Seasonal Forecasting of Extreme Wind and Precipitation Frequencies in Europe

Matthew J. Swann

A thesis submitted for the degree of Doctor of Philosophy
at the University of East Anglia, Norwich
November 2008

Climatic Research Unit,
School of Environmental Sciences,
University of East Anglia,
Norwich, NR4 7TJ, UK

Abstract

Flood and wind damage to property and livelihoods resulting from extreme precipitation events and windstorms in Europe accounts for billions of pounds worth of damage annually. In many cases the interannual variability of these extreme events can be closely related to the large-scale atmospheric circulation. For example the summer 2007 floods in the UK are thought to be at least partly related to an anomalously weak and southward displaced North Atlantic jet stream. This raises the question of whether any useful predictability for these phenomena exists on seasonal timescales.

An exploratory empirical analysis into the potential predictability of 90th and 95th percentile threshold exceedance counts of daily precipitation and peak wind gusts is conducted across a pan-European domain, considering twelve overlapping three-month seasons for each predictand. Models are trained over the period 1958 to 1995, and the period 1996 to 2005 constitutes an independent validation period.

It is widely recognized that European seasonal predictability is inherently low owing to the dominant role of internal atmospheric variability in the midlatitudes. However, a number of studies have identified potential sources of predictability, including the El Nino Southern Oscillation (ENSO) and North Atlantic sea surface temperatures. However in general the observed empirical relationships are not yet fully supported by theory or modeling. The research is conducted within these constraints, such that the principle objective is to identify novel, potentially useful sources of predictability which might in future lead to operationally useful forecasts given the verification of the observed relationships by means of theoretical explanation and numerical modeling.

An initially large set of predictors is tested for field-significant responses in the predictand spatial domain using Monte Carlo resampling. Those predictors which are associated with a field-significant response are included in a model selection algorithm

which selects models based on the mean absolute error of the cross-validated fit. The models are then tested on the validation data.

It is found that appreciable levels of skill exist during the model training period. This skill is attributable to a wide range of predictors. Substantial degradation in skill is observed over the validation period, indicating that either the models are over-fitted, or that nonlinear or nonstationary relationships are identified. However, in some cases skill is retained throughout the validation period. In particular, two regions reaching from the North Sea to the Baltic States, and from the Pyrenees to the Balkans respectively show potentially useful skill for the wind predictands during the early winter. This results from a combination of predictors, but predominantly featuring indices of stratospheric temperature from the preceding summer. Some indices associated with ENSO are also found to be potentially useful, as are some local SST anomalies for the precipitation predictands.

Acknowledgements

A large number of people were instrumental to the completion of this thesis, and I wish to thank them all. Jean Palutikof and Tom Holt put the project proposal together, and guided me through the first year of the thesis. On their departure from CRU, I was most grateful to have Phil Jones take over as my principal supervisor, with additional input from Adrian Matthews and some continued input from Tom Holt. I thank them all for their help throughout their involvement. My particular thanks go to Phil, firstly for taking on the role of my supervisor, and secondly for his invaluable suggestions, and meticulous and prompt reading and commenting of my written work. I am also grateful to Phil and the whole department for their flexibility in allowing me to take on a job and effectively write up the thesis on a part time (or almost part-time) basis! Within CRU, Mike Salmon and Ian 'Harry' Harris deserve particular thanks for their assistance in IT and Matlab-related matters – both have provided help well beyond the call of duty!

My fellow students of climate within UEA – including especially Oli Charlesworth, Carol McSweeney, Kate Willett and Jade Johnstone – have been a great support throughout, particularly in their insistence on discussing nothing but climbing during many lengthy coffee breaks, despite my best attempts to bring up the subject of work. I thank them all for their help and friendship. They all finished before me – it must be the climbing.

Richard Washington really got the ball rolling on this one, and I am privileged to have had him as a tutor in my undergraduate studies, and to have worked with him subsequently. Before I took on this project he suggested that the nature of the topic was such that results may be hard to come by. The following document is a testament to his prescience, and if I have met the challenge with the required rigour it will be due to him in no small part.

My employers at Hiscox have been most patient and accommodating during the completion of this project. Thanks in particular to Rob Caton for his understanding, and Mike Palmer and Matt Harrison, for lending a sympathetic ear!

All my family have offered essential encouragement moral support throughout, as well as welcome breaks in Australia – thank you to all of you.

Finally, Anna – the love and patience you have given me throughout this long journey means the most. This is a landmark for me, which I dedicate to you with all my love and thanks.

Contents

List of Figures

List of Tables

1 Introduction

This study seeks to develop empirical models using linear statistical techniques, which identify potential predictability in extremes of wind and precipitation over Europe at seasonal timescales. The objective is not to develop an operational forecasting product, but rather to explore the possibility that a thorough empirical analysis can offer insights into whether any aspects of the predictands can be skilfully predicted. The research findings are presented sequentially starting with the development of predictand datasets, through the selection and refinement of a potential predictor set, the development of the models, and finally model validation and further discussion on the potential predictive skill – or lack thereof – which results. This introduction presents a brief discussion of the rationale for seasonal forecasting in general in Section 1.1; the issues pertaining to seasonal forecasting in Europe including this study in particular in Section 1.2, and finally sets out the structure of the thesis as presented in the remainder of this document in Section 1.3. Further detail on Sections 1.1 and 1.2 are presented in the context of Chapter 2, comprising a review of the existing literature on seasonal forecasting.

1.1 Introduction to seasonal forecasting

The variability of climate on seasonal to interannual scales has been a central problem for society throughout recorded history, as discussed for example in Troccoli *et al.* (2008). Vulnerability to changes in temperature, precipitation, and wind have been of greater or lesser importance to the means of food production, travel and trade, communications, and a range of other activities, and civilisations have prospered and declined as a direct result of climate variability. In the present day, this relationship still holds, although different societies in different parts of the globe experience very different priorities when it comes to adaptation to climate variability and change. The challenge of seasonal to interannual forecasting – whereby some sort of skilful prediction of the climate can be made at lead times sufficient for society to make adaptations – is therefore an important one, and has been the focus of a great deal of research particularly in the last two decades.

The scientific framework within which seasonal forecasting is carried out is based on the principle that although the atmosphere itself is not thought to be predictable beyond a maximum of two weeks, other more slowly varying components of the climate system may be predictable at longer timescales. The effect of these slowly varying processes upon the atmosphere may be sufficient to distinguish with some level of confidence a departure in the mean behaviour of the atmosphere – or more specifically a particular variable such as precipitation – at seasonal or longer timescales.

The late Nineteenth and early Twentieth Centuries saw the first attempts to understand climate variability in a scientific context as we would recognise it today (e.g. Walker and Bliss, 1932). A particular feature of interest in this context is the recognition that climate variability in many cases is associated with large-scale processes which take the form of teleconnections – linking remote areas through the transfer of atmospheric and oceanic mass and energy. The most important and well known of these is the El Niño Southern Oscillation (ENSO) – an interaction between the trade winds and the ocean currents in the tropical Pacific – which affects the variability of climate throughout the tropics, and beyond, with research in recent years suggesting that its influence may extend as far as Europe (for example Fraedrich, 1994; Broennimann, 2007). Due to the strongly coupled interaction between the slowly varying ocean and the overlying atmosphere in the tropical Pacific, since the 1980s extensive and successful efforts have been made to predict ENSO variability and its associated impacts at seasonal timescales (for example Cane and Zebiak, 1986; Folland *et al*., 2001). These efforts form a large part of the basis of seasonal forecasting as it is known today.

More recently, a number of empirical studies have identified other features of the climate system, including snow cover on land (for example Cohen and Entekhabi, 1999), and the stratosphere (for example Baldwin *et al*., 2003) which may be potentially useful predictors of seasonal climate variability. These and many other research efforts in recent decades have been driven by the need to improve our understanding of what the climate holds in store over the next seasons and years.

Fundamentally there are two methods by which predictive models may be developed – firstly, empirical models which use observations of the climate system to generalise and quantify potentially useful predictive relationships, and secondly numerical techniques, where dynamical models of the climate system are integrated forward in time from as close as possible a representation of the present to express a possible future climate conditioned on the persistence or evolution of the same kind of slowly varying fields that inform the empirical models, but as a function of the model physics. Both methods have their advantages and disadvantages, which can be briefly summarised as follows: for empirical methods the physical processes which cause predictability are not explicitly realised in the models – that is, any statistically significant relationship may be found between predictor and predictand, but it does not follow that real structure (as opposed to noise) in the predictand is explained by the predictor. However, empirical models are by definition conditioned on observed behaviour of the climate system and the full extent of this observed behaviour may not be captured by dynamical models. Therefore, although dynamical models in describing potential predictability, describe a system where all the physical processes may in theory be explicitly linked to the predictability, they may not capture all the features of the climate system relevant to this problem. These issues are discussed more fully in Chapter 2.

1.2 European seasonal forecasting and this research

Seasonal forecasting is a branch of climate science undergoing major development. The theory underpinning the forecasting process, the technology and data that allow methodological advances and the increasing ability of society to facilitate adaptation to climate variability and change are all moving in a direction so as to encourage this process. In most respects the science of seasonal forecasting and its applications still have much further to develop, and any study which seeks to advance or quantify an aspect of the current state of climate predictability must be placed firmly within the broader context of the science.

For Europe specifically, seasonal forecasting is not as advanced as it is for example in certain regions of the tropics, where the influence of ENSO is much more clearly defined and understood. The midlatitude atmosphere evolves as a function of internal variability (as opposed to variability forced by slowly varying boundary conditions) to a much greater extent than does the tropical atmosphere, and therefore predictability at seasonal timescales is thought to be inherently lower than in the tropics. This is the central obstacle to the predictability of the European climate at seasonal timescales (for example as reviewed in Rodwell and Doblas-Reyes, 2006). In this respect, the context within which the work for this thesis is carried out is primarily exploratory. To date no comprehensive attempt to identify potential seasonal predictors for indices of precipitation and wind extremes in Europe exists, although such a body of work, should it exist would be of great value in light of the observed impact of climate extremes in Europe in recent years. Furthermore, it is likely that this impact will increase in the future as rainfall events are expected to become more extreme (Fowler and Hennessy, 1995; Palmer and Raisanen, 2002), and the threat of coastal flooding due to storm surges increases with sea level rise. The main objective is therefore the search for and statistical validation of empirical relationships which may afford some predictive skill for extremes of precipitation and wind in Europe.

However, the exploratory context notwithstanding, there is substantial evidence for potentially skilful predictors of the European climate. These predictors are reviewed extensively in Chapter 2 and Chapter 5, and include large-scale sea surface temperature (SST) anomalies, particularly relating to ENSO, and also to the North Atlantic, where some empirical and theoretical evidence for coupling between the ocean and the atmosphere exists (Rodwell *et al*., 1999). Evidence also exists to support the notion that both land surface snow cover (Cohen *et al*, 2001) and the relatively slowly evolving stratosphere (Baldwin and Dunkerton, 2001) might offer additional predictability.

1.3 Thesis outline

Following this introduction, the study is presented in the form of a literature review, a brief methodology outlining the process by which the predictors are selected and related to the predictands, two data chapters dealing with predictands and predictors respectively, and three results chapters detailing the sequential development and validation of the models. The final chapter summarises and concludes the study. The chapter structure is as follows:

1. Introduction. A brief outline of the problem is presented, and the thesis structure is explained.
2. Literature review. The research draws on a number of topics, including the science of seasonal forecasting, and climate extremes. The full range of topics relevant to this research is discussed in light of the existing literature, including a more comprehensive introduction to seasonal forecasting and extremes.
3. Methodology. A brief illustration of the methods by which an initially large predictor set is refined and models are developed and tested, is presented. The emphasis is on the requirement that this study is treated as exploratory in nature, and rather than precisely quantifying observed skill, seeks rather to offer insights into what features of the climate system might offer predictability.
4. Predictands. The development of precipitation and wind predictand datasets comprising seasonal counts of exceedances over the 90^{th} and 95^{th} percentile thresholds is presented.
5. Predictors. Potential predictors are discussed with reference to the literature, and a full set of predictors is defined, with the aim of substantially reducing this set as part of the model selection process.
6. Initial predictor selection. The full set of predictors is refined using field-significance testing by resampling within the model training period. Only those predictors which are associated with a field-significant response in the predictand data are retained.
7. Model selection. Models are developed using an all-subsets selection algorithm, accounting for multicolinearity, and using cross-validation to assess the best model in each case.
8. Model validation. Models are tested on an independent validation period and the results are reported, paying particular attention to the cases where skill appears to be retained in the validation period. These cases comprise a small minority of the total.
9. Conclusions and suggestions for further work. The findings are summarised and conclusions are drawn. Suggestions for further work are made.

2 Literature Review

2.1 Introduction

This literature review presents a survey of the literature relevant to the PhD thesis 'Seasonal forecasting of extreme wind and precipitation frequencies in Europe'. A broad range of research is covered, but the review focuses particularly on studies addressing the following questions:

- What is to be predicted, and why?
- How has the problem traditionally been approached?
- Finally, what empirical and theoretical material is available to aid the development of a seasonal forecasting model?

The review is therefore structured as follows:

- Section 2.1 comprises the introduction.
- Section 2.2 is a review of the literature on weather extremes relevant to the thesis, that is, the predictands.
- Section 2.3 is a broad introduction to the seasonal forecasting problem, comprising:
    - A brief history of the science and general background.
    - Specific work on European seasonal forecasting.
    - The seasonal forecasting of extremes, on which a relatively small body of literature exists.
- Section 2.4 focuses on the seasonal forecasting of extreme events.
- Section 2.5 deals with the potential predictors for the seasonal forecasting model. Literature is examined on a number of boundary forcing processes: including sea

surface temperature (SST), land surface processes – primarily snow cover, and atmospheric parameters, including analyses of stratospheric data.

- Section 2.6 summarises the review.

Kushnir *et al*, (2004) describe the evolution of the science of climate prediction as a three stage process. Firstly, empirical evidence describes potentially predictable relationships within the climate system. Secondly, attempts are made to understand the physical basis for these empirical relationships, and thirdly the replication of these relationships in climate models allows a more complete understanding of the dynamics and potential predictability. With respect to European seasonal forecasting, the science is largely restricted to stage one in this framework, and observed potentially predictable relationships at the seasonal timescale do not provide high levels of skill. There are a few well documented processes where potentially predictable relationships have a tentative theoretical basis, but there is no clear consensus in numerical model studies of these processes.

Currently, several factors inhibit understanding of the theory underlying potentially predictable processes and, thus, impede the development of skilful operational forecasts. Primarily, in many cases the apparently nonlinear nature of such relationships not only makes them difficult to identify, but also difficult to test using meaningful measures of statistical significance. This is because sample sizes derived from the observed record are small, and the range of possible outcomes increases greatly when nonlinearities are considered. The current inability of dynamical models to capture the full range of interactions between components of the climate system (as discussed for example in Broenniman, 2007) means that the sample size cannot be increased to the point where a more robust assessment of significance can be made.

2.2 Predictands

Operational seasonal forecasts attempt to predict climate anomalies (deviations from the mean), for a given season, location and variable. Traditionally the variable under consideration is temperature or precipitation, with the prediction indicating if mean conditions in the given season and location will be warmer/colder or wetter/drier than average. More recently, and particularly with the development of dynamical ensemble forecasting techniques (for example, see Palmer *et al.*, 2004), dynamical probabilistic forecasts have been possible, with a probability assigned to the forecast anomaly based on the ensemble spread.

However, when considering the interannual variability and forecasting of extreme events, a number of problems arise, requiring choices to be made. These choices can be split into two categories. First, the nature of the predictand must be considered; that is, how the phenomena are best represented for the required purpose. Second, the source of the raw data – which variables to use, and the temporal and spatial resolution of the data – is of great importance. This section will review the extremes literature, and then address these two fundamental requirements.

2.2.1 Extreme Weather in Europe

Storms and flood events account for the bulk of severe weather occurrences (that is, extreme events having a societal impact) in Europe, although extremes of temperature have also had considerable impacts, not least in the summer of 2003. Severe weather events have a considerable human and financial cost in Europe. From 1970 to 2002, the total cost of windstorm damage in Europe was $21 billion, while the cost of floods was $7 billion (Murnane (2004), compiled from Zanetti *et al.* (2003)). More recently, the UK floods of July and August 2007 are estimated by the Association of British Insurers (ABI), among others, to have cost in the region of £3bn, while insured damage from windstorm Kyrill, in January 2007 is estimated by Swiss Re to be in the region of €3.5bn across Europe. All values are given as they appeared at the time of reporting.

Much recent research has been motivated by the need to understand the likely impacts of climate change on the frequency and magnitude of extremes. Recently, a number of studies have documented several measures of extreme events throughout Europe and elsewhere. A special issue of the Journal Climatic Change (volume 42 issue 1, 1999) focuses on the need to improve understanding of the observed and potential effects of climate change on extreme events. Much of this work (e.g. Jones *et al*., 1999; Heino *et al.,* 1999) focuses on temperature extremes, particularly since in general they are found to provide the least ambiguous trend signal in the observed record. However, precipitation and storms also come under consideration. Broadly speaking there is less evidence of trends in the observed record for precipitation and storms, particularly since the spatial coherence of these variables is lower than that of temperature. Nevertheless, this work provides a useful framework for the study of interannual variability of extremes (which must also be considered in the context of climate change). While the seasonal forecasting of extremes must be more directly concerned with interannual variability, significant trends would contribute to predictability, and in some cases for the European sector may be the most important contributor to predictability. Frich *et al*. (2002) continue this work with a study of observed changes in extremes in the last half century. Again, while temperature extremes show some fairly clear trends, results for precipitation are less well defined. Mudelsee *et al*. (2003) provide further evidence for a lack of upward trends in precipitation with their study on trends in flooding in central Europe. However, there is evidence of a decrease in winter flooding in the two basins studied. Klein Tank *et al*. (2002) and Klein Tank *et al*. (2003) also examine trends in daily extremes in Europe as part of the European Climate Assessment & Dataset project (ECA). On closer examination of indices capturing particular properties of precipitation, some trends are evident, although these are spatially variable. The application of these results to seasonal forecasting models is potentially of interest, when an optimal climate normal (OCN) approach is used, and possibly combined with other regression-based methods. The OCN technique, described in van den Dool (2007) uses a segment of the recent past whose length is determined empirically as the predictor, essentially optimising the information in a trend or a decadally varying process to make a forecast. More recently, Luterbacher *et al*. (2004) find that a 500 year reconstruction of monthly and seasonal temperatures reveals marked warming in the last few decades.

As well as the basic development of a predictand dataset for the seasonal forecasting model, there is a growing volume of additional information on large-scale and long-term patterns in extremes that can potentially be used as predictors.

2.2.2 Interannual Variability of Extremes and the Large-Scale Circulation

As well as trends in indices of extremes, perhaps the most important measure of variability is at interannual timescales. All of the literature shows significant interannual variability in the frequency and magnitude of extremes. This raises the question of how predictable is the European climate at interannual timescales, when much of the literature finds that trends and decadal variability offer better scope for predictability? It seems to be the case, for example, that the NAO, which accounts for a large proportion of the climate variability in north western Europe, is more predictable on decadal timescales than interannual (e.g. Hurrell, 1995; Livezey *et a.l*, 1999; Sutton and Hodson, 2003). One approach to studying the interannual variability of extremes, which can give a useful indication of the spatial and temporal coherence of extreme events, is to consider the relationship between extremes and the large-scale atmospheric circulation. Most precipitation and wind extremes in Europe are embedded in the prevailing westerly circulation. Particularly with respect to Europe, large-scale modes such as the NAO are known to relate to the interannual variability of extreme events (e.g Scaife *et al.,* 2008) implying some spatial and temporal coherence in their occurrence.

Fraedrich (1992, 1994) provided some early work on the relationship of ENSO to European extremes, showing that ENSO affects the European circulation, and thus the occurrence of extreme events. Away from Europe, Cavazos (1999) makes a detailed study of the relationship between extreme precipitation in Mexico and Texas and its relationship to large-scale circulation patterns. Of particular interest here is the use of artificial neural networks (ANN) in a self-organising map (SOM) configuration to model this relationship, with some success.

A series of papers (e.g. Schmith, 2001 and Quadrelli *et al.*, 2001) from the Atmospheric Circulation Classification and Regional Downscaling (ACCORD) project (Jones *et al.*, 2001) has added to the understanding of European extremes and the general circulation. Schmith (2001) identifies a global warming signature in observed winter precipitation in north western Europe, using station-based records of precipitation. Principal Component Analysis (PCA) of mean sea level pressure (MSLP) is used to construct a multiple regression model to predict precipitation as a function of the circulation, giving some idea of the role of the wider atmospheric circulation, and the spatial cohesiveness of precipitation in the study area. Quadrelli *et al*. (2001) focuses on links between Alpine precipitation and the large-scale circulation, finding that the December-March season has the strongest links with circulation, and hence the greatest spatial cohesion. Chaboureau and Claud (2003) explore the variability of North Atlantic oceanic precipitating systems, relating individual systems to the large scale circulation – in particular the AO/NAO, and finding significant links between the large-scale circulation and individual weather systems. Kaczmarek (2003) relates the risk of flooding in Poland directly to the NAO, finding a significant relationship between winter and spring runoff and the NAO. Given the highly variable and complex nature of precipitating systems in the midlatitudes, studies on downscaling and the relationship of individual weather systems with the large-scale circulation should be of considerable interest to the seasonal forecaster. Haylock and Goodess (2004) study the links between European extremes of rainfall and the mean large-scale circulation. PCA of station rainfall records identifies coherent patterns in the occurrence and interannual variability of extremes, as well as a significant trend in the data. The NAO was identified as an important descriptor of precipitation extremes, as were a number of other MSLP patterns. While the coherence of extreme rainfall was found to be less than that of mean rainfall, there is still information to be gained from such an analysis beyond the sole influence of local processes.

Wang and Rogers (2001) study cyclogenesis in the North Atlantic, where many of the cyclones formed go on to affect Europe. Two regions of cyclogenesis are compared, and both are found to have different relationships with the atmospheric circulation. The implications of this for seasonal predictability are of importance when considering model design. For instance, predictors may have to be highly localised, and as far as the predictands are concerned, it is possible that a detailed climatology of cyclones may be necessary to successfully predict interannual variability. Further work on cyclones has been carried out by Paciorek *et al*. (2002), who use a series of indices to assess time trends of cyclone activity in the Northern Hemisphere. Further studies of extratropical cyclone climatologies and variability in the North Atlantic/European sector are also of interest. Hanson *et al*. (2004) develop a technique to represent the cyclone climatology of the North Atlantic in two reanalysis datasets, ERA-40 (Uppala *et al*., 2005) and NCEP (Kalnay *et al*., 1996). In this instance, ERA-40 is shown to give a better representation of the observed climatology than NCEP, probably because of its much superior spatial resolution.

2.2.3 Indices of Extremes

This section introduces the literature on extreme precipitation and wind relevenat to this thesis, and the data available to support empirical work on this subject. Further practical detail is provided in Chapter 4.

2.2.3.1 Precipitation

Data sources for extreme precipitation and wind speeds will be considered separately. The bulk of work on precipitation extremes has been carried out on station data (e.g.

