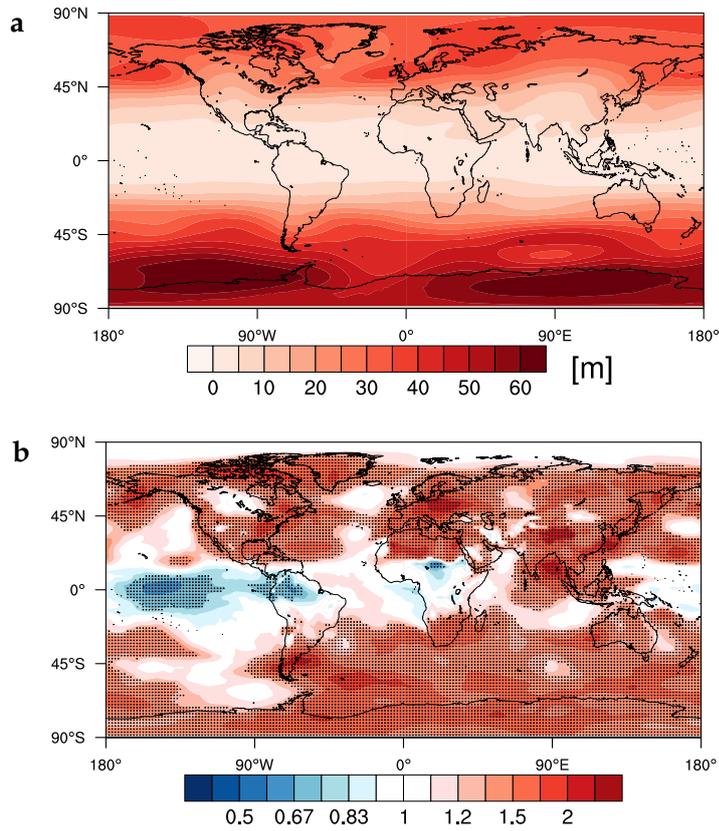


Figure 1.3: **a** The spread of the 30 ensemble members generated by the MPI-ESM-MR depicted as the average spread of all summers (July-August mean) in 1982-2016 for 500 hPa geopotential height (Z_{500}) anomaly. **b** Spread-error-ratio for Z_{500} anomaly in summer, derived as the ratio between the average spread of all ensemble members and the root-mean-square-error of the mean over all 30 ensemble members in comparison to the ERA-Interim reanalysis in 1982-2016. Optimally, this ratio is equal to one. Dots indicate areas where the ratio is significantly different from one at the 5% level.

physical pathway resulting in an ensemble spread that is clearly distinguishable from the probability density function (PDF) of the observed climatology (Fig. 1.2a). However, in many forecasts, the individual ensemble members predict several different physical pathways resulting in an ensemble prediction with a larger spread that is not necessarily distinguishable from the PDF of the observed climatology (Fig. 1.2b). Forecasts are usually assessed by the mean over all ensemble members. But, if a mean is taken over those members with a PDF close to the climatological PDF, the ensemble mean prediction results in a mean value close to the observed climatological mean.

Comparing the spread of the seasonal forecast to the skill of the ensemble mean prediction (Figs. 1.3a and 1.1), we see that most areas of significant hindcast skill overlap with areas of low spread in the ensemble. Therefore, significant hindcast skill is mainly achieved in areas in which ensemble members do not spread out much and thus presumably agree on the physical pathways. Since the spread of an ensemble is a measure of the uncertainty of the forecast, in a reliable forecast this uncertainty is about equal to the error of the forecast (Ho et al., 2013), which is assessed as the difference between the ensemble mean and the observations. Thus, the spread-error-ratio of the ensemble is ideally equal to one.

Analysing this ratio shows that, especially over Europe, values larger than one are obtained, demonstrating that the spread of ensemble predictions over Europe is too high (Fig. 1.3b), as was shown by previous studies as well (e.g., Ho et al., 2013; Eade et al., 2014; Dobrynin et al., 2018). Seasonal ensemble hindcasts for European summer climate thus show a spread that is higher than its error (cf. Fig. 1.3b) and a hindcast skill that is not significant (cf. Fig. 1.1).

Previous studies showed that seasonal European summer climate is influenced by various physical processes (e.g., Cassou et al., 2005). Individual ensemble members could thus predict the different physical processes for European summers. A mean taken over all ensemble members thus averages over different physical processes. Such a prediction is not physically consistent and the signal of the physical processes in the prediction is lost.

In this dissertation, I refine the ensemble prediction for seasonal European summer climate by concentrating on the different physical processes that influence this region on seasonal timescales. I assess the high spread of the ensemble and the physical pathways that are predicted by individual ensemble members. I propose a physically consistent prediction by grouping the ensemble members according to the physical process they represent.

This cumulative thesis is structured into two individual articles attached in Chapters 2 and 3. In the first article, I analyse the processes that drive European summer climate on seasonal time scales and assess if those processes are represented in the considered seasonal climate prediction system. I investigate if the seasonal climate prediction skill of the model is influenced by the dom-

inant physical processes and how the skill is affected if the governing process is considered in the analysis. The second article then focuses on one of those processes, which is the second leading mode of summertime pressure variability over Europe. I assess how this process influences the seasonal climate prediction skill of MPI-ESM-MR.

In the following, I give an overview over those two articles and guide through my research by asking four questions, answering these questions along the way. I put them in the scientific context and finish with an outlook concerning the findings of this dissertation.

1.2 PHYSICAL PROCESSES DRIVING EUROPEAN SUMMER CLIMATE

The lack of skill in seasonal climate predictions often results from a lack in understanding the relevant physical processes (Doblas-Reyes et al., 2013). In the last decades a lot of emphasis has been put into the understanding of the physical processes that influence seasonal climate variability (Wang et al., 2009). For Europe, most previous studies agree that the seasonal European winter climate is primarily driven by the North Atlantic Oscillation (NAO) (e.g., Hurrell, 1995), which explains up to 50% of variability and thus dominates almost all winters (e.g., Hurrell et al., 2003). In summer, on the other hand, low-frequency variability modes influencing Europe on seasonal timescales have been less extensively studied (e.g., Hannachi et al., 2017), partly because they are not as easily separable as in winter and appear to be less pronounced (e.g., Hurrell and Deser, 2010; Cattiaux et al., 2013). In summer, the NAO is much less distinct, explaining around 30% of variability (e.g., Bladé et al., 2012). The second mode of atmospheric variability that explains about 20% of the total variance (e.g., Saeed et al., 2014; Wulff et al., 2017; Neddermann et al., 2019) is thus almost as pronounced as the NAO in summer. Finding the mechanism that is dominant in an individual summer is therefore a more complex problem than in winter and has not yet been investigated throughout the entire 20th century.

However, assessing the dominant process per summer in observational data is necessary to later relate the processes of European summers to their prediction and to conduct a physically consistent forecast. Therefore, I concentrate on the following first research question:

1. Which physical processes are dominating European summer climate in individual summers throughout the entire 20th century?

I use the ERA-20C reanalysis data set to address this question for the period 1900-2010. As opposed to many previous studies that focus on mechanisms influencing European in the more recent decades, I investigate the entire 20th

century to be able to better differentiate between the different physical processes in individual years.

Most previous studies further assume the spatial patterns of the processes influencing Europe on seasonal timescales to be constant over time (e.g., Folland et al., 2009). A recent study by Wang et al. (2012), however, discloses that the spatial pattern of the winter NAO shows multidecadal variations. To account for such pattern variations in summer as well, I apply a cluster analysis that allows for the spatial patterns of the mechanisms to change over time.

Unlike previous work that investigated cluster analysis for European summers based on daily data (e.g., Cassou et al., 2005; Cattiaux et al., 2013), I analyse seasonal means and can thus identify the one dominating cluster per summer. With this analysis, I find that the dominant seasonal European summer process can be described by the two governing mechanisms of European summer climate in their positive and negative phase - a meridional pressure gradient, known as the NAO, and a more zonal pressure gradient (PD), with similarities to the East Atlantic Pattern (e.g., Wallace and Gutzler, 1981; Barnston and Livezey, 1987). Both mechanisms exhibit asymmetries in their positive and negative phases, resulting in four different clusters. For every summer from 1900-2010, I identify which of those clusters dominates a particular summer. I find that the first half of the 20th century is dominated by the NAO in its positive and the PD in its negative phase, while the second half of the century is dominated by the NAO in its negative and the PD in its positive phase (Fig. 2.2a).

1.3 PHYSICAL PROCESSES IN AN ENSEMBLE-BASED SEASONAL CLIMATE PREDICTION SYSTEM

The correct representation of the observed processes and their spatial patterns and persistence properties is essential for a model to properly simulate the climate variability and its long term changes (Cattiaux et al., 2013). The representation of mechanisms in prediction models has largely been discussed on the weather timescale and shown, that the prediction skill and ensemble spread on those timescales depends on the mechanism that is predicted by the model (e.g., Ferranti et al., 2015; Matsueda and Palmer, 2018). However, the representation of such patterns on seasonal timescales and in a seasonal climate prediction model has been less intensively studied.

Here, I work with the ensemble-based fully-coupled seasonal climate prediction system based on the Max Planck Institute Earth System Model at mixed resolution (MPI-ESM-MR) to answer the second research question:

2. Does the MPI-ESM-MR represent the processes that dominate European summer climate on seasonal timescales?

Most former studies restrict their hindcast analysis to the last decades. Some recent studies consider hindcasts throughout the entire 20th century, but focus solely on the winter season (e.g., Weisheimer et al., 2017; O'Reilly et al., 2017; Weisheimer et al., 2018). Here, I consider hindcast runs in the fully coupled model MPI-ESM-MR in the summer season. For this, 10 ensemble members were generated starting every May from 1902-2008.

I assign every ensemble member in every year to one of the four observed clusters using a pattern matching algorithm in terms of the root mean square error. I show that the MPI-ESM-MR is able to represent the spatial pattern of the four identified clusters (Fig. 2.1). The model also agrees with the overall frequency of the four clusters if all ensemble members in all years are considered. However, the predicted cluster only agrees with the observed cluster in one third of the analysed years. Thus, the model predicting the dominant cluster per year only slightly outperforms a prediction by chance. And while I show in the reanalysis that one cluster is dominating a particular summer, the 10 ensemble members generated by the model for every summer disagree on the cluster and thus the physical process they predict for individual summers (Fig. 2.2b). Therefore, the MPI-ESM-MR is able to reproduce the overall spatial patterns and the frequency of the four dominant clusters of seasonal European summers, but is not able to predict the timing of the clusters for individual summers.

Seasonal climate predictions are usually assessed by taking the mean over all ensemble members. However, if different ensemble members predict different physical processes, a mean taken over those members means to take an average over different processes. Therefore, taking the mean over the full ensemble is not physically consistent and results in a prediction in which the signal of individual processes is lost (Fig. 1.4a,b). Since each cluster influences the pressure and temperature over Europe in a different way (Figs. 2.1 and 2.6a-d), the values predicted by the individual ensemble members differ strongly if they predict a different cluster for a particular summer. The disagreement of the predicted clusters then results in highly spread predicted values over Europe for all years. A mean taken over those values then produces an ensemble mean prediction close to the climatological mean with a variability that is considerably smaller than observed (Fig. 2.3b).

For a physically consistent prediction I thus suggest to group the ensemble members according to the physical process they predict in individual years and to take a new mean only over those members that predict the same process. Such a mean then results in a prediction in which the signal of the individual processes is much more pronounced with a magnitude close to that of observations (Fig. 1.4c).

I thus find that the assessed model is predicting different physical processes for individual summers and that those processes are better represented in the

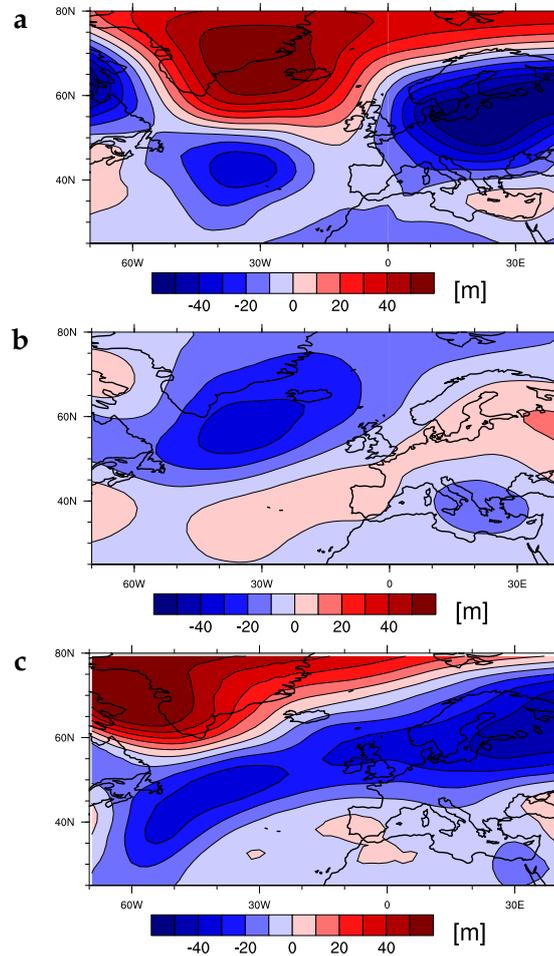


Figure 1.4: 500 hPa geopotential height (Z_{500}) anomaly with respect to the climatological mean for the summer (July-August mean) in 1978. **a** In the ERA-20C reanalysis the anomaly shows the spatial pattern of the North Atlantic Oscillation in its negative phase (NAO-), which was the observed dominant mechanism in that year (Fig. 2.2a). **b** The mean over all 10 ensemble members generated by MPI-ESM-MR is predicting a different spatial pattern of lower magnitude. The low magnitude likely results from the individual members predicting different mechanisms for that year (Fig. 2.2b). Especially over Europe this results in anomaly values of low magnitude close to the climatological mean which disagree with the observed value. **c** The mean over only the ensemble members that predict a NAO- and thus the dominant mechanism that is observed results in a spatial pattern and predicted values over Europe that are much closer to the observation.

MPI-ESM-MR if a mean is not taken over all ensemble members, but only over those members that predict the dominant process in the observations.

1.4 HINDCAST SKILL USING A PROCESS-BASED APPROACH

The skill of a model is usually derived by comparing the ensemble mean prediction to observations. This approach does not lead to significant seasonal skill for European summers. I showed that the assessed model is predicting multiple physical processes for one individual summer, such that the ensemble mean is a mean over multiple processes that results in values for European summers that differ from those observed. A physically more consistent prediction thus averages over members reproducing the same physical process, which is also a way of refining the ensemble.

Most previous studies assume that the skill of a model increases with an increasing number of ensemble members (e.g., Murphy, 1990; Scaife et al., 2014). Taking a mean over only the ensemble members that represent one physical process, however, means to decrease the ensemble size, raising the question if such an approach can lead to meaningful hindcast skill. I thus ask the third research question:

3. How is the hindcast skill of MPI-ESM-MR affected by refining the ensemble using a process-based approach?

The dominant cluster per summer can be either found in the model, through the cluster that is predicted by the model, or in the observations, through the cluster that is observed in ERA-20C in each summer. I find that the cluster that is predicted by the model does not agree with the observed cluster in two thirds of the years. A new ensemble mean over only those ensemble members in the predicted dominant cluster is thus not leading to skill (Fig. 2.3c). However, if an ensemble mean is taken over the members in the observed dominant cluster, significant hindcast skill can be obtained in large parts over Europe and the North Atlantic (Fig. 2.3e). Additionally, the spread of the ensemble is increased for values over Europe resulting in a variability comparable to that in observations (Fig. 2.3f).

I thus show that the low skill and low variability of seasonal European summer predictions can be related to the ensemble members disagreeing on the dominant cluster for individual years. I demonstrate that including the dominant mechanisms of seasonal summer climate into the seasonal hindcast analysis of the North-Atlantic-European sector is key to improving seasonal hindcast skill and variability of hindcasted European summers.

The two considered mechanisms, the NAO and the PD both in their positive and negative phase, explain about 50% of the variance of seasonal European

summers. Potentially, further mechanisms could be investigated and included in the prediction analysis. However, it is unclear if those mechanisms would be represented by the assessed model and if the hindcast skill could be further improved. Including more mechanisms in the analysis would also require more ensemble members than the 10 members considered so far. I find, however, that even though further mechanisms influence European summers on seasonal timescales, the hindcast skill can already be remarkably improved if only the first two leading mechanisms are considered in a hindcast analysis.

1.5 PROCESS-BASED PREDICTION IN A REAL FORECAST SETUP

I showed that it is possible to improve seasonal hindcast skill by shaping the processes into the prediction through grouping of ensemble members according to the process they predict. However, I so far only examined hindcast analysis and thus retro-perspective forecasts, in which the dominating mechanism was known in advance through observations and could therefore be included in the analysis of the ensemble members. To apply this method in a real forecast setup, I would need to be able to predict the dominant mechanisms, leading me to the fourth and last research question:

4. **Can process-based seasonal climate prediction assessment be operated in a real forecast setup?**

So far I concentrated on two mechanisms in their positive and negative phase: the NAO and the PD. Many previous studies analyse the NAO and try to predict its phase (e.g., Butler et al., 2016; O'Reilly et al., 2017; Weisheimer et al., 2017; Dobrynin et al., 2018). While most of those studies concentrate on the NAO in winter, Düsterhus et al. (under review) show that the phase of the NAO in summer can be predicted as well. However, little attention has been paid to the PD, the second mode of seasonal variability in Europe, and its predictability in summer.

Seasonal climate predictability arises from persistent and slowly-evolving boundary forcing, most often found in the tropics (e.g., Palmer and Anderson, 1994). Prediction skill in extra-tropical regions like Europe then emerges from those predictable regions through atmospheric teleconnections (e.g., Shukla et al., 2000). In the North-Atlantic-European sector, the North Atlantic is the major source of low-frequency climate variability (e.g., Marshall et al., 2001; Cattiaux et al., 2013). In summer, warm sea surface temperature (SST) in the tropics excite a Rossby wave in the northern hemisphere that is known as the circumglobal wave-train (CGT, Branstator (2002)). In accordance with recent studies by Saeed et al. (2014) and Wulff et al. (2017), I show in the ERA-Interim reanalysis for 1982-2016 that this CGT acts as a teleconnection between the tropical SST in the

North Atlantic and the PD, that in turn influences seasonal European climate (Fig. 3.1). Unlike previous work that only suggests that this connection could lead to prediction skill (e.g., Ding and Wang, 2005; Gastineau and Frankignoul, 2015), I make use of this connection by predicting the phase of the PD in summer through the phase of the tropical SST observed in ERA-Interim in April.

I apply this knowledge to the 30 ensemble members generated by MPI-ESM-MR initialised every May for 1982-2016. I showed that different ensemble members predict different physical processes to be dominant in individual summers. Using the tropical SST observed in April, I anticipate one phase of the PD to be the dominant mechanism in a particular summer and thus constrain the hindcast analysis to one relevant mechanism only. Instead of including all ensemble members in a prediction I instead only include the members that predict the anticipated phase of the PD. To derive which ensemble member is predicting which phase of the PD, I analyse each member individually and assess if it predicts a chain of known physical relations connected to the PD. This chain of connections is based on the observed relation between the PD and the wave-train, as well as its influence on European summer temperature. For the prediction I then select only the ensemble members that predict all parts of this chain and form an ensemble mean over only those selected members. This selection results in significant seasonal hindcast skill over central Europe (Fig. 3.6). Such a prediction only includes observations from April, which is before the initialisation of the ensemble. Therefore, such a prediction is feasible in a real forecast setup.