Klein Tank *et al.*, 2002, 2003; Moberg and Jones, 2005), with some climate change studies using GCM output (e.g. Palmer and Raisanen, 2002). However, station data are not without problems. For example, Karl (1999) highlights issues of data quality – most notably temporal homogeneity in station records, and differences in observing methods. While it is possible to account for some of these problems, there are still reliability issues, and the network of quality controlled station datasets is sparse. This is a major limitation on the study of the interannual variability and predictability of extreme precipitation, not least because the geographical configuration of north west Europe means that large areas of ocean have no coverage at all. Gershunov (1998) and Gershunov and Cayan (2003) use station data to assess the predictability of heavy precipitation over the contiguous United States. However they use a more extensive network than that available in Europe, partly due to differences in data protection policies in the US, and partly because the study area is mostly land. The limitations of station coverage and availability in north west Europe raise a series of questions for this study. Primarily, the detection of patterns of precipitation variability associated with large-scale boundary forcing processes may be difficult to assess, particularly in terms of their field significance across the whole study domain. For the reasons presented here, precipitation from reanalysis datasets has been considered. There are two reanalysis products currently available with quite long records, from the US National Centers for Environmental Prediction (NCEP), and the ERA-40 reanalysis from the European Centre for Medium Range Weather Forecasts (ECMWF), (Uppala *et al*., 2005). While the reanalysis precipitation is modelled rather than assimilated, it does provide a series of advantages over the station-based datasets. Spatial and temporal coverage is better – and while stations give a good representation the point-source precipitation, they may not capture the full spatial extent and intensity of an extreme rainfall event. Additionally, the ERA-40 reanalysis data divides precipitation into convective precipitation, large-scale or stratiform precipitation, and snowfall. These types of precipitation have different causes, particularly the convective as opposed to large-scale rainfall or snowfall, and so the separate consideration of the extremes may be instructive. A number of studies consider precipitation extremes in reanalysis datasets. Zolina *et al.* (2004) compare four different reanalysis products (NCEP1, NCEP2, ERA15 and ERA-40) and a set of observational (station) data with respect to precipitation extremes in Europe. They fit gamma distributions to derive statistical characteristics of the daily precipitation, including percentile thresholds. A principal component analyses (PCA) of the shape and scale parameters, and percentiles of the Gamma distributions for reanalysis gridboxes, is used to study the interannual variability (and secular changes) of extremes on the basis of the coefficients of the principal components and of the Gamma distributions. The authors conclude that, in general, NCEP products show higher estimations of extreme precipitation than ECMWF products, and are thought to be closer to the observed data. In particular, the NCEP2 data has some assimilation of Xie and Arkin (1997) pentadal precipitation, which is thought to account for much of this correction. However, the NCEP2 record is too short (1979-present) for the purposes of this research. Zolina *et al.* (2004) also conclude that the reanalyses demonstrate acceptable skills in the simulation of the variability of heavy and extreme precipitation in the cold season. However, this does not imply confidence in the actual amounts of precipitation which, in the case of extremes, are thought to be considerably underestimated. However, for the purpose of studying the interannual variability of extremes, this deficiency is not particularly important. Other studies documenting the quality of precipitation data in reanalyses datasets include Kanamitsu *et al.* (2002); Arpe *et al.* (1999) and Hagemann *et al.* (2002). Overall, the implication is that the use of reanalysis datasets for precipitation research must be treated with caution, and rigorous testing and comparison with observed precipitation is required to ascertain the suitability of the data for the analysis in question.

2.2.3.2 Wind Speed

The availability of maximum daily wind speed is very limited. A number of studies use peak wind speed data, but the data tend to be localised in space and time, and concerned with applications such as wind stress on buildings, rather than large-scale variability in extremes (e.g. Dukes and Palutikof, 1995; Zwiers, 1987). In a study of the large-scale variability of high wind speeds, Enloe *et al.* (2004) find a series of significant relationships between peak wind gust speeds and ENSO using a network of stations across the contiguous United States.

With respect to observed maximum wind speeds, ready availability of station data over the period of observation required is limited to Germany, the Netherlands and the UK. Although other data exist, and are obtainable from national meteorological agencies, the cost far exceeds the budget available for this thesis. In all cases, concerns over the quality of station data apply. In the case of wind speeds, which are closely linked to other reanalysis variables such as pressure, the reliability of reanalysis data is likely to be better than precipitation, with orography probably the main constraint on the accuracy of extreme wind speeds. Moreover, as with precipitation, when considering interannual variability the magnitude of extreme events is of secondary concern. For these reasons, reanalysis datasets are a strong candidate for use in deriving the wind predictands.

2.2.4 Summary

In summary, a growing body of literature exists on climate extremes, much of it with a focus on secular changes, and some looking at interannual variability. In almost all cases (other than climate change projections), station data are used to derive indices. The applicability of station data to studies of interannual variability with a view to developing a seasonal forecasting model is questionable. This is particularly the case when one considers that reanalysis datasets – superior in temporal and spatial coverage to station data and less likely to suffer from inhomogeneities – may be able to capture the interannual variability of extremes reasonably well. This warrants further investigation, which will be undertaken as part of this study.

2.3 *Seasonal Forecasting*

2.3.1 Introduction

Research on seasonal climate forecasting has increased considerably since the mid to late 1980s. This increase is highlighted in several recent review papers on seasonal forecasting (Carson, 1998; Goddard *et al*., 2001; Rodwell and Doblas-Reyes, 2006). Furthermore, a search of the peer-reviewed literature (using the ISI Web of Science database) reveals that the number of seasonal forecasting papers published has increased greatly in absolute terms, and also relative to the collective body of climate literature.

There are a number of reasons for this increase in interest in seasonal forecasting. Perhaps the most important facilitator is the huge and continuing increase in available computing power. This expands the capability to analyse large quantities of data rapidly and allows the development of ever more complex and accurate models of the earth’s atmosphere and ocean. Growth in the atmospheric and ocean sciences over the last few decades has also fuelled progress in seasonal forecasting, providing an ever-expanding body of scientific knowledge from which to develop prediction systems. Underlying these developments is the increasing awareness of society’s vulnerability to the interannual variability of climate, and to climate change.

From an end-user point of view, therefore, seasonal forecasting is seen as increasingly important and useful for a variety of applications, including the agricultural and energy sectors, and natural disaster forecasting. Importantly, this is recognised by the seasonal forecasting community, where the process of operational forecasting has moved from being a purely academic exercise to one with the requirements of the user-community very much in mind (Goddard *et al*., 2001).

This literature review will present a brief historical account of the development of seasonal forecasting, focusing on a few key early papers, and charting the subsequent development, particularly over the last 20 years. Initially, the review will focus exclusively on predictability and forecasting of ENSO, where virtually all the early seasonal forecasting efforts were focused. It will then address seasonal forecasting in the Atlantic basin, with a focus on work in the North Atlantic/European (NAE) sector, where predictability is thought to be largely independent of ENSO.

2.3.2 ENSO and Early Seasonal Forecasting

The relevance of the early work to seasonal forecasting in the NAE sector is indirect. It is instructive to follow the history of these early developments in the tropical Pacific, and compare and contrast with current efforts in the NAE sector, which are just beginning successfully to incorporate coupled modeling (e.g. Palmer *et al*., 2004). Currently, for this regions, empirical techniques are still of at least comparable skill to dynamical methods.

ENSO is a globally unique phenomenon. It has far more wide-reaching effects than any other single mode of climate variability. Due to the size of the ocean basin within which it occurs, and to the location of that basin with respect to the overlying atmospheric circulation, it is uniquely important, and currently, uniquely predictable. It is therefore unsurprising that early seasonal forecasting efforts focused on ENSO, and its effects. The development of a mechanistic understanding of ENSO, originating in seminal papers by Bjerknes (1966, 1969) and Wyrtki (1975, 1979) on atmospheric and oceanic components of the system, respectively, fuelled a major effort in the prediction of El Nino events. This was particularly the case in the years following the major El Nino event of 1982-83 (e.g. Cane *et al*., 1986). The importance of this early work cannot be understated and, in the field of seasonal forecasting, constitutes perhaps the most important single development in realizing potential predictability. A subsequent paper by Wyrtki (1985) highlights the importance of research into mechanisms of variability in driving forecasting efforts; examining, in particular, the quasi-cyclical nature of ENSO. The next key development in ENSO forecasting was the introduction of coupled models, combining the theories of Bjerknes and Wyrtki (Cane *et al*., 1986; Zebiak *et al*., 1987). At the time these models were more successful than contemporaneous and earlier statistical models (e.g. Hasselman and Barnett, 1981; Barnett, 1981; Graham *et al*., 1987) in terms of providing longer lead times for predictability, and subsequently most ENSO forecasting efforts have used dynamical techniques in the form of coupled models. Currently, a number of institutions provide skillful forecasts of ENSO based on dynamical and statistical models, including the National Centers for Environmental Prediction (NCEP), (US), the International Research Institute for Climate Prediction (IRI), (US), and the European Centre for Medium Range Weather Forecasting (ECMWF). These forecasts are applied across the tropics and subtropics (e.g. Folland *et al*., 2001), and to some extent in extratropical North America (e.g. Gershunov *et al*., 1998). The continued operational use of statistical models is a testament to that fact that with the appropriate care taken in their development, they can often match the dynamical models in skill, and are orders of magnitude more easy to develop and maintain. Examples of operational statistical schemes include the National Centers for Environmental Prediction (NCEP)/Climate Prediction Center's (CPC) CCA scheme (Barnston and Ropelewski, 1992) and the National Oceanographic and Atmospheric Administration Climate Diagnostics Center (NOAA CDC) linear inverse model (Penland and Magorian, 1993).

2.3.3 ENSO and European Seasonal Forecasting

Most seasonal forecasting efforts are focussed on ENSO related predictability, for the simple reason that ENSO is the largest and most predictable mode of climate variability with a well documented set of teleconnections. Compared to ENSO and the tropics, the mid-latitudes, particularly the Euro-Atlantic region, have low predictability. However, interest in predictability and forecasting in this region is increasing, as demonstrated by a growing body of literature on the subject in recent years. Research is based partly on potential predictability as represented by numerical models (e.g. Martineau *et al*., 1999; Friederichs and Frankignoul, 2003), and partly on observed predictability (e.g. Rodwell and Folland, 2002). Additionally, there is some interest in ENSO-related predictability in the NAE sector. For example, Mathieu and Sutton (2004) find that the influence of individual El Nino and La Nina events is discernable on the North Atlantic/European climate. However, the nature of the forcing is different in each case, and sensitive to specific details of the tropical Pacific anomalies. As far as traditional seasonal forecasting models are concerned, the implications of this are clear – new techniques need to be developed to deal with this level of complexity in the climate system, and ideally the physical mechanisms – if any – associated with the apparent teleconnections need to be reproduced by climate models for a better assessment of the predictability associated with these teleconnections. Such an undertaking is beyond the scope of this study. Wu and Hsieh (2004) examine nonlinear forcing of the Northern Hemisphere (NH) atmosphere by ENSO. A neural network model based on nonlinear canonical correlation analysis (NLCCA) is used to study the effect of ENSO on NH 500-mbar geopotential height, with significant results being obtained for regions including Europe. In particular, the NAO and PNA patterns exhibit a quadratic response to SSTA.

An interesting development from the early coupled model forecasts of ENSO is the recognition that sea surface temperatures (SSTs) are of great value owing to their persistence over a period of months or longer, and to their spatial coherence. This application has been transferred to the North Atlantic with mixed success. There is certainly no consensus on the issue of seasonal forecasting skill in Europe based on SST anomalies. Indeed, some recent work has found little evidence of coupling and suggests that it is, in fact, the atmosphere which exerts a more important influence on the ocean (e.g. Wu and Gordon, 2002). Conversely, a growing body of research finds evidence for coupling between the North Atlantic (and other basins) and the North Atlantic/European sector atmosphere (e.g. Peng *et al*., 2002, 2003).

2.3.4 North Atlantic/European Sector Seasonal Forecasting

Barnston (1994) provides a relatively early assessment of NH predictive skill using canonical correlation analysis (CCA). Wide variations in forecast skill are found, with most skill being associated with areas directly affected by the tropical Pacific. Johansson *et al*. (1998) concentrate on potential skill in northern Europe, again using CCA, and applying the technique to surface air temperature. They find that some skill is present in the winter season, using geopotential height as a predictor. Further work by Johansson (2007) examines predictability of the NAO and PNA patterns, using a dynamical model and finds possible evidence for nonlinearity in the predictability, with strong positive events in the NAO being apparently more predictable – albeit at intraseasonal rather than seasonal timescales. In their review of the prospects for seasonal forecasting, Palmer and Anderson (1999) recognise the nonlinear and chaotic nature of extratropical variability, and the limitations this imposes on seasonal predictability. Their review focuses on the advances necessary in the field of numerical modelling to improve dynamical seasonal forecasting. This focus is of particular relevance to European seasonal forecasting, where it is far from clear whether numerical or empirical techniques offer the most promising solutions in the medium term. Currently, it is widely accepted that both approaches are equally skilful (that is to say, both have rather low skill), but that dynamical methods offer the greater long-term potential, whilst in the medium term both methods should be used. This thesis will deal exclusively with empirical statistical techniques, and will seek to include a wide range of potential predictors. Additionally, it is hoped to capture more complex interactions that cannot be represented in the current generation of numerical models. For example, little is currently known about the possible linkages between ENSO and the stratosphere, and how this may in turn affect the Northern Hemisphere extratropical climate. Preliminary research in this area has been undertaken by Broennimann (2007), and Garcia-Herrera *et al* (2006). Currently, complex nonlinear and/or threshold-dependent predictable relationships between European predictands and boundary forcing conditions are beyond the reach of empirical study owing to the relatively small sample sizes of available highly-detailed climate data. For example, Broennimann *et al* (2004) find evidence for a Eurasian winter response to the strong

ENSO of 1940-41, which they are able to reproduce in a climate model. They observe that it was the particular features of this ENSO – i.e. unusually strong and sustained – which seemed to cause such a notable response in Eurasia. This supports the notion that there may be thresholds at which particular anomalies in the climate system might force responses in a nonlinear fashion. However, this study represents one event, and, as mentioned above, in order for a robust empirical appraisal of such nonlinear components of seasonal climate variability and potential predictability, a much greater sample size is required. These current shortcomings in our understanding and in the available data underpin the need for accurate, high-resolution models which could capture these nonlinearities. One of the main aims of this thesis is to identify some potential nonlinear predictable relationships.

Since Barnston (1994) and Palmer and Anderson (1994) no major increases in European seasonal forecast skill have been made, although a number of studies report incremental improvements in certain areas, using both empirical and dynamical techniques. There is also increased understanding of relevant boundary forcing processes, which offer some encouragement that further advances will be made. The remainder of this section will review a selection of seasonal forecasting studies relevant to the thesis. It should be noted at this point that there is recognition amongst potential end-users of European seasonal forecasts that even marginal improvements in forecast skill have practical value. Saunders and Qian (2002) identify levels of skill corresponding to R^2 values of ~0.5 to 0.6 as being 'marginal but useable' in the prediction of the winter NAO index. Such levels of skill would likely be very useable for any industry which is able to apply a multi-year approach to using seasonal forecasts. For example, energy companies, and insurance and other financial sector companies might stand to gain over the longer term from smaller improvements in skill than for example the agricultural or tourism sector, where there is less opportunity to hedge against the risk of an incorrect forecast.

Davies *et al.*, (1997) use the Hadley Centre Atmospheric Model (HADAM1) to assess predictability in the NAE sector, finding that it is greatest in winter and spring. Furthermore, in years with extreme SST anomalies in the North Atlantic, it is possible to obtain useful predictive skill for the NAO. Colman (1997) employs empirical techniques to investigate the relationship between North Atlantic SST anomalies and summer central England temperature (CET), obtaining a correlation skill of about 0.5 at a four month lead time. While this study is not directly relevant to the thesis, its results may have some bearing on the precipitation model in particular. Feddersen (2000) studies temperature forecasting in summer and winter throughout Europe, employing worldwide SSTs, and using the ECMWF GCM. A relationship between temperature at 850hPa (T850) and ENSO is identified, although the nature of the forcing is not stationary. Additionally, when the results are compared with a reanalysed dataset, this link is not found.

A series of three papers by Pavan *et al*. (2000a, b, c) highlight some of the problems of nonlinearity faced in European seasonal forecasting. Firstly (Pavan *et al*., 2000a), in a paper on the seasonal prediction of blocking frequency, it is found that the skill of a GCM in predicting the large-scale evolution of blocking patterns is a function of the pattern considered. In the second paper (Pavan *et al.*, 2000b), the results of a set of multi-model seasonal hindcasts are considered. It is found that skill is highly variable over Europe, and is also dependent on the variable and the season. Systematic error in the model does not improve the hindcast skill. Finally (Pavan *et al*., 2000c) considers the interannual variability of large-scale flow over Europe. A number of important patterns are identified using an EOF analysis. The relationships between these modes are nonstationary, and in addition, there are areas where the model does not represent the observed very well. The implication of these papers is that a more comprehensive empirical understanding of boundary forcing is required, in parallel with improved representation by models. These shortcomings notwithstanding, it is found that the use of large ensembles for seasonal forecasting does improve skill., See, for example, Piedelievre (2000); Richardson (2000); Shukla *et al*. (2000); Doblas-Reyes *et al*. (2003); Feddersen (2003); Palmer *et al*. (2004).

A number of studies focus on the observed record alone. For example, Lloyd-Hughes and Saunders (2002) find that spring precipitation forecast skill is significant based on ENSO and North Atlantic SSTs, with model skill 14-18% better than climatology. Wedgbrow *et al*. (2002) find that a range of relationships exist between large-scale atmospheric circulation patterns such as the NAO and Polar-Eurasian (POL) patterns, SST anomalies in the North Atlantic and summer rainfall and river flow anomalies in England. However, the forecasting rules can only be applied to specific predictor configurations, and further research (and more data) is required to extend these forecasts to cover a greater range of years. A similar study by Wilby (2001) finds a relationship between the NAO index and summer monthly mean flows for a number of UK rivers. Qian and Saunders (2003) use northern hemisphere summer snow extent to forecast the following winter NAO index, and find that this provides reasonable skill when applied as a measure of North Atlantic winter storminess. Blender *et al*. (2003) use linear regression to develop a seasonal forecasting model predicting monthly mean temperature anomalies for North Western Europe. Again, observed North Atlantic SSTs are used as predictors, in conjunction with a selection of European climate variables. They find marginal improvements over climatology in the skill, which, in turn is significantly better than in a simulated forecasting experiment.

Also worthy of note is the importance of decadal scale variability in the NAE sector, both in the ocean circulation, and especially the NAO. The enhanced predictability of the NAO at decadal rather than interannual timescales makes a strong case for the use of the Optimal Climate Normals (OCN) method used in several operational schemes, notably that at the US Climate Prediction Centre (CPC) (van den Dool, 2007). OCN simply takes the climatology for the number of years that optimise forecast skill, and applies this to the forecast – either directly as the forecast itself, or conceivably as a coefficient for a forecast derived by other methods.

As well as the sample of experiments outlined above, a number of meteorological agencies in Europe carry out operational seasonal forecasting programs, including the UK Met Office (UKMO), the European Centre for Medium Range Weather Forecasting (ECMWF), and MeteoFrance. In addition to this, forecasts are made on a commercial basis, the results of which are not freely available.

2.3.5 Seasonal forecasting using Artificial Neural Networks (ANN)

Although this study is limited to the application of linear statistical techniques, is it of interest to note that nonlinear methods, such as artificial neural networks have been applied with apparently useful results. A relatively small number of studies have applied ANN methods to seasonal forecasting. Specifically for Europe, Bodri and Cernak (2000) and Bodri (2001) apply ANN models to the prediction of extreme precipitation, with encouraging results as far as seasonal predictability is concerned. In this case, the neural network model used is a back-propagation network. Dawson and Wilby (1999) Shamseldin and O'Connor (2001) and Campolo *et al*. (2003) apply neural network methods to river flow forecasting at much shorter timescales, with considerable success. Worth noting here is that a variety of ANN model structures exist, and it is important to pick a model suitable for the application. For example, Dawson and Wilby (1999) found that a multi-layer perceptron model was considerably more effective than a radial basis function model.

Papers by Hsieh and various co-authors identify key areas where ANN models may be applied to improve understanding of climate variability and to implement climate forecasting. Hsieh and Tang (1998) explore the application of ANNs to prediction in meteorology and oceanography, identifying several techniques which simplify the analysis and interpretation processes, including the use of principal component analysis (PCA) to reduce dataset dimensionality. A further series of papers (Hsieh, 2000; Hsieh, 2001a, b; and Hsieh and Wu 2002) develops nonlinear forms of canonical correlation analysis (NLCCA), PCA (NLPCA), and singular spectrum analysis (NLSSA) for application to climatology. In each case, these methods offer new insights into climate variability in the tropical Pacific. It is possible that these methods could be applied to the NAE sector climate. However, the lack of signals as strong as those in the tropical Pacific may be an obstacle to increased understanding using these techniques. Where linear techniques are used successfully in the NAE sector – such as PCA or SVD to represent major modes of oceanic or atmospheric variability – it may well be the case that ANN methods offer new insights into variability, and possibly could be used to provide predictor timeseries for a seasonal forecasting model.

2.3.6 Summary

The implication from many of the European seasonal forecasting studies is that the predictor-predictand relationships are highly complex, and much work remains to be done. As is to be expected, much of the interannual variability of the European climate, particularly in winter, can be described by the NAO. It is therefore of real importance to be able to improve the predictability of the NAO at seasonal timescales. The predictability of the European climate appears to be highly dependent on the variables and seasons considered. For instance, temperature generally has higher predictability than precipitation. This has some implications for the thesis, both in terms of the lower expectations of precipitation forecast skill, and the nature of the precipitation predictands with respect to spatial and temporal aggregation.

There is evidence to support the fact that predictability of the European climate may be nonlinear. However, for the purposes of this study it is thought that an insufficiently long observational record is present to allow the reliable use of nonlinear techniques. Furthermore, it is felt that there is particular danger in applying nonlinear computer-learning algorithms to data where the nature of the physical links between the predictors and the predictand (if any) are not fully understood.

2.4 *Seasonal forecasting of extreme events*

2.4.1 Introduction

This thesis is primarily concerned with the prediction of variations in the frequency of extreme events at the seasonal scale. For this to be possible, it is self evident that there should be some seasonal predictability at least of the mean state of the climate (given that the occurrence of extreme events is related in some way to the boundary forcing processes that affect the seasonal mean), or more desirably that the frequency of the extremes themselves should be predictable. Additionally, the occurrence of extreme events within the seasonal mean climate should exhibit interannual variability. The first requirement – that of predictability of seasonal means - has been addressed elsewhere. This section will address extreme events in more detail, both with respect to their interannual variability, and their potential seasonal predictability.

Owing to the impact of extreme weather events in Europe, and current concerns about the impact of global warming on their frequency and magnitude (e.g. Fowler and Hennessy, 1995; CLIVAR/GCOS/WMO workshop on indices and indicators for climate extremes, 1999; Palmer and Raisanen, 2002), the study of the observed record of extremes has increased in recent years. Consequently, extreme wind and rainfall/flooding events in Europe are relatively well documented.