I show here a way of refining the ensemble that is based on a known successive chain of physical relations. With this method, I assume the PD to be the dominant mechanism throughout all summers in 1982-2016. This is opposed to my analysis of the dominant mechanisms in ERA-20C for 1900-2010, in which I find that either PD or NAO in its positive or negative phase is dominant. An analysis that combines predicting both the phase of the NAO and the phase of the PD, and which of those mechanisms dominate a certain summer, would thus be useful.

Unlike most previous studies that assume a larger ensemble to result in higher prediction skill (e.g., Scaife et al., 2014; Butler et al., 2016), I show an alternative approach in which the ensemble size is refined on the basis of known physical relations. Refining a predicted ensemble through a process-based approach is a method that combines observed statistic relations with a dynamical model. Unlike other statistical methods, a process-based approach allows to form a new ensemble mean over selected ensemble members for any of the variables simulated by the model. Since entire ensemble members are selected and included in the analysis, the fields derived out of those are dynamically self-consistent in space and time (Dobrynin et al., 2018). This process-based approach thus makes use of statistical connection while maintaining the advantages of a dynamical model.

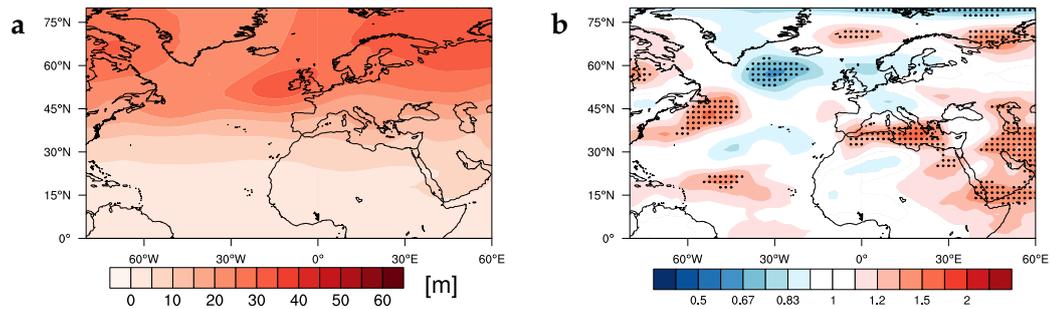


Figure 1.5: Same as Fig. 1.3, but for the refined ensemble and zoomed in for the investigated North-Atlantic-European region.

Selecting individual members using the proposed method decreases the ensemble to on average one fourth of its original size that are eventually analysed for their hindcast skill (Table 3.1). Such a reduction of the ensemble size strongly influences the spread of the ensemble over Europe (cf. Figs. 1.5a and 1.3a). In a reliable ensemble forecast, the spread is comparable to the error of the forecast (Ho et al., 2013). Through the refinement of the ensemble and thus reduction of the ensemble spread, the ratio of the spread to the error of the ensemble over Europe is close to one (Fig. 1.5b) and thus more reliable than the prediction in which the ensemble mean is taken over the full ensemble (Fig. 1.3b).

The optimal size of an ensemble is a widely discussed and an ongoing topic (e.g., Scaife et al., 2014), to which this dissertation constitutes an alternative approach in which a refined ensemble is leading to higher prediction skill than using the full ensemble. Nevertheless should the ensemble from which the refinement is conducted be large enough such that the resulting number of selected ensemble members eventually results in a large enough spread, ideally comparable to the error of the prediction.

1.6 PROCESS-BASED PREDICTION IN THE FUTURE

In this dissertation my study area is Europe and my focus is on mechanisms influencing this area in summer. Seasonal European summer climate is influenced by several processes and its ensemble forecast shows a spread that is considerably larger than the error of the prediction. I show that maintaining the physical consistency of the ensemble forecast by focusing on physical processes in a hindcast analysis leads to improved seasonal hindcast skill in the areas that the processes are influencing, also resulting in a lower spread of the ensemble and thus a more reliable ensemble forecast than a prediction with the full ensemble. I therefore demonstrate that the knowledge of the physical processes is important for a skilful prediction and show that, if the physical processes that influence European summer climate on seasonal time scales are considered in the seasonal

climate prediction, the assessed model is actually capable of predicting European summer climate.

This process-based approach could be extended to further regions that are also influenced by several processes and for which the ensemble prediction shows a too large spread as well. This approach could potentially lead to improved predictions in such regions and I encourage future work to test such an approach.

Predicting European summers is considered to be one of the most complex issues in seasonal climate prediction (e.g., Beverley et al., 2019). The focus of this dissertation is on understanding the physics behind such a prediction in an ensemble-based setup, which is the first step for dissecting the complexity of the issue to predict seasonal European summer climate. While the derived method still needs further refinement to be applicable in an operational seasonal climate prediction system, I show in this dissertation that this method has potential to lead to increased seasonal prediction skill for European summer climate in the future.

THE INFLUENCE OF DIFFERENT PROCESSES ON SEASONAL
PREDICTABILITY OF EUROPEAN SUMMER CLIMATE
THROUGHOUT THE 20TH CENTURY

Nele-Charlotte Neddermann^{1,2}, Wolfgang A. Müller³, André Düsterhus¹,
Holger Pohlmann^{3,4}, Johanna Baehr¹

¹Institute for Oceanography, CEN, Universität Hamburg, Germany

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

³Max Planck Institute for Meteorology, Hamburg, Germany

⁴Deutscher Wetterdienst, Hamburg, Germany

AUTHOR CONTRIBUTIONS N.C.N. designed the research, performed the analysis and drafted the manuscript including all text and figures with guidance from J.B. and W.M.. All authors contributed through discussions on the interpretation of the results. H.P. and W.M. performed the model simulations.

ABSTRACT

We improve seasonal hindcast skill of European summer climate in an idealised ensemble-based hindcast analysis through a refinement of the ensemble by grouping the ensemble members according to the physical process they reproduce. For this, we assess the dominant mechanisms of European summer climate in a coupled seasonal climate prediction system and examine how the different mechanisms influence the hindcast skill throughout the entire 20th century. With a cluster analysis that allows for patterns to vary over time, we analyse seasonal summer means of 500 hPa geopotential height. In the ERA-20C reanalysis we find that the dominant mechanism for individual summers in 1900-2010 can be explained by one of the first two leading modes of seasonal summer climate variability in the North-Atlantic-European sector - a meridional pressure difference, known as the North Atlantic Oscillation, and a zonal pressure difference with similarities to the East Atlantic pattern, both in their positive and negative phases. With this analysis we examine how well the fully coupled MPI-ESM in the mixed resolution setup is able to represent the different clusters. We analyse the hindcasts for 1930-2008 using 10 ensemble members, initialised every year in May. By identifying the different phases of the mechanisms in individual ensemble members, we find that the model is able to reproduce the clusters and their overall frequency. Yet, no hindcast skill for European summers is obtained for the mean over all ensemble members. We show that this low skill results from the individual members predicting different mechanisms for the same summer. As a result, a mean over all members is a mean over different physical processes. We group the ensemble members according to the cluster they reproduce and refine the ensemble in an idealised hindcast analysis in which for every summer only the members in the dominant cluster are considered. With such a refinement, significant hindcast skill can be obtained over large parts of the North-Atlantic-European sector, showing that the considered model is generally capable of predicting European summers.

2.1 INTRODUCTION

Most current state-of-the-art prediction systems show only limited skill for seasonal prediction of European climate in summer (e.g., Mishra et al., 2018; Neddermann et al., 2019). The lack of skill in seasonal climate prediction often results from a lack in understanding the relevant physical processes (Doblas-Reyes et al., 2013) and a recent study by Beverley et al. (2019) showed that the poor skill of a seasonal climate prediction model can be linked to poor representation of the physical processes in the model. Seasonal European summer climate is affected by a range of physical processes (e.g., Cassou et al., 2005) and in this study we assess if this is the reason we still lack seasonal hindcast skill for European summer climate. For this, we examine the ability of an initialised seasonal climate prediction system to reproduce the different mechanisms that influence European summer climate and analyse how the different processes affect the hindcast skill over the 20th century.

Mechanisms influencing European climate and their representation in prediction models for the North-Atlantic-European sector has largely been discussed on the weather timescale (e.g., Corti et al., 2003; Ferranti and Corti, 2011; Dawson et al., 2012). On those synoptic timescales several studies show that the prediction skill and the ensemble spread depend on the mechanism that is predicted by the model (e.g., Ferranti et al., 2015; Matsueda and Palmer, 2018). However, the representation of such patterns in seasonal climate prediction models has been less intensively studied. Especially, how the seasonal climate prediction skill is affected by the model predicting different physical processes over Europe has not yet been assessed. Here, we analyse how the physical patterns that influence European summer climate are represented in an initialised ensemble based seasonal climate prediction system, and if the seasonal hindcast skill itself depends on the different mechanisms.

The skill of an ensemble prediction is usually assessed by the mean over all ensemble members. However, recent studies by Dobrynin et al. (2018) and Neddermann et al. (2019) show that selecting ensemble members according to the mechanism they represent can lead to improved seasonal hindcast skill. Analysing the different mechanisms in a seasonal climate prediction system allows us to group the ensemble members according to the mechanism they represent. Instead of taking an ensemble mean over all ensemble members, a new mean can be formed over the ensemble members reproducing the same mechanism. A mean over ensemble members reproducing different mechanisms averages over different physical processes, and therefore the signal of individual processes is lost. Therefore, a mean over members reproducing just one mechanism is physically more consistent.

The skill of a model is assumed to increase with an increasing number of ensemble members (e.g., Murphy, 1990; Scaife et al., 2014). Taking a mean over only the ensemble members that represent one physical process however means to decrease the ensemble size. Here we analyse if a physically more consistent analysis by a refinement of an ensemble based on the physical process the model represents, is leading to improved seasonal skill for European summer climate.

Former studies that assess the physical mechanisms influencing European summer climate usually focus on the North Atlantic Oscillation (NAO), which is the dominant mechanism of pressure variability in the North Atlantic sector in summer and explains around 30% of the total variance (e.g., Hurrell et al., 2003; Bladé et al., 2012). The spatial structure of the NAO consists of a meridional pressure gradient with one pressure center over Greenland and the other over north-western Europe (Folland et al., 2009). Recent studies also focus on the second leading mode of atmospheric variability over Europe and the North Atlantic that explains up to 20% of the total variance in summer (e.g., Saeed et al., 2014; Wulff et al., 2017; Neddermann et al., 2019). Its spatial pattern reveals a zonal pressure gradient with centers over the North Atlantic and Europe, while the spatial structure appears to be asymmetric for its positive and negative phase (Cassou et al., 2005). Further mechanisms influencing summer climate over Europe on seasonal time-scales have received little attention in the literature so far.

Most previous studies are restricted to the more recent decades and assume the spatial patterns of the mechanisms to be constant in time. Nevertheless, some recent studies show that the center of highest variability of the winter NAO is shifted on decadal timescales (e.g., Jung et al., 2003; Wang et al., 2012). In our investigation of the mechanisms in the 20th century, we allow for the spatial patterns to change over time.

Other recent studies investigate if the seasonal hindcast skill of the winter NAO is time dependent (e.g., Weisheimer et al., 2017; O'Reilly et al., 2017; Weisheimer et al., 2018). They show that the hindcast skill of the NAO during winter strongly varies over time in dependence on its dominant phase. Hindcast skill is found to be higher during positive NAO phases, indicating that the hindcast skill does depend on the dominating mechanisms, at least for European winters. Time dependency of the seasonal hindcast skill however has not yet been investigated for European summers. We thus assess how the hindcast skill for European summers changes throughout the 20th century and how the skill depends on the dominating mechanism.

To address the questions whether seasonal climate prediction in summer depends on different mechanisms influencing Europe and to further also examine the variability of the forecast skill over time, we have performed a long coupled

seasonal forecast experiment with 10 ensemble members covering all summers from 1902-2008.

The used data are described in section 2.2. We determine which mechanisms dominate European summer climate in individual summers throughout the entire 20th century and whether those mechanisms are represented in the assessed initialised ensemble based seasonal climate prediction model in section 2.3. In section 2.4, we investigate how those mechanisms influence the hindcast skill and the variability of the skill over time. The results are discussed in section 2.5, followed by the summary and conclusions in section 2.6.

2.2 MODEL AND DATA

2.2.1 *Reanalysis Data*

We use the European Centre for Medium-Range Weather Forecast (ECMWF) reanalysis of the 20th-century (ERA-20C, Poli et al., 2016) monthly-mean fields from 1900 to 2010. To eliminate long-term trends, we linearly detrend the data at each grid point. We consider only monthly anomaly with respect to the climatological mean, that is taken over the same period.

2.2.2 *Model Setup*

The coupled global Max Planck Institute Earth System Model is used in its mixed-resolution configuration (MPI-ESM-MR) to generate initialized seasonal hindcasts for the period 1902-2008. The atmospheric component ECHAM6 (Stevens et al., 2013) has a horizontal resolution of 200 km (1.875°) and 95 vertical levels up to 0.01 hPa, which is coupled to the ocean component MPI-OM (Jungclaus et al., 2006) with a horizontal resolution of 40 km (0.4°) and 40 vertical layers. External forcing is taken from CMIP5 (Giorgetta et al., 2013).

The assimilation experiments are performed by the coupled model, with full-field nudging by Newtonian relaxation towards all atmospheric and ocean levels except in the boundary layer. The atmosphere conditions of vorticity, divergence, three-dimensional temperature and two-dimensional pressures are taken from ERA-20C. In the ocean, three-dimensional daily mean salinity and temperature anomalies are nudged at a relaxation time of approximately 10 days. The ocean state is derived in an ocean run performed with MPI-OM that is forced with the atmospheric variables from ERA-20C (2 m air and dewpoint temperature, precipitation, cloud cover, downward shortwave radiation, 10 m wind speed and surface wind stress). For a spin up, the ocean model performed 5 prior cycles.

The initial conditions are taken from the three-dimensional atmospheric and ocean fields of the assimilation experiments. 10 ensemble members of 6 month

hindcast simulations are initialised in May every year from 1902-2008. The ensemble members are generated by small perturbations of the atmospheric state by disturbing the diffusion coefficient in the uppermost layer.

A similar set of experiments with the NOAA 20th century data has been performed to study the multidecadal variations in the North Atlantic (Müller et al., 2015).

2.2.3 Analysis

In this study "summer" is defined as the July-August (JA) mean, which is in accordance with previous studies by Folland et al. (2009) and Bladé et al. (2012), who point out that for Europe the temporal variability in June deviates from the variability in July and August.

We perform a cluster analysis on the JA mean of ERA-20C 500 hPa geopotential height (Z₅₀₀) anomaly in the domain [25° – 80°N, 70°W–40°E] for the time-period 1900-2010. We chose to assess the mean of each summer instead of daily or monthly data to reduce within-season variations and to identify the one dominant mechanism per summer.

With our cluster approach we allow for the spatial pattern of the clusters to vary over time and can identify the one cluster that dominates each year. For this, a k -means cluster algorithm (Michelangeli et al., 1995) is performed for the first 30 analysed summers (1900-1930) in the ERA-20 reanalysis. The clustering is carried out in the reduced phase space that is defined by the leading 6 Empirical Orthogonal Functions (EOF, North et al., 1982), explaining about 80% of the total variance of the dataset. We reduce the phase space before clustering, because this work focuses on large scale structures and higher order EOFs account mainly for smaller spatio-temporal scales. In agreement with previous works, the optimal partition is obtained for $k = 4$ (e.g., Cassou et al., 2005; Cattiaux et al., 2013). Higher order clustering results in similar regimes with the four dominant patterns split into more than one cluster per mechanism. Each of the four clusters groups the summers with similar spatial patterns. Therefore, their averages represent the most common spatial patterns for 1900-1930.

For the subsequent summers after 1930, every JA mean of ERA-20C Z₅₀₀ anomaly in the region [25° – 80°N, 70°W–40°E] is attributed to the closest cluster via a pattern matching algorithm in terms of the root mean square error (RMSE). For every subsequent summer this pattern matching is only based on the cluster composites of the prior 30 summers, which allows for the spatial pattern of the ERA-20C clusters to vary over time.

For the analysis of the hindcasts, we use this yearly deviation of the different clusters in ERA-20C to assign each of the 10 ensemble members generated by the model to the closest ERA-20C cluster by the same pattern matching algorithm

based on the RMSE. The pattern matching for the ensemble members is also derived against the ERA-20C cluster composites of only the prior 30 years. Since cluster classification in ERA-20C is done starting 1900, the first year that can be considered in the hindcast analysis is 1930. All years prior to 1930 are thus excluded in the hindcast analysis. This allows also for the spatial pattern of the predicted clusters to vary over time and no future information of the spatial pattern of the observed clusters is included in the assigning of the ensemble members to the different clusters.

The hindcast skill of the model is assessed against ERA-20C with the point-wise detrended Anomaly Correlation Coefficient (ACC, Collins, 2002). We derive significance via bootstrapping with 500 samples at the 95% confidence level.

To diagnose multidecadal variability of the hindcast skill throughout the 20th century, we obtained a time varying ACC by deriving the ACC for single time series between the model output and ERA-20C for a moving 30-year window. The resulting ACC for each 30-year window is then depicted in the center time step of the window, resulting in a time series of ACC.

2.3 SPATIO-TEMPORAL VARIABILITY OF THE CLUSTERS

Assessing the mechanisms that influence European summer climate on seasonal timescales using the evolved cluster analysis with ERA-20C data for 1900-2010 results in four clusters (Fig. 2.1a-d). The Z₅₀₀ composites show the spatial pattern of the four dominant mechanisms over the entire 20th century. Two of the clusters depict a more meridional pressure gradient, that can be associated with the summer NAO in its positive and negative phase (Fig. 2.1a (NAO+) and 2.1b (NAO-)). A slight spatial asymmetry can be detected between the two phases in the southern node of the two NAO clusters. NAO+ has its highest variability over Scandinavia and NAO+ and over Great Britain. The other two clusters exhibit a more zonal pressure gradient with similarities to the East Atlantic pattern (EA, Wallace and Gutzler, 1981; Barnston and Livezey, 1987), the Atlantic Low (Cassou et al., 2005) and the summer East Atlantic mode (SEA, Wulff et al., 2017). Those zonal pressure difference (PD) patterns show strong spatial asymmetries in both the magnitude and spatial distribution. While in its positive phase, the PD depicts two nodes that are located over the eastern North Atlantic and northern Europe (Fig. 2.1c (PD+)), the negative phase of the PD results in three alternating nodes, two over the North Atlantic and one over Europe (Fig. 2.1d (PD-)).

We assess the model output by the mean over all ensemble members assigned to one of the clusters (Fig. 2.1e-h). Compared to the observed clusters (cf. Fig. 2.1a-d), the spatial patterns of those averaged modelled clusters demonstrate that the model is able to represent each cluster, both in the overall spatial structure, as

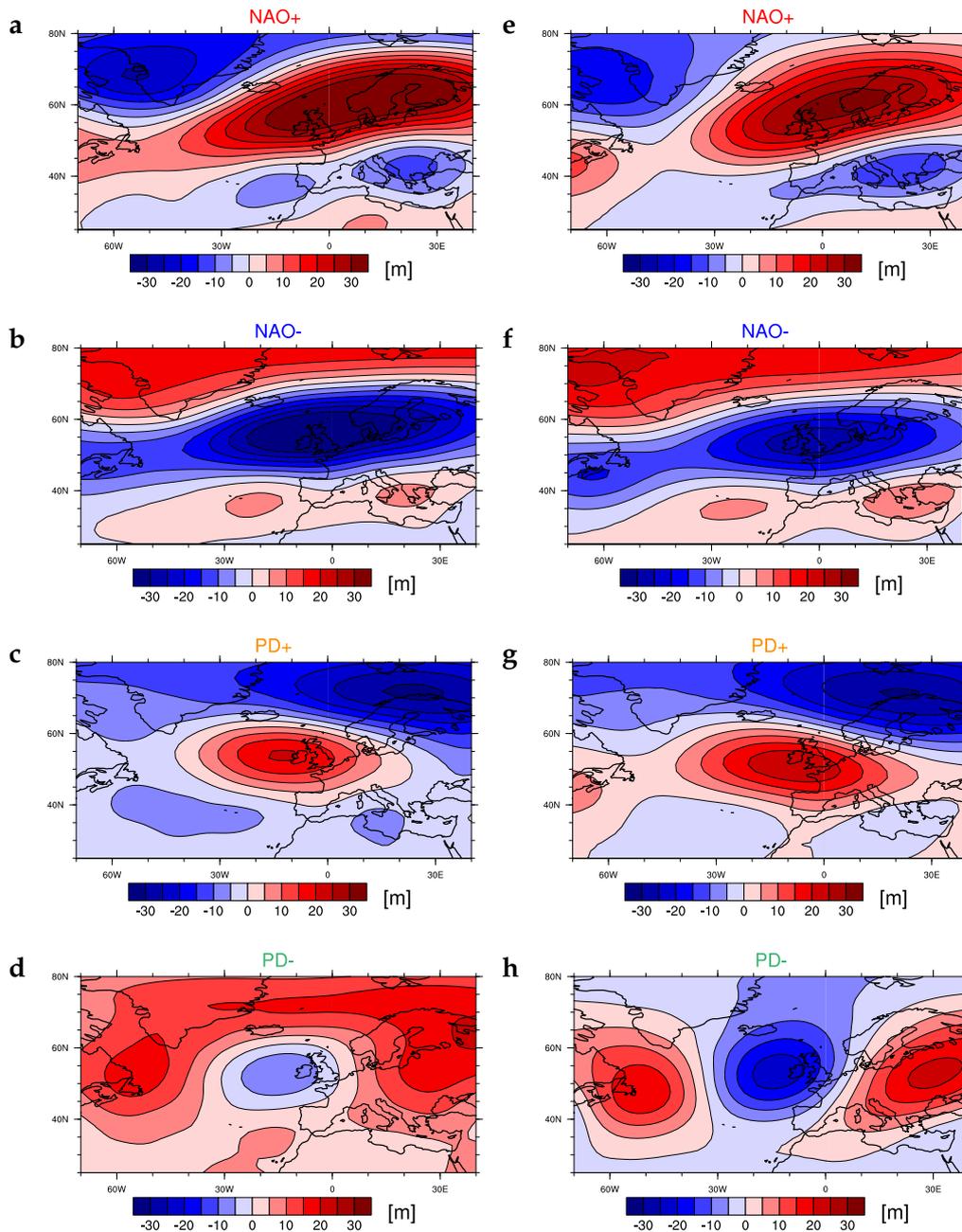


Figure 2.1: **a-d** The spatial patterns of the four clusters derived with the cluster analysis in ERA-20C for 1900-2010, depicted as the composites over the yearly occurrence of the observed clusters for July-August (JA) means of Z₅₀₀ anomaly. **e-h** Mean over the Z₅₀₀ anomaly of all ensemble members generated by MPI-ESM-MR that are assigned to the corresponding cluster in 1930-2008.

well as in magnitude. Only for PD- the mean over all ensemble members results in slightly more pronounced alternating nodes than for the observed cluster.

We further consider the temporal distribution of the clusters throughout the 20th century. We investigate this in ERA-20C by the yearly distribution of the clusters (Fig. 2.2a) and in the model output by the number of ensemble members assigned to each cluster (Fig. 2.2b). The yearly distribution of the clusters observed in ERA-20C shows that NAO+ and PD- are more dominant in the first half of the century, while NAO- and PD+ occur more often in the second half. The temporal distribution of the different ensemble members assigned to the one of the clusters (cf. Fig. 2.2b) reveals that, in most years, every cluster is represented by the model, but the number of ensemble members assigned to the different clusters varies over time. NAO+ is reproduced by more ensemble members in the beginning and mid of the century, and NAO- more in the second half of the century. In the model, PD+ is represented by more ensemble members in the first half of the century, and only a low number of ensemble members reproduce PD- over the whole timespan.

However, while we show in the observations that each summer can be assigned to one dominant observed cluster (cf. Fig. 2.2a), the model predicts several different clusters per summer (cf. Fig. 2.2b). To evaluate the model prediction, we thus have the option to either consider multiple clusters per summer and therefore all ensemble members in all clusters (option 1), or to restrain the ensemble to one dominant cluster per summer. We can find the one dominant cluster per summer in the model (option 2) or in the observations (option 3). For option 2, to find the dominant cluster in the model, we use the cluster that is predicted by the model through the most ensemble members in a particular summer. For example, in the first considered year 1930, the NAO+, PD+ and PD- clusters are each reproduced by one ensemble member, while seven ensemble members are representing NAO- (cf. Fig. 2.2b). The model is therefore predicting NAO- to be the dominant cluster in 1930. For option 3, the dominant cluster is chosen in comparison to observations. For 1930 this is the PD- cluster, which is reproduced by one ensemble member in the model (cf. Fig. 2.2a). The number of ensemble members in the dominant predicted (option 2) or dominant observed (option 3) cluster varies for every summer as depicted in Fig. 2.2b, while the number of ensemble members in all clusters (option 1) is for every summer equal to ten.

If we now consider the ensemble members for those three options, meaning for every summer either all members (option 1), the members in the predicted cluster (option 2), or the members in the observed cluster (option 3), we have three options for which we can analyse the frequency of occurrence of ensemble members in particular clusters. We do this in comparison to the frequency of the observed clusters, which is derived by the number of years that are assigned to a particular cluster (Fig. 2.2c). We thus compare the frequency of ensemble

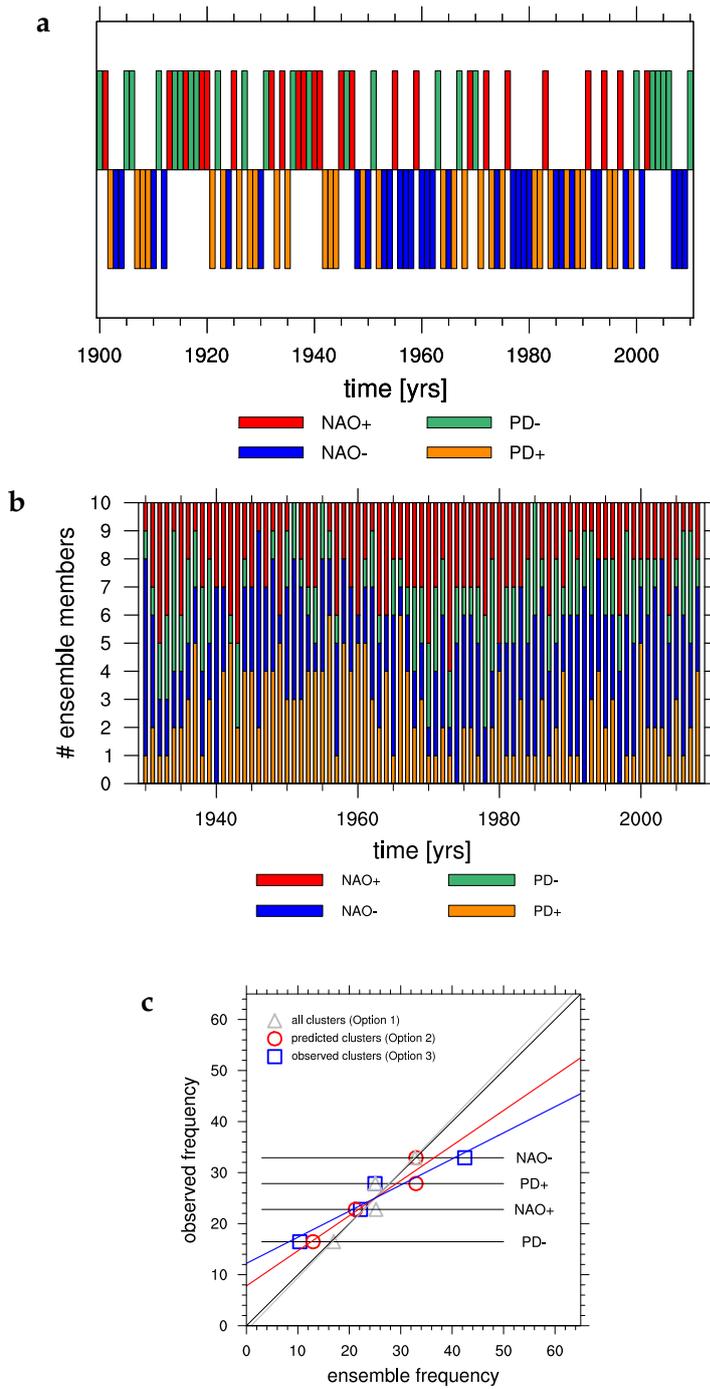


Figure 2.2: **a** Occurrence of the four different clusters in ERA-20C for 1900-2010 indicated by vertical bars. **b** Counted numbers of the ensemble members generated by MPI-ESM-MR that are assigned to one of the ERA-20C clusters by a pattern-matching algorithm for 1930-2008. **c** Frequency of the ensemble members in each cluster compared to the frequencies of years the clusters are observed in ERA-20C, including linear regression. The frequencies of the ensemble members are derived for the ensemble members in all clusters (option 1, grey), for the ensemble members in the predicted dominant clusters, which are only the members in the cluster with the highest number of ensemble members in each year (option 2, red), and for the ensemble members in the observed dominant cluster, which are only the members in the cluster that is observed in each year (option 3, blue).

members in particular clusters to the frequency of years that those clusters are observed.

Regarding the number of ensemble members in all clusters (option 1), most ensemble members represent NAO-, the least PD-, and about equal NAO+ and PD+ (Fig. 2.2c, grey). This frequency of all ensemble members in all clusters approximately agrees with the frequency of the yearly distribution of the observed clusters. Assessing the frequency of only the ensemble members in the predicted dominant clusters (option 2), the proportions deviate from the observed yearly frequency for both PD clusters, where PD+ is slightly over- and PD- slightly under-represented (Fig. 2.2c, red). The frequency of ensemble members in the observed dominant clusters (option 3) shows the frequency of the ensemble members that agree with the observed clusters for individual summers (Fig. 2.2c, blue). This reveals that the ensemble members in the NAO- cluster most often agree with the observations, while PD- is the cluster for which predicted and observed dominant clusters match the least.

Comparing the frequencies of the different options of ensemble groups to the observed yearly frequencies of the clusters, we can see that the data points for all three mentioned options are approximately located on one line (cf. Fig. 2.2c). This demonstrates that the ordering of the frequencies of the different clusters between modelled and observed clusters agree. The incline, however, demonstrates that the frequencies are closest to observation for the ensemble members in all clusters (option 1), while for the members in the predicted clusters (option 2) and in the observed clusters (option 3) the frequencies deviate stronger from the observed yearly distribution of the clusters.

If we assume that the model predicts one cluster per year through the cluster that is represented by the highest number of ensemble members in that year (option 2), we can further compare the predicted to the observed cluster in every year. This results in model and observations agreeing on the dominant cluster in 28 out of 79 investigated years, according to about one third of all years. Thus, the model is able to represent both the spatial appearance and the overall frequency of occurrence of the observed clusters, but is not able to predict the timing of the clusters for individual summers.

2.4 SEASONAL HINDCAST SKILL

Hindcast skill of an ensemble-based prediction system is usually assessed by taking the mean over all ensemble members. Comparing this ensemble mean to reanalysis data in the North-Atlantic-European sector for the summer season results in significant skill in the tropical region and some parts of the North Atlantic (option 1, Fig. 2.3a). However, significant seasonal summer hindcast skill is not achieved over Europe.

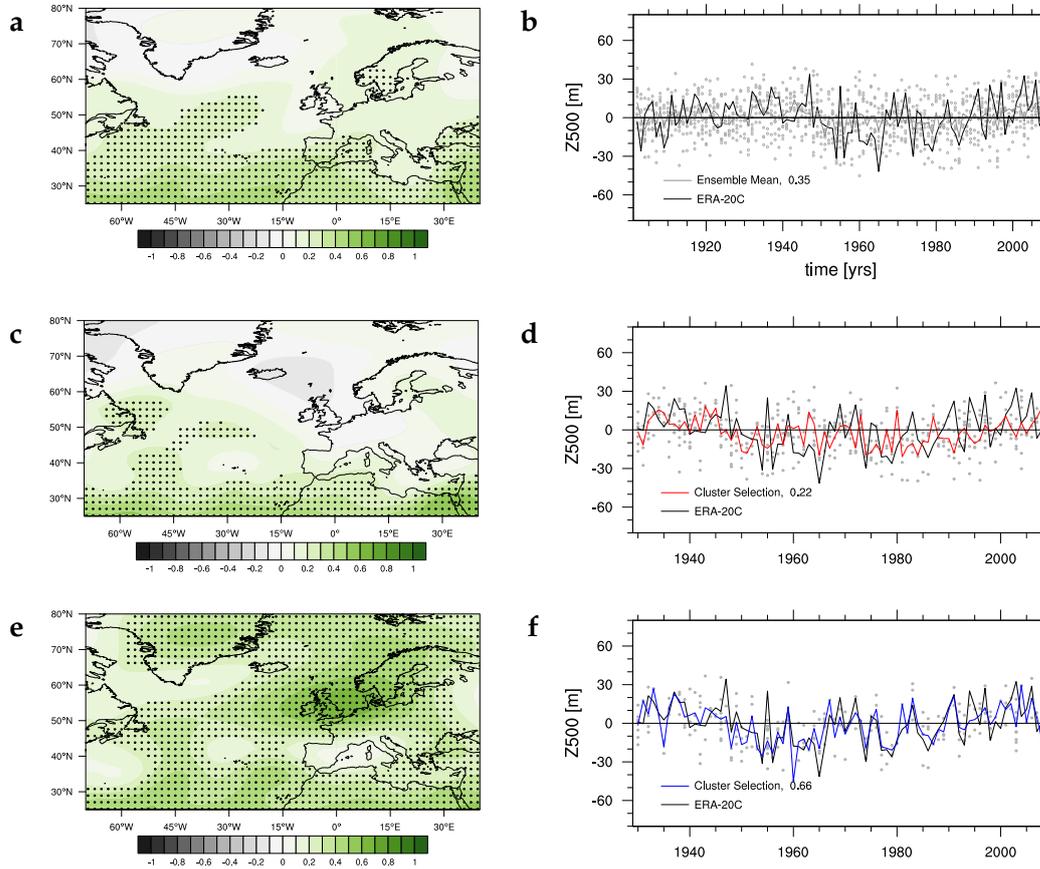


Figure 2.3: (left column) Hindcast skill depicted by the Anomaly Correlation Coefficient (ACC) for Z500 anomaly in summer (JA) comparing the model predictions of MPI-ESM-MR to ERA-20C for different ensemble means. Black dots show significance at the 95% confidence level. (right column) Averaged Z500 anomaly in Europe in the area $[35^{\circ} - 70^{\circ}\text{N}, 10^{\circ}\text{W} - 30^{\circ}\text{E}]$ for all ensemble members (grey circles), selected members (filled grey dots), and their mean (coloured lines) in comparison to ERA-20C (black line), including correlation values. Ensemble means are taken over **a-b** all ensemble members in all clusters for 1902-2008 (option 1), **c-d** ensemble members selected for the predicted dominant clusters for 1930-2008 (option 2), and **e-f** ensemble members selected for the observed clusters for 1930-2008 (option 3).

To investigate the reason for this low skill, we analyse the hindcast skill over Europe in more detail by evaluating the hindcasts for averaged Z₅₀₀ anomaly in the area [35 – 70°N, 10°W-30°E] for all individual ensemble members (Fig. 2.3b). The values for Z₅₀₀ that are predicted by the individual members are spread out over a large range of possible values of both positive and negative sign in each year. If, as in usual predictions, a mean is taken over all ensemble members to assess the hindcast skill, the impact of the high spread of the ensemble results in a low variability of the ensemble mean (std=6.6 m) compared to the variability in the reanalysis (std=16.1 m, Fig. 2.3b). Nevertheless, most values of the reanalysis lie within the spread of all ensemble members, indicating that the range of observations is contained within the range of the ensemble. However, the observed values for Z₅₀₀ are not achieved by taking the mean over all ensemble members.

Different ensemble members predict different clusters for one summer and the different clusters have different imprints on Z₅₀₀ over Europe (Figs. 2.2b and 2.1a-d). To avoid averaging over different clusters that individually have a different impact on Z₅₀₀ over Europe, we now consider only ensemble members in one cluster in each year. For every year considered, we predict the one dominant cluster by the cluster that contains the highest number of ensemble members (option 2). We then form a new ensemble mean over only the members in this cluster. Such a yearly selection does not result in significant hindcast skill over Europe (Fig. 2.3c), but in less ensemble spread and thus higher variability of the mean over the ensemble members (std=10.4 m, Fig. 2.3d).

If we instead select only those ensemble members that are assigned to the observed dominant cluster for every year (option 3), the ensemble mean formed over only those ensemble members results in significant hindcast skill over large parts of Europe and the North Atlantic (Fig. 2.3e). The average over Europe then shows a significant correlation ($r=0.66$) that is much higher compared to the mean over all ensemble members ($r=0.35$) and a variability (std=14.2 m) that is comparable to the variability of the reanalysis (std=16.1 m, Fig. 2.3f). The model is therefore able to achieve significant seasonal hindcast skill for European summers, if the cluster predicted by the model agrees with the observed dominant cluster in every summer.