2.4.2 Seasonal forecasting of extreme events

Compared to the forecasting of seasonal averages, attempts at forecasting extremes are relatively rare, and there are fewer publications. A number of studies focus on the USA (e.g. Gershunov and Cayan, 2003). In a series of papers by Gershunov, Cayan and others, the effect of ENSO on extreme rainfall in the USA is studied. It is found that regional predictability in the USA is a function of the strength and stability of the ENSO signal in that region; although not all clear ENSO signals result in predictive skill. Cayan *et al*. (1999) find that ENSO has a significant effect on the frequency distribution of daily precipitation and stream flow for selected seasons and regions in the USA. They identify these shifts with changes in the likelihood of flooding (with the effect of ENSO on streamflow being amplified over that on precipitation), and conclude that the prospects for seasonal forecasting are hopeful. However, the identification of differences in the response to each ENSO event mean that care must be taken, even when using a powerful signal like ENSO as a predictor. More recent work (Gershunov and Cayan, 2003) considers ENSO, as well as other predictors, for extreme rainfall in the USA. While the majority of the forecast skill is due to ENSO, there are also potentially useful non-ENSO predictors, including north Pacific SST forcing. Additionally, it is found that predictions of mean, or less extreme, precipitation are considerably more skillful than those for the 90th percentile or lower probability events. A pure statistical model is shown to outperform a hybrid dynamical-statistical model (where SST drives atmospheric predictors in a GCM).

As well as US precipitation forecasts, some work has been done on the seasonal forecasting of east coast storm systems. Notably, DeGaetano *et al*. (2002) devise a nonparametric statistical model for predicting US east coast winter storm frequency. A pool of potential predictors is reduced using a chi-squared screening procedure, and used to develop a series of discriminant functions, relating them to above or below normal storm activity. A high degree of skill is obtained for the forecasts, although the physical mechanisms behind this skill are not clear, and further work remains to be done in support of the results. The application of similar methods to the problem of European storm forecasts may well be of use, although similar problems are likely to be encountered concerning the mechanistic explanations for any useful skill.

Qian and Saunders (2003) develop a model to forecast wintertime storminess over the North Atlantic. This model is based on the relationship between the NAO and northern hemisphere summer snow extent, putting into practice some of the theory discussed in the section of this review on snow cover and seasonal predictability. Results are encouraging, and a plausible physical mechanism for the relationship between snow cover and the NAO is included, although the short length (29 years) of the snow cover records used means caution should be exercised in the application of these results. Incidentally, the ERA-40 dataset – including the operational analysis – provides close to 50 years of snow cover data, and might well be of use when applied to this problem.

A report by Holt *et al*. (2001) for the TSUNAMI project includes a seasonal forecast model for wind exceedances using a variety of regression techniques and a set of predictors taken from North Atlantic SSTs and large-scale atmospheric circulation indices. Partial least squares regression was found to provide the most accurate forecasts, as well as being computationally fast, from a practical point of view. The aim of this thesis is partly to build upon this work, applying new methods, and possibly including new predictors.

2.4.3 Summary

Overall, while the number of seasonal forecast models to predict extremes is limited in the academic literature, there is the basis for the development of valid models for extreme rain and wind events in Europe. However, a central tenet of any interpretation of forecast results and skill should be that without plausible or demonstrated physical mechanisms, results should be treated with caution. In the longer term, and beyond the scope of this study, the development of numerical models which recognise a more complete suite of forcing mechanisms relevant to this problem will be necessary. In the meantime, empirical methods, while potentially providing forecast skill, should also add to our understanding of where to look for the physical mechanisms which explain interannual climate variability in Europe.

2.5 Potential sources of predictability

2.5.1 Atlantic Ocean

2.5.1.1 Introduction

Current seasonal forecasting efforts are largely predicated on the persistence of ocean mixed layer temperature anomalies, the predictability of these anomalies, and their coupling with the overlying atmosphere. It is therefore to be expected that a major goal in the development of seasonal forecasting systems for Europe is the understanding of oceanic variability in the North Atlantic, and its interaction and possible coupling with the atmosphere. This section will focus on oceanic variability at a range of timescales, and consider key relationships between the ocean and atmosphere, and the development of understanding of Atlantic (and in particular North Atlantic) climate variability and potential predictability. Variability of the North Atlantic Oscillation (NAO) and other modes of atmospheric variability will play a role in this discussion, and the relationship of these phenomena to European climate will be considered in more detail in a separate section. Other potential predictors of the NAO and other circulation modes will also be considered separately.

The North Atlantic Ocean comprises the boundary conditions upstream of north western Europe. It is clearly a candidate for sources of potential seasonal predictability, although the picture that has emerged from some three decades of research is highly complex, and far from being resolved. The key difficulties are in unpicking the nature and strength of air-sea coupling in the North Atlantic, and to some extent the relationship between the tropical and extratropical Atlantic regions. Both are complicated by the stochastic nature of atmospheric variability in the extratropics, in which any potential signal (linear or otherwise) imparted by the more slowly varying ocean is likely to be significantly smaller than internal atmospheric variability (e.g. Weng and Neelin, 1998). The latter (tropical-extratropical coupling of the Atlantic) is further complicated by the incomplete understanding of ocean variability in the tropical Atlantic (e.g. Marshall *et al.*, 2001). However, considerable progress is being made in understanding all these problems, and in particular, the development of more sophisticated global climate models (GCMs) has allowed much interesting work to be carried out in recent years, considerably enhancing our understanding of the problem (e.g. Huang *et al*., 2004; Rodwell and Folland, 2002).

Marshall *et al*. (2001) identify three key interrelated phenomena that comprise climate variability in the tropical and North Atlantic basins:

- Tropical Atlantic variability (TAV), comprising
  - A covarying fluctuation of tropical Atlantic sea surface temperature (SST) and
  - Trade winds straddling the Intertropical Convergence Zone (ITCZ)
- The NAO, which is the primary mode of climate variability in the North Atlantic and surrounding regions, consisting of a dipolar exchange of atmospheric mass between centres of action located roughly over the Azores and Iceland (e.g. van Loon and Rogers, 1978).
- The Atlantic Meridional Overturning Circulation (MOC), which is a measure of the intensity of the global Thermohaline Circulation (THC) in the North Atlantic, driven by convection in the sub-Arctic, and affecting the amount of heat transported northwards from the Equator.

While these three phenomena are interrelated to some extent, they all have distinctive signatures in the Atlantic basin, and can be identified as varying individually in time and space. The NAO and its relationship with the North Atlantic basin will be considered in most detail here, as it has the most direct and obvious effect on European climate. In addition, there is some evidence to suggest that tropical Atlantic variability may be an important source of potential predictability for the North Atlantic/European sector (e.g. Peng *et al.*, 2005).

2.5.1.2 North Atlantic ocean-atmosphere variability

As discussed previously, the North Atlantic Oscillation (NAO) is the primary mode of atmospheric variability affecting European climate. Patterns of variability connected to the NAO have been recognised for centuries, and in 1932 Walker and Bliss provided perhaps the earliest description of the coherent phenomenon at sea level over the Atlantic. Much more recently, the question of whether the NAO is in fact a local expression of a hemispheric mode – the Arctic Oscillation (AO), or Northern Hemisphere Annular Mode (NAM) – has been raised (Thompson and Wallace, 1998). While important, this question is deemed to be of peripheral relevance in this discussion, and will be addressed in more detail elsewhere. With respect to seasonal forecasting of the European climate, the key issue is identified as predictability of the NAO at a range of timescales. It is therefore important that this research focuses on air-sea interactions in the North Atlantic - a feature it shares in common with most attempts to understand the seasonal predictability of European climate, of which a case in point is CLIVAR Atlantic research program which has co-ordinated much of the recent work on this subject. (http://www.clivar.org/science/atlantic.htm). The earliest work on air-sea interactions in the Atlantic was carried out by Bjerknes (1959, 1964), Frankignoul (1978) on the generation of SST anomalies, and Frankignoul and Reynolds (1983) on mid-latitude SST anomalies in a dynamical model. Numerical modelling has become indispensable to the development of our understanding of North Atlantic climate variability. Indeed Wallace *et al*. (1990) write of the impossibility of distinguishing the direction of forcing between the atmosphere and ocean from observed data alone. This highlights the importance of a physical understanding of the processes, and this in turn has increased along with the improvement of climate models.

Prior to the 1990s, most work on the North Atlantic focussed on atmospheric forcing of the ocean (Wallace *et al*., 1990), as was apparent in the structure of SST anomalies and lagged relationships with the atmospheric circulation (this is in contrast to the North Pacific, where evidence for atmosphere-ocean coupling was uncovered during the 1980s (Namias and Cayan (1981) and Namias *et al*. (1988)). However, during the 1990s a parallel strand of research found evidence for atmosphere-ocean coupling in the North Atlantic (e.g. Rodwell and Folland, 2002). Both of these strands will be discussed, but the latter is of particular interest owing to the associated implications for seasonal predictability.

2.5.1.3 Uncoupled variability in the Atlantic

It is widely accepted that the main driver of SST anomalies in the North Atlantic on interannual timescales is atmospheric forcing. The leading mode of SST variability in the North Atlantic is a tripole structure – the North Atlantic tripole (Friederichs and Hense, 2003; Cayan, 1992a,b), with centres of the same sign located east of Newfoundland just south of the Labrador Basin, and spanning the tropical/subtropical North Atlantic centred on about 15°N. The centre of action with the opposing sign lies in between these two, in the Gulf Stream region, at about 35°N, 60°W, concentrated more heavily in the west of the basin. Cayan (1992a) finds that the observed SST anomalies that make up this leading mode are driven primarily by latent and sensible heat flux anomalies. These in turn are driven by the overlying atmospheric circulation. As the ocean integrates atmospheric forcing over a weekly-monthly period, the seasonal-scale characteristics of the atmospheric circulation are captured. In the case of the North Atlantic tripole, the SST anomalies can be expressed as being driven primarily by the NAO at seasonal timescales, with a positive NAO index (NAOI) resulting in negative anomalies in the subpolar and subtropical regions, and a positive anomaly in the Gulf stream region. This atmospheric forcing relationship also holds for the second and third modes of SST variability in the North Atlantic – driven by lesser modes of variability in the atmosphere (Cayan, 1992a). These relationships are explored in further detail in Cayan (1992b). Halliwell and Mayer (1996) explore the processes by which apparently stochastic wind forcing generates SST anomalies by driving the heat flux between the ocean and atmosphere. This acts as both a forcing and damping mechanism, giving rise to temporally coherent SST forcing over timescales ranging from a few months to eight years (Halliwell and Mayer, 1996). Zorita and Frankignoul (1997) develop the concept of a ‘red’ spectral response to the ‘white’ atmospheric forcing at interannual to decadal timescales. Their results are based on coupled model output, using the ECHAM1/LSG coupled ocean-atmosphere GCM (CGCM), and they show that stochastic forcing by the atmosphere can contribute to oceanic variability at decadal timescales.

As well as direct localised heat exchange driven by the atmosphere, oceanic advection processes are also important in driving SST anomalies. Delworth and Mehta (1998) explore the role of advective processes in observed data and in data from a CGCM. While this study is confined to the tropical and subtropical Atlantic, the authors find that on interannual timescales anomalous heat fluxes drive SST variability, while advection becomes increasingly important at decadal timescales. This has implications for longer range predictability of the ocean-atmosphere system over the North Atlantic, and is related to TAV and MOC variability, which will be explored in more detail later. Saravanan (1998) explores the relationship between atmospheric low-frequency variability and midlatitude SST variability in the National Center for Atmospheric Research (NCAR) Climate System Model (CSM).

In the late 1990s coupled modelling was beginning to develop to the stage where models could represent the variability and mean state of the climate system quite well, without flux adjustments. The NCAR CSM is one of this cohort of models, and is thus suitable for investigation of ocean –atmosphere interaction in the North Atlantic. Saravanan (1998) found that SST variability in the North Atlantic is primarily driven by atmospheric forcing at the interannual timescale – in accordance with studies based on the observed data, for example Wallace *et al*. (1990) – and that the main modes of SST variability are driven by the main modes of variability in the overlying atmosphere. Saravanan's study makes a strong case for the use of coupled models to develop our understanding of North Atlantic climate variability, based on internal nondeterministic variability of the midlatitude atmosphere, and its effect on the ocean, which renders an atmospheric model forced with SSTs inadequate.

In general, atmospheric forcing of the heat flux between atmosphere and ocean is thought to be the main driver of North Atlantic SST anomalies and variability at interannual to decadal timescales. A large body of literature exists on this. For example, using observed data, Zorita *et al*. (1992) use canonical correlation analysis (CCA) to demonstrate that the leading modes of sea level pressure (SLP) in the North Atlantic lead winter SST. Deser and Blackmon (1993) describe North Atlantic Climate variability in the 20th century and find that the association between an SST dipole (similar to the northern two centres of the North Atlantic tripole), and surface winds. Anomalously strong winds occur over anomalously cool SSTs and vice versa. Furthermore, quasi-biennial and quasi-decadal peaks in the power spectra of these covarying timeseries are observed. This implies that ocean-atmosphere coupling may be present at timescales relevant to the seasonal prediction problem. The results indicate an important role for oceanic advection as well as for local wind forcing. Halliwell (1997) expands on this, using individual anomalies of surface pressure between 1950 and 1992. The influence of these anomalies on SST is noted, but the key point of the Halliwell analysis is that the atmospheric variability is nonstationary when studied using this methodology. The associated consequences of this are important, particularly if air-sea coupling is present, as far as seasonal predictability is concerned.

Studies based on model output are also numerous. Miller (1992), in one of the earliest coupled model studies, found that atmospheric forcing is predominant, although the findings are inconclusive due to limitations of the model physics. Battisti, Bhatt and Alexander (1995) study an ocean model coupled to observed atmospheric parameters. In the absence of any oceanic forcing, the model reproduces the observed temporal and spatial structure of North Atlantic SST anomalies well, with some regional exceptions attributed to advective processes (not included in the model). Delworth (1996) uses the Geophysical Fluid Dynamics Laboratory (GFDL) coupled GCM to investigate zonal bands of SST anomalies in the North Atlantic, with similar findings to Battisti *et al* (1995) and Cayan (1992a,b) (using model and observed data respectively). That is, that the SST anomalies are primarily driven by anomalous surface heat fluxes, which are in turn driven largely by the atmosphere. Blade (1997) finds that dominant atmospheric modes lead equivalent oceanic modes by about one month; however, there is evidence that coupling between the ocean and atmosphere can result in a discernable signal in the 500mb geopotential height field. The study identifies a negative thermal feedback on the atmosphere associated with heat flux anomalies. Coupling reduces this feedback, and therefore enhances the persistence of atmospheric anomalies. This result is based on a comparison between coupled and uncoupled realisations of the model.

A suggestion for future research made by Blade (1997) is to compare long coupled runs with long uncoupled runs using fixed SST data provided by the coupled runs. Visbeck *et al*. (1998) also study the response of an ocean model to atmospheric forcing derived from a boundary layer atmospheric model coupled to an OGCM. Results indicate that the subtropical cell of the Atlantic tripole SST mode is directly influenced by atmospheric forcing related to the NAO. In the Gulf Stream region, advection and re-emergence, in addition to atmospheric forcing, play a role in determining SST anomalies. In the subpolar gyre region, the strongest response of the ocean is on a decadal timescale, raising the possibility of a feedback loop where negative SST anomalies and a positive NAO phase are coupled. In this scenario, the positive feedback results in a persistent negative SST anomaly to the point where advection related to the resulting enhanced thermohaline circulation warms the subpolar region, reversing the cycle at decadal timescales. Other studies propose similar mechanisms (e.g. Zorita and Frankignoul, 1997; Timmerman *et al*., 1998), with the result that decadal variability and predictability are reasonably well understood in the North Atlantic. Further studies find no or very little discernable influence of the midlatitude ocean on the overlying atmosphere (e.g. Lau, 1997; Saravanan, 1998). The particular benefit of these model studies has been in furthering an understanding of physical modes of variability in the North Atlantic Ocean, and to a lesser extent in the overlying atmosphere.

2.5.1.4 Coupled ocean-atmosphere interaction

A branch of research running parallel to the study of uncoupled atmospheric forcing of the North Atlantic – and focussing on coupled variability – has grown in recent years, although the idea of coupled variability in the North Atlantic has been around for some time (e.g. Palmer and Sun (1985), who were perhaps the first to apply numerical techniques to this particular problem). One of the key drivers of this recent development has been the increasing sophistication and accuracy of GCMs, allowing relatively long integrations of realistic atmospheric variability. As well as in model data, the observed data has also provided evidence for coupling in the North Atlantic sector (e.g. Wallace, 1990; Rodwell and Folland, 2002). In particular, the high resolution satellite data that has been available in recent years has provided new and important insights into this phenomenon (see Xie, 2004). This will be discussed in more detail below.

This section provides a review of work that has identified evidence for coupled variability in the North Atlantic sector, with an appraisal of the associated implications for predictability at the seasonal to interannual timescale. The implications of this work are of considerable importance to European seasonal forecasting, as SST predictors provide one of the most important areas of research concerning potential seasonal predictability of the European climate and in the UK Met Office seasonal forecasts of the NAO, one of the few instances where operational forecasts with appreciable levels of skill are produced for Europe (Rodwell *et al*., 1999).

2.5.1.4.1 Model studies

Miller (1992) is one of the earliest studies to use a coupled model realisation of the North Atlantic sector. The research indicated the presence of some coupling, although this did not stand up to detailed scrutiny, and was considered inconclusive. Increasingly, throughout the 1990s, more detailed coupled models, and further coupled and atmospheric model validation against the observed has reduced these uncertainties. In fact, a significant proportion of the research in this area has focussed on the forcing of AGCMs with observed and idealised SST anomalies. Peng *et al.* (1995) show that an atmospheric model can react to prescribed SST anomalies in the mid-latitudes, including shifts in the location of the storm track over the North Atlantic. In common with other results (e.g. Peng, 2002) it is found that the atmospheric response is nonlinear, being significant only in the case of positive SST anomalies. Other relatively early studies finding evidence for coupling or ocean to atmosphere forcing include Power *et al*. (1995), Ting and Peng (1995) and Kushnir and Held (1996). Rodwell *et al*. (1999) used an AGCM (Hadley Centre 2nd generation Atmospheric Model, HadAM2) to examine the influence of North Atlantic SST anomalies on the NAO index. Using an ensemble of six runs, each comprising the latter part (1947-97) of 128 years of model output, and forced with observed SST, it is found that the ensemble reproduces multi-annual to decadal variability in the NAO. Correlations with the observed record demonstrate that this result is significant at the 98% level for the unfiltered NAO indices, while significance increases when the indices are filtered to remove high frequency (interannual) variability. Contrary to much of the earlier work showing that North Atlantic SST has little or no effect on the overlying atmosphere, results from this study indicate that the SST tripole mode – which is believed to be generated by anomalous surface heat flux driven by the atmosphere – in fact contributes in turn to driving atmospheric variability (in the form of the NAO), at least in this model. This raises a series of interesting questions about feedback processes, the spatial and temporal scales at which they may operate, and what sort of resolution in observed and modelled data is required to improve understanding of these processes.

The increasing availability of high resolution satellite data of SSTs and ocean winds offers potentially useful insights, as discussed by Xie (2004). Further findings in Rodwell *et al*. (2002) include NAO responses to idealized Atlantic tripole SST patterns that are similar to the observed. The mechanism put forward to explain these results is driven by anomalies in evaporation associated with anomalous SSTs. The local effects of anomalous evaporation (precipitation and atmospheric heating) act to reinforce the thermal and geopotential structure of the NAO (Rodwell *et al*., 1999). Although anomalous evaporation should act as a negative feedback on SST anomalies, Rodwell *et al*. (1999), find that anomalous advection driven by Ekman transport provides a countering positive feedback, particularly north of about 45°N. Implications of the Rodwell *et al*. (1999) study are encouraging for potential predictability of the North Atlantic/European climate, particularly when taken in conjunction with studies that show useful potential predictability for North Atlantic SSTs up to several years in advance (e.g. Sutton and Allen, 1997). A similar study carried out by Mehta *et al*. (2000) using the NASA Seasonal-to-Interannual Prediction Project (NSIPP) model supports the findings of Rodwell *et al*. (1999). Robertson *et al*. (2000) use the University of California, Los Angeles (UCLA) AGCM, and obtain similar results, including a five-fold increase in NAO variability when observed SSTs are used, as compared to climatological SSTs. It should be noted that model representations of the North Atlantic climate are not uniform from one model to the next. This model dependence is discussed in some detail in Robertson (2000), and Kushnir *et al*. (2002). Peng *et al*. (2002) use an AGCM forced with climatological SSTs, and an idealised SST tripole anomaly added or subtracted from climatology, similar to the method used by Rodwell *et al*. (1999), and find that there is a 500hPa response to SST variability, although this is weaker than in the observed response identified by Czaja and Frankignoul (2001). It is thought that the atmospheric response is maintained by eddy vorticity fluxes. Later results based on observed data support the notion that cool ocean SSTs can drive vertical momentum mixing, forcing surface wind adjustments to SST gradients (Xie, 2004).

Several recent studies use AGCMs forced with observed, climatological or idealised SST anomalies, and focussing on the North Atlantic. These include Drevillon *et al*. (2003); Frankignoul *et al*. (2003); Friederichs and Hense (2003); Friederichs and Frankignoul (2003); Lin and Derome (2003); Paeth *et al*. (2003); Peng *et al*. (2003); Robinson *et al*. (2003); Rodwell and Folland (2003); Sutton and Hodson (2003); Cassou *et al*. (2004); Deser *et al*. (2004); Kvamsto *et al*. (2004); and Magnusdottir *et al*. (2004). Almost invariably in these studies, a discernable influence of the North Atlantic on the overlying atmosphere is found. Some of the findings including – among others – Robinson *et al*. (2003) and Deser *et al*. (2004) suggest evidence for nonlinearity in the atmospheric response to oceanic forcing. Sutton and Hodson (2004) note an atmospheric model response to observed SSTs in the North Atlantic that is nonstationary, perhaps a function of decadal or multidecadal oceanic processes. More recently, Mosedale *et al.* (2005) find an ocean-to-atmosphere influence in the northeastern Atlantic using a time series modelling framework, while Ferreira and Frankignoul (2005) find a positive feedback resulting from anomalous diabatic heating of the atmosphere in the process of SST anomalies being damped. This feedback can have an adjustment time of several months. Wang and Chang (2004) identify advective processes that significantly enhance predictability in certain regions of the North Atlantic, out to lead times of up to five months. With respect to Tropical Atlantic forcing of the North Atlantic, Peng *et al.* (2005) extend work carried out by e.g. Czaja and Frankignoul (2002) on the observed North/Tropical Atlantic horseshoe pattern and its relationships with the NAO using a large ensemble of AGCMs and AGCMs coupled to a mixed layer ocean. They find little evidence in the model to support the Czaja and Frankignoul hypothesis that the summer North Atlantic Horseshoe (NAH) forces the early winter NAO, finding that the Tropical influence alone is more important in forcing the winter NAO. Interestingly, Frankignoul and Kestenare (2005**)** develop this further in a study on the observed record, confirming the findings of Czaja and Frankignoul (2002), but de-emphasizing the role of the northern part of the NAH (southeast of Newfoundland). Clearly there is still work to be done in this area, and as indicated in Frankignoul and Kestenare (2005), atmospheric models are improving in their representation of North Atlantic air-sea coupling.