2.4.1 Cluster dependency of the hindcast skill

We assess the hindcast skill of the model in the individual clusters by separating the ensemble members into all four clusters in every year and form a mean over all ensemble members in each cluster. We further assign each year to one cluster predicted by the model (option 2) and consider for each cluster only those predicted years. Such a prediction results in strong spatial differences for the hindcast skill of the individual clusters and reveals the areas in which the

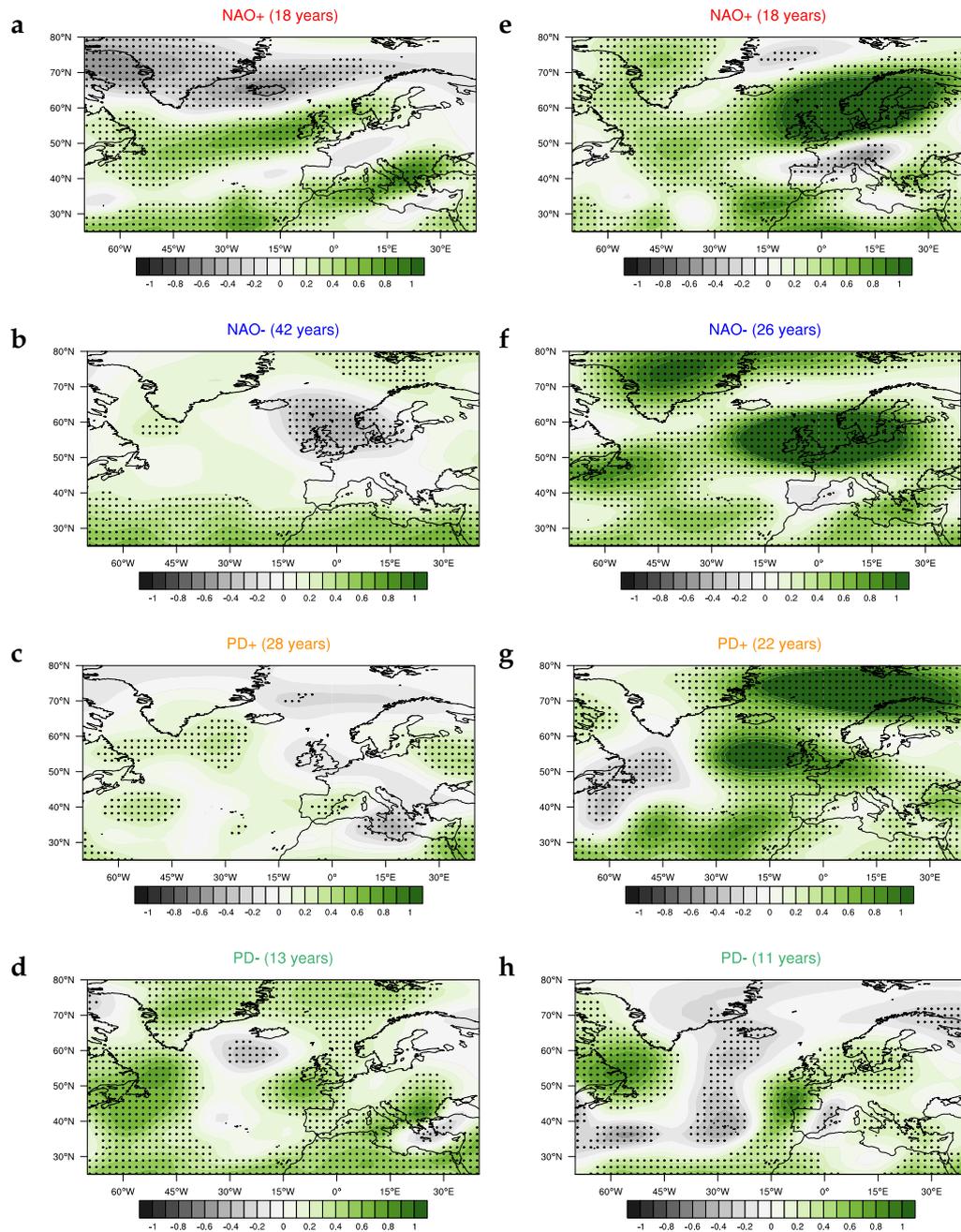


Figure 2.4: ACC for JA mean Z500 anomaly comparing the model predictions of MPI-ESM-MR to ERA-20c for the ensemble members grouped for the different clusters for 1930-2008. ACC is shown by yearly deviation for the ensemble members in the predicted dominant clusters (option 2, left column), and in the observed dominant clusters (option 3, right column). The numbers in brackets indicate in how many of the considered years the respective cluster was chosen to be dominant. Black dots show significance at the 95% confidence level.

model most likely achieves skill if it predicts a certain cluster (Fig. 2.4a-d). If the model predicts a NAO+, significant skill is achieved over large parts of the North Atlantic and in northern Europe, while for NAO- skill is mainly obtained in the tropics. For PD+ hindcast skill is only visible in single spots in the western North Atlantic and parts of eastern Europe. With PD- hindcast skill is achieved in large parts of the North Atlantic and in parts of northern Europe.

We can compare these results to the skill of the ensemble mean of the individual clusters divided into the years in which the clusters are observed in ERA-20C (option 3), which represent the skill that the model can maximally achieve in the years that are dominated by one of the mechanisms. Additionally they show the regions in which each cluster influenced the hindcast skill. Generally the hindcast skill is highest in the areas of highest variability of the different clusters (cf. Figs. 2.4e-h and 2.1). For NAO+ and NAO- these areas are over Greenland, the central North Atlantic and northern Europe, while NAO- also achieves significant skill over central Europe, which is in accordance with the location of the southern node of NAO-. With PD+, significant skill is also obtained in the two nodes of highest variability, resulting in hindcast skill over central Europe and large parts of the North Atlantic. For PD-, skill is only achieved in small parts in the North Atlantic and over Europe, not agreeing with the location of the nodes of PD-. All in all, these results reveal that the model can most likely achieve hindcast skill over Europe in the years that are dominated by NAO- or PD+, while hindcast skill over the North Atlantic can be achieved in years dominated by any of the clusters, except for PD-.

2.4.2 Temporal variability of the hindcast skill

We investigate the temporal variability of the hindcast skill over Europe for the three prior analysed cases (Fig. 2.5a). This analysis reveals that the hindcast skill changes over time with an abrupt skill change in 1970. The mean ACC over the ensemble members in all clusters (option 1) is decreasing until 1970, where the skill suddenly increases and stays roughly constant. The skill of the ensemble members in the predicted clusters (option 2) also decreases until 1970. This contrasts with the skill of the ensemble members in the observed clusters (option 3), which stays about constant until 1970. After 1970, the hindcast skill for the ensemble members of the predicted clusters is thus lower, while the skill for the ensemble members in the observed clusters is higher after 1970. This demonstrates that the model is more capable of predicting the observations before 1970 than afterwards.

However, analysing the spatial variance of the skill before and after the skill change in 1970 separately (Fig. 2.5b-e) shows that the model by itself is not able to achieve significant hindcast skill over Europe in either of the time periods.

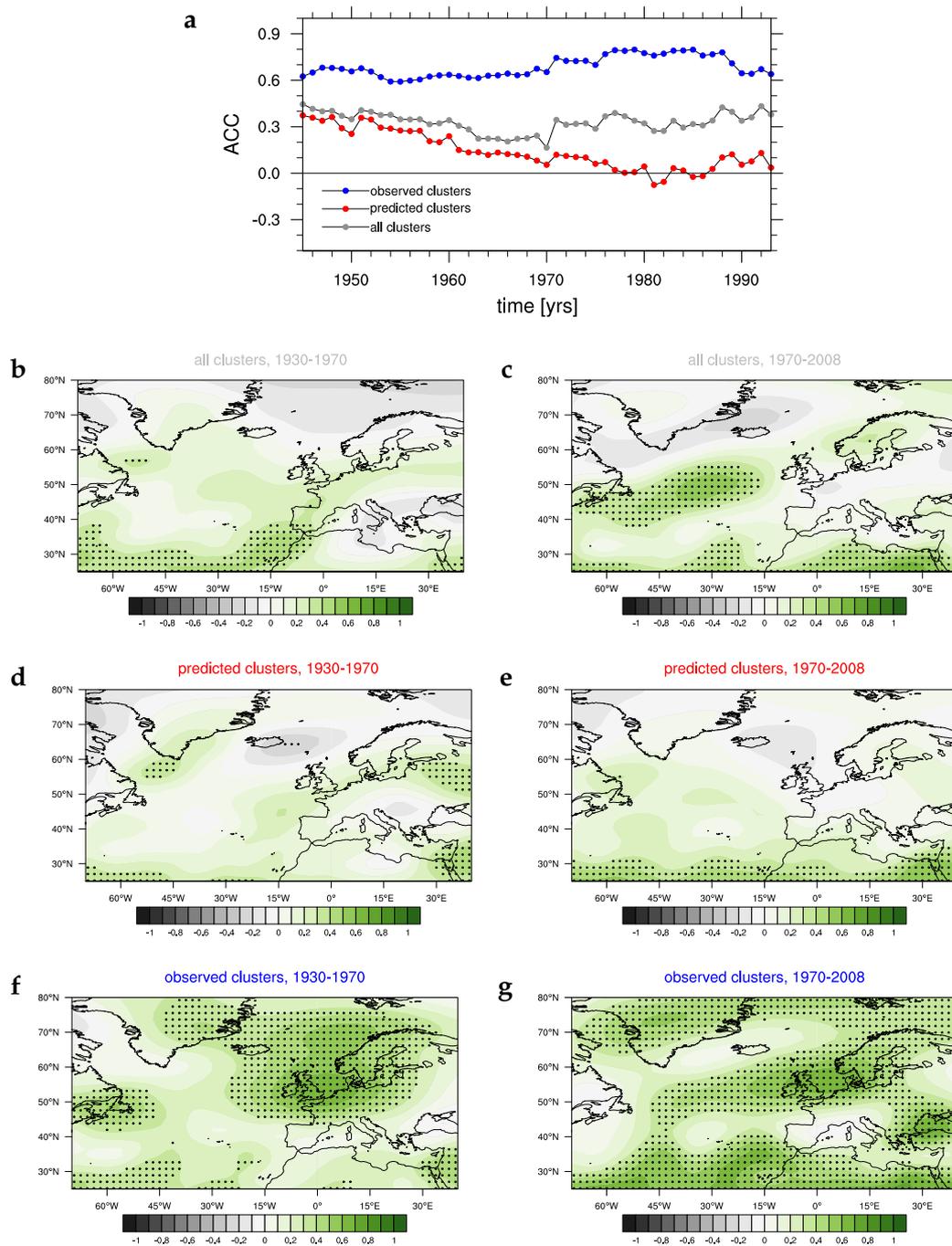


Figure 2.5: **a** Time dependency of the hindcast skill of averaged Z500 anomaly in Europe in the area $[35^{\circ} - 70^{\circ}\text{N}, 10^{\circ}\text{W} - 30^{\circ}\text{E}]$ shown by the evolution of the ACC in 1930-2008 for the mean over the ensemble members in all clusters (option 1, grey), the mean over the members in the predicted clusters (option 2, red), and the mean in the observed clusters (option 3, blue). Values are derived using a 30-year moving window and depicted in the center of the considered timespan. **b-g** ACC for each of those ensemble means in 1930-1970 (left column) and 1970-2008 (right column). Black dots show significance at the 95% confidence level.

Significant hindcast skill over Europe is only achieved for the ensemble mean in the observed clusters (Fig. 2.5f-g).

All plots reveal spatial differences in hindcast skill before and after 1970, confirming that the hindcast skill in some regions is time-dependent. For the ensembles in all clusters (option 1, Fig. 2.5b-c) this difference is strongest in the North Atlantic, where the model is able to achieve significant hindcast skill after, but not before 1970. For the predicted clusters (option 2, Fig. 2.5d-e), significant skill is obtained in eastern Europe, but only before 1970. After 1970 significant skill can only be achieved in the tropical region. Compared to the hindcast skill of the predicted clusters individually (cf. Fig. 2.4a-d), the spatial appearance before 1970 is closest to the hindcast skill of the PD+ cluster, while after 1970 the spatial pattern resembles the skill of the NAO- cluster. These results concur with the distribution of the ensemble members in the different clusters over time (cf. Fig. 2.2b), since PD+ is the predicted dominant cluster in most years before 1970, and NAO- after 1970.

Assessing the skill of the ensemble members in the observed dominant clusters (option 3, Fig. 2.5f-g), the patterns of significant skill also vary over time. The distribution of the observed cluster over time (cf. Fig. 2.2a) shows that before 1970, NAO+ and PD- were the dominating clusters, and NAO- and PD+ afterwards. In comparison to Fig. 2.4e-h that depicts the areas where these clusters have the highest impact on the hindcast skill, we see that the hindcast skill before 1970 resembles the skill structure of NAO+ and PD-, and after 1970 the skill structure of NAO- and PD+. The higher skill over Europe after 1970 (cf. Fig. 2.5a, blue) can thus be explained by the dominance of NAO- and PD+ in this period of time, that both show a higher hindcast skill over central Europe than NAO+ and PD- (cf. Fig. 2.4e-h), which dominate the time before 1970. The temporal variability of the hindcast skill is thus strongly influenced by the clusters that dominate a certain period of time.

2.4.3 *Grouping of clusters*

Examining the imprint of the different clusters on European summer temperature shows that both NAO phases influence Europe via a north-south dipole, while both PD phases show an east-west dipole (Fig. 2.6a-d), which is both in agreement with findings of Bladé et al. (2012) for the NAO and Neddermann et al. (2019) for the PD. A spatial asymmetry can be detected for both the NAO and PD clusters. NAO+ shows a north-south gradient of higher amplitude than NAO-, especially in southern Europe. PD- has a higher imprint on the temperature in eastern Europe than PD+, but influences all of Europe with the same temperature anomaly, while PD+ shows a temperature imprint of different sign on western than on eastern Europe.

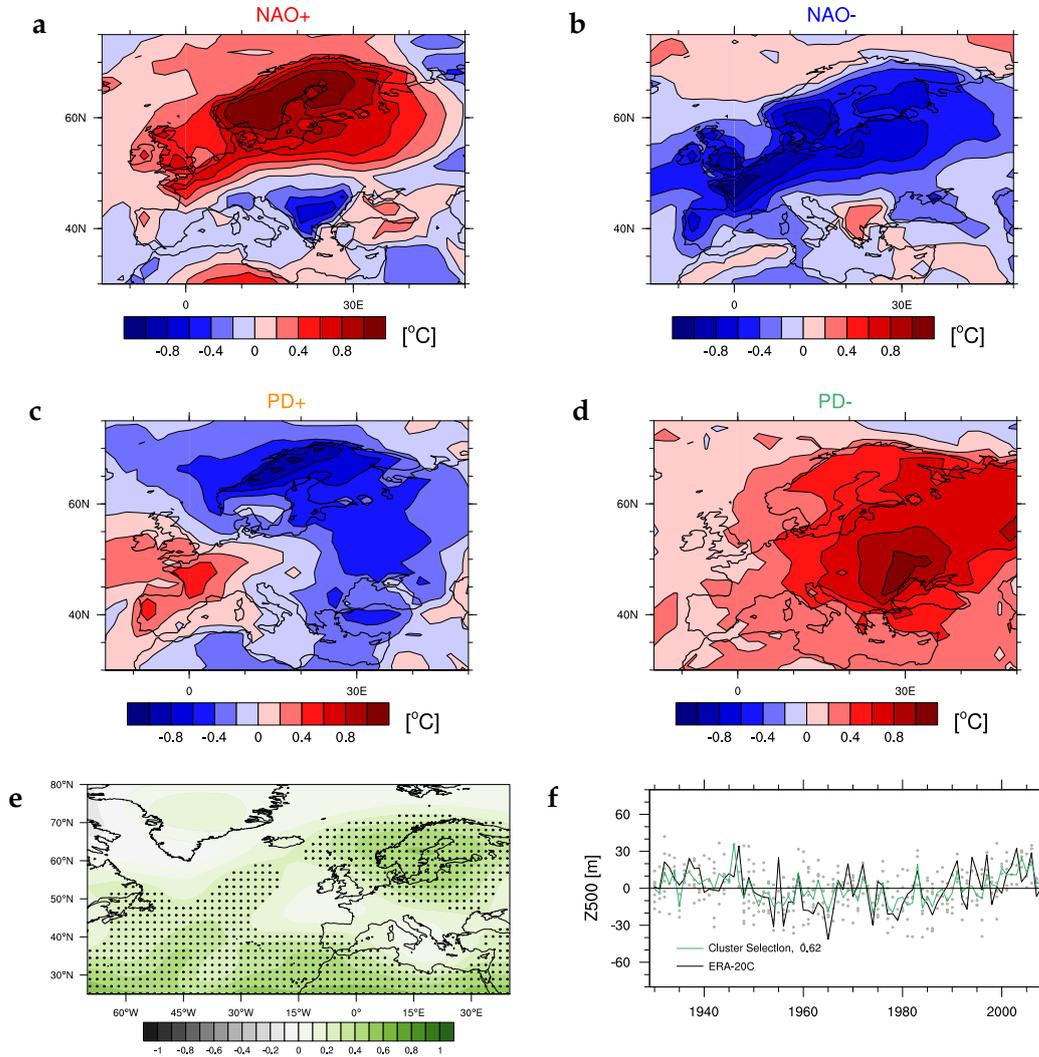


Figure 2.6: **a-d** The spatial patterns of the four clusters derived with the cluster analysis in ERA-20C for 1900-2010, depicted as the composites over the yearly occurrence of the observed clusters for JA means of temperature anomaly. **e** ACC for JA Z500 anomaly comparing the model predictions of MPI-ESM-MR to ERA-20C for ensemble members selected for the two groups of clusters (i.e. NAO+ & PD- and NAO- & PD+). Black dots show significance at the 95% confidence level. **f** Averaged Z500 anomaly in Europe in the area [35° – 70°N, 10°W – 30°E] for the selected ensemble members (grey circles) and their mean (green line) in comparison to ERA-20C (black line), including the correlation between both.

Comparing the imprint of all clusters on the temperature and Z500 anomalies in northern Europe (cf. Fig. 2.6a-d and Fig. 2.1a-d) and their temporal distribution over the whole century (cf. Fig. 2.2a) for ERA-20C, two groups of clusters seem to coincide with each other. While NAO+ and PD- both show a positive Z500 and temperature anomaly imprint on northern Europe and are more dominant in the first half of the century, NAO- and PD+ influence Europe with negative Z500 and temperature anomalies and dominate the second half of the century. Comparing the temporal distribution of our four clusters to the temporal variability of the averaged temperature and Z500 over Europe (not shown), we find a similarly strong agreement between warm temperature and positive Z500 anomaly for NAO+ and PD- and the opposite for NAO- and PD+. Both the temperature and Z500 time series also show the multidecadal trend identified in the temporal distribution of the clusters.

This agreement of the two cluster groups discloses the potential of combining two clusters for the hindcast analysis, implying the possibility of a fourth option of grouping the ensemble members for predicting European summers. In such a prediction the observed group of dominant clusters is known in advance and only ensemble members from the two dominant clusters are chosen for the hindcast analysis (option 4). Such a prediction results in a hindcast skill that is significant over large parts of northern Europe (Fig. 2.6e). For the mean over Europe a similar correlation and variability is achieved as for the prediction that takes all four clusters individually (cf. $r=0.62$, $\text{std}=11.2$ m, Fig. 2.6f and $r=0.66$, $\text{std}=14.2$, Fig. 2.3f). Through this grouping of two clusters, the problem of predicting clusters can therefore be reduced to two instead of four options, while resulting in similar hindcast skill for European summer climate.