2.5.1.4.2 Observed studies

A number of observed studies have been carried out in parallel with the model studies, which also reveal evidence of coupling between the ocean and atmosphere. A selection of the more recent papers is discussed here. Czaja and Marshall (2001) use an index of SST variability to show that similarities exist in the power spectra for this index and the northern part of an SLP dipole reminiscent of the NAO. A broad spectral peak is observed at 10-20 years. Czaja and Frankignoul (2002) use the NCEP-NCAR reanalysis to investigate covariability between oceanic (North Atlantic) and atmospheric variables at a range of leads and lags. While atmospheric forcing of the ocean is found to dominate, there is significant covariance when the atmosphere (Z500) lags by several months, implying an oceanic effect on the atmosphere. Rodwell and Folland (2003) give evidence that the North Atlantic atmosphere responds to the ocean throughout the annual cycle, in particular, responding to the tripole pattern in winter, and affecting anticyclonicity in the UK downstream of summer SST anomalies. When the observed results are compared with those from an atmospheric model (HadAM3) forced with the same SSTs, the atmospheric response is not as strong, although of a similar nature. While this provides physical support for the features identified in the observed study, it is clear that significant improvements are required in GCM representations of the North Atlantic climate. Cassou *et al*. (2004) identifies asymmetries between the two phases of the NAO, with the positive phase showing a displacement towards Europe, as well as nonstationarity in the variability of this system. It is found that SSTs in the tropics and extratropics affect the North Atlantic regime, and ENSO is thought to have an effect also. Pozo-Vazquez *et al.* (2005) examine winter SLP anomalies in the Northern Hemisphere, finding a potentially predictable relationship between autumn cold ENSO events and NAO-like SLP patterns; although the mechanism for this is thought to owe more to standing wave trains that propagate from the North Pacific than Atlantic SST anomalies.

2.5.1.5 Summary of Atlantic Ocean studies

The modelled and observed evidence for oceanic forcing of the North Atlantic/European atmosphere is well established. However, whether this coupling can produce useful predictability is a question that requires further research. In the light of model inadequacies in representing these complex relationships, a useful step towards better understanding may well be the exploration of statistical relationships as a guide to model development – although these should take care not to contradict the fundamental physical principles of the models.

There are many more aspects of Atlantic climate variability which may have a significant effect on the European climate – including tropical processes (e.g. Sutton and Hodson, 2003), the meridional overturning circulation (e.g. Delworth and Mann, 2000), and the re-emergence hypothesis (e.g. Deser and Blackmon, 1993). All of these will be considered in an exploratory analysis of potential predictors for the seasonal forecasting model to be developed in this thesis.

A recent workshop of the CLIVAR Atlantic group (May 2004), exploring climate variability and predictability throughout the Atlantic makes a series of recommendations for future work on research, the observing system, and the development of prediction systems (Sutton, personal communication). Some of these will be briefly summarised here. The extension of advances in ENSO prediction to the Atlantic basin is seen as a priority, bearing in mind the differences in forcing that are apparent from one ENSO event to another. In the extratropical North Atlantic, there is a need to better understand the interannual variability of the NAO, primarily through furthering the development of coupled models. The influence of coastal SSTs should also not be discounted, particularly for their effects on local climate. Additionally, the roles of land surface and stratospheric processes should be explored in more detail. A selection of these potential predictors will be discussed below.

2.5.2 Stratospheric variability and potential predictability of the surface climate

2.5.2.1 Introduction

A relatively recent addition to the field of potential predictors of European climate is the stratosphere. While stratospheric interannual variability has been studied and understood for some years (e.g. Labitzke, 1982), and the study of the interaction between the stratosphere and the troposphere likewise, the recognition that there may be processes with useful predictability in the stratosphere that affect the surface climate is quite new. Traditionally, the tropospheric forcing of the stratosphere has been more widely studied and understood (e.g. Hu and Tung, 2002). These insights are due in no small part to greater data availability, particularly from satellite and reanalysis sources.

2.5.2.2 Early work

Kodera and Yamazaki (1994) examine polar stratospheric forcing of the troposphere, noting that forced changes in the stratospheric polar night jet propagate into the troposphere. Kodera (1995) explores this forcing in more detail. Composite analysis of the stratospheric circulation based on a series of observed forcing reveals that the stratosphere displays an internal mode of variability that can be triggered by external processes (for example solar activity, the quasi-biennial oscillation (QBO), volcanic aerosols), and that extends into the troposphere. Baldwin *et al*. (2001) explore external forcing of the stratospheric polar vortex by the QBO, identifying a notable forcing impact, which results in changed surface weather patterns. This provides a mechanism for the QBO to have an effect at the earth's surface. The predictability of the QBO is largely low with respect to phase changes at monthly resolutions. Baldwin *et al*. (2001) identify a periodicity averaging 28 months, although the standard deviation (of the order of 4 months) imposes limitations on the potential utility of this index. Brankovic *et al*. (1994) identify links between ENSO and the QBO, and potential predictability of the QBO at the seasonal timescale. To date it seems that models do not represent this relationship, and it would certainly be desirable to understand more, given the potential implications for transmission of the ENSO signal outside the tropical Pacific basin.

2.5.2.3 Stratospheric processes and the AO/NAO

Baldwin and Dunkerton (2001) were the first to specifically identify potential predictability of the surface climate based on stratospheric variability. This predictability derives from large stratospheric anomalies that persist for a period of several months, and force an Arctic Oscillation (AO) response in the troposphere, as well as affecting the storm track and spatial distribution of midlatitude storms. Ambaum and Hoskins (2002) identify forcing in the other direction – of the NAO on the stratosphere – with the consequences that the stratosphere may act as an integrator of the NAO. However Kuroda (2002) does not support the idea of upward forcing by the northern hemisphere annular mode (NAM).

In recent years the study of stratosphere-troposphere coupling has received more attention. Taguchi and Yoden (2002) use a GCM to study coupling, finding that stratospheric interannual variability is close to a red noise spectrum, which is of interest when compared to the internal variability/external forcing hypothesis of Kodera (1995). Thompson and Baldwin (2002) find a link between the stratosphere and troposphere in the northern winter, that forces surface weather. They conclude that there may be potential predictability at the seasonal timescale based on knowledge of the stratosphere, in particular the strength of the polar vortex, and to a lesser extent the QBO. Zhou *et al*., (2002) find that the downward propagation of stratospheric anomalies does take place, but is conditional on a very large initial wave forcing in the stratosphere, and a reversal of the polar westerly wind. In these cases, there are discernable effects on the major northern hemisphere surface mode. Baldwin *et al*. (2003) continue on the theme of stratospheric-related predictability, demonstrating some skill in seasonal-scale forecasts of the Arctic oscillation (AO). An interesting recent development (Broennimann *et al.*, 2004) present strong evidence for forcing of the high latitude Northern Hemisphere stratosphere by a strong ENSO event from 1940-42, and subsequent anomalously cold temperatures in Europe.

2.5.2.4 Summary of Stratospheric Data

It appears that stratospheric processes may offer potential predictability of the seasonal climate, and more work is warranted to explore these relationships further. In particular, in the case of the NAO/AO the combined and independent forcing of the ocean and the stratosphere may well improve predictability. Another consideration is the apparent inability of models to capture the full range of interactions between different components of the climate system involving the stratosphere. While this offers grounds for cautious optimism as far as predictability is concerned, the physical significance of potentially predictable signals must be treated with care, at least until they can be modelled more accurately.

2.5.3 Snow cover as a potential predictor for European climate

2.5.3.1 Introduction

The role of snow cover as a potential seasonal predictor developed roughly simultaneously with extratropical seasonal forecasting efforts and also in conjunction with research into variability of the Asian monsoon (e.g. Hastenrath and Greischar, 1993). In the Northern Hemisphere, snow cover extent has been shown to affect atmospheric temperature and circulation (e.g. Cohen and Rind, 1991), and its properties as a relatively persistent boundary condition are therefore of potential value in seasonal forecasting efforts. A number of studies have applied snow cover as a predictor in seasonal climate forecasts. These are outlined below. The potential for improved understanding of the effects of snow cover on European climate and seasonal forecasts will be discussed.

2.5.3.2 Early Work

Early work on snow cover relationships with the atmosphere include Kukla (1979); Walsh (1984) on snow cover and atmospheric variability; Namias (1985), examining the relationship between snow cover and temperature and precipitation; Walsh (1984) and Ross and Walsh (1986), who studied the influence of snow cover on the overlying atmosphere at synoptic scales; Barnett *et al*. (1988, 1989) studying the effect of Eurasian snow cover on global climate and global climate variations; Iwasaki (1991) studied the effects of snow cover on interannual climate variability, this is closely related to the seasonal forecasting problem. Cohen and Rind (1991) studied the effects of idealised snow cover anomalies on the atmosphere in a GCM. This study indicated that there were discernable if short-lived effects of snow cover on the atmosphere. These results are supported by Walland and Simmonds (1996, 1997), among others. Randall *et al*. (1994) study the snow-atmosphere feedbacks in the 14 Atmospheric Model Intercomparison Project (AMIP), and find that in all cases there is some feedback between snow cover and the atmosphere, although the strength and sign of the feedback is model dependent. In general it is accepted that snow cover has a discernable effect on the atmosphere, although the detailed nature of this relationship is not fully understood. Further work on the observed record is required, and further model development and analysis, at least to the point where there is broad agreement between models in an AMIP-type study.

2.5.3.3 Snow cover and climate variability

Based on the prior work discussed above, the relationships between snow cover and climate variability continue to be developed in the literature. There are a number of distinct areas of interest, for example the relationship between Himalayan snow cover and the south Asian monsoon. This review will be restricted to covariability directly relevant to the northern midlatitudes, in particular focusing on the relationship between Eurasian (and, to a lesser extent North American) snow cover, and the North Atlantic

Oscillation/Arctic Oscillation (NAO/AO), or circulation patterns related to weather conditions in north western Europe.

Starting with Walland and Simmonds (1996), atmospheric responses to Northern Hemisphere (NH) snow cover were identified in idealised model runs. Of note are the responses of SLP and 500hPa height over land areas. Notably, changes over the North Atlantic and Pacific basins are noted, as well as a reduction in the strength of the storm track and cyclogenesis. Walland and Simmonds (1997) extended this work to look at the major modes of variability in snow cover, identifying the major components, and correlating them with atmospheric modes. No simple interactions were found between snow cover and atmospheric modes, but rather the relationships were complex, indeed the potential for nonlinear or threshold dependent relationships is of interest here (see for example Campbell, 2005). Cohen and Entekhabi (1999) explicitly identified a possible role for snow cover in enhancing European climate predictability, finding evidence of coupling between the atmosphere and the snow cover boundary layer, while Frei and Robinson (1999) note the 'significant month-to-month persistence' of snow cover in certain areas during winter and spring. Ye (2000) found a link between decadal patterns in snow cover variability over Russia and tropical Atlantic SST anomalies, and subsequently extended the research to include the Pacific basin**.** However, the bulk of the research in this field has focused on snow cover and atmospheric variability. Cohen *et al*. (2001), Cohen and Entekhabi (2001) and Saito *et al*. (2001) explore the relationship between Eurasian snow cover and NH climate variability, with the aim of assessing climate predictability based on snow cover forcing. Gong *et al*. (2002) develop this further by examining the link between snow cover and the NAO/AO using model output, finding that snow cover can modulate the variability of this mode.

Subsequently, a number of studies focused on the snow cover-NAO link, with some potentially useful results with respect to climate predictability. Bojariu and Gimeno (2003a, b) explore this relationship, again finding that snow cover is important, but at decadal/multi-annual timescales rather than interannual. However, Saito and Cohen (2003) identify potentially useful links between snow cover and atmospheric variability at interannual timescales. Gong *et al*. (2003a, b) study the model atmospheric response to snow cover anomalies over North America and Eurasia, finding that Eurasia has a significantly greater influence on the atmosphere, providing 'a physical explanation for how regional land surface snow anomalies can influence winter climate on a hemispheric scale', and the associated implications for predictability are discussed. Cohen and Saito (2003) explore the predictability of US winter climate based on snow cover alone, bypassing the need to forecast the NAO, with some positive results. Schlosser and Mocko (2003) use a range of dynamical forecasting models to study the impact of winter snow on spring temperature and circulation, concluding that dynamical models need to be improved before applying them successfully to seasonal forecasts using snow as a predictor. Qian and Saunders (2003a, b) and Saunders *et al*. (2003) study the relationship between snow cover and the NAO, with the explicit aim of improving seasonal predictability, with some potentially useful results.

Kumar and Yang (2003) are less encouraging in their findings, noting that although snow cover exerts a discernable influence on the troposphere, is has no discernable effect on the spatial organisation of the major NH atmospheric modes. While they are not dismissive of the possibility of useful predictability from snow cover anomalies, they stress the need for further research on the snow cover-atmosphere variability relationship. Similarly, Schlosser and Mocko (2003) use a range of dynamical forecasting models to study the impact of winter snow on spring temperature and circulation, concluding that dynamical models need to be improved before applying them successfully to seasonal forecasts using snow as a predictor. Gong *et al*. (2004) explore the role of orography in modulating the relationship between snow cover and atmospheric variability, finding that it is an important factor. Gong *et al*. (2004) carry out extensive model studies of snow anomaly effects, including snow depth as well as snow cover, and finding that both these variables have a bearing on the results. Saito *et al*. (2004) explore sub-decadal relationships between snow cover and the NAO/AO in the observed record, finding evidence for nonstationarity in the structure and strength of the lead-lag relationships – something that has an important bearing on the application of this data to seasonal forecasting models. More recently, Shongwe *et al* (2007) find a link between snow cover and anomalously cold spring seasons in Eastern Europe.

2.5.3.4 Summary of Snow Cover Data

A large amount of work still remains to be done on quantifying the viability of snow cover as a useful predictor for the European climate. In particular, clarifying aspects of the snow-atmosphere forcing mechanism(s), and exploring the relationship between snow cover and the NH atmospheric circulation with respect to other forcing modes, such as soil moisture and SST variability. These requirements notwithstanding, it may well be that, based on current research; snow cover can be used as a predictor. While it is beyond the scope of this thesis to study mechanistic details of the relationship, there is certainly scope to focus on the interaction of snow cover with other forcing processes, and their subsequent effects on the atmospheric circulation, with the associated implications for predictability.

2.6 Summary and conclusions

This literature review discusses a sample of the literature published in recent years that is relevant to the seasonal forecasting of extremes in North Western Europe. This necessarily covers a broad spectrum of topics within climate research, from papers directly relating seasonal forecasting studies, through those concerned with the predictability and variability of various components of the atmosphere-ocean-land system, to studies concerned with extreme events. Each of these research areas constitutes a major body in its own right.

From the existing literature, it is possible to draw a number of conclusions on the current prospects for seasonal forecasting of extremes, and on the areas in which it may be possible to advance knowledge further, and utilize existing knowledge more effectively.

Oceanic forcing from the North Atlantic is most likely to provide improvements to predictability, given the recent volume of research that has been dedicated to this field, and the increasing understanding of the physical mechanisms driving this variability. Land based and stratospheric processes offer some hope for enhanced predictability, but are somewhat constrained by the lack of well established physical mechanisms to support observed predictability.

To date there are no seasonal forecasting studies for Europe which utilize the full range of potential predictors explored here. There is clearly potential to develop empirical seasonal forecasting models using a wide range of predictors and techniques.

3 Methodology

3.1 Introduction

This chapter presents a description of the key methods employed during the research of this thesis. A range of standard statistical techniques is employed in the process of this research and where necessary each data and results chapter includes a detailed description of the method relevant to that chapter, and refers where necessary to the content of this chapter. The main objective of this chapter is to set out the statistical tools whereby the predictor data are related to the predictand data, the implications of this process in a statistical context, and the experimental design which will aim to ensure that the process is rigorous and the results meaningful.

The experimental design is shaped by the objective of developing statistical models of seasonally predictable relationships between a range of predictor variables, and four sets of predictand variables, describing extreme precipitation and wind events across a European domain. Of primary concern here is the lack to-date of a comprehensive understanding of any physical basis for potential predictability. This informs the experimental design throughout, to the extent that the primary focus is on the exploration and validation of statistical relationships, with an associated discussion of the potential causal mechanisms underlying the observed relationships, rather than the precise quantification of parameter error and uncertainty, which is of secondary importance to this problem.

This testing and validation process takes place in four stages:

1. Significance testing – typically at the 95% confidence level – for every local assessment of statistical relationships using Poisson regression
2. Field-significance testing using resampling techniques
3. Model selection using all possible predictor subsets within predefined constraints, and using cross-validation to choose the best model

4. The first three steps above all apply in the context of a model training period. Further validation against an independent sample of data comprises the final assessment of model fit.

The chapter is further divided into four sections, each addressing the points outlined above. Section 3.2 addresses point 1, above, on generalised linear models (GLM); Section 3.3 addresses field-significance testing as a means to define a set of predictors for consideration in the models; Section 3.4 describes the procedure by which the predictive models are selected and Section 3.5 explains the requirement for independent validation and how it is carried out in this case. Finally Section 3.6 summarises the methodology.

3.2 Generalised Linear Models

The data used in this study is of the form that lends itself to analysis using Poisson regression – a member of the family of Generalised Linear Models (GLM) (McCullagh and Nelder 1989). Poisson regression is applicable specifically to count data, with the key features of such data being that they are comprised of non-negative values, and are Poisson distributed. In this case the predictands – counts of seasonal exceedances of percentile thresholds – meet the first criteria, and in most cases satisfy the second criteria, that is, they approximately follow a Poisson distribution which can be described as follows:

$$\Pr(N_t = k) = f(k;\lambda t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}. \qquad (3.1)$$

where $\Pr(N_t = k)$ is the probability of observing k events over a given time t. In this case, an event is a day on which precipitation or peak wind gusts exceed a stated percentile threshold, and t is a three month season. Therefore k is the count of the number of events in a given season. The mean and variance of the distribution are described by λt which in this case is equivalent to the climatology. Further illustration of the suitability of the predictand data to the application of Poisson regression is given in Chapter 4.

Before describing Poisson regression, it is instructive to illustrate the case of standard linear regression, in which the behaviour of a predictand variable as a function of a number of predictor variables leads to an estimate (in this case a forecast or hindcast) of the predictand. The standard linear regression equation is as follows:

$$f_i = \sum_{j=1}^{p} x_{ij}\beta_j \quad \text{for} \quad y_i \tag{3.2}$$

where the observations of the predictand Y are given by y_i for season i, and observations of p the predictor X_j are given by x_{ij}. The regression weights β_j are found from the observed values of Y and X_j over a training period by minimising the mean squared error (MSE) such that the forecast f_i is as close as possible a match to the Y.

Poisson regression is simply a variant of this case, where non-negative, Poisson distributed count data are used, and for which 3.2 above becomes:

$$\ln(f_i) = \sum_{j=1}^{p} x_{ij}\beta_j \tag{3.3}$$

where *ln(.)* is the natural logarithm. The weights in this case are then found using the iteratively weighted least squares method (IWLS) as defined in Lee *et al*. (2006), such that they converge on the least squares solution – where the MSE is minimised.

For each estimate of model fit, the significance of each of the covariates, and the model as a whole are then estimated using the Student's T distribution.

In this study, each elementary assessment of predictor-predictand fit is carried out using the method presented above, where statistical significance at 95% confidence satisfies the first criteria for further consideration as detailed in the remainder of this chapter. The IWLS procedure and other statistics such as measures of model fit can be found using a standard software package - in this case Matlab. An additional aspect of the use of Poisson regression is that the measures of model fit can be weighted according to the dispersion of the predictand data. That is, if the variance of the data is greater than the mean (where the theoretical Poisson distribution describes both the mean and variance using a single statistic), the data is overdispersed. In this case a slightly more conservative estimate of the model error, and significance, are made.

The following example illustrates the application of Poisson regression in this study. The predictand Y is the number of days in the season December-February (DJF) with peak wind gusts exceeding the 90th percentile level for this season at a gridbox over the Netherlands. The predictors are the 3rd principal component of stratospheric temperature at 50hPa, during the July-September (JAS) season (X_1), and the standardised SLP anomaly at Darwin (one of the components of the SOI) from the preceding August (X_2).

In this specific case, the Poisson regression can be written:

$$\ln(f_i) = 0.2579 X_{Strat} + 0.2899 X_{Darwin} + 1.6224 \qquad (3.4)$$

to obtain the forecast (f_i) of Y such that given positive values of the 3rd PC of 50hPa temperature during JAS, and positive pressure anomalies at Darwin during August (which might more generally be associated with El Niño conditions), we might expect increased counts of days with high wind speeds in the specified gridbox the following winter.

3.3 Field-significance testing using Monte Carlo resampling

The technique presented in 3.2 is adopted as the basic tool to determine local statistical significance between predictors and predictands in this study. However, when conducting multiple tests of local statistical fit, at a given threshold of significance, it follows that a number of tests will fulfil the requirements of local significance by chance.

If we consider each test to be independent of the others, the number of tests passing due to chance can be described by the binomial distribution. This problem of multiplicity in the analysis of climate fields is discussed in detail in Livezey and Chen (1983), and Wilks (1995), for example. Further detail is presented here in Section 6.3.1.1.

An additional complication when conducting multiple local significance tests on climate fields arises when we consider that there is a degree of spatial correlation between the locations being tested, which reduces the degrees of freedom by a considerable amount - depending on the spatial correlation scale of the field. It is therefore necessary to determine the minimum spatial extent at which the predictand response will be deemed to have field-significance, as a function of the spatial autocorrelation properties of the predictand. To date the most effective way to assess field-significance in statistical climate analysis is presented in Livezey and Chen (1983), who used resampling, or Monte Carlo techniques to empirically quantify the probability distribution of the spatial extent of the predictand response. That is, by repeatedly measuring the response of the whole predictand field to randomly generated timeseries which have the same temporal autocorrelation structure as the predictor, it is possible to determine the field-significance of the predictand response as a function of the spatial extent of the locally significant responses. Wang and Shen (1999) conduct a series of tests to estimate the spatial degrees of freedom (dof) of a climate field by four methods, including fitting the sum of the squared differences between each realisation of the field, and the climatology at every location to a chi-squared distribution (e.g. Fraedrich *et al.*, 1995); secondly, a measure derived from the variance of the distribution of the pattern correlation coefficients between realisations of a normalised climate field (e.g. Sachs, 1984); thirdly, the dof is estimated as the ratio of the variance of the field's average to the average of the variance field (e.g. Smith *et al*., 1994); and finally, the method outlined above, due to Livezey and Chen (1983). They find that the latter method is the most accurate of the four. A more detailed formulation of this technique with application to this study is presented in Section 6.3.1.1 and 6.3.1.2.

Field significance testing is used here as a tool to determine which of the potential predictors might usefully be applied to the model selection process. In a sense it acts as a filter conditioned on the spatial extent of the predictand response to the predictor, although in the absence of any other basis for predictor selection based on established physical relationships it is an important step. Therefore for each potential predictor, in order for it to be further considered as part of the final model selection process it must satisfy the condition of field-significance at the 95% confidence level.

3.4 All subsets model selection using cross-validation

Having defined a predictor set with each predictor satisfying the criteria of a field significant response in the predictand domain, a subset of the initial predictor set is then available to supply the model selection phase. For each gridbox over the predictand domain, those field-significant predictors which are locally significant at that gridbox are considered in an all-subsets model selection algorithm with a maximum of two predictors per model. This maximum is set on the basis that the record length is relatively short, and it is important to minimise overfitting. Also, given the inherently low (at best) levels of predictive skill of wind and precipitation extremes in the mid-latitudes, it is not thought likely that any linear combination of three or more predictors would add useful skill, and in fact would be more likely to tend towards overfitting the model – that is, fit the model to noise in the data, rather than any underlying structure due to large-scale predictable forcing.

Each valid subset of predictors is then tested further using leave-one-out cross-validation (e.g. as described in Wilks, 1995), in which the model is fit n times to $n - 1$ of the observations, the remaining observation is predicted, and the residuals for each of the n forecast observations are summarised using the mean absolute error (MAE). The final model is that for which cross-validation minimises the MAE. Further details on this method are provided in 7.2

3.5 Model validation against independent data

In any exploratory statistical investigation the importance of the independent validation of the prediction experiment is paramount. For example see discussions in Lloyd-Hughes and Saunders (2002) and van den Dool (2007) among others. In this case the model selection process – as outlined conceptually above – is carried out over the observation period 1958-1995, and the validation period is 1996-2005. Both of these periods are necessarily short owing to the availability of data and the relative length of each is the result of a compromise between the importance of conditioning the models on sufficient data and the requirement to assess statistical significance in the validation period.

The model selection is achieved by taking the best cross-validated fit of all subsets of models having locally significant fit to the predictand gridbox, in the context of a field-significant response. The validation process then tests the ability of the model to replicate observed skill levels completely independently of the model training process.

3.6 Summary of methodology

In summary, a method is presented to carry out an analysis of potential linear predictability of European extremes of precipitation and wind – two climate variables which are not known for their predictability, but for which the interannual variability in extremes carries important consequences for society. The method should be treated strictly as exploratory, and any potentially useful predictability requires corroboration with physical mechanisms, and possibly further quantification in light of this, since precise parameter estimation is impossible on such a relatively short record, even assuming perfect linearity and stationarity apply – which is also unlikely.

Other methods are employed in the course of the research, including the fitting of gamma distributions to the predictand data, and principal components analysis (PCA) to derive some of the predictor indices. As these are not directly pertinent to the broad experimental design discussed here, they are outlined in more detail in Chapters 4 and 5 respectively.

The results of each of the steps outlined in 3.2 to 3.5 above are presented sequentially in chapters 6, 7 and 8 as follows. Chapter 6 presents the results of the field-significance testing on the full predictor set. Chapter 7 uses the predictors retained in Chapter 6, and presents the results of the model selection process using all subsets selection and training-period cross-validation. Chapter 8 presents the results of the model testing on the validation dataset.

4 Predictands

This chapter presents the datasets available to derive predictands for the forecasting model. The requirements for predictands are discussed, and results of dataset validation and comparisons are presented, in order to determine the most suitable predictands. Section 4.1 comprises an introduction, which will outline the main questions and objectives of the chapter. Section 4.2 presents a range of datasets, describing their characteristics, qualities and limitations. Section 4.3 presents a range of methods to derive indices of extremes. Section 4.4 compares raw data and indices from the range of available datasets. Section 4.5 summarises the chapter and defines the predictand datasets to be used in the forecasting model.

4.1 Introduction

Predictands in seasonal forecasting are traditionally some measure of the average conditions expected, for example, in the case of statistical forecasting, the model will be trained on seasonal mean conditions, and used to predict these accordingly (e.g. Feddersen, 2003). In the case of dynamical forecasting, mean conditions are derived from the model output, broadly speaking (e.g. Gueremy et al, 2005).

The seasonal forecasting of extreme events presents additional challenges in the selection of the predictands. Firstly, what data should be used, since daily frequencies (at least) are required, and data availability is limited in space and time at this resolution? Secondly, what measure should be used to represent interannual variability in extreme events, in order that the predictand is useful in that it represents both events of interest to the end user, as well as interannual climate variability? Both these challenges will be addressed in this chapter with the objective of defining a meaningful – and practical – measure of extreme events to train the forecasting model, and ultimately to be operationally useful.

Given the applied nature of the end-product in this case, it is first considered whether predictands should be restricted solely to meteorological variables, and whether other measures, such as flood records and windstorm damage should be considered. There are arguments for and against the inclusion of impact-based predictands. From the end user point of view, it is useful to be able to define exposure using these measures, but the limited availability of useful data imposes a formidable practical constraint on this approach. For example flood records can often be affected by the floods themselves, resulting in inhomogeneous data. Additionally, the event-based detail of available data is more encouraging for a case-study approach, where individual events can be linked to the large-scale circulation, and meteorological conditions. This might be extended to provide a more general understanding of the meteorological conditions required to cause a damaging event, but is considered beyond the scope of this thesis. Some local efforts to link impacts directly with the large-scale circulation include Kaczmarek (2003) for flood risk in Poland, although in this case precipitation data are used extensively to assess linkages with the large-scale circulation. Kim and Barros (2001) use a neural network method to forecast floods using local and large-scale meteorological conditions at daily rather than seasonal timescales.

Haylock and Goodess (2004) use daily station records of extreme rainfall to assess links with the large-scale circulation, while Gershunov (1998) and Gershunov and Cayan (2003) use US station records of precipitation and temperature to assess links with ENSO, and hence seasonal predictability of extreme events. Similarly, in this study, it is the large-scale conditions themselves that are of interest, since they provide a more direct link to potential predictability derived from the boundary forcing variables. Potential predictands are therefore restricted to meteorological variables, in order to gain a more general understanding of what might drive interannual variability in extreme events across the study region. Given this decision, the question remains as to what magnitude of event should constitute an extreme. On the one hand, the low frequency, high impact extremes are of most interest to end users, but on the other hand, these are difficult to link to predictable components of the climate system due to both the small sample size, and the range of causes that combine to produce an event of this magnitude. Approaches to these threshold decisions will be discussed more fully in 4.3, supported by results in Section 4.4.

4.2 Potential Predictand Datasets

The main criteria for potential predictand datasets include faithful representation of extremes and in particular the interannual variability of extremes, and ideally a continuous and homogenous record in space and time. The study area is defined as the domain 40-68N, 11W-27E.

4.2.1 Precipitation data

A number of sources of precipitation data exist for the European domain, providing adequate periods of record for model development. These are outlined in Table 4.1. The station records from the European Climate Assessment and Dataset Project (ECA, http://eca.knmi.nl/, Klein Tank et al, 2002), comprise 208 stations in total, of which 138 contribute a useful record within the selected spatial domain (Figure 4.1). Temporal resolution is daily. The data are provided with quality control flags from the ECA. Further checks were necessary, and these were carried out by comparing the monthly totals of the station records with a monthly gridded product (CRU TS2.1) provided by the Climatic Research Unit, School of Environmental Sciences, University of East Anglia (CRU), to check for unrealistic daily totals, as well as with neighbouring stations from the ECA network. Further station data are available, for example from the British Atmospheric Data Centre (BADC), covering further UK stations, and from the National Climatic Data Centre (NCDC), covering the whole study region. However, the necessary quality control procedures, and the extent of missing records, on first inspection are prohibitive. With the additional quality control carried out, the ECA station precipitation records constitute a useful, although spatially incomplete record of point source data.

The European Centre for Medium Range Weather Forecasts (ECMWF) 40 year reanalysis (ERA-40, Uppala et al, 2005) is the other daily dataset considered. Other reanalyses include ERA15 (Gibson et al, 1996) – a precursor to ERA-40, and those from the National Centers for Environmental Prediction, in association with the National Center for Atmospheric Research (NCAR), and the US Department of Energy (DoE), referred to as NNR1 (Kalnay et al, 1996) and NNR2 (Kanamitsu et al, 2002) respectively. Both ERA15 and NNR2 cover too short a period to be useful in this context, while ERA-40 was chosen over NNR1 due to its higher spatial resolution (available at approximately 1.25°x1.25° resolution, compared to NNR resolution of 2.5°x2.5°).

A reanalysis dataset comprises spatially and temporally complete output from a numerical weather prediction model, forced with observed data, which is assimilated at regular time intervals. The volume of data assimilated varies over time, as it is available, although the assimilation system remains the same throughout. However, not all the available observed data is assimilated, including precipitation. Instead precipitation is a forecast product, and therefore subject to model errors and biases, which are particularly apparent in the case of convective precipitation. The reanalysis precipitation therefore requires validation against observed datasets. The CRU dataset (CRU TS 2.1, Mitchell and Jones, 2005) is a high resolution (0.5°) monthly dataset derived from an extensive database of station records, while the US Climate Prediction Center Merged Analysis of Precipitation (CMAP) dataset (Xie and Arkin, 1996) is also a gridded monthly product at 2.5° resolution, derived from satellite and station measurements of precipitation, and augmented where necessary with reanalysis (NNR) precipitation. Both of these datasets have been used to assess the suitability of the ERA-40 data for use in the predictand dataset.

Dataset	Description	Temporal resolution	Spatial resolution	Period of Recording
ECA	Station records	Daily	208 stations	Various
ERA-40	Reanalysis	Daily	Approx.1.25°x1.25°	09/1957-08/2002
CRU TS 2.1	Gridded station records	Monthly	0.5°	1901-2002
CMAP	Merged (satellite, station, reanalysis) gridded analysis	Monthly	2.5°	1979-present

Table 4.1 Precipitation datasets

4.2.2 Wind data

Daily data for wind speeds, particularly daily maximum wind speeds, are relatively sparse for the European region, over the timescales necessary. Selected station records of daily maximum wind speed are available from national meteorological services, but in many cases at a considerable expense, even when only handling fees are charged. Thirty seven station records were obtained from the Deutsche Wetter Dienst (DWD) website (http://www.dwd.de/de/FundE/Klima/KLIS/daten/online/nat/index_standardformat.htm), (although missing periods reduced this to 31), and 6 were obtained from the Royal Netherlands Meteorological Institution (KNMI) website (http://www.knmi.nl/klimatologie/daggegevens/download.cgi). In all cases the records of daily maximum wind speeds extend back only as far as 1971.

ERA-40 was selected as the reanalysis source of wind data for the same reasons outlined in Section 4.2.1 above. It includes several wind variables, both reanalysis and forecast products. Since wind speed is not cumulative it is necessary to define extreme winds over a short timescale. The shortest time-step at which reanalysed data is available in ERA-40 is 6 hourly mean wind speeds. While this is likely to capture the context of the event within which wind damage actually occurs, a variable which expresses the actual peak wind speed is more desirable. The 10 metre peak wind gust variable is a forecast as opposed to reanalysed product, and as such, requires further comparison, in this case with the available station data. However, many of the problems associated with precipitation in the reanalysis are less important in the case of wind speed, since it is closely related to easily measured and reanalysed variables such as surface pressure. Additionally, local effects such as topography are an important factor in determining peak wind gust speeds, and station records can be problematic in this sense. Since wind extremes in north-western Europe are almost always embedded within large-scale weather systems, it is reasonable to suppose that extremes of interest are as likely to be picked up by a gridded product as a point source product such as a station record. Nevertheless, the station and reanalysis data are compared, at least in order to provide insights into the magnitude of the differences between the two datasets.

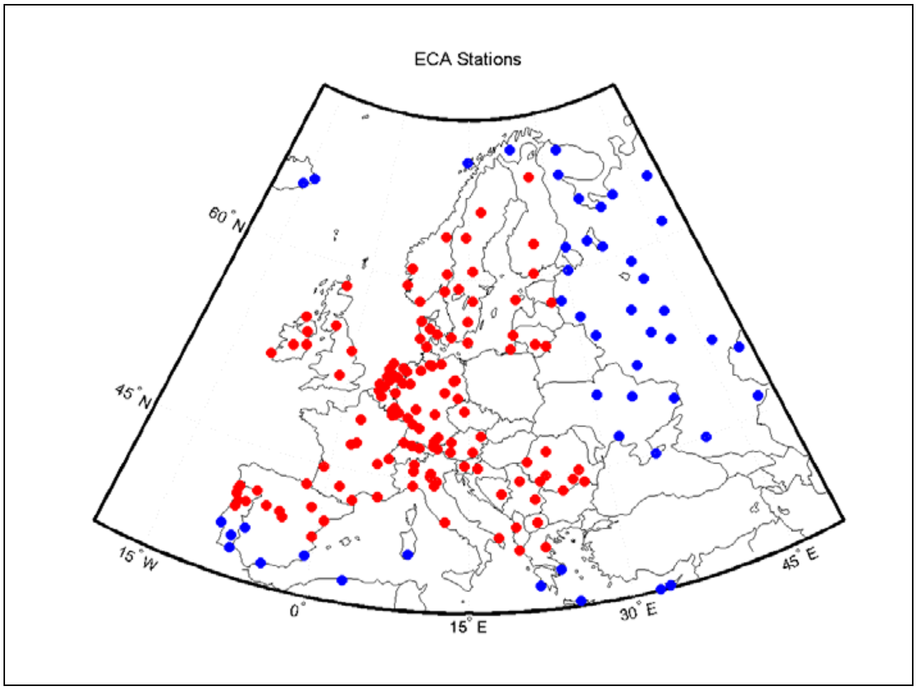

Figure 4.1 ECA precipitation stations. Those in red are within the study domain

4.3 Indices

While the cost of weather extremes in Europe has run, on average, into hundreds of millions of US dollars annually since 1970 (Murnane, 2004), these costs are skewed towards a small number of very large events. Nevertheless, minor events still cause damage to property and loss of life and, from an applications perspective, are still of some interest to end users. This raises the question of how best to sample the raw data for extremes. Typically, an end user for example in the insurance industry is interested in events with a return period in the order of years (if not tens of years) to centuries. By definition this will not provide an adequate sample to represent interannual variability. It is therefore necessary to compromise by substantially reducing the threshold at which events are included as extremes.

Much of the work done on extremes is concerned with formulating indices of those extreme events best representing the processes being studied. With respect to this thesis, one of the difficulties highlighted by Nicholls and Murray (1999) is national/regional differences in precipitation monitoring standards and quality control procedures. The same argument applies to wind speed data. Nicholls and Murray identify the need to develop simple, uniform indices of precipitation extremes. For the European region this need has subsequently been met to a great extent by work on two specific projects – the European Climate Assessment (ECA) project (Klein Tank *et al*2002, 2003), and the Statistical and Regional Dynamical Downscaling of Extremes for European Regions project (STARDEX) (e.g. Haylock and Goodess, 2004). Both of these focus on temperature and precipitation, with precipitation being of particular interest here. A series of indices were developed through the STARDEX project, using the STARDEX Diagnostic Extremes Indices Software, available at www.cru.uea.ac.uk/cru/projects/stardex. These indices were suggested by the Expert Team on Climate Change Detection Monitoring and Indices (ETCCDMI). With respect to representing the interannual variability of extremes, a subset of these indices is of particular interest, including the number of events in a given season exceeding a certain threshold measured as a percentile. Gershunov (1998) and Gershunov and Cayan (2003) use this measure to assess the predictability of heavy daily precipitation over the contiguous United States at seasonal timescales. Also of interest are measures of daily intensity – reflecting the relationship between total precipitation and the number of days with precipitation in any given period, and also the percentage of precipitation within a defined period that falls on days stratified by percentile thresholds. These indices are all discussed in Nicholls and Murray (1999) and within the documentation for the STARDEX software. Nicholls and Murray also discuss the possibility of calculating indices from monthly data, although this is not ideal for the purpose of seasonal forecasting of extreme events. Additionally, it is recommended that a gamma distribution be fit to the daily precipitation values to determine percentile thresholds, although the STARDEX indices use a simple ranking approach.

Much work has been undertaken on the statistics of wind speed distributions. In particular, Palutikof *et al.* (1999) and Brabson and Palutikof (2000) explore the use of the Generalised Extreme Value distribution (GEV) and the Generalised Pareto Distribution (GPD) to calculate extreme wind speeds. In addition, this work provides a useful discussion of methods to ensure that the datasets under consideration fulfil the necessary criteria of being independent and identically distributed. In particular, the serial correlation present in daily wind speed time series must be accounted for. A number of different approaches are suggested, including a peaks over threshold (POT) approach using the GPD. Given the appropriate choice of threshold, a sufficiently large sample size can be obtained to represent interannual variability while still focusing on relatively extreme events. A further consideration is the minimum separation distance between the extremes. Palutikof *et al*. (1999) recommend a minimum of 48 hours to ensure that the events are not related. More recently, Harris (2005) finds that the POT approach using the Weibull distribution as the parent distribution does not conform to the GPD model. However, much of the work on wind extremes is carried out with applications sensitive to long return periods in mind, for example design loads for buildings. In the case of this thesis, 'softer' extremes are of interest, for reasons outlined above, and an approach more similar to that taken by Haylock and Goodess (2004), using purely empirical techniques can be considered, although some of the selection criteria outlined by Palutikof *et al*. (1999), such as those involving serial correlation must be considered.

Following a similar method to Haylock and Goodess (2004), a threshold exceedance method was used in order to represent the interannual variability of extreme events. Long-term 90th and 95th percentile thresholds for twelve three-month seasons (JFM, FMA, etc) at each location were calculated by fitting a gamma distribution to the daily data for the period 1958-2001. Percentiles were estimated using the gamma distribution fit rather than a simple ranking procedure due to Zhang *et al*. (2005) who find that the ranking method can result in inhomogeneities outside the sampling period. The exceedance count for each season was then calculated. In the case of precipitation data, a wet-day threshold of 0.1mm was set for both the station and reanalysis data, and only wet-days were considered in the estimation of the thresholds. For the peak wind speed data, serial correlation in the daily timeseries was accounted for by removing the lesser exceedance should two fall on consecutive days, or the middle/alternate exceedance(s) should three or more occur consecutively. The decision to use twelve three-month overlapping seasons is made to satisfy two criteria: firstly the use of seasons rather than single months increases the sample size of available extremes; secondly, the use of twelve overlapping seasons rather than four may help to identify predictable relationships that are sensitive to the annual cycle. A similar method is used by Gershunov (1998), for the same purpose of developing a predictand dataset for seasonal forecasting of extremes.

4.4 Results

Results are presented from a series of dataset comparisons. The objective is to assess the suitability of the available datasets to derive predictand indices. For peak wind speeds, the choice of dataset is essentially restricted to ERA-40, since the spatial and temporal coverage of the station data is insufficient. Nevertheless, results of comparisons with those station data will be presented.

4.4.1 Precipitation results

It is assumed that stations provide an accurate representation of point-source extreme events. Station data are compared with ERA-40 and gridded monthly data in order to assess how well ERA-40 matches the station records, and how useful the station records may be in representing areal extremes.

While it is not possible to compare observed gridded extreme precipitation with the reanalysis, monthly totals can be compared to give some indication of how well ERA-40 and the station records capture interannual variability with respect to the CRU TS2.1 data. Figure 4.2 shows that the station records are more frequently highly correlated with the gridded data, although the mean is lower for all months except August (not shown). However a number of the stations are not well correlated, and by implication may not be suitable to represent extreme events due to apparent local effects. A further reduction in the density of the station network is not desirable. Furthermore, there appears to be a national bias in the correlations, perhaps reflecting differences in station quality control procedures at the national level, or station densities provided for the CRU dataset. Correlations with both datasets are generally lower in July, reflecting the greater importance of convective precipitation at small spatial scales.

The lower correlations between the monthly ERA-40 and CRU datasets are illustrated in Figure 4.3, showing the annual cycle of anomaly correlations between the two datasets. Correlations are best in the winter and spring, and lowest in August. Again, this is likely to be due to the increased importance of convective precipitation in the summer months.

Figure 4.4 shows the correlations between CRU, ERA-40, CMAP and station monthly mean precipitation and the North Atlantic Oscillation (NAO, Jones et al, 1997) in January, for the period 1979-2001. Figure 4.5 shows the same for July. Although the July correlations are weak, they show strikingly similar spatial patterns in all datasets, as do the stronger January correlations. In general simultaneous correlations with the NAO and all four datasets are very similar throughout the year.

Comparisons between monthly mean precipitation indicate that in general the datasets compare well with respect to interannual variability, and that the ERA-40 precipitation responds well to large-scale circulation variability in the case of the NAO.

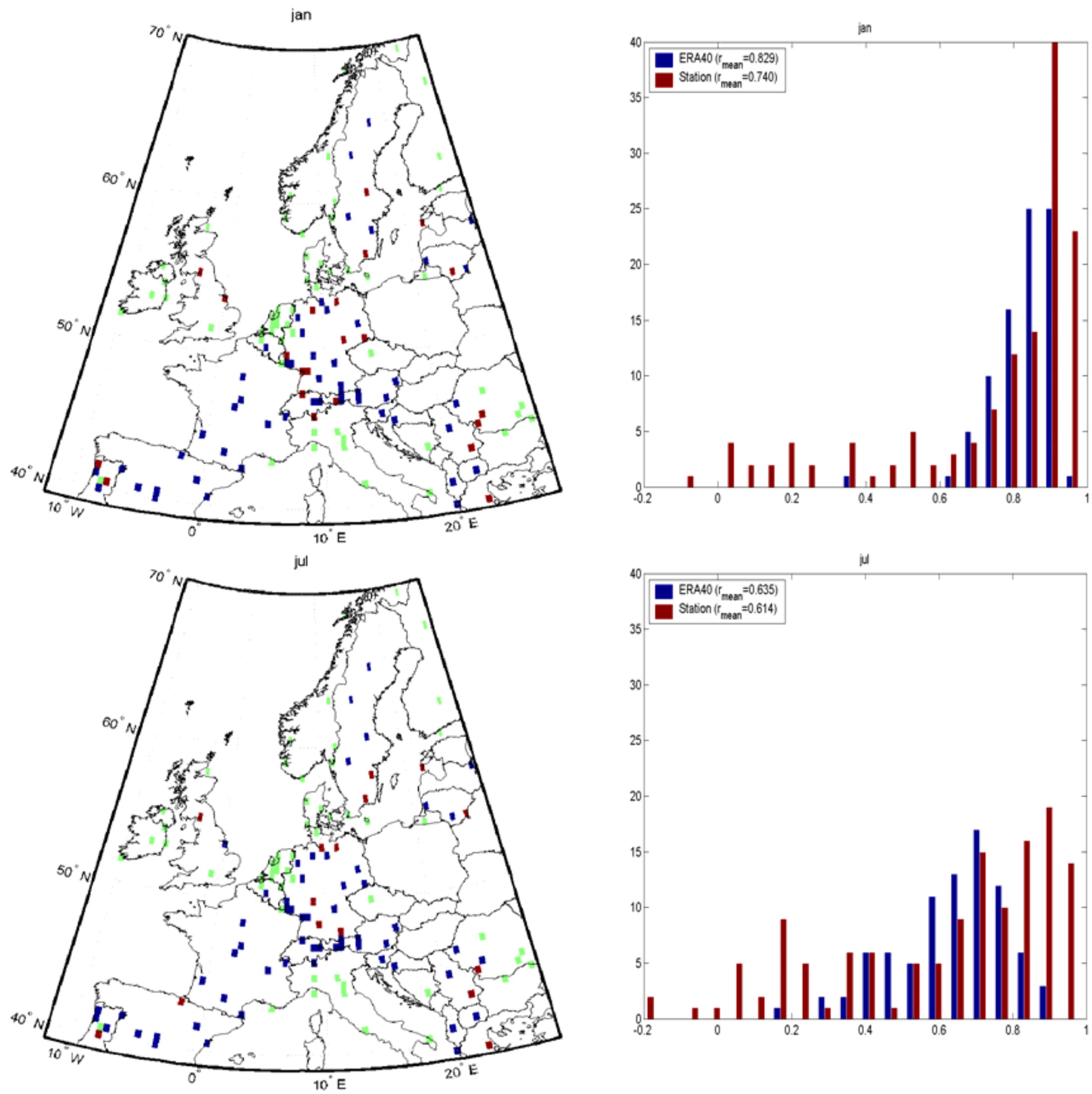

Figure 4.2 Differences in ERA-40 and station correlations with CRU TS2.1 monthly precipitation. January (top), July (bottom). The ERA-40 data has been regridded onto the CRU grid for comparison, using bilinear interpolation. In the maps, gridboxes where the ERA-40-CRU correlation is significantly higher than the station-CRU correlation (based on 95% confidence intervals for r) are shown in red. Those that are lower are in blue, and those with no significant difference at 95% confidence are in green. The histograms show the distributions of these comparisons. ERA-40 correlations with the CRU data are represented by the blue bars, and station-CRU correlations by the red bars.