2.5 DISCUSSION

We show here that, despite the fact that various mechanisms are influencing European summers, the dominant mechanism of European summer climate can be explained by either of the four considered mechanisms - a meridional pressure gradient, known as the North Atlantic Oscillation, and a zonal pressure gradient, both in their positive and negative phases.

We consider only large scale structures by using a cluster analysis of data filtered with the first 6 dominant EOFs. For smaller scale structures, higher order EOFs would have to be considered, as well as higher order clusters. However, we find that higher order clustering results in similar clusters with the four dominant patterns split into more than one cluster per mechanism. The diversity in the spatial appearance of the patterns in higher order clustering that show the same mechanism probably resemble the variety of spatial appearances of the clusters over time. The variation of the spatial appearance of the different

clusters over time are considered through our cluster analysis approach, that allows for such spatio-temporal variations of the considered mechanisms. Thus, we find that four clusters represent the dominating seasonal summer patterns in the North-Atlantic-European sector.

While previous work examine cluster analysis for European summers with daily data (e.g., Cassou et al., 2005; Boé et al., 2009; Cattiaux et al., 2013), we show here that a cluster analysis in this area is possible with seasonal summer means. Such an analysis discloses the possibility of finding the one dominant mechanism per season. Compared to the cluster analysis with daily data by Cassou et al. (2005), all clusters show strong similarities, where our PD+ cluster agrees to their Atlantic Ridge and our PD- to their Atlantic Low pattern.

Since our cluster approach considers only JA means to concentrate on the one dominant cluster per European summer, we assume that only one mechanism is present in every summer. Nevertheless, we show here that even with this assumption, significant hindcast skill can be achieved over large parts of Europe and the North Atlantic (Fig. 2.3e). This demonstrates that the problem of predicting European summers could be simplified to predicting which of the four clusters is going to dominate the following summer.

Comparing the cluster analysis in ERA-20C and in the different ensemble members of the model, we show here that the model is able to represent the spatial appearance of the four different clusters (Fig. 2.1) and also resembles the overall frequency of the different clusters observed over time (Fig. 2.2c). Thus, the model is able to reproduce the statistics of the observed climate patterns.

However, the timing of the prediction of the clusters only agrees with the observations in about one third of the years, thus being only slightly better than predicting the dominant cluster by chance. If ensemble members are then selected for the dominant predicted cluster (option 2), the predicted cluster disagrees with the observed clusters in the majority of the years. Such a prediction therefore leads to Z500 values that disagree with the observations and is not resulting in hindcast skill over Europe (Fig. 2.3c). Nevertheless, significant hindcast skill would be achieved if the timing of the predicted cluster agrees with the timing of the observed cluster (Fig. 2.3e). We thus find that the model is generally able to represent the physical mechanisms necessary to achieve significant hindcast skill over large parts of the North-Atlantic-European sector.

This cluster analysis further allows us to divide the hindcast skill into the skill of the model for the different clusters (Fig. 2.4). Our analysis can thus be used to determine in which areas the seasonal hindcast skill is influenced by the individual physical processes. For the patterns analysed here, we show that the hindcast skill is mainly influenced in the areas of highest variability of the clusters (cf. Fig. 2.1), which holds for all clusters, but PD-. The deviation for PD- could result from differences of the spatial structure between the observed and

predicted PD- (cf. Fig. 2.1d and 2.1h) and from the low number of ensemble members representing PD- (cf. Fig. 2.2b).

Most previous studies assume that the hindcast skill is constant over time. We show here that the hindcast skill of Z500 in the North-Atlantic-European sector varies over time for all the three considered options of ensemble refinement (Fig. 2.5a). Future work that analyse hindcast skill for time periods that cover multiple decades should thus consider hindcast skill independently for different time periods. The analysis of the time dependency further reveals that the spatial pattern of significant hindcast skill depends on the ability of the model to predict certain mechanisms and how pronounced these mechanisms were in the considered period of time.

Generally, the skill for the ensemble mean is strongly influenced by the dominant mechanism present in the observations and significant skill over Europe is only obtained when the ensemble members are selected to match this dominant mechanism (option 3 and 4). Therefore, in a real forecast setup, the observed dominant mechanism would have to be known in advance to achieve hindcast skill over Europe. This means that precursors for each cluster would be required. Recently, Totz et al. (2017) show that a clustering approach can be used to find precursors for certain precipitation regimes in winter. A similar approach could be applied to detect precursors for the analysed mechanisms here.

Another option to finding precursors for all clusters is to reduce the number of possible choices. We show here that even if only the group of clusters is known in advance (i.e. NAO+ & PD- and NAO- & PD+), significant skill can be achieved over northern Europe (option 4, Fig. 2.6e). This demonstrates that the problem of predicting one of four clusters can be reduced to predicting the group of clusters and thus lower the number of possible choices from four to two. Note that we only consider the sign and not the magnitude of the clusters. Therefore, such a prediction would mean to predict the sign of either of the two mechanisms for the following summer, so either if the NAO or if the PD is going to be in its positive or negative phase. Recent studies show that such a prediction is possible for both the NAO (Düsterhus et al., under review) and the PD in summer (Neddermann et al., 2019) and thus discloses the possibility of a skilful prediction with this approach.

We further find that the sign of the Z500 and temperature anomalies over Europe also agrees with the occurrence of the different clusters over time. Most years with a positive temperature and Z500 anomalies over Europe coincide with NAO+ and PD- and negative anomalies with NAO- and PD+. Cassou et al. (2005) found similar results for their cluster analysis, based on daily data. Their counterparts to our NAO+ and PD- both favour extremely warm days over Europe, while the counterparts to NAO- and PD+ favour extremely cold days. We thus find that also on seasonal timescales two clusters seem to have a similar

impact on European summers, both in their Z_{500} and temperature imprint. This agreement of Z_{500} and temperature anomalies with the cluster groups discloses the possibility that predicting the sign of the Z_{500} or temperature anomalies over Europe could also be used to anticipate the group of clusters for the following summer. For such a prediction, precursors would be needed for either of the two variables, while only the sign of the anomalies and not the magnitude would have to be known in advance.

However, the assumption of the similar imprint of two cluster groups on the temperature and Z_{500} over Europe is only valid for some regions in Europe, mainly the northern part (Fig. 2.6a-d). Therefore, the achievement of significant skill through such a selection is constrained to those regions (Fig. 2.6e). Nevertheless, for the mean over Europe a similar correlation and variability is achieved as for the prediction that takes all four clusters into account (cf. Figs. 2.6f and 2.3f).

Most previous studies assess the skill of an ensemble hindcast by the mean over all ensemble members. However, such an approach is not leading to skilful seasonal predictions for European summer climate. We show here that a mean over all ensemble members over Europe is a mean over different physical mechanisms that predict different Z_{500} values for European summers, thus leading to a high spread of the ensemble and an ensemble mean of lower variability as the observations (Fig. 2.3b). The problem of a too large ensemble spread demonstrates that the signal of the ensemble mean is too low, while the noise is too high. This has been shown in previous studies for different ensemble prediction systems as well (e.g., Ho et al., 2013; Eade et al., 2014). Opposed to previous work that assumes that the skill of a model increases with an increasing number of ensemble members (e.g., Murphy, 1990; Scaife et al., 2014), we present an approach in which we concentrate on a physically consistent analysis by refining the ensemble with the physical process the individual members are representing. In this approach a new mean is formed over only those ensemble members that represent the same physical process. This approach combines clustering with ensemble predictions and offers the possibility to compromise between the advantage of condensing forecast into using a few ensemble members against the disadvantage of losing information associated with the full ensemble. Our approach thus amplifies the signal, while suppressing the noise, and leads to significant seasonal hindcast skill of seasonal summer climate over Europe and the North Atlantic, given that the dominant cluster is known a priori.

2.6 SUMMARY AND CONCLUSIONS

We assess the ERA-20C reanalysis to investigate the dominant mechanisms in European summers on seasonal timescales. We use a cluster analysis that allows for the spatial pattern of each cluster to vary over time and assign the dominating

cluster to each year. We compare this analysis to 10 ensemble members generated using MPI-ESM-MR by assigning each ensemble member in every year to one of the examined clusters.

From this analysis we conclude:

- Seasonal European summers are dominated by four different clusters throughout the 20th century (Fig. 2.2a): (i) The North Atlantic Oscillation in its positive and negative phase (NAO+ and NAO-), with a meridional pressure gradient and an imprint on northern and southern European temperature (Figs. 2.1a-b and 2.6a-b) and (ii) a zonal pressure gradient with a stronger asymmetry in its positive and negative phase (PD+ and PD-) with an imprint on central European temperature (Figs. 2.1c-d and 2.6c-d).
- The model is able to represent the spatial pattern (Fig. 2.1e-h) and the overall observed distribution of the different clusters (Fig. 2.2c).
- The model shows no hindcast skill over Europe, if the skill is examined for the mean over all ensemble members (Fig. 2.3a). Individual ensemble members predict different clusters for one summer (Fig. 2.2b) and thus a variety of different Z500 values for Europe, which may cancel each other out. The high spread of the ensemble members results in a much lower variability of the ensemble mean than of the observed values (Fig. 2.3b).
- The ensemble mean variability can be adapted through grouping the ensemble members according to the cluster they represent and forming a new ensemble mean over members in one cluster (Fig. 2.3d,f). Yet, a skilful prediction is only achieved if the chosen predicted cluster agrees with the observed one. Such a prediction, however, results in significant hindcast skill over large parts of Europe and the North Atlantic (Fig. 2.3c,e).
- The different clusters influence the hindcast skill in different regions in the North-Atlantic-European sector (Fig. 2.4), mainly agreeing with the regions of highest variability of the individual clusters (Fig. 2.1).
- NAO+ and PD-, as well as NAO- and PD+ show a similar influence on Z500 and temperature anomalies in northern Europe and have a similar multidecadal appearance (Figs. 2.1a-d, Fig. 2.6a-d and 2.2a). Grouping those clusters leads to improved hindcast skill over large parts of northern Europe (Fig. 2.6e-f).
- Hindcast skill over the North-Atlantic-European sector differs over time with a sudden change in skill in 1970, resulting in different spatial appearances of the skill before and after 1970 (Fig. 2.5).

We show here that the ensemble generated by the seasonal climate prediction system of the MPI-ESM-MR is overall able to represent the dominant mechanisms

of seasonal European summers. So far, hindcasts were not skilful for European summers, because the dominant cluster predicted by the model mostly disagrees with the observed cluster in time. However, we find that if the timings agree with each other, significant seasonal hindcast skill for European summers is achieved. We therefore demonstrate that the model is capable of skilfully predicting European summers, if the dominant physical process that influences European summers is considered.

ACKNOWLEDGEMENTS

Many thanks go to the *Climate Modeling* 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, HP); by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2037 'CLICCS - Climate, Climatic Change, and Society' – Project Number: 390683824, contribution to the Center for Earth System Research and Sustainability (CEN) of Universität Hamburg (AD, JB); and by the Deutscher Wetterdienst (HP). The model simulations were performed using the high-performance computer at the German Climate Computing Center (DKRZ).

SEASONAL PREDICTABILITY OF EUROPEAN SUMMER
CLIMATE RE-ASSESSED

Nele-Charlotte Neddermann^{1,2}, Wolfgang A. Müller³, Mikhail Dobrynin¹,
André Düsterhus, Johanna Baehr¹

¹Institute for Oceanography, CEN, Universität Hamburg, Germany

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

³Max Planck Institute for Meteorology, Hamburg, Germany

AUTHOR CONTRIBUTIONS N.C.N. designed the research, performed the analysis and drafted the manuscript including all text and figures with guidance from J.B. and W.M.. All authors contributed through discussions on the interpretation of the results. M.D. performed the model simulations.

ABSTRACT

We improve seasonal hindcast skill of European summer climate in an ensemble based coupled seasonal climate 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 anomaly (SST) is 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 temperature, 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 SST 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 setup, this approach can potentially be applied in operational ensemble prediction systems.

3.1 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 climate 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 timescales (e.g., Cassou et al., 2005). As shown by 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 climate 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 (e.g., Marshall et al., 2001). The high persistence of tropical sea surface temperature (SST) 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 SST 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 SST in the Gulf of Mexico is 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 temperature 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 anomaly in the Caribbean and in the tropical North Pacific by showing that the tropical SST is leading the extra-tropical patterns by a few months. This temporal lag arises from the persistence of the SST 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 SST 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 SST 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 SST 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 climate 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 climate 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 timescales. Therefore, individual ensemble members in dynamical seasonal climate 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 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. (under review) 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 (e.g., 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 climate 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 climate 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 section 3.2. 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 section 3.3, followed by a description of the subsampling method in section 3.4. The effect of ensemble subsampling on seasonal hindcast skill is then accessed in section 3.5. Section 3.6 provides the discussion, followed by the summary and conclusions in section 3.7.

3.2 MODEL AND DATA

3.2.1 *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.

3.2.2 *Model Setup*

We use the dynamical seasonal climate 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 timescale of one day. In the ocean, the ECMWF Ocean reanalysis System (ORA-S4, Balmaseda et al., 2013) is used for nudging of temperature and salinity with a relaxation timescale 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.

3.2.3 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 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 analyzed period 1982-2016. The ACC is calculated for every cross-validated iteration between the reforecasted and reanalyzed 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.

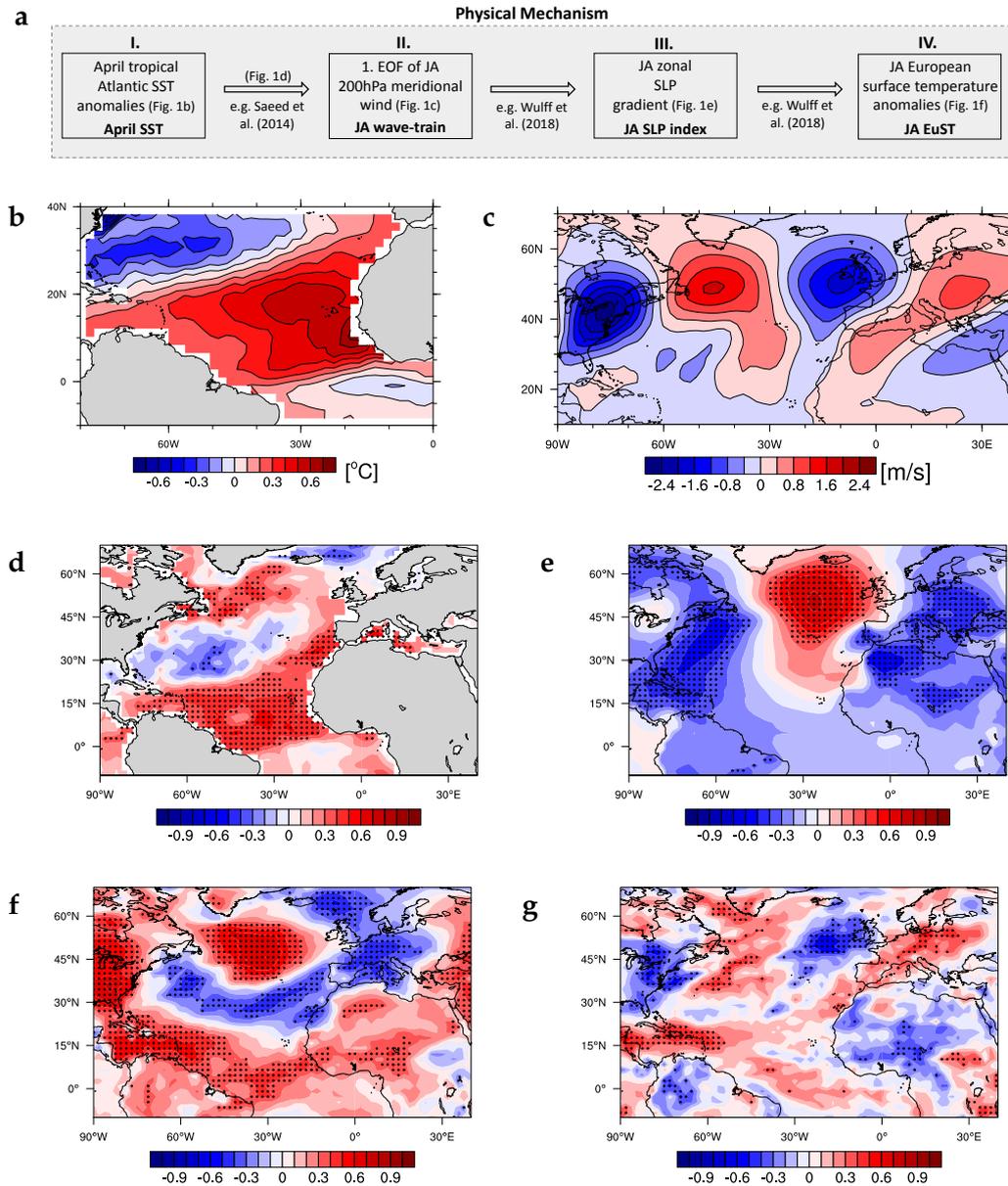


Figure 3.1: **a** Schematic representation of the proposed physical mechanism analysed in the ERA-Interim reanalysis in 1982–2016. **b** (I.) The signal of the mechanism starts in the tropical North Atlantic in spring, where SST anomaly is the source of strong convection, depicted by the first EOF of SST anomaly in April in the region [10°S–15°N, 80°W–20°W], which explains 41.2% of the total SST variance. **c** (II.) The strong convection induces a summertime wave-train, specified by 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 SST in April confirms this relation. **e** (III.) The wave-train is accompanied by a zonal SLP gradient, indicated by 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 temperature in JA and **g** total precipitation (TP) in JA. Dots represent significance at the 95% confidence level.

3.3 PHYSICAL MECHANISM

We describe a physical mechanism connecting SST in the tropical North Atlantic in spring and surface temperature over Europe in summer (Fig. 3.1a). SST in the tropical North Atlantic show high persistence, such that a signal is apparent in the SST from spring until summer. In the tropical North Atlantic, the main SST variability in spring lies in the latitudinal band between $0^{\circ} - 20^{\circ}\text{N}$ (Fig. 3.1b). Warm SST in this area is 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 by narrowing it to the region of interest [$10^{\circ} - 70^{\circ}\text{N}$, $90^{\circ}\text{W} - 40^{\circ}\text{E}$] (Fig. 3.1c). 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 SST in the North Atlantic (Fig. 3.1d), resulting in a significant correlation in the same region as the highest variability can be seen in the leading EOF of SST in the tropical North Atlantic in April (cf. Fig. 3.1b).

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. 3.1e), 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 temperature and precipitation over central Europe (Fig. 3.1f and 3.1g).

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. Figs. 3.1d and 3.1f). The strong SST anomaly in the western tropical Pacific in JA is then accompanied by low pressure and high total precipitation anomalies in this region (Figs. 3.1e and 3.1g), 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. 3.1e) is related to the second EOF of SLP in JA (Fig. 3.2a), 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)

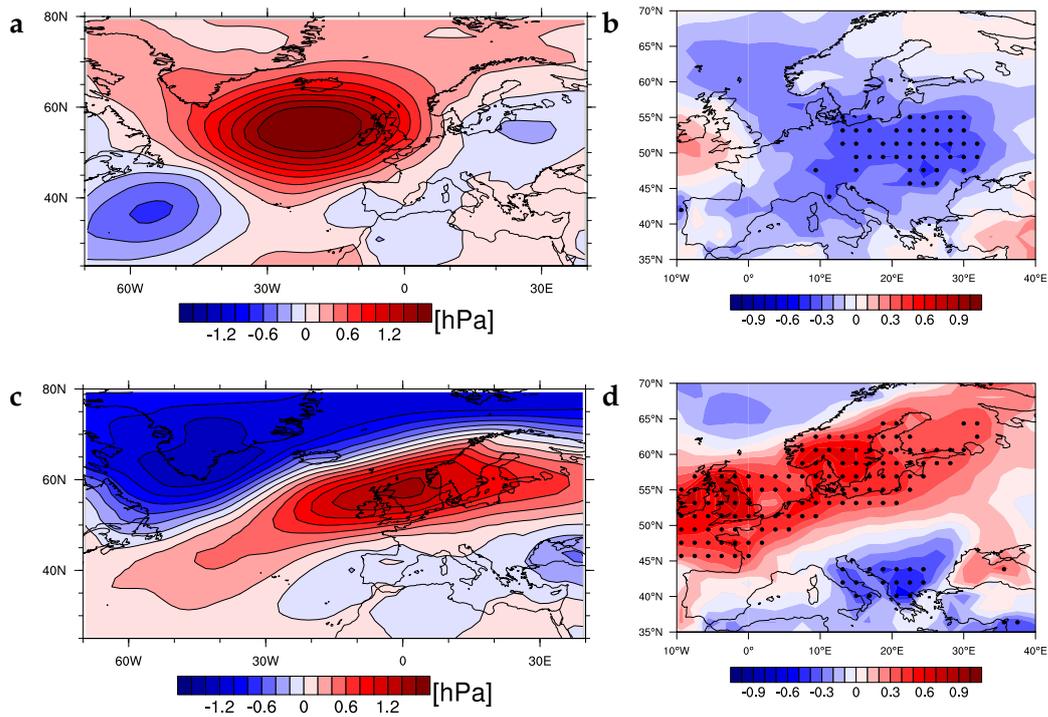


Figure 3.2: **a** The second EOF of SLP in JA in the region $[25^{\circ}\text{N}–80^{\circ}\text{N}, 70^{\circ}\text{W}–40^{\circ}\text{E}]$, explaining about 19.2% of the total variance. **b** The associated principal component of the second EOF of SLP in JA, pointwise correlated with surface temperature in JA. **c** The first EOF of SLP in JA in the region $[25^{\circ}\text{N}–80^{\circ}\text{N}, 70^{\circ}\text{W}–40^{\circ}\text{E}]$, 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 temperature in JA. The patterns are computed using the ERA-Iterim reanalysis in 1982–2016. Dots represent significance at the 95% confidence level.

or summer East Atlantic mode (SEA; Wulff et al., 2017). The influence of the zonal EOF pattern on European summer temperature is in strong agreement with the influence of the proposed wave-train mechanism (cf. Figs. 3.2b and 3.1f).

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. 3.2c) 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. 3.2d), which resembles the findings of Folland et al. (2009) and Bladé et al. (2012).

3.4 ENSEMBLE SUBSAMPLING

Based on the analysis of the proposed mechanism in the ERA-Interim reanalysis, we characterise the mechanism by four individual parts, namely (I.) the SST in the tropical North Atlantic in April and (II.) the wave-train, (III.) the zonal pressure gradient and (IV.) the temperature anomalies over central Europe in JA (Fig. 3.1a), 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 is either anomalous high or low SST anomaly in the tropical North Atlantic in April. Based on the correlation analysis between the wave-train and the SST anomaly in spring (Fig. 3.1d), SST is averaged in the area $[0^\circ - 15^\circ\text{N}, 80^\circ\text{W} - 20^\circ\text{W}]$.
- II. **JA wave-train** The wave-train in JA, which we so far defined by an EOF pattern (Fig. 3.1c). 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 $[10^\circ\text{N} - 70^\circ\text{N}, 90^\circ\text{W} - 40^\circ\text{E}]$ for the investigated 35 years. The analysis is conducted with $k = 2$, resulting in a positive and negative wave-train cluster (Figs. 3.3a and 3.3b, respectively), while the positive cluster occurred in 17 and the negative one in 18 years (Fig. 3.3c). There is a good agreement between the cluster and the EOF analysis, both for the patterns (cf. Figs. 3.3b and 3.1c, pattern correlation of $\pm 87.1\%$) and occurrences (Fig. 3.3c), 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-

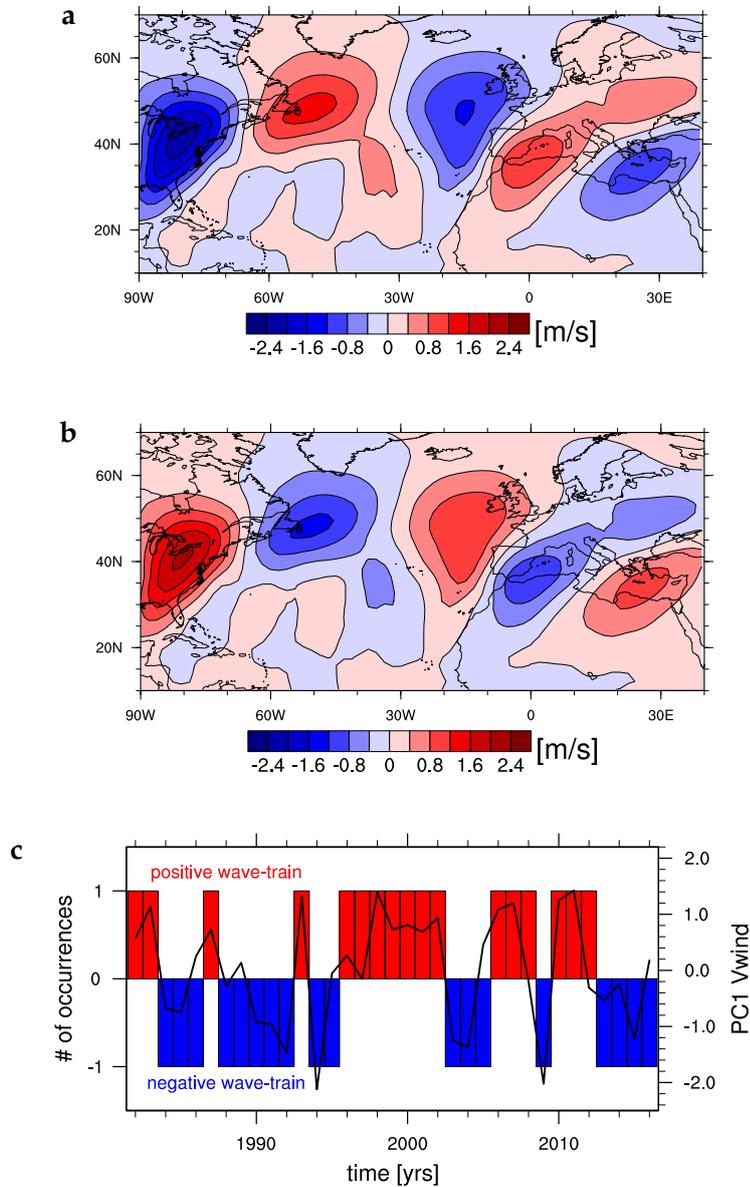


Figure 3.3: **a** Positive and **b** negative wave-train cluster derived by the *k*-mean algorithms for ERA-Interim 200 hPa meridional wind in the area [10°N–70°N, 90°W–40°E] in 1982–2016. The pattern correlation with the wave-train EOF pattern (Fig. 3.1c) is $\pm 87.1\%$. **c** Frequency of the cluster over the whole period for the positive (red, frequency 48.6%) and the negative cluster (blue, frequency 51.4%), compared the principal component (PC) associated with the wave-train (black line).

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. 3.2a) and on the correlation pattern of the wave-train with the SLP (Fig. 3.1e), 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 [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. 3.1f), 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. 3.4. 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. 3.1), 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 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

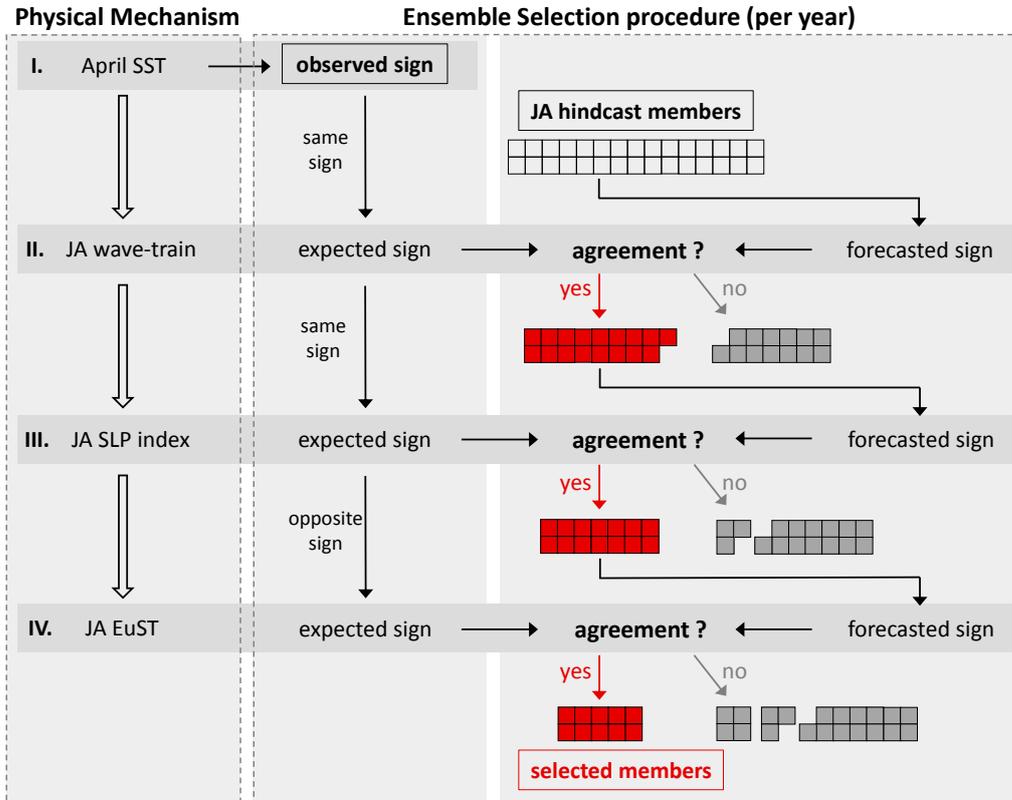


Figure 3.4: Schematic representation of the subsampling method for 30 ensemble members selected for the proposed mechanism in 1982-2016 in one of the considered years. The proposed mechanism as depicted in Fig. 3.1a and described in section 3.3, is based on observations in 1982-2016 and is depicted here 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 SST (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 members 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 SST. The remaining ensemble members are then in a last step checked for their sign of JA European surface temperature (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 3.1. The ensemble mean formed over the selected members in all years is termed "selected ensemble" (see e.g. Fig. 3.5).

SST in the tropical North Atlantic in April and the mechanism in summer (Fig. 3.1d), 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 SST and the mechanism (Fig. 3.1d), we assume that for a positive (negative) SST anomaly in April the mechanism in the considered year is positive (negative).

Following this procedure (cf. Fig. 3.4), 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 tropical SST that is 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 3.1), 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. 3.5) 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 (Figs. 3.5a-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 (Figs. 3.5c-d). While the linear regression of the full ensemble results 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 temperature.

As shown in section 3.3, 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 SST, 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 timescales and could still be important for the representation of those mechanisms.

year	predicted phase	# ensemble members	wind cluster hit	SLP index hit	hit with both
1982	-	10	0	0	0
1983	+	10	x	x	x
1984	-	10	x	0	0
1985	-	8	x	x	x
1986	-	10	x	0	0
1987	+	7	x	x	x
1988	+	6	0	0	0
1989	-	8	x	0	0
1990	+	4	0	0	0
1991	-	10	x	x	x
1992	-	4	x	x	x
1993	+	11	x	x	x
1994	-	7	x	x	x
1995	+	11	0	0	0
1996	+	7	0	x	0
1997	+	4	x	0	0
1998	+	3	x	x	x
1999	-	13	0	x	0
2000	+	10	x	x	x
2001	-	12	0	x	0
2002	-	13	0	0	0
2003	-	8	x	x	x
2004	+	11	0	0	0
2005	+	2	0	x	0
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	0	0
2015	-	12	x	x	x
2016	+	0	0	0	0
# hits			24	23	19

Table 3.1: Overview over the phase of the mechanism in the analysed years (second column) anticipated by the sign of the April SST anomaly in the area $[0^\circ - 15^\circ\text{N}, 80^\circ\text{W} - 20^\circ\text{W}]$ taken from ERA-Interim in 1982-2016. "+" 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.

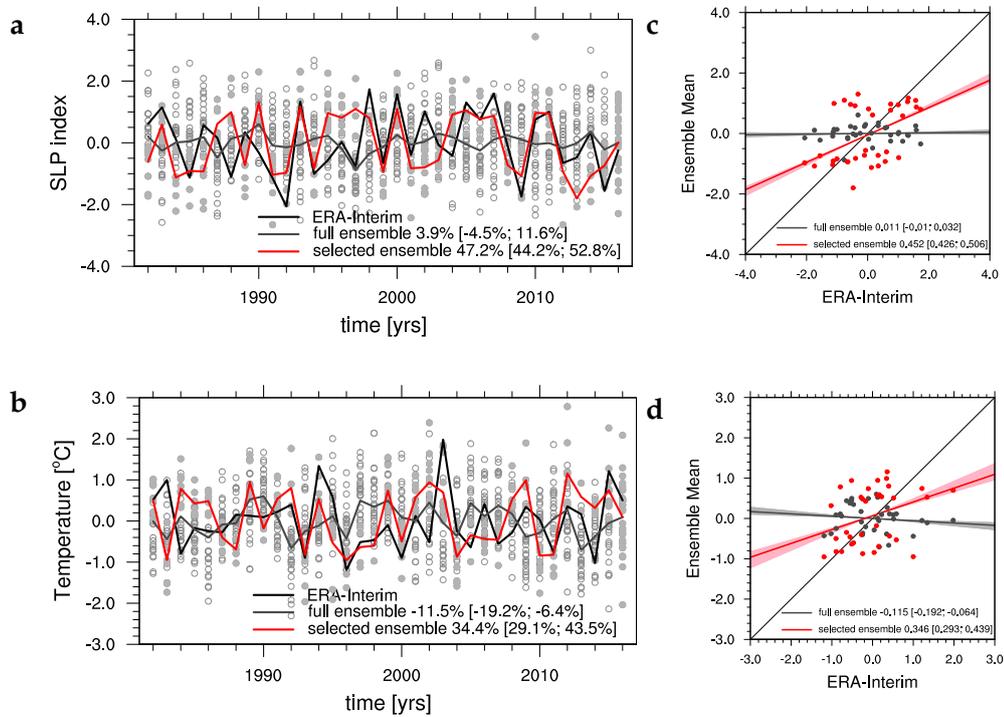


Figure 3.5: **a-b**: Comparison between ERA-Interim (black lines) and the ensemble mean over the full (grey lines) or the selected ensemble (red lines) from MPI-ESM in 1982-2016 including the mean and range (indicated in parentheses) leave-one-out cross-validated correlation values. Dots indicate the 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 [$40^{\circ}\text{N}-60^{\circ}\text{N}$, $40^{\circ}\text{W}-10^{\circ}\text{W}$] and Europe [$40^{\circ}\text{N}-60^{\circ}\text{N}$, $10^{\circ}\text{E}-40^{\circ}\text{E}$] and **b,d** the temperature anomaly in Europe in JA averaged in the area [$35^{\circ}\text{N}-55^{\circ}\text{N}$, $0^{\circ}-20^{\circ}\text{W}$].

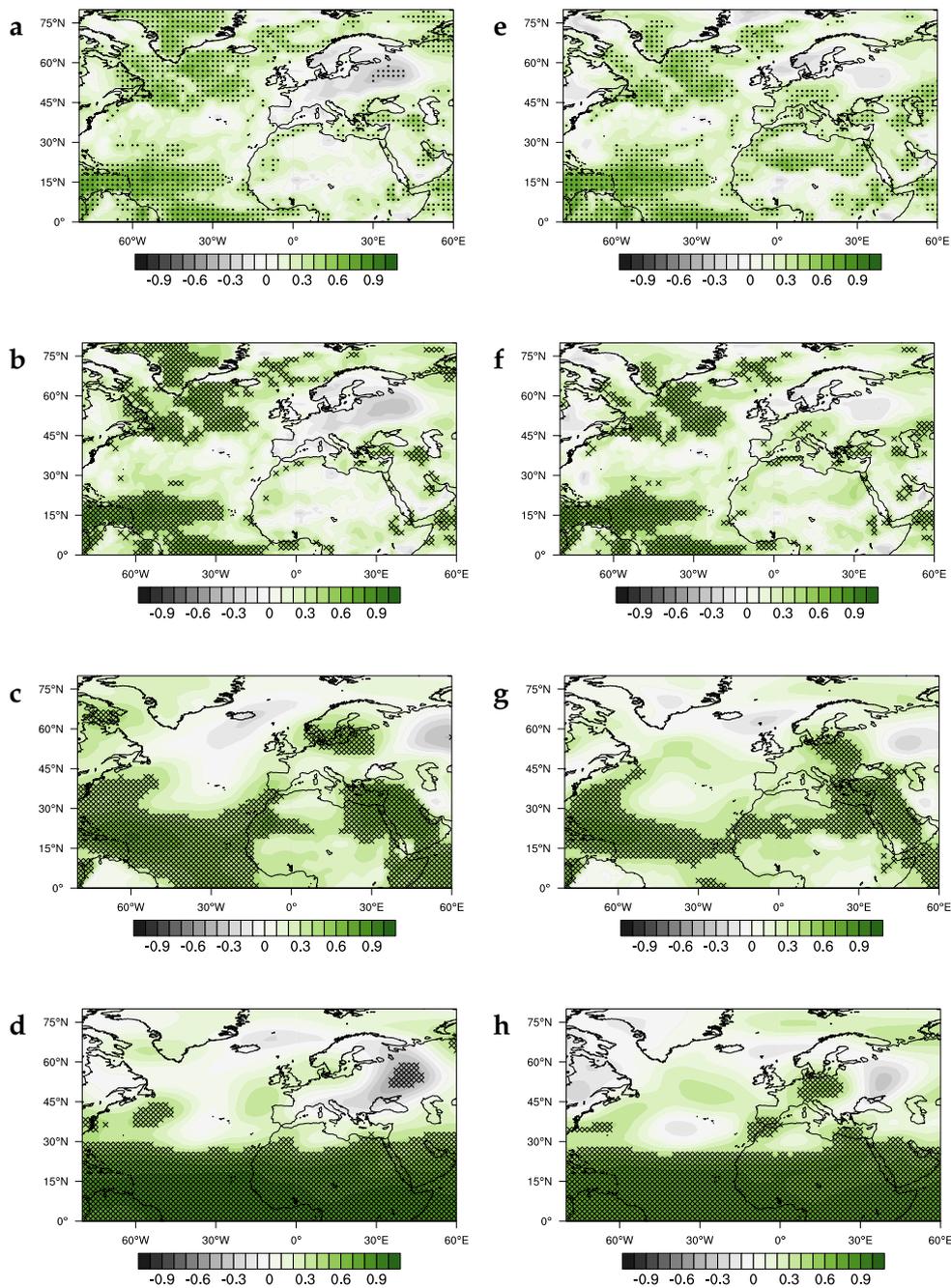


Figure 3.6: Anomaly Correlation Coefficient (ACC) derived without cross-validation for (first row) surface temperature and derived as the mean over all leave-one-out cross-validated correlation values for (second row) surface temperature, (third row) SLP and (fourth row) 500 hPa geopotential height (Z_{500}) 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.