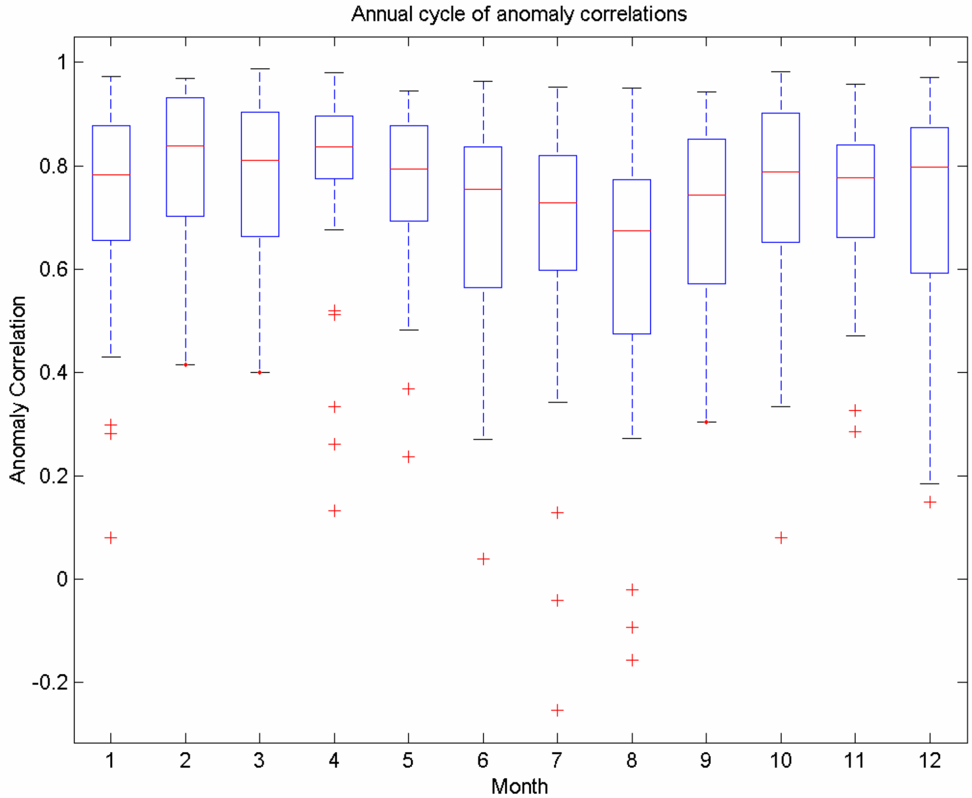

Figure 4.3 Annual cycle of anomaly correlations between the CRU TS2.1 and ERA-40 datasets over the land areas of the European domain for the period 1958-2001. Red line denotes mean, box denotes quartiles, and whiskers denote range of data, not including outliers, which are identified by a red +.

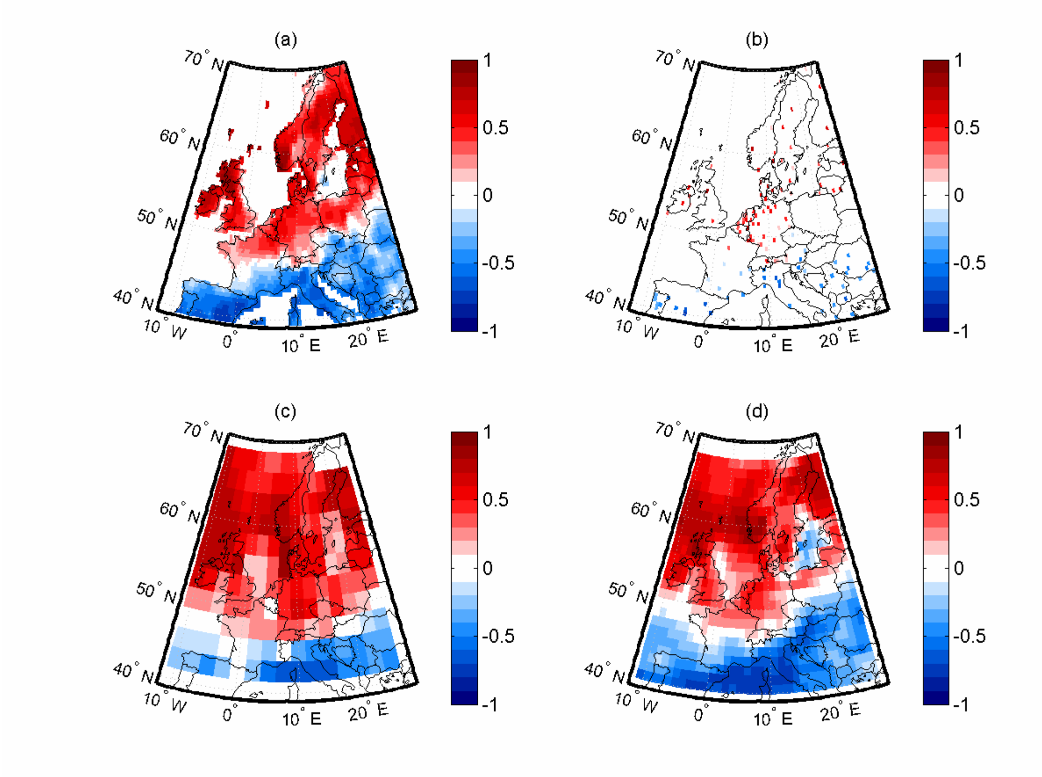

Figure 4.4 January 1979-2001 correlations between the NAO and (a) CRU; (b) stations; (c) CMAP and (d) ERA-40.

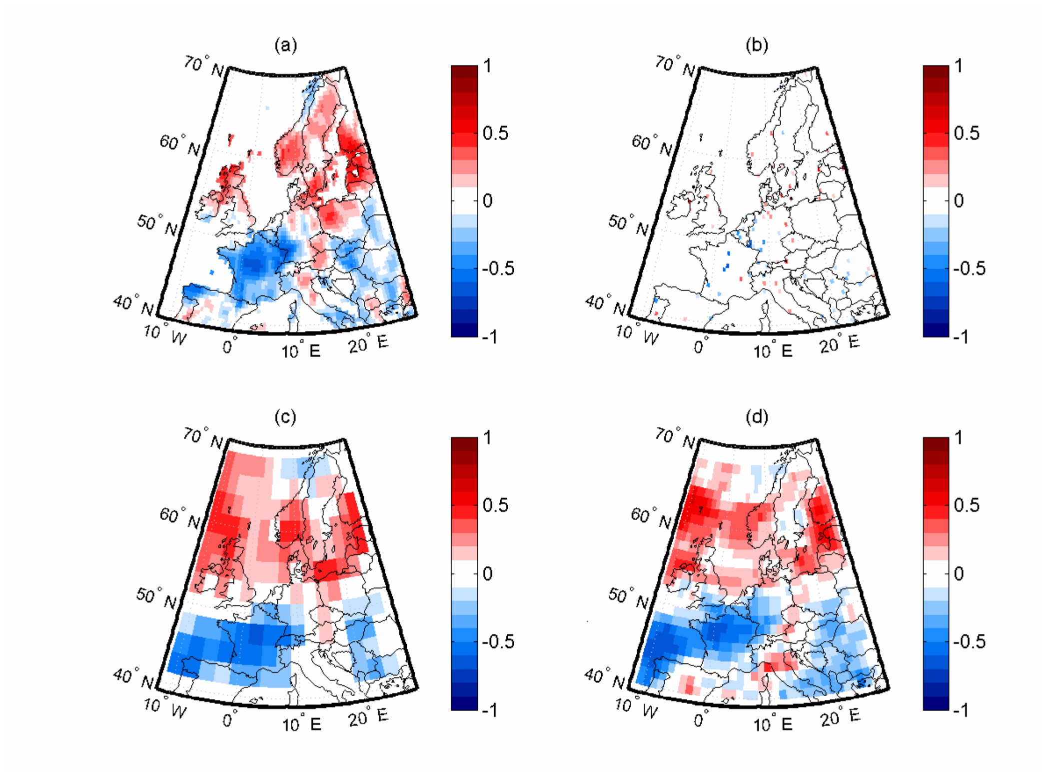

Figure 4.5 July 1979-2001 correlations between the NAO and (a) CRU; (b) stations; (c) CMAP and (d) ERA-40.

The seasonal exceedance indices derived from station records and the nearest ERA-40 gridbox were compared for all seasons. The correlations were inversely related to the percentile threshold, and followed a similar seasonal cycle to that observed in Figure 4.2 and Figure 4.3, that is, stronger in winter/spring, and weaker in late summer. Owing to considerable differences in the frequency of light precipitation in the station and ERA-40 datasets, the number of wet-days per season is greater in ERA-40. In order to assess the optimum wet-day threshold for ERA-40, thresholds were selected for each gridbox and each season using monthly counts of rain days as provided with the CRU TS 2.1 data. These results indicated that in order to comply with the wet-day frequency in the CRU data, unrealistically high thresholds would have to be set. Using these thresholds substantially reduced the correlations between the seasonal exceedance counts of the station and ERA-40 precipitation. It was therefore decided to use the same threshold (0.1mm) in both datasets.

Figure 4.6 summarises the correlation between station and ERA-40 90th and 95th percentile exceedances for December-February (DJF) and June-August (JJA). There is considerable geographic variation in the strength of the correlations, which tend to be stronger in north-western Europe. The weakness of the correlations, particularly in the summer months indicates that both datasets should be treated with caution, and the possibility of using the large-scale precipitation component from ERA-40 as a separate predictor may be helpful.

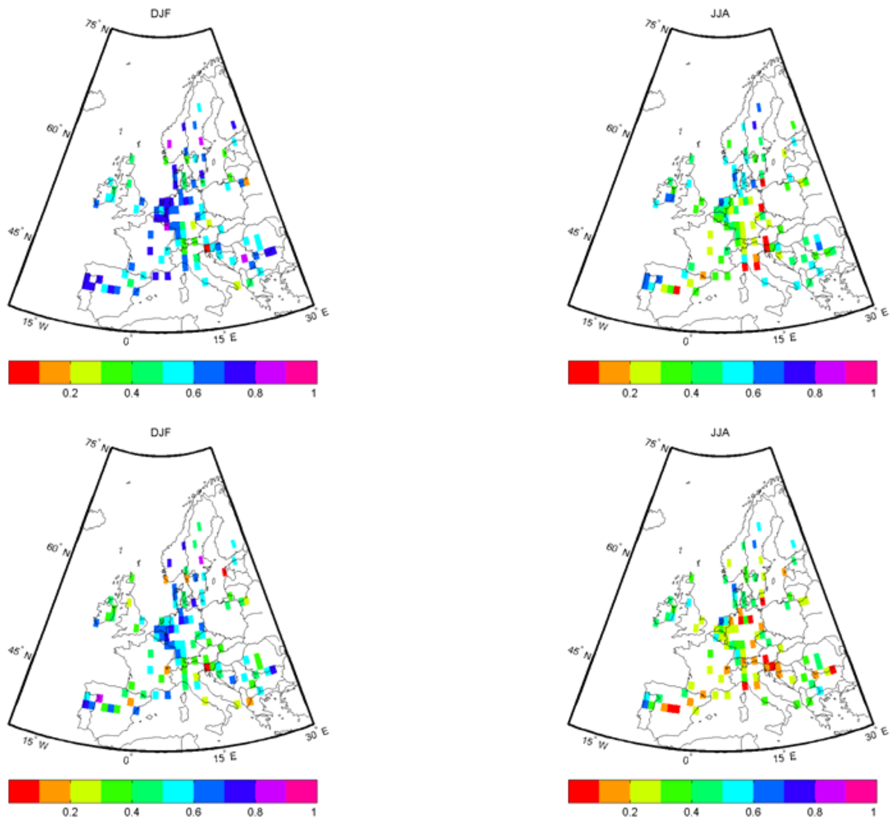

Figure 4.6 Correlation between stations and nearest ERA-40 gridbox for DJF (left) and JJA (right) 90th (top) and 95th (bottom) percentile exceedance counts.

Figure 4.7 illustrates the time series of 95th percentile exceedances for Bremen (Germany) and Eskdalemuir (Scotland) for DJF and JJA. In the case of Bremen, the agreement is strong in the winter, but disappears in the summer months. The converse is true to a lesser extent for Eskdalemuir. For the winter season in Bremen, while the counts do not always compare well, the sign of the anomalies from year to year (the first difference) is generally correct. This is not the case for Eskdalemuir, and is may be attributable to local topographic effects in relation to the areal mean precipitation, as well as model bias due to unrealistic topography.

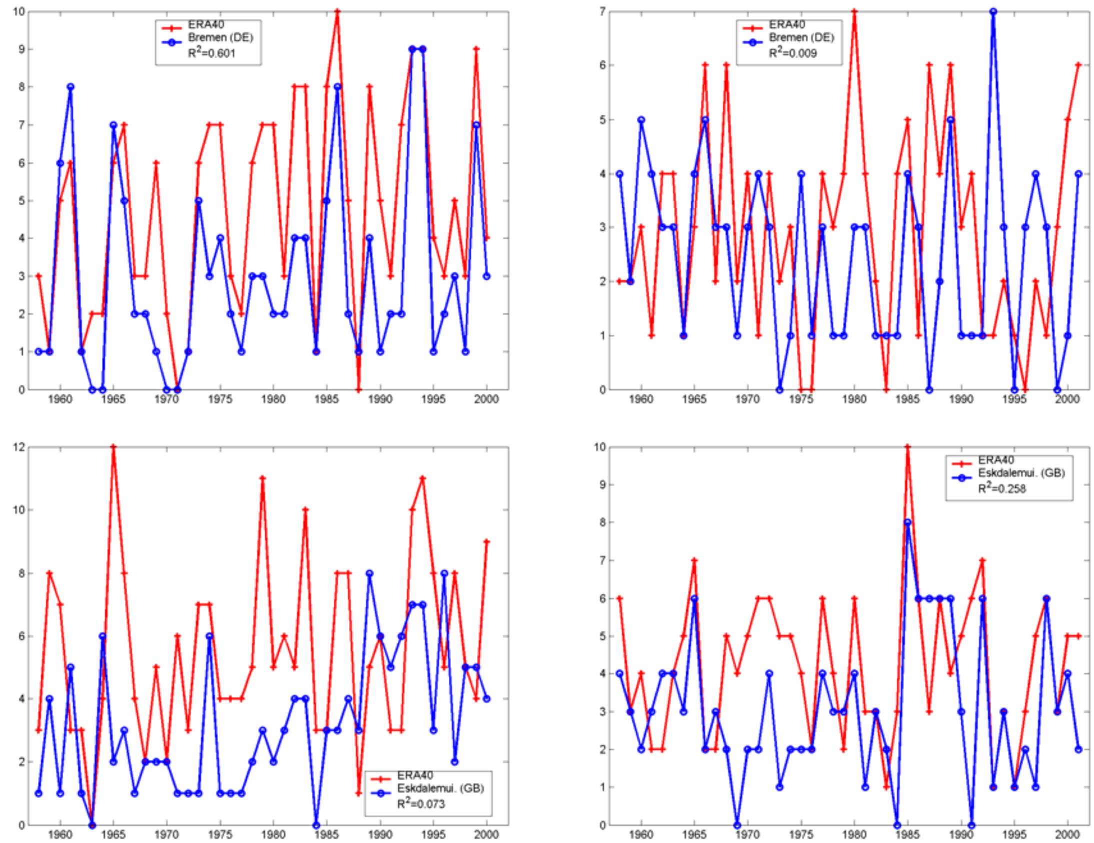

Figure 4.7 Comparison of seasonal 95th percentile exceedance counts for the ERA-40 peak wind gust and station data for Bremen, Germany (top) and Eskdalemuir, Scotland (bottom) for DJF (left) and JJA (right).

4.4.2 Wind results

Owing to the scarcity of peak wind speed data, a limited range of comparisons were carried out between seasonal exceedance counts for the available wind stations and the corresponding ERA-40 gridbox. Figure 4.8 summarises the correlations between the 90th and 95th percentile exceedance counts for DJF and JJA. As with the precipitation indices, correlations are markedly lower in the summer months. However, this is not of great practical importance owing to the relatively low impact of wind speeds in the summer. The differences in correlations between the 90th and 95th percentile exceedances are less marked than with the precipitation data.

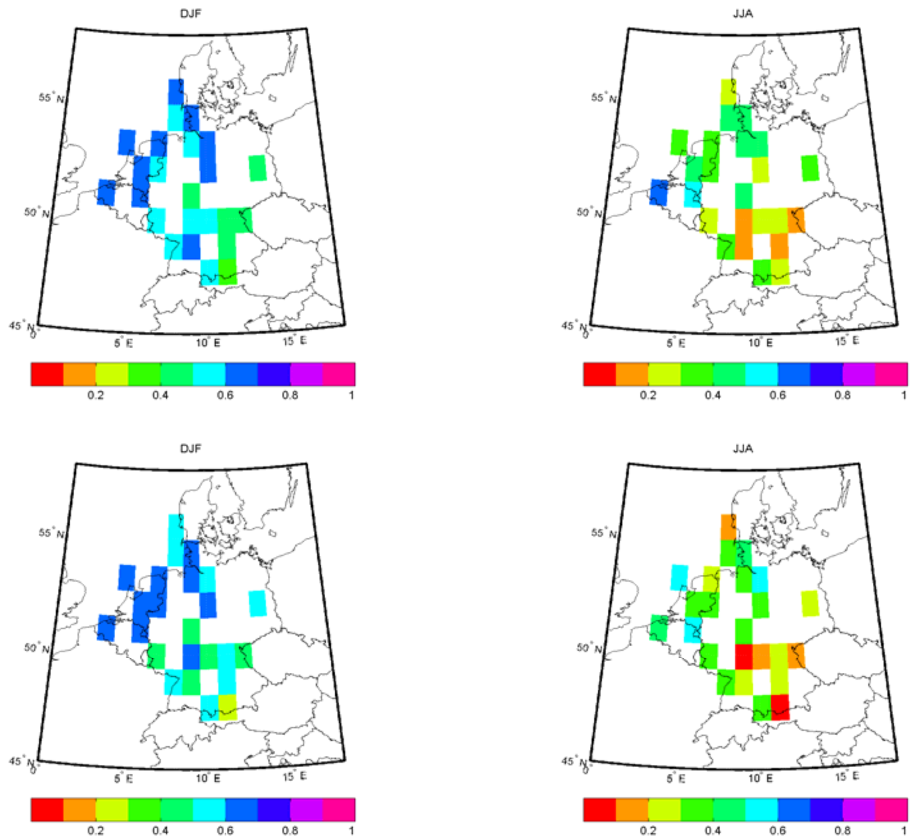

Figure 4.8 Correlation between stations and nearest ERA-40 gridbox for DJF (left) and JJA (right) 90th (top) and 95th (bottom) percentile exceedance counts.

Figure 4.9 extends the comparison to individual station timeseries, again for DJF and JJA 95th percentile exceedance counts. The stations shown are De Bilt (Netherlands) and Stuttgart (Germany). Although there is a considerable difference in the correlations between winter and summer for both stations, the difference is less marked in the case of De Bilt, which is likely due to less of an influence from topographic effects.

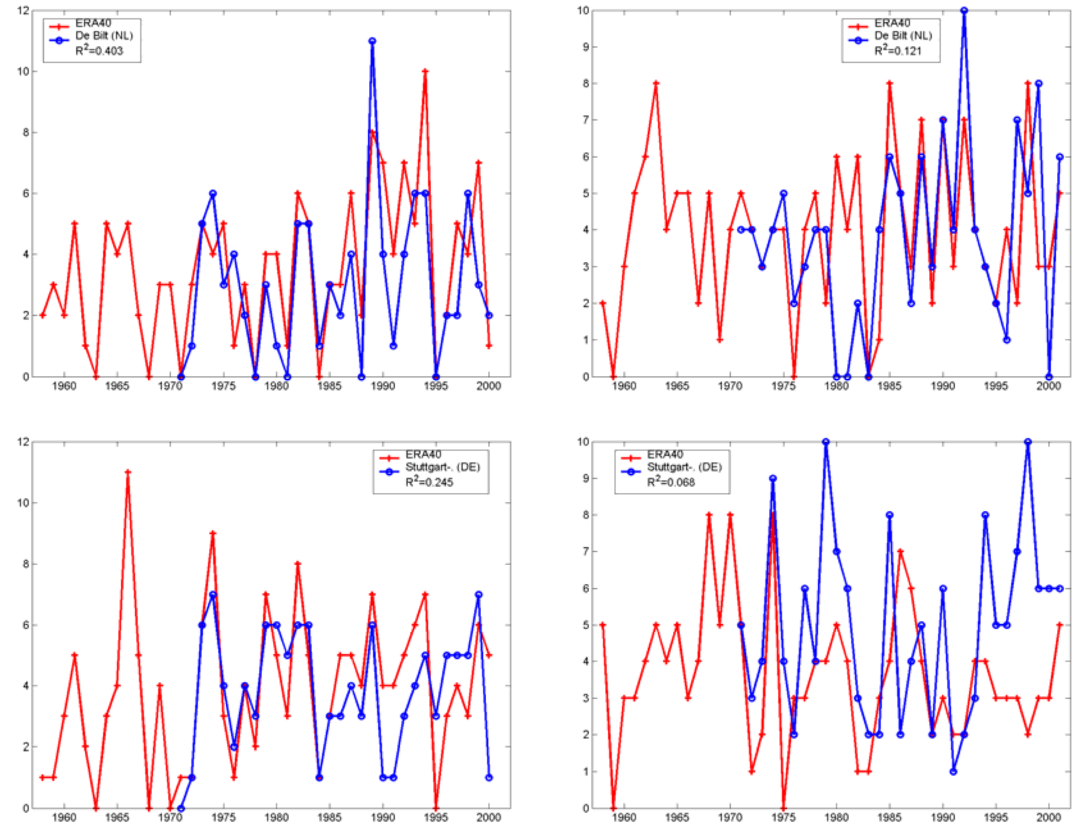

Figure 4.9 Comparison of seasonal 95th percentile exceedance counts for De Bilt (Netherlands) and Stuttgart (Germany) for DJF (left) and JJA (right).