3.5 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. 3.6). For surface temperature, the full ensemble shows hindcast skill mainly over the North Atlantic and Greenland (Fig. 3.6a-b). The SLP prediction for the full ensemble shows significant skill in the tropical North Atlantic and over parts of northern Europe and Arabia (Fig. 3.6c), while for Z500 hindcast skill can only be achieved in the tropical belt (Fig. 3.6d). 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. 3.6e-h), agreeing with the area where the proposed mechanism shows significant imprint on seasonal European summer climate (cf. Figs. 3.1e-g and 3.2b). For SLP and Z500, improved skill can also be achieved over the areas of the North Atlantic where the zonal pressure gradient is located (Fig. 3.1e). 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 areas where the SNAO has its biggest influence (cf. Fig. 3.2d and 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. 3.6a and 3.6e), also derive the hindcast skill that is significant in all cross-validated iterations (Fig. 3.6b and 3.6f). 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. 3.7). So far, we determined the phase of the mechanism by the observed sign of spring SST 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 SST 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. 3.7a and 3.6d).

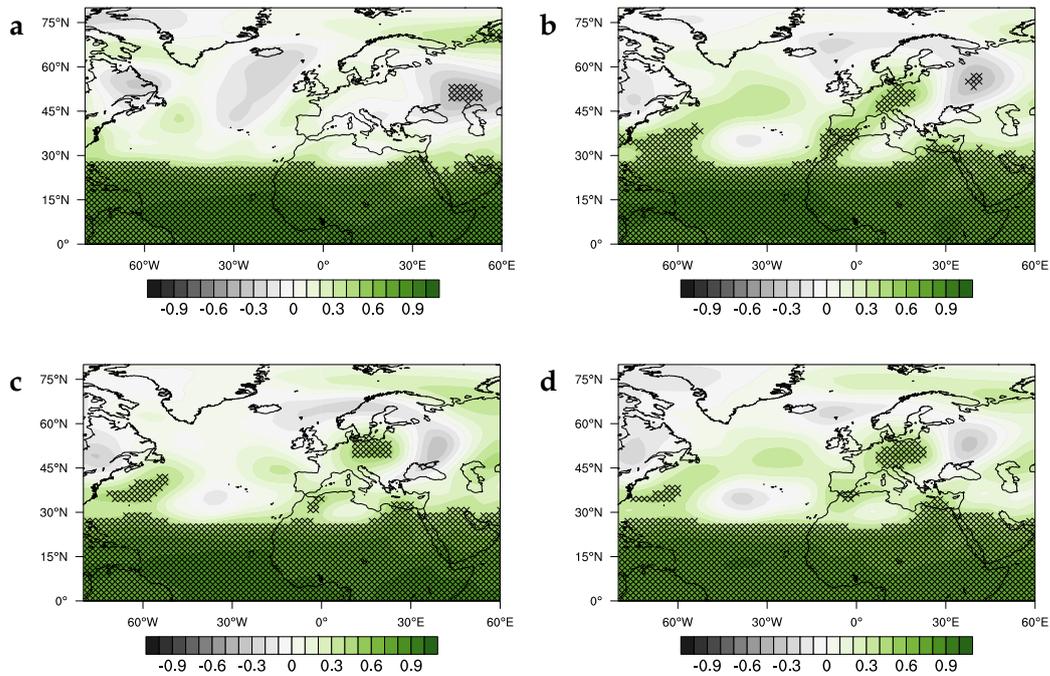


Figure 3.7: Anomaly Correlation Coefficient (ACC) for 500 hPa geopotential height (Z_{500}) in summer (JA) comparing the model predictions of MPI-ESM to ERA-Interim in 1982-2016 derived as the mean over all leave-one-out cross-validated correlation values. **a** The ensemble members are selected by all 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 SST 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.

The largest 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. 3.7a), 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 3.2). This "hit rate" improves if the phase of the mechanism is anticipated by observed spring SST 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 3.1).

If the phase of the mechanism is determined by spring SST, we can also evaluate the influence of the different selection criteria, namely the wave-train cluster, the SLP index or the European temperature anomaly, on the ACC (Figs. 3.7b-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 such an 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. 3.7b). If the temperature over Europe is the only included criterion, then the hindcast skill is stronger improved over central Europe (Fig. 3.7c), while a combination of the wave-train cluster and the European temperature criterion shows improved hindcast skill over larger areas of Europe (Fig. 3.7d). Including the SLP index as a third criterion then only slightly improves the hindcast skill over Europe (Fig. 3.6h). The small changes on ACC between the embedding of the different criteria is consistent with our analysis in section 3.3, 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. 3.8). 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. 3.5, 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 possible forecasts probability by the full ensemble deviates rarely from the average value of a forecasts. The selected ensemble mean on the other hand also shows 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.

year	predicted 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

Table 3.2: As in Table 3.1, but with the phase of the mechanism anticipated by the phase that contains the majority of ensemble members in MPI-ESM.

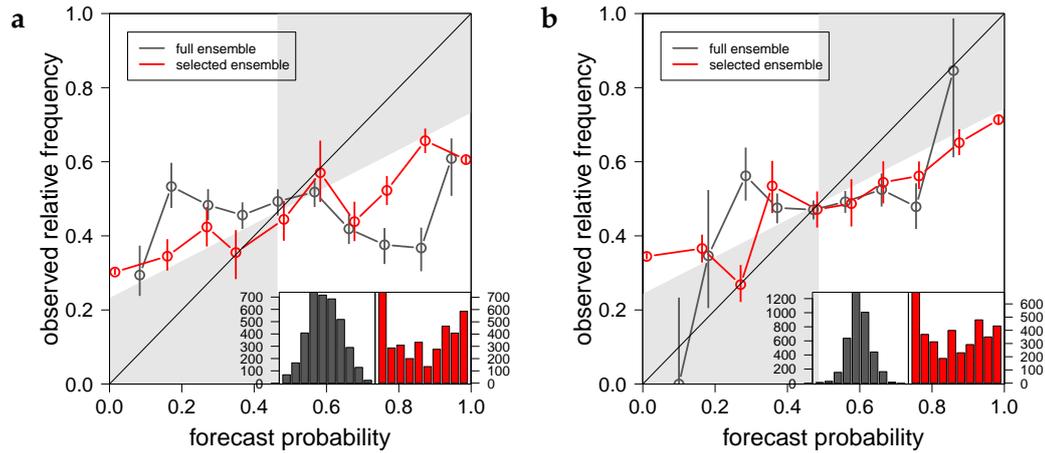


Figure 3.8: Reliability diagrams comparing the mean over the full (grey) and the selected (red) ensemble of MPI-ESM to ERA-Interim in 1982–2016 for **a** temperature and **b** Z500 in the area $[35^{\circ}\text{N}–55^{\circ}\text{N}, 0^{\circ}–20^{\circ}\text{W}]$. The diagonal line indicates perfect 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.

According to our analysis in Fig. 3.7, 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 3.1). 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 spring SST, 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 setup, 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. 3.9a–b and 3.6f,h), 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. 3.1e), 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.

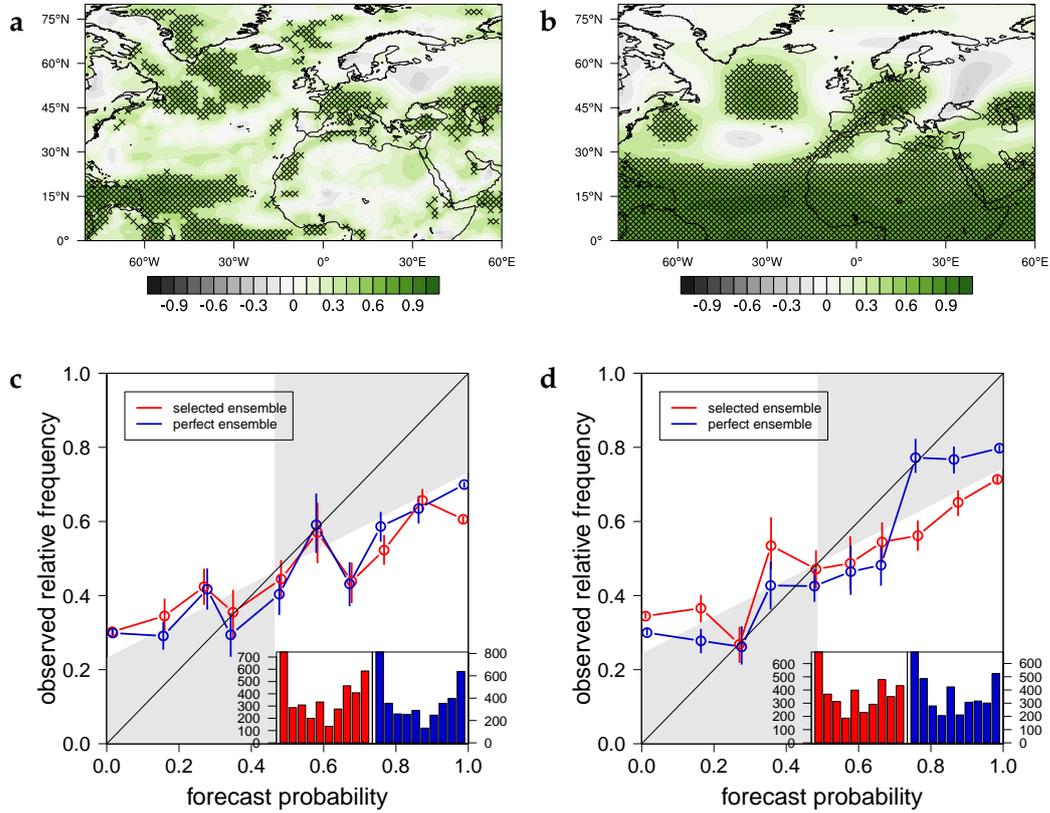


Figure 3.9: Hindcast skill for the "perfect" ensemble selection derived with the known state of the SLP index. **a-b** Anomaly Correlation Coefficient (ACC) for **a** surface temperature and **b** 500 hPa geopotential height (Z_{500}) in summer (JA), comparing the model predictions of MPI-ESM to ERA-Interim 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** Z_{500} 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.

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. 3.9c-d). For the temperature, the reliability diagram of the perfect ensemble is almost identical to the one of the selected ensemble. For Z_{500} , 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 European summer climate on seasonal timescales. The mechanism is included into the prediction by subselection of ensemble members through successive criteria based on the physical variables that define the proposed mechanism.

3.6 DISCUSSION

Our ensemble selection shows improved hindcast skill that is consistent throughout all analysed fields (Fig. 3.6e-h). Such a consistency in hindcast skill is not achieved in a hindcast analysis that uses the mean over all ensemble members (Fig. 3.6a-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 timescales, the model only predicts the chain of processes of the proposed mechanism 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. 3.5). 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. 3.8 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 timescales, 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 timescales. 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. 3.2). While we find decreased hindcast skill over Greenland and parts of northern Europe, Düsterhus et al. (under review) 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 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), SST in different ocean basins are linked by an "atmospheric bridge", such that SST anomalies in the North Atlantic in spring and summer are lead by SST anomalies in the tropical Pacific, which does not disagree with our findings. On the contrary, since ENSO is leading spring SST in the North Atlantic, which is used here 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 temperature 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 temperature over Europe (cf. Figs. 3.1f and 3.2b). 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. 3.8). 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 interdecadal timescales and associate it with the Atlantic Multidecadal Oscillation (AMO). This is in agreement with the findings of Gastineau and Frankignoul (2015) who show that the SST, that is connected to the zonal pressure gradient, is 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 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. 3.9). This perfect analysis results in improved skill in those areas, that are expected to be influenced by the proposed mechanism (cf. Fig. 3.1). 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.

3.7 SUMMARY AND CONCLUSIONS

We assess the summer seasonal hindcast skill of the MPI-ESM-MR ensemble based seasonal climate prediction system over the North-Atlantic-European sector with regard to a mechanism that is influencing this region on seasonal timescales. The proposed mechanism, here analysed in the ERA-Interim reanalysis has its origin in the tropical North Atlantic in spring, where persistent SST anomaly is the source of a Rossby wave-train that is accompanied by a zonal pressure gradient and has in turn 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 the 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 SST in April, which is before the initialisation of the prediction system and can thus 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 the 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.
- 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. 3.2).
- The individual parts of the summertime mechanism can be characterized by the wave-train cluster, the SLP index and the averaged European summer temperature. 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. 3.5).
- 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 3.2).

- 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 climate prediction system (Fig. 3.9). The areas of improved hindcast skill coincide with the areas in which the mechanism is influencing the North Atlantic-European climate on seasonal timescales (Fig. 3.1).
- Due to the demonstrated influence of tropical spring SST on European summers (Fig. 3.1), we make use of the high persistence and predictability of tropical regions and use the observed SST 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. 3.6).

We show an alternative approach, in which the ensemble size, instead of being further increased, is decreased through ensemble selection based on a chain of known physical relations. We make use of the influence of high persistent and predictable spring SST 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 timescales and demonstrate that including a mechanism into the prediction through subsampling of the ensemble by using successive physical relations is an effective method. This approach could be extended with further mechanisms or also be applied for other mechanisms 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 Modeling* 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), and by the German Science Foundation (DFG) through the University Hamburg's Cluster of Excellence Integrated Climate System Analysis and Prediction (CliSAP; AD, JB). The model simulations were performed using the high-performance computer at the German Climate Computing Center (DKRZ).

BIBLIOGRAPHY

- Alexander, M. A., I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. D. Scott (2002). "The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans." In: *Journal of Climate* 15.16, pp. 2205–2231.
- Arribas, A., M Glover, A Maidens, K Peterson, M Gordon, C MacLachlan, R Graham, D Fereday, J Camp, A. Scaife, et al. (2011). "The GloSea4 ensemble prediction system for seasonal forecasting." In: *Monthly Weather Review* 139.6, pp. 1891–1910.
- Baehr, J., K Fröhlich, M. Botzet, D. I. Domeisen, L. Kornblueh, D. Notz, R. Piontek, H. Pohlmann, S. Tietsche, and W. A. Mueller (2015). "The prediction of surface temperature in the new seasonal prediction system based on the MPI-ESM coupled climate model." In: *Climate Dynamics* 44.9-10, pp. 2723–2735.
- Baehr, J. and R. Piontek (2014). "Ensemble initialization of the oceanic component of a coupled model through bred vectors at seasonal-to-interannual timescales." In: *Geoscientific Model Development* 7.1, pp. 453–461.
- Baker, L., L. Shaffrey, R. Sutton, A. Weisheimer, and A. Scaife (2018). "An Inter-comparison of Skill and Overconfidence/Underconfidence of the Wintertime North Atlantic Oscillation in Multimodel Seasonal Forecasts." In: *Geophysical Research Letters*.
- Balmaseda, M. A., K. Mogensen, and A. T. Weaver (2013). "Evaluation of the ECMWF ocean reanalysis system ORAS4." In: *Quarterly Journal of the Royal Meteorological Society* 139.674, pp. 1132–1161.
- Barnston, A. G. and R. E. Livezey (1987). "Classification, seasonality and persistence of low-frequency atmospheric circulation patterns." In: *Monthly weather review* 115.6, pp. 1083–1126.
- Beverley, J. D., S. J. Woolnough, L. H. Baker, S. J. Johnson, and A. Weisheimer (2019). "The northern hemisphere circumglobal teleconnection in a seasonal forecast model and its relationship to European summer forecast skill." In: *Climate Dynamics* 52.5, pp. 3759–3771.
- Bjerknes, J. (1966). "A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature." In: *Tellus* 18.4, pp. 820–829.
- Bladé, I., B. Liebmann, D. Fortuny, and G. J. van Oldenborgh (2012). "Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region." In: *Climate dynamics* 39.3-4, pp. 709–727.

- Boé, J., L. Terray, C. Cassou, and J. Najac (2009). "Uncertainties in European summer precipitation changes: role of large scale circulation." In: *Climate Dynamics* 33.2-3, pp. 265–276.
- Boer, G. J., D. M. Smith, C. Cassou, F. Doblas-Reyes, G. Danabasoglu, B. Kirtman, Y. Kushnir, M. Kimoto, G. A. Meehl, R. Msadek, et al. (2016). "The decadal climate prediction project (DCPP) contribution to CMIP6." In: *Geoscientific Model Development (Online)* 9.10.
- Branstator, G. (2002). "Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic Oscillation." In: *Journal of Climate* 15.14, pp. 1893–1910.
- Branstator, G. and H. Teng (2017). "Tropospheric Waveguide Teleconnections and Their Seasonality." In: *Journal of the Atmospheric Sciences* 74.5, pp. 1513–1532.
- Bretherton, C. S., C. Smith, and J. M. Wallace (1992). "An intercomparison of methods for finding coupled patterns in climate data." In: *Journal of climate* 5.6, pp. 541–560.
- Butler, A. H., A. Arribas, M. Athanassiadou, J. Baehr, N. Calvo, A. Charlton-Perez, M. Déqué, D. I. Domeisen, K. Fröhlich, H. Hendon, et al. (2016). "The Climate-system Historical Forecast Project: do stratosphere-resolving models make better seasonal climate predictions in boreal winter?" In: *Quarterly Journal of the Royal Meteorological Society* 142.696, pp. 1413–1427.
- Cantelaube, P. and J.-M. Terres (2005). "Seasonal weather forecasts for crop yield modelling in Europe." In: *Tellus A: Dynamic Meteorology and Oceanography* 57.3, pp. 476–487.
- Cassou, C., L. Terray, and A. S. Phillips (2005). "Tropical Atlantic influence on European heat waves." In: *Journal of climate* 18.15, pp. 2805–2811.
- Cattiaux, J., B. Quesada, A. Arakélian, F. Codron, R. Vautard, and P. Yiou (2013). "North-Atlantic dynamics and European temperature extremes in the IPSL model: sensitivity to atmospheric resolution." In: *Climate dynamics* 40.9-10, pp. 2293–2310.
- Collins, M. (2002). "Climate predictability on interannual to decadal time scales: The initial value problem." In: *Climate Dynamics* 19.8, pp. 671–692.
- Corti, S., S. Gualdi, and A. Navarra (2003). "Analysis of the mid-latitude weather regimes in the 200-year control integration of the SINTEX model." In: *Annals of Geophysics* 46.1.
- Council, N. R. et al. (2010). *Assessment of intraseasonal to interannual climate prediction and predictability*. National Academies Press.
- Dawson, A., T. Palmer, and S Corti (2012). "Simulating regime structures in weather and climate prediction models." In: *Geophysical Research Letters* 39.21.
- Dee, D. P., S. Uppala, A. Simmons, P. Berrisford, P. Poli, S Kobayashi, U Andrae, M. Balsameda, G Balsamo, P Bauer, et al. (2011). "The ERA-Interim reanalysis: Configuration and performance of the data assimilation system." In: *Quarterly Journal of the royal meteorological society* 137.656, pp. 553–597.

- Ding, Q. and B. Wang (2005). "Circumglobal teleconnection in the northern hemisphere summer." In: *Journal of Climate* 18.17, pp. 3483–3505.
- Doblas-Reyes, F. J., J. García-Serrano, F. Lienert, A. P. Biescas, and L. R. Rodrigues (2013). "Seasonal climate predictability and forecasting: status and prospects." In: *Wiley Interdisciplinary Reviews: Climate Change* 4.4, pp. 245–268.
- Dobrynin, M., D. I. Domeisen, W. A. Müller, L. Bell, S. Brune, F. Bunzel, A. Düsterhus, K. Fröhlich, H. Pohlmann, and J. Baehr (2018). "Improved teleconnection-based dynamical seasonal predictions of boreal winter." In: *Geophysical Research Letters*.
- Domeisen, D. I., A. H. Butler, K. Fröhlich, M. Bittner, W. A. Müller, and J. Baehr (2015). "Seasonal predictability over Europe arising from El Niño and stratospheric variability in the MPI-ESM seasonal prediction system." In: *Journal of Climate* 28.1, pp. 256–271.
- Duchez, A., E. Frajka-Williams, S. A. Josey, D. G. Evans, J. P. Grist, R. Marsh, G. D. McCarthy, B. Sinha, D. I. Berry, and J. J. Hirschi (2016). "Drivers of exceptionally cold North Atlantic Ocean temperatures and their link to the 2015 European heat wave." In: *Environmental Research Letters* 11.7, p. 074004.
- Düsterhus, A., M. Dobrynin, D. I. Domeisen, H. Pohlmann, W. A. Müller, and J. Baehr (under review). "A statistical-dynamical seasonal prediction of the Summer North Atlantic Oscillation." In:
- Eade, R., D. Smith, A. Scaife, E. Wallace, N. Dunstone, L. Hermanson, and N. Robinson (2014). "Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?" In: *Geophysical research letters* 41.15, pp. 5620–5628.
- Ferranti, L., S. Corti, and M. Janousek (2015). "Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector." In: *Quarterly Journal of the Royal Meteorological Society* 141.688, pp. 916–924.
- Ferranti, L. and S Corti (2011). "New clustering products." In: *ECMWF Newsletter* 127.6-11, pp. 1–2.
- Fetterer, F, K Knowles, W Meier, and M Savoie (2002). "Sea ice index." In: *National Snow and Ice Data Center, Boulder, CO, digital media.*[Available online at <http://nsidc.org/data/g02135.html>.]
- Folland, C. K., J. Knight, H. W. Linderholm, D. Fereday, S. Ineson, and J. W. Hurrell (2009). "The summer North Atlantic Oscillation: past, present, and future." In: *Journal of Climate* 22.5, pp. 1082–1103.
- Gastineau, G. and C. Frankignoul (2015). "Influence of the North Atlantic SST variability on the atmospheric circulation during the twentieth century." In: *Journal of Climate* 28.4, pp. 1396–1416.
- Giorgetta, M. A., J. H. Jungclaus, C. H. Reick, S. Legutke, J. Bader, M. Böttinger, V. Brovkin, T. Crueger, M. Esch, K. Fieg, et al. (2013). "Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled

- Model Intercomparison Project phase 5." In: *Journal of Advances in Modeling Earth Systems* 5.3, pp. 572–597.
- Hannachi, A., D. M. Straus, C. L. Franzke, S. Corti, and T. Woollings (2017). "Low-frequency nonlinearity and regime behavior in the Northern Hemisphere extratropical atmosphere." In: *Reviews of Geophysics* 55.1, pp. 199–234.
- Ho, C. K., E. Hawkins, L. Shaffrey, J. Bröcker, L. Hermanson, J. M. Murphy, D. M. Smith, and R. Eade (2013). "Examining reliability of seasonal to decadal sea surface temperature forecasts: The role of ensemble dispersion." In: *Geophysical Research Letters* 40.21, pp. 5770–5775.
- Hodson, D. L., R. T. Sutton, C. Cassou, N. Keenlyside, Y. Okumura, and T. Zhou (2010). "Climate impacts of recent multidecadal changes in Atlantic Ocean sea surface temperature: A multimodel comparison." In: *Climate Dynamics* 34.7-8, pp. 1041–1058.
- Hoskins, B. J. and T. Ambrizzi (1993). "Rossby wave propagation on a realistic longitudinally varying flow." In: *Journal of the Atmospheric Sciences* 50.12, pp. 1661–1671.
- Hoskins, B. J. and D. J. Karoly (1981). "The steady linear response of a spherical atmosphere to thermal and orographic forcing." In: *Journal of the Atmospheric Sciences* 38.6, pp. 1179–1196.
- Hurrell, J. W. (1995). "Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation." In: *Science* 269.5224, pp. 676–679.
- (1996). "Influence of variations in extratropical wintertime teleconnections on Northern Hemisphere temperature." In: *Geophysical Research Letters* 23.6, pp. 665–668.
- Hurrell, J. W. and C. Deser (2010). "North Atlantic climate variability: the role of the North Atlantic Oscillation." In: *Journal of Marine Systems* 79.3-4, pp. 231–244.
- Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck (2003). *An overview of the North Atlantic oscillation*. Wiley Online Library.
- Iglesias, I., M. N. Lorenzo, and J. J. Taboada (2014). "Seasonal predictability of the East atlantic pattern from sea surface temperatures." In: *PloS one* 9.1, e86439.
- Jung, T., M. Hilmer, E. Ruprecht, S. Kleppek, S. K. Gulev, and O. Zolina (2003). "Characteristics of the recent eastward shift of interannual NAO variability." In: *Journal of Climate* 16.20, pp. 3371–3382.
- Jungclaus, J. H., N. Keenlyside, M. Botzet, H. Haak, J.-J. Luo, M. Latif, J. Marotzke, U. Mikolajewicz, and E. Roeckner (2006). "Ocean circulation and tropical variability in the coupled model ECHAM5/MPI-OM." In: *Journal of climate* 19.16, pp. 3952–3972.
- Kim, H.-M., P. J. Webster, and J. A. Curry (2012). "Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter." In: *Climate Dynamics* 39.12, pp. 2957–2973.

- Lau, N.-C. and M. J. Nath (2001). "Impact of ENSO on SST variability in the North Pacific and North Atlantic: Seasonal dependence and role of extratropical sea-air coupling." In: *Journal of Climate* 14.13, pp. 2846–2866.
- Li, J. and C. Ruan (2018). "The North Atlantic–Eurasian teleconnection in summer and its effects on Eurasian climates." In: *Environmental Research Letters* 13.2, p. 024007.
- Lledó, L., V. Torralba, A. Soret, J. Ramon, and F. J. Doblas-Reyes (2019). "Seasonal forecasts of wind power generation." In: *Renewable Energy* 143, pp. 91–100.
- Lowe, R., M. García-Díez, J. Ballester, J. Creswick, J.-M. Robine, F. Herrmann, and X. Rodó (2016). "Evaluation of an early-warning system for heat wave-related mortality in Europe: implications for sub-seasonal to seasonal forecasting and climate services." In: *International journal of environmental research and public health* 13.2, p. 206.
- Marshall, J., Y. Kushnir, D. Battisti, P. Chang, A. Czaja, R. Dickson, J. Hurrell, M. McCARTNEY, R Saravanan, and M. Visbeck (2001). "North Atlantic climate variability: phenomena, impacts and mechanisms." In: *International Journal of Climatology* 21.15, pp. 1863–1898.
- Matsueda, M. and T. Palmer (2018). "Estimates of flow-dependent predictability of wintertime Euro-Atlantic weather regimes in medium-range forecasts." In: *Quarterly Journal of the Royal Meteorological Society* 144.713, pp. 1012–1027.
- Michelangeli, P.-A., R. Vautard, and B. Legras (1995). "Weather regimes: Recurrence and quasi stationarity." In: *Journal of the atmospheric sciences* 52.8, pp. 1237–1256.
- Mishra, N., C. Prodhomme, and V. Guemas (2018). "Multi-model skill assessment of seasonal temperature and precipitation forecasts over Europe." In: *Climate Dynamics*, pp. 1–19.
- Müller, W. A., J. H. Jungclaus, T. Mauritsen, J. Baehr, M. Bittner, R. Budich, F. Bunzel, M. Esch, R. Ghosh, H. Haak, et al. (2018). "A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1. 2-HR)." In: *Journal of Advances in Modeling Earth Systems* 10.7, pp. 1383–1413.
- Müller, W. A., D. Matei, M. Bersch, J. H. Jungclaus, H. Haak, K. Lohmann, G. Compo, P. Sardeshmukh, and J. Marotzke (2015). "A twentieth-century reanalysis forced ocean model to reconstruct the North Atlantic climate variation during the 1920s." In: *Climate Dynamics* 44.7-8, pp. 1935–1955.
- Murphy, J. (1990). "Assessment of the practical utility of extended range ensemble forecasts." In: *Quarterly Journal of the Royal Meteorological Society* 116.491, pp. 89–125.
- Neddermann, N.-C., W. A. Müller, M. Dobrynin, A. Düsterhus, and J. Baehr (2019). "Seasonal predictability of European summer climate re-assessed." In: *Climate Dynamics* 53.5, pp. 3039–3056.

- North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng (1982). "Sampling errors in the estimation of empirical orthogonal functions." In: *Monthly Weather Review* 110.7, pp. 699–706.
- O'Reilly, C. H., J. Heatley, D. MacLeod, A. Weisheimer, T. N. Palmer, N. Schaller, and T. Woollings (2017). "Variability in seasonal forecast skill of Northern Hemisphere winters over the twentieth century." In: *Geophysical Research Letters* 44.11, pp. 5729–5738.
- Palmer, T. N. and D. L. T. Anderson (1994). "The prospects for seasonal forecasting—A review paper." In: *Quarterly Journal of the Royal Meteorological Society* 120.518, pp. 755–793.
- Poli, P., H. Hersbach, D. P. Dee, P. Berrisford, A. J. Simmons, F. Vitart, P. Laloyaux, D. G. Tan, C. Peubey, J.-N. Thépaut, et al. (2016). "ERA-20C: An atmospheric reanalysis of the twentieth century." In: *Journal of Climate* 29.11, pp. 4083–4097.
- Robine, J.-M., S. L. K. Cheung, S. Le Roy, H. Van Oyen, C. Griffiths, J.-P. Michel, and F. R. Herrmann (2008). "Death toll exceeded 70,000 in Europe during the summer of 2003." In: *Comptes rendus biologiques* 331.2, pp. 171–178.
- Russo, S., A. Dosio, R. G. Graversen, J. Sillmann, H. Carrao, M. B. Dunbar, A. Singleton, P. Montagna, P. Barbola, and J. V. Vogt (2014). "Magnitude of extreme heat waves in present climate and their projection in a warming world." In: *Journal of Geophysical Research: Atmospheres* 119.22, pp. 12–500.
- Saeed, S., N. Van Lipzig, W. A. Müller, F. Saeed, and D. Zanchettin (2014). "Influence of the circumglobal wave-train on European summer precipitation." In: *Climate dynamics* 43.1-2, pp. 503–515.
- Scaife, A., A. Arribas, E. Blockley, A. Brookshaw, R. Clark, N. Dunstone, R. Eade, D. Fereday, C. Folland, M. Gordon, et al. (2014). "Skillful long-range prediction of European and North American winters." In: *Geophysical Research Letters* 41.7, pp. 2514–2519.
- Shukla, J., J. Anderson, D. Baumhefner, C. Brankovic, Y. Chang, E. Kalnay, L. Marx, T. Palmer, D. Paolino, J. Ploshay, et al. (2000). "Dynamical seasonal prediction." In: *Bulletin of the American Meteorological Society* 81.11, pp. 2593–2606.
- Shukla, J. and J. Kinter III (2006). "Predictability of seasonal climate variations: A pedagogical review." In: *Predictability of weather and climate*. Ed. by T. Palmer and R. Hagedorn.
- Soares, M. B. and S. Dessai (2015). "Exploring the use of seasonal climate forecasts in Europe through expert elicitation." In: *Climate Risk Management* 10, pp. 8–16.
- Stevens, B., M. Giorgetta, M. Esch, T. Mauritsen, T. Crueger, S. Rast, M. Salzmann, H. Schmidt, J. Bader, K. Block, et al. (2013). "Atmospheric component of the MPI-M Earth System Model: ECHAM6." In: *Journal of Advances in Modeling Earth Systems* 5.2, pp. 146–172.

- Tietsche, S, D Notz, J. H. Jungclauss, and J Marotzke (2013). "Assimilation of sea-ice concentration in a global climate model-physical and statistical aspects." In: *Ocean science* 9.1, p. 19.
- Totz, S., E. Tziperman, D. Coumou, K. Pfeiffer, and J. Cohen (2017). "Winter precipitation forecast in the European and Mediterranean regions using cluster analysis." In: *Geophysical Research Letters* 44.24.
- Wallace, J. M. and D. S. Gutzler (1981). "Teleconnections in the geopotential height field during the Northern Hemisphere winter." In: *Monthly Weather Review* 109.4, pp. 784–812.
- Wang, B., J.-Y. Lee, I.-S. Kang, J Shukla, C.-K. Park, A Kumar, J Schemm, S Cocke, J.-S. Kug, J.-J. Luo, et al. (2009). "Advance and prospectus of seasonal prediction: assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980–2004)." In: *Climate Dynamics* 33.1, pp. 93–117.
- Wang, Y.-H., G. Magnusdottir, H Stern, X Tian, and Y Yu (2012). "Decadal variability of the NAO: Introducing an augmented NAO index." In: *Geophysical Research Letters* 39.21.
- Weisheimer, A., D. Decremmer, D. MacLeod, C. O'Reilly, T. N. Stockdale, S. Johnson, and T. N. Palmer (2018). "How confident are predictability estimates of the winter North Atlantic Oscillation?" In: *Quarterly Journal of the Royal Meteorological Society*.
- Weisheimer, A., N. Schaller, C. O'Reilly, D. A. MacLeod, and T. Palmer (2017). "Atmospheric seasonal forecasts of the twentieth century: multi-decadal variability in predictive skill of the winter North Atlantic Oscillation (NAO) and their potential value for extreme event attribution." In: *Quarterly Journal of the Royal Meteorological Society* 143.703, pp. 917–926.
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*. Vol. 100. Academic press. Chap. 8, pp. 334–340.
- Wu, B., J. Lin, and T. Zhou (2016). "Interdecadal circumglobal teleconnection pattern during boreal summer." In: *Atmospheric Science Letters* 17.8, pp. 446–452.
- Wulff, C. O., R. J. Greatbatch, D. I. Domeisen, G. Gollan, and F. Hansen (2017). "Tropical forcing of the Summer East Atlantic pattern." In: *Geophysical Research Letters* 44.21.
- Yasui, S. and M. Watanabe (2010). "Forcing processes of the summertime circumglobal teleconnection pattern in a dry AGCM." In: *Journal of Climate* 23.8, pp. 2093–2114.
- Zanetti, A, R Enz, P Heck, J Green, and S Suter (2004). *Natural Catastrophes and Man-Made Disasters in 2003*.

ACKNOWLEDGMENTS

This work would not have been possible without the contribution and support of many people and I would like to thank all of them here.

First of all, I want to thank Johanna Baehr for supervising this work in the best way I can possibly imagine. Thank you for providing guidance when I needed it, while still giving me just the right amount of space to develop my own ideas. Thank you for always seeing sense in what I was doing and teaching me how to see this sense myself, thereby keeping me motivated. Without you and your never-ending ideas, this work would not have developed the way it did and without you, I probably would not have developed into the person who was capable of writing this thesis. I also owe a debt of gratitude to Wolfgang Müller for co-supervising this work. Thank you for the discussions we had. Your ideas and your honest opinion improved this work in so many ways. I further want to thank Eduardo Zorita for his calm way of reviewing this work throughout the years.

I am very thankful to the members of the Climate Modelling research group at Universität Hamburg. I enjoyed the group meetings and discussions we had on not only my work, but also on yours. Thank you for your honest feedback and for the ideas that developed out of it. A special thanks goes to André for the patience and time you took to answer all my questions.

I want to thank Antje, Conny and Michaela for their constant guidance and support which make the life of the IMPRS doctoral researchers much easier. Antje, I admire you for the passion you put into helping every one of us through this rough PhD time. You make our time as part of the IMPRS a really special one and I very much enjoyed being part of it.

Thank you to many of my fellow IMPRS students. Without you, I would not have enjoyed my everyday life at the institute as much as I did. I owe a special thanks to my current office mate David and my former office mate Leo. Thank you for lighting up so many of the days I spent in the office. Leo, thank you for numerous discussions and talks we had and for the lunch breaks in which you tried to understand my unusual taste of food. And a special thanks for helping me to see sense in what I was writing during the process of writing this thesis, and sometimes understanding my research better than I did. David, thank you for cheering me up whenever something went wrong and for helping me clear my thoughts on so many occasions. I hope you will continue learning German, I really enjoyed trying to teach you. Apart from Leo and David, I also want to

thank my friends Clara and Veit for helping me through the rough last weeks before submission by reviewing the numerous drafts of this thesis.

I am more than grateful to my family, especially to my parents, and to Claudia and Andy, whom I consider my second family. Thank you all for your support throughout the years, for helping me follow my dreams, for always believing in me, and for cheering for and with me.

And last but not least I owe my biggest gratitude to my partner Max for supporting me throughout the entire time and for always, always believing in me, especially in the times I couldn't believe in myself. In a way, you were the one who kept me going and without you, I would not be where I am today. I consider myself more than lucky to have you in my life.

VERSICHERUNG AN EIDES STATT – AFFIRMATION ON
OATH

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation mit dem Titel: „Seasonal prediction of European summer climate: a process-based approach“ selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen – benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Dissertation oder Teile davon vorher weder im In- noch im Ausland in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

Hamburg, den 5. August 2019

Nele-Charlotte Neddermann

Hinweis / Reference

Die gesamten Veröffentlichungen in der Publikationsreihe des MPI-M
„Berichte zur Erdsystemforschung / Reports on Earth System Science“,
ISSN 1614-1199

sind über die Internetseiten des Max-Planck-Instituts für Meteorologie erhältlich:
<http://www.mpimet.mpg.de/wissenschaft/publikationen.html>

*All the publications in the series of the MPI -M
„Berichte zur Erdsystemforschung / Reports on Earth System Science“,
ISSN 1614-1199*

*are available on the website of the Max Planck Institute for Meteorology:
<http://www.mpimet.mpg.de/wissenschaft/publikationen.html>*