For the purpose of further analysis, as outlined in Chapter 3, it is a requirement that the predictand data are approximately Poisson distributed. Figure 4.10 and Figure 4.11 illustrate comparisons of the observed frequency distribution and the theoretical Poisson distribution for selected gridboxes across the study domain. JJA 95th percentile precipitation exceedances and DJF 90th percentile wind exceedances are shown, respectively. It can be seen from Figure 4.10 and Figure 4.11, and also more generally for all gridboxes, seasons and predictands (not shown) that the frequency distributions approximate the Poisson distribution, although in some cases there is a tendency for unexpectedly high numbers of seasons with high rates of occurrence of extreme events. This is likely attributable to the small sample size, since the aggregation of gridboxes into still larger regions tends to remove this effect (not shown).

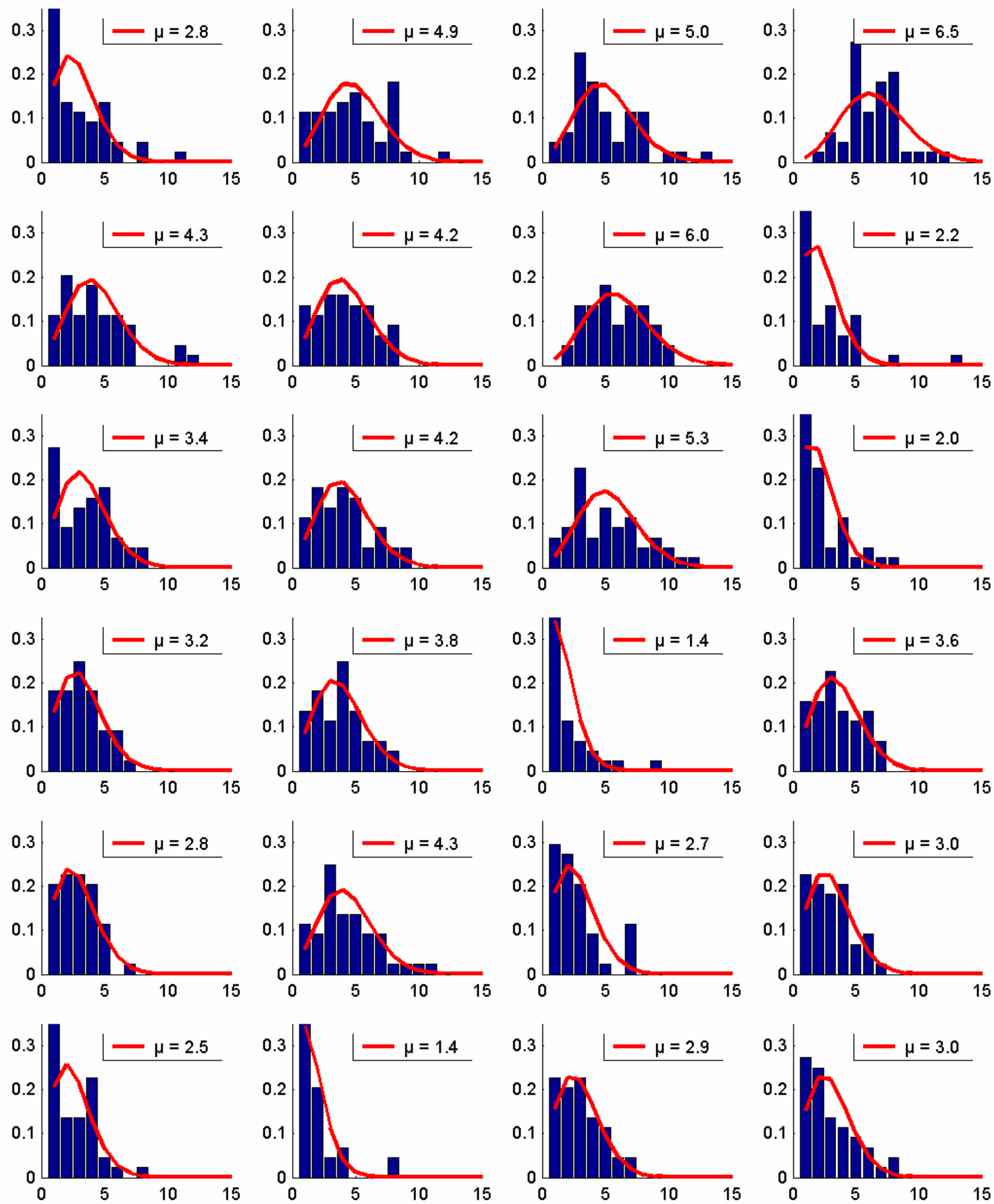

Figure 4.10 A sample of empirical frequency distributions of JJA 95th percentile precipitation exceedance counts from the ERA-40 dataset, compared to the theoretical Poisson distribution with mean μ equal to the observed mean at each gridbox. ERA-40 gridboxes are aggregated to groups of four and the predictands re-derived from the raw data.

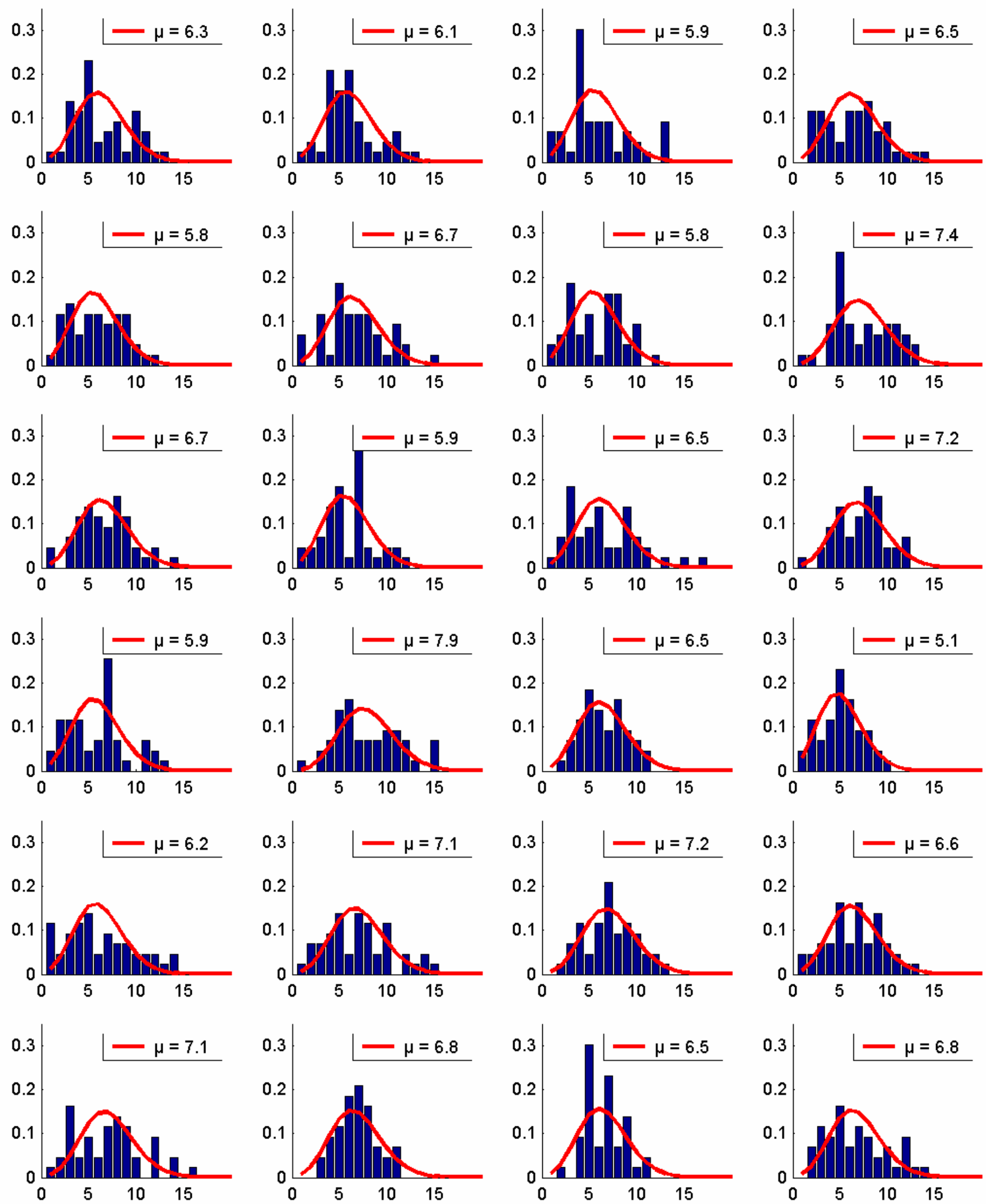

Figure 4.11 A sample of empirical frequency distributions of DJF 90th percentile wind exceedance counts from the ERA-40 dataset, compared to the theoretical Poisson distribution with mean μ equal to the observed mean at each gridbox. ERA-40 gridboxes are aggregated to groups of four and the predictands re-derived from the raw data.

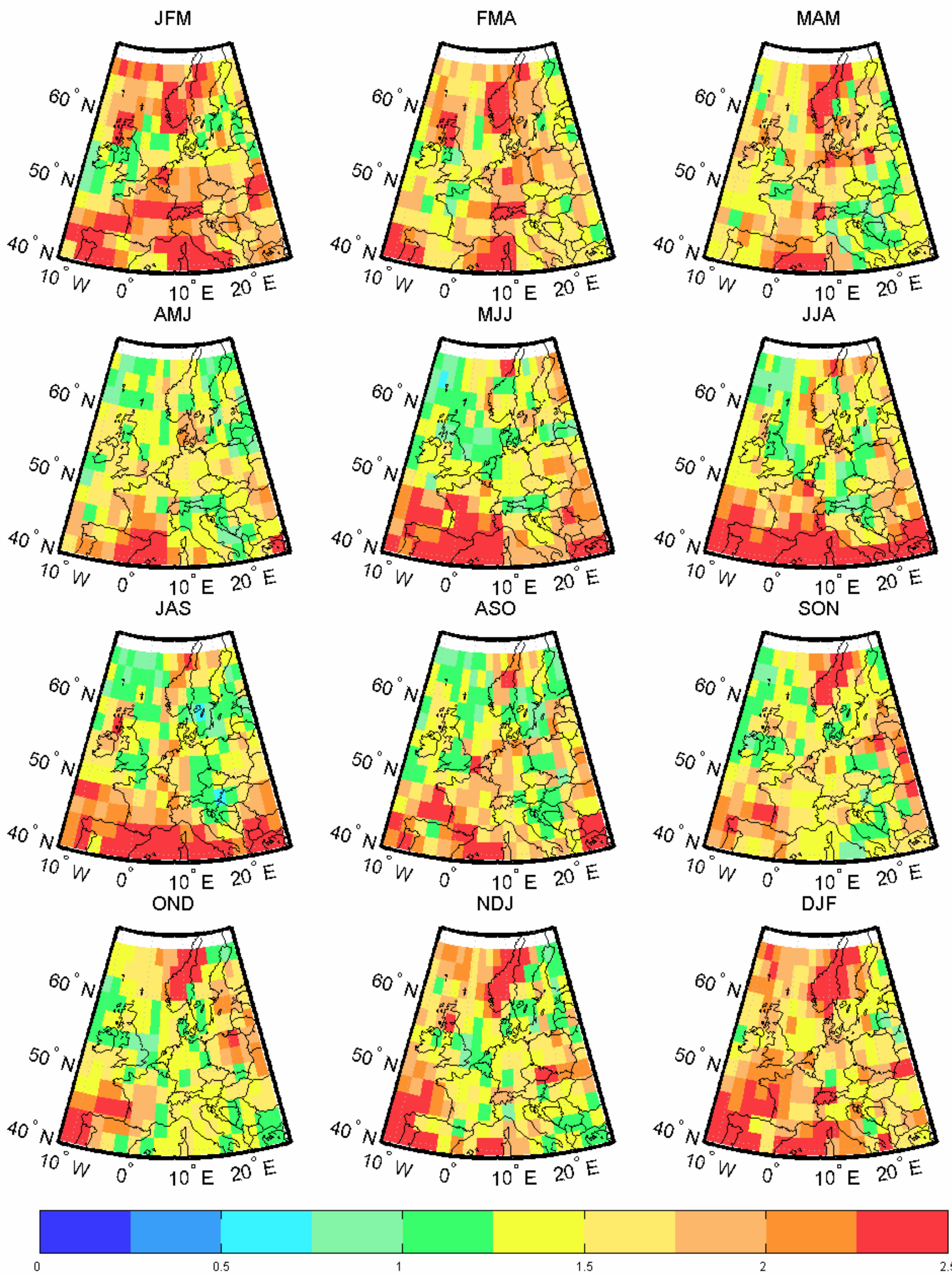

Figure 4.12 Illustration of the dispersion parameter for ERA-40 90th percentile precipitation exceedance counts by season. Values greater than 1 indicate overdispersion, where the variance of the exceedance counts is greater than the mean.

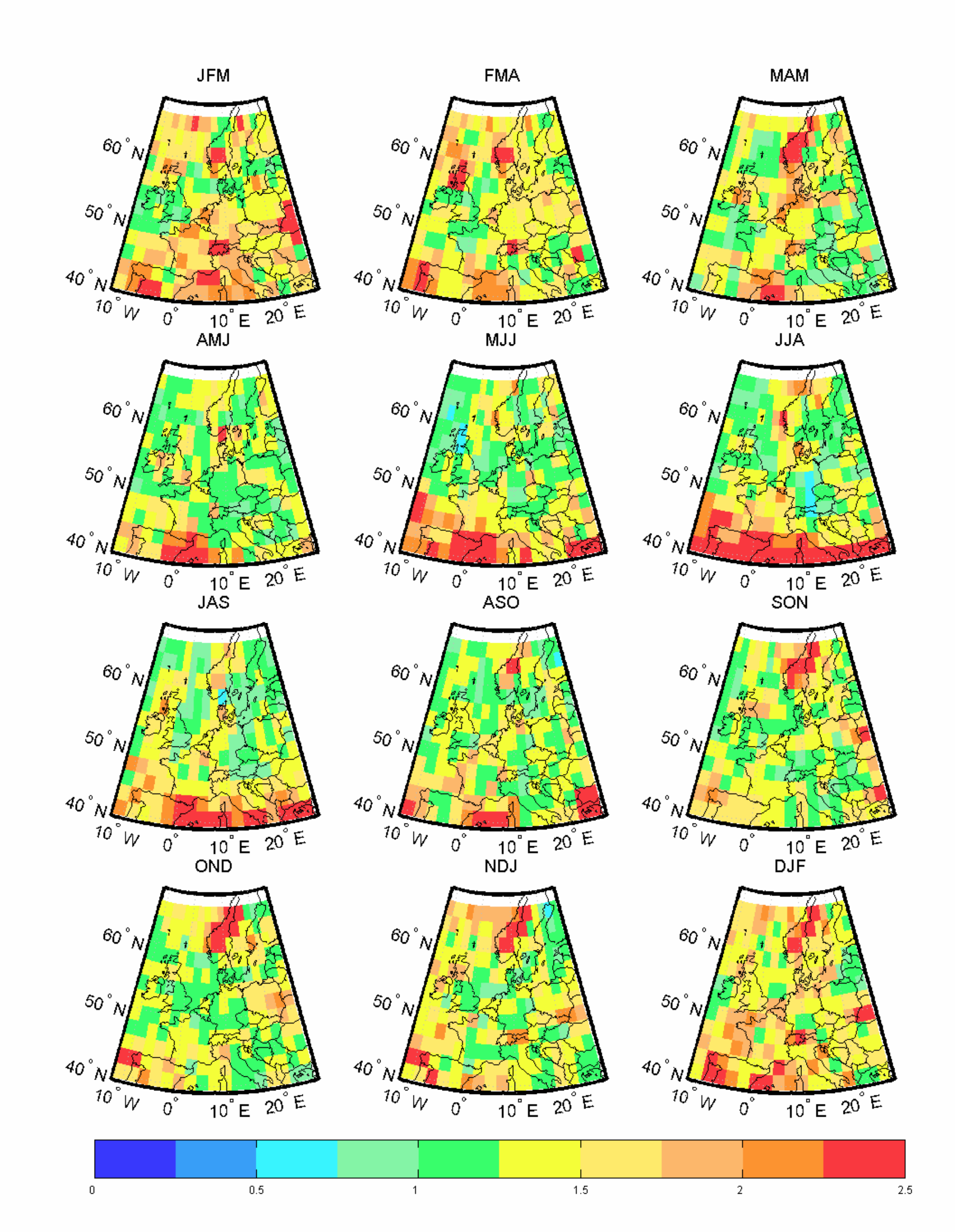

Figure 4.13 Illustration of the dispersion parameter for ERA-40 95th percentile precipitation exceedance counts by season. Values greater than 1 indicate overdispersion, where the variance of the exceedance counts is greater than the mean.

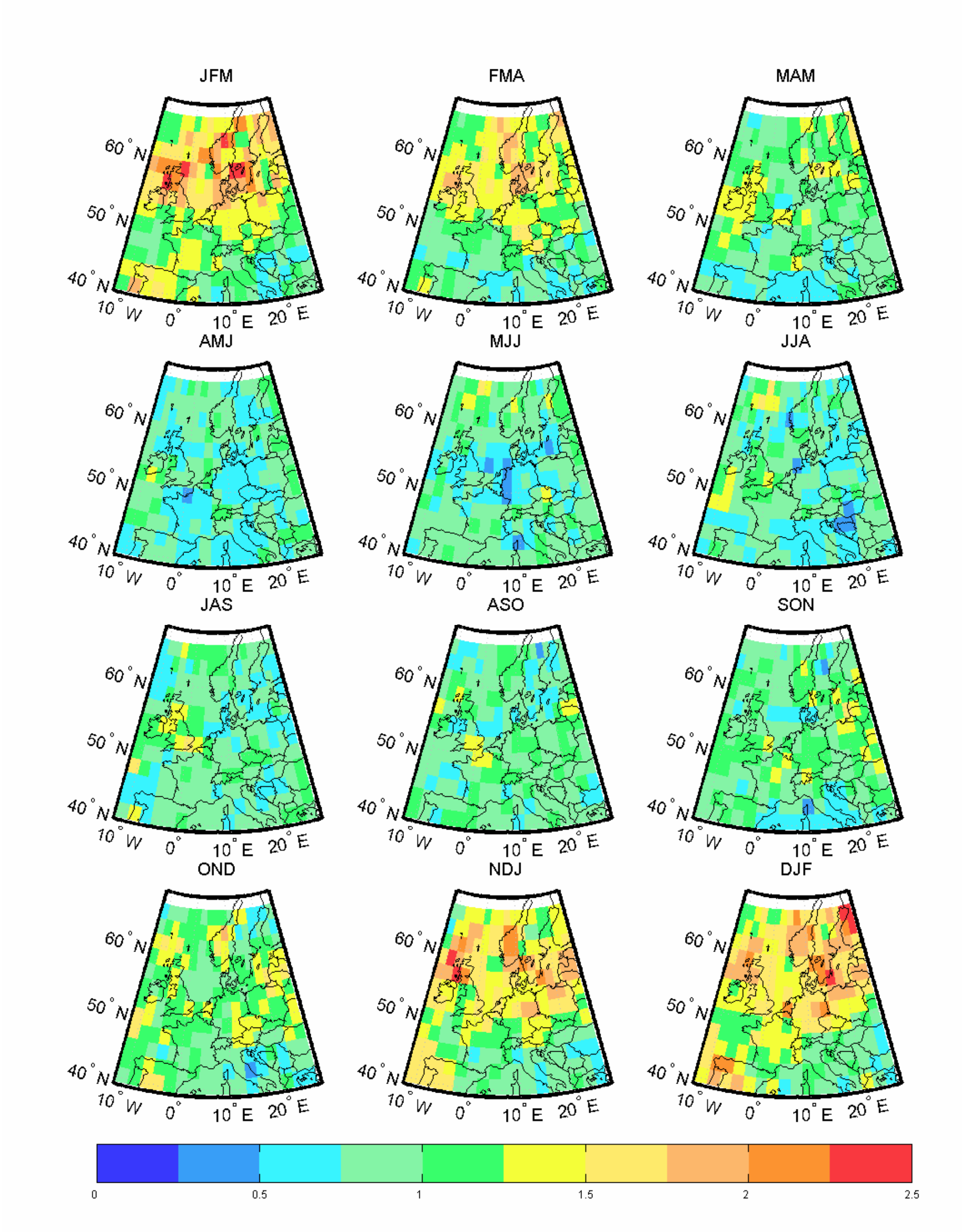

Figure 4.14 Illustration of the dispersion parameter for ERA-40 90th percentile wind exceedance counts by season. Values greater than 1 indicate overdispersion, where the variance of the exceedance counts is greater than the mean.

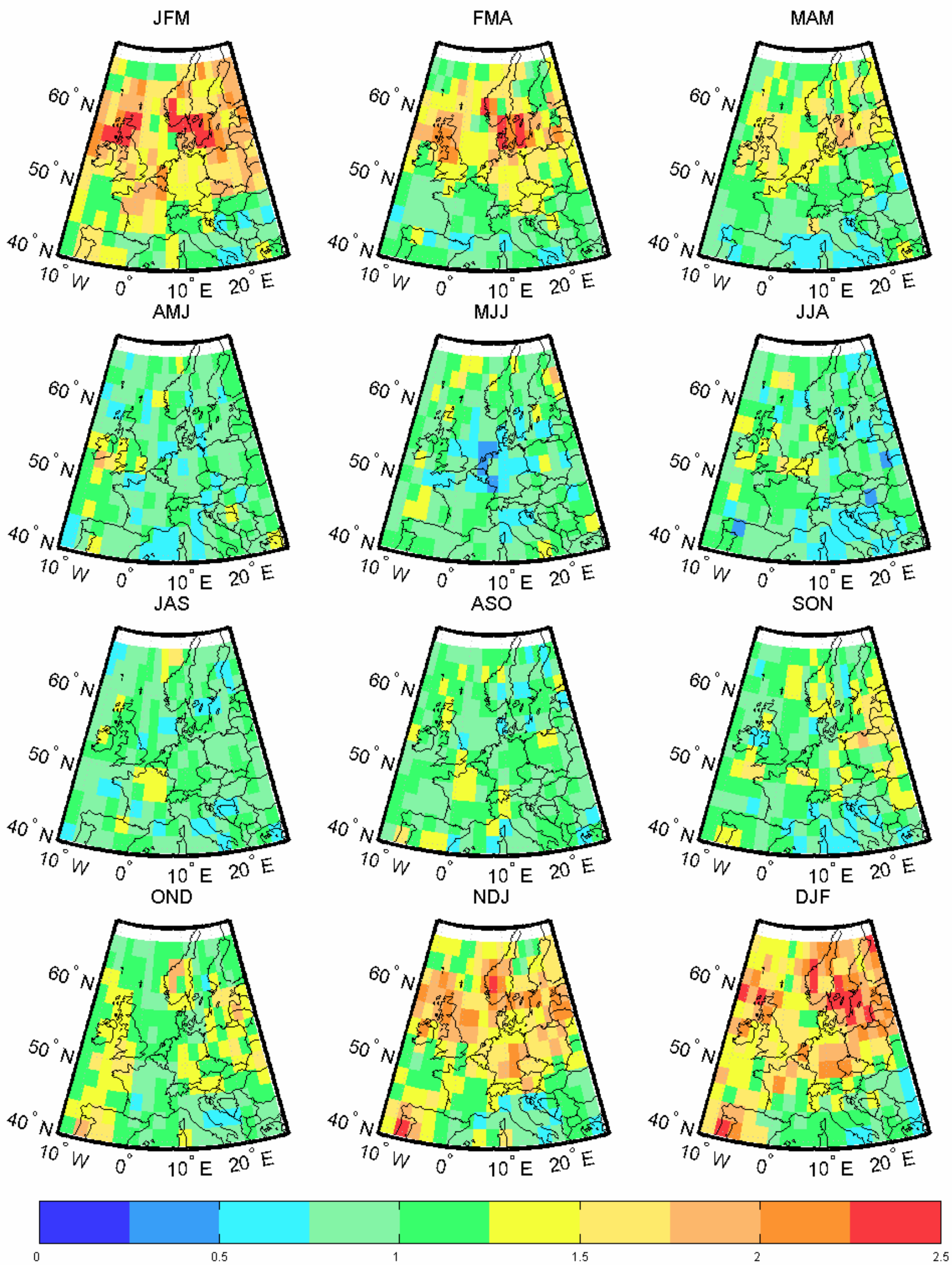

Figure 4.15 Illustration of the dispersion parameter for ERA-40 95th percentile wind exceedance counts by season. Values greater than 1 indicate overdispersion, where the variance of the exceedance counts is greater than the mean.

The Poisson distribution is described by a single parameter μ, representing both the mean and the variance of the distribution. When applying a Poisson regression model to data which approximates a Poisson distribution it is desirable to assess the degree to which the empirical data satisfies the condition of equivalent mean and variance. The standard approach is simply to measure the ratio of the observed mean and variance to give a metric of the dispersion. Values greater than one, where the variance is greater than the mean are said to indicate overdispersion, and values less than one underdispersion. Figure 4.12 illustrates the dispersion statistics for the ERA-40 90th percentile precipitation exceedance counts, for each of the 12 three-month seasons. The general pattern is for overdispersion, which is particularly strong in the south of the domain, and during the summer months. This indicates that data might for example be better described using a negative binomial distribution, which has two parameters, describing the mean and variance. The 95th percentile exceedance counts for precipitation are a closer approximation to the Poisson distribution as shown in Figure 4.13. Here a similar geographical and seasonal pattern is observed, but generally the values of the dispersion statistic are closer to one. The implication to be drawn from this is that the 90th percentile data may suffer from the occurrence of precipitation events or dry spells which are not entirely independent in time. For example due to extended periods of heavy rainfall resulting from the same synoptic weather system, or extended periods of drought in southern Europe.

The wind predictands generally have dispersion statistics much closer to one, with the exception of regions of northern Europe during winter, as illustrated in Figure 4.14 and Figure 4.15, for the 90th and 95th percentile exceedance counts respectively. Where the data appear to be substantially overdispersed, the implication is that there is some effect due to the clustering of extra-tropical cyclones, as explored in for example Mailier *et al* (2006). However, overall it is thought that both the wind and precipitation predictand data are sufficiently closely approximated to a Poisson process for Poisson regression to be used to model the predictor predictand relationships. Some overdispersion is unavoidable, and as discussed in Chapter 3, estimates the model error and significance of fit are made more conservatively given this overdispersion.

4.5 Summary

The definition of predictands for a seasonal forecasting model designed to forecast extreme events is a challenging task. The predictands must be useful with respect to the events it predicts, and their frequency of occurrence must be sufficient to discern interannual variability related to the large-scale circulation of the atmosphere. µ

In the case of precipitation, both the station dataset provided by the ECA and the forecast precipitation from ERA-40 seem to capture interannual variability in the monthly means quite well. The realistic relationship between ERA-40 and the NAO is particularly encouraging. There are some marked differences in the representation of extreme precipitation. However the pattern of interannual variability is well represented, particularly in the winter months. Given the available station network, it is impossible to make an objective appraisal of which dataset represents interannual variability of extreme precipitation more accurately. Further considerations are the possible inhomogeneities and the missing data in the station records. Kharin *et al*. (2002) comment on the utility of reanalysis precipitation in respect of its temporal and spatial homogeneity, while Zolina *et al*. (2004) compare European extreme precipitation in all four reanalysis datasets (ERA-40, ERA15, NNR1 and NNR2) and conclude that with respect to moderate extremes, the reanalyses capture temporal variations effectively. Given the requirements of this thesis, temporal and spatial continuity are highly important, as is a realistic representation of interannual variability. ERA-40 will therefore be used as the predictand component of the model. The 90th and 95th percentile exceedance threshold counts will be considered. Owing to the computational expense of the predictor selection algorithms as described in Chapter 3 and Chapter 6, the gridboxes are combined in groups of four as the mean value of the raw data to reduce the spatial resolution – that is, four neighbouring gridboxes at the native resolution are averaged to form a new gridbox, where geographically each of the old gridboxes comprises a quadrant of the new gridbox. The predictands are derived from the aggregated gridboxes in exactly the same way as described above for the data comparisons, and correlate very highly with the original data. A spatial illustration of the new aggregated data is illustrated in Figure 4.12. As well as reducing the computational expense of the predictor and model selection algorithms, the spatial scale of the extremes, particularly the wind and winter precipitation extremes are such that the interannual variability is still adequately represented by the aggregated gridboxes, and the great majority of the observed extremes are greater than the gridbox scale. Indeed it is speculated by Lloyd-Hughes and Saunders (2003) that for seasonal forecasts of mean precipitation it may be necessary to aggregate the predictand data to much larger (pan-European) domains in order to gain any forecast skill. This approach does not necessarily translate to the prediction of extremes, given the extent to which extreme events would be smoothed over a very coarse domain.

Finally, in order to maximise the model training and validation period, data from the ECMWF Operational Analysis are used to extend the predictand data to 2005. This data is also obtained from the BADC, who provide a version of the Operational Analysis that is consistent with the ERA-40 reanalysis. Although some changes were made to the methodology in developing the operational analyses – for example the transition from 3D-Var to 4D-Var data assimilation, as well as the addition of some new data sources (Uppala *et al*, 2005), it is found that for the period in which the two datasets (ERA-40 and the ECMWF Operational Analysis) overlap (2000-2001), the data are virtually identical (not shown). However, the forecast data from ERA-40 (i.e. both the precipitation and the wind fields) are derived based on the operational forecasting system in place from June 2001 to January 2002, after which a new forecast scheme was used to generate the operational data (Uppala, *et al*, 2005). There is therefore no opportunity to compare the ERA-40 data with operational based on data post-2002 forecast schemes. It is therefore possible that there are significant inhomogeneities in the data. However, it is thought that given the substantial similarities that do still exist between the two datasets, it is justifiable to include the operational data in order to maximise the record length.

5 Predictors

This Chapter introduces the procedure by which potential predictors are selected as candidates in the first stage of the model selection process. Section 5.1 comprises the introduction. The approach to potential predictor selection is outlined, including a discussion of the nature of the predictor selection problem in model design in more general terms. Section 5.2 gives details of the predictors, including justification for their inclusion, data sources for the predictors, and the development of predictor indices. Section 5.3 summarises the predictor information, and the implications for further model development.

5.1 Introduction

As discussed previously, there are few useful predictors of wind and precipitation extremes in Europe. Some tentative statistical relationships have been identified, focussing on seasonal precipitation totals. For example Colman and Davey (1999) use the first principal component (PC) from a principal component analysis (PCA, also referred to as empirical orthogonal functions, or EOF analysis) of January-February North Atlantic SST anomalies, over the spatial domain 20-80°N as a predictor for summer temperature, rainfall and pressure in Europe. Some very limited skill exists for July-August rainfall over areas of Western Europe. Feddersen (2003) identifies Tropical and North Atlantic SSTs, as well as the NAO as potentially useful predictors of Scandinavian precipitation. Muñoz-Díaz and Rodrigo (2006) use regional sea-level pressure (SLP) fields and indices of the El Niño-Southern Oscillation (ENSO) phenomenon, respectively, as statistical predictors of precipitation over Spain. Lloyd-Hughes and Saunders (2002) use ENSO and local SST fields to develop a seasonal forecasting system for central European Spring precipitation. Wilby *et al.* (2004) suggest that the search for potential predictors (with specific reference to summer hydrometeorological conditions of the River Thames, UK) should not be limited to the Northern Hemisphere. Predictors included in the Wilby *et al.* (2004) study include sea-level pressure modes such as the NAO, indices of North Atlantic SST, and Northern

Hemisphere sea-ice cover indices. General descriptors of the climate such as the NAO can also be used as predictands, from which mean conditions can be inferred (e.g. Qian and Saunders, 2003). For example the UK Met Office (UKMO) bases its operational forecasts of the winter NAO (from which climatological variables such as temperature and precipitation can be inferred) on the first PC of late summer SST anomalies in the North Atlantic. Other studies have found relationships between Northern Hemisphere or Eurasian snow cover anomalies and the NAO, or other measures of large-scale circulation variability, using both the observed record (e.g. Qian and Saunders, 2003; Saunders *et al*, 2003; Bojariu and Gimeno 2003a, b; Cohen *et al.,* 2001; Cohen *et al.*, 2002) and model studies (Kumar and Yang, 2003).

Potential predictors for extremes of precipitation and wind have not been investigated to the same extent as those for seasonal mean conditions. Much of the work relevant to precipitation focuses on simultaneous relationships between the large-scale circulation and the likelihood of extremes, often for the purpose of developing precipitation downscaling schemes. Haylock and Goodess (2004) use canonical correlation to identify links between PCs of extreme winter rainfall in Europe, and mean sea level pressure (MSLP) over the north Atlantic, finding significant links, in particular with the NAO, where a positive NAO leads to increased frequencies of extreme precipitation over Northern Europe and a decrease over Southern Europe, matching the canonical pattern for mean rainfall shown by e.g. Hurrell (1995) and Trigo *et al.* (2002). Additionally, a mode centred over the North Sea, which appears to be linked to local SSTs is found to fit well with precipitation extremes over the same spatial domain. Simultaneous SSTs were found to have a weaker fit to the rainfall data. Hellstrom (2004) uses a system based on the Lamb classification (Lamb, 1950) to assess the relationship between circulation type and extreme (≥40mm) summer precipitation in Sweden, finding that extreme precipitation events tend to be associated with lower westerly flow, and higher southerly flow. Plaut and Simonnet (2001) show that the frequency of wet days in the top decile of precipitation amount (equivalent to the 90^{th} percentile exceedance in this study) over south east France, is linked to so-called weather regimes defined by a cluster analysis, and describing large-scale circulation over the North Atlantic-European sector. Moreover, the clusters tend to act as attractors for the precipitation, and as such they are more effective descriptors of the extreme climatology than the mean. Gallego *et al.* (2005) find a strong and virtually simultaneous association between the winter NAO index, and the frequency and intensity of daily precipitation over the Iberian Peninsula.

Wind and storms in particular have been the focus of substantially less research than precipitation. To many extents forecasts of the NAO have implications for the likelihood of wind extremes in Europe, relating to the strength of the prevailing westerlies and the position of the storm track, particularly in Winter and Spring (e.g. Trigo *et al.,* 2002; Bojariu and Gimeno, 2003; Qian and Saunders 2003). DeGaetano (2002) examines potential predictors for seasonal forecasts of East-coast USA windstorms. A wide range of potential predictors are considered, based on work by Hirsch *et al.* (2001), and other relationships described in the literature. These predictors include Niño 4.3 SST, and SST from the tropical Atlantic, and Eastern Seaboard of the USA. Also, US Land temperature, and indices such as the NAO, and Southern Oscillation Index (SOI) are included. The initial predictor set is refined using a screening procedure, to yield a more manageable subset of statistically useful predictors. Significant skill over climatology is reported. Enloe *et al.* (2004) find a significant ENSO influence on peak wind gust magnitudes over the United States, manifest as an asymmetric response to cold-phase events (La Niña) more than warm-phase (El Niño). With respect to the predictability of windstorms in Europe, little work has been done. Palutikof *et al.* (2002) find little evidence for predictability, although Qian and Saunders (2003) identify a statistically significant link between Eurasian summer snow cover and the winter NAO, from which they derive a model to predict European winter storminess, with some skill. Yan *et al.* (2002) studied the simultaneous variability of European wind extremes with indices such as the NAO using a GLM.

With respect to this thesis, the papers referred to above serve to illustrate some of the predictor selection issues faced here. Namely, they suggest that developing a precipitation or windstorm prediction system for seasons throughout the annual cycle, and over an area such as northwestern Europe requires a diverse set of potential predictors, with little guidance from the literature as to which predictors are most skilful or physically plausible. Additionally, skill is expected to be very low. While some studies have focused on a narrow family of predictors, for example snow cover to predict the NAO (Saunders *et al*., 2003), others use a wider range of potential predictors (for example Wilby *et al.,* 2004); De Gaetano 2002). The approach taken here is to consider a wide range of potential predictors, including SSTs, atmospheric indices and land surface indices such as snow cover, but also including solar forcing, which according to some recent analyses (e.g. Haigh *et al*., 2005) may offer some predictive skill at seasonal timescales.

However, there is danger in considering a large number of potential predictors without a clear physical mechanism by which they might each contribute predictability, especially when a relatively short record of the order of 40 samples is considered. The possibility of overfitting the model – that is – fitting the predictors to noise in the data, rather than signals which may be present becomes almost a certainty, even when the number of covariates in the model is low. It is therefore necessary to have rigorous methods by which the initially large set of potential predictors can be reduced and validated. In this case, field significance tests using resampling will be used to select only those predictors which have a significant (95%) and spatially coherent impact on the predictand. This allows an initially large set of predictors to be considered and then significantly reduced in number. It should be emphasised that this approach, in the absence of physically explicable predictability must be essentially exploratory. Details of the field-significance testing will be given in Chapter 6.

5.2 Potential Predictors

5.2.1 Large scale SST

Sea surface temperature is typically the primary source of long-range predictability in statistical seasonal forecasting (Goddard *et al.,* 2001). Early seasonal forecasting efforts were driven by the need to understand the impacts associated with the ENSO phenomenon in the Tropical Pacific, during the 1980s (e.g. Ropelewski and Halpert 1987). Since then most skilful forecasts have relied to some extent on the persistence or prediction of tropical SST anomalies (e.g. Colman and Davey (2003); Folland *et al.,* 1991; Ward and Folland 1991). Since ENSO has been linked to European climate variability widely in the literature (Fraedrich 1994, Pozo-Vazquez *et al.,* 2005a, b), Pacific SST variability is considered here as providing a set of potential predictors. Work by Bader and Latif (2005); Goswami *et al.,* (2006) and Li *et al.,* (2006) suggest that the Indian Ocean can also be causally linked to the European climate, in particular through the NAO. Li *et al.* (2006) identify a positive feedback whereby warming in the Indian Ocean induces a positive NAO, which is maintained by positive feedback with the local (North Atlantic) SST anomalies. More recently, the causal nature of the link between Indian Ocean warming and a positive NAO has been questioned (Copsey *et al*., NCAS/NERC conference, 2006). In this case, a GCM forced with observed SST showed a decreasing SLP trend over the Indian Ocean, contrary to observations, bringing into question the mechanism proposed by Li *et al*. (2006). The Atlantic Basin does not have a dominant mode with far-reaching teleconnections in the same manner as the Pacific, nor is it clearly linked to ENSO variability (Sutton *et al.,* 2000). However, several clear modes of variability have been identified. In the tropics, where ocean-atmosphere coupling is stronger, owing to the warmer SSTs (Sutton *et al.,* 2000; Frankignoul and Kestenare 2005), considerable research has attempted to link persistent or predictable SST modes with European climate. Some studies have found links (e.g. Czaja and Frankignoul 2002; Peng *et al.,* 2005) between tropical Atlantic variability and the North Atlantic/European climate, although the exact nature of the physical modes, and the forcing mechanism are as yet unclear. In the extratropics, modes such as the North Atlantic tripole (e.g. Rodwell and Folland 2002, 2003) are primarily integrations of atmospheric forcing – in this case by the NAO. However, there is some evidence of coupling which in the case of the North Atlantic Tripole results in a positive feedback relationship between the tripole and the NAO (Rodwell and Folland, 2002, 2003; Czaja and Frankignoul 2002). It should be noted here that there are still substantial differences between the observed and modelled representation of these relationships. For example Rodwell and Folland (2002) find that the observed relationship is significantly stronger than the modelled relationship. Hurrell *et al.,* (2006) provide an excellent review of work to date on Atlantic climate variability and predictability. They emphasise, among other things, the importance of the continued development of the observing system. In particular, the measurement of subsurface temperatures is thought to offer potentially enhanced insights into the nature and predictability of Atlantic Ocean variability. While the nature and strength of ocean-atmosphere coupling in the Atlantic as a whole, and the North Atlantic in particular are subject to much ongoing research, useful predictive skill has been found and used operationally by groups such as the UK Met Office (UKMO). The whole Atlantic basin is considered here as a source of potential predictability.

Further studies utilise SSTs more specifically to assess predictability of European climate. For example, Colman and Davey (1999) devise a forecast model for European summer precipitation, temperature and pressure based on the leading mode of North Atlantic winter SST. Benestad and Melsom (2002) assess the effect of North Atlantic SSTs on Norwegian precipitation, finding possible evidence of an influence driven by the warming trend in the North Atlantic. McGregor and Phillips (2004) explore the predictability of rainfall in southwest England, and find that much of the apparent predictability is due to North Atlantic SSTs. Predictive skill is typically low, but a slight improvement on climatology.

5.2.1.1 Large Scale SST Data and Predictor Indices

The Hadley Centre (UKMO) ice and SST dataset HadISST1.1 (Rayner *et al.* 2003) is used. The dataset covers the period 1870 to the present, although only the period 1958-2005 is considered here. Data is provided by the British Atmospheric Data Centre (BADC, http://badc.nerc.ac.uk).

Broadly speaking there are three methods by which predictors from multivariate SST data can be derived. Firstly, correlation or regression can be used to identify regions which correlate highly with the predictand (e.g. Peng and Mysak, 1993). Given that the predictand is also a field, and that the assessment of significant predictor-predictand relationships must include some form of field-significance testing to assess the spatial degrees of freedom in the SST data, this approach will either result in a very large number of predictors, each of which specifically apply to one predictand timeseries (gridbox), or a smaller number of regional indices will have to be selected subjectively. Either way, the requirement for field significance testing imposes a considerable computational expense. Given the availability of alternative techniques, this approach is therefore not considered further.

Secondly, instead of identifying key regions which are correlated with the predictand, it is possible, using a technique such as PCA, to identify the most important modes of variability in the SST data, and to use the expansion coefficients – scores or timeseries – of these principal components as the predictor indices. In this way, SST regions can be defined objectively, and it is ensured that only the most important modes are considered, reducing the possibility of including too many noisy predictors. However, the question of assessing spatial degrees of freedom applies also to the predictand data, and predictor-predictand relationships must therefore be assessed with this in mind. Typically, a field significance test based on Monte-Carlo resampling (Livezey and Chen, 1983) is used. The application of this method is discussed in more detail in Chapter 6.

Thirdly, a multivariate technique such as canonical correlation analysis (CCA) (e.g. Barnett and Preisendorfer, 1987), or maximum covariance analysis (MCA) (e.g. Rodwell and Folland, 2003), can be used. These methods identify dominant patterns in both fields that are correlated in time. In the case of CCA, it is necessary first to reduce the dimensionality of the two datasets by some method such as PCA. Haylock and Goodess (2004) show that useful links can be identified between extremes of precipitation and features of the large-scale circulation, including for example the NAO. The use of CCA in this case is dependent upon an initial PCA decomposition, where a small number of precipitation 'modes' are retained for the CCA, describing approximately between 40% and 50% of the variability, depending on the precipitation index under consideration. This method is useful in that it both expresses coherent variability in the predictand, and removes likely noise from the predictand data. However, some considerations, such as the non-Gaussian nature of the predictand data may affect the validity of the results, although this is thought to be of limited importance (Haylock, personal communication). In addition, the initial PCA decomposition of the predictor is similar to the second approach outlined above, except that in the second approach, the predictor PCs are applied directly to the predictand data, with some additional constraint such as the field-significance test. The second approach is adopted here, for the reasons outlined above, and additionally since there is a large exploratory component to the work, it is desirable to simplify the interpretation of results as far as possible – that is to say, the combination of PC modes in the CCA method is less easy to interpret than considering the PC modes separately.

Since the spatial correlation scales in monthly SST data are large, the HadISST 1.1 data is regridded onto a regular 5° by 5° grid, using a triangle-based linear interpolation, and four different spatial domains are defined: a 'global' basin, with latitude limits 60°S to 70°N, the Atlantic basin, the Pacific basin including the Maritime Continent, and the Indian Ocean basin. Spatial domains of the latter three basins are given in Figure 5.1, with the global domain comprising all three of these sub-domains, plus the Mediterranean basin. In all cases, areas with significant periods of seasonal sea-ice coverage are excluded from the spatial domain. The annual cycle is removed from the data, which are then reweighted by the square-root of the cosine of the latitude, in order to account for the decrease in the size of gridboxes towards the poles (e.g. Lin and Derome, 2003). PCA is performed on twelve three-month overlapping seasons (with the data kept at monthly resolution) for each of the four spatial domains. Following suggestions by von Storch (Chapter 13, von Storch and Navarra, 1999), the correlation matrix is used. Based on an informal assessment of the fraction of variability accounted for by the leading modes, the first five PCs are retained from each analysis. Typically, the first five PCs account for 50-60% of the total variance in the individual basins, and 42-45% over the global domain, depending on the season. The use of twelve three-month seasons is motivated partly by the exploratory nature of the work. Interactions between SST and the overlying atmosphere are sensitive to the phase of the annual cycle, and it is thought that restriction to the more traditional four-season annual cycle may not optimise the potential for identifying predictable relationships. Recently, Okumura and Xie (2006) find that a flexible approach to defining seasons allows new SST patterns and forcings in the Tropical Atlantic to be identified. This prompts the use of twelve seasons, rather than the more traditional four. While PCA carried out on a single month may provide sharper temporal resolution, it is reasonable to assume that the phase-locking to the annual cycle of ocean-atmosphere interaction is not sufficiently rigid to ensure that a potential forcing by an SST mode takes place within the confines of one month. Therefore three month seasons are preferred.

The data are not detrended before carrying out the PCA, but the PC time series are detrended and low-pass filtered using a 20 year window to remove variability at timescales incompatible with the period of observation considered here. It is thought that by preserving modes which may be associated with trends, the potential impact of the interannual variability of that trend may be assessed.

Given four basins on which PCA is carried out, and five retained PC time series, a total of 20 timeseries are derived. Each timeseries is considered at 10 different lead times, so that the JFM predictand season uses predictors from OND (shortest lead time), to the preceding JFM (longest lead time). In total, this gives 200 predictor timeseries for large-scale SSTs.

Figure 5.2 shows the spatial loadings of PCs one and two of DJF and JJA SSTs over the Atlantic domain. The dominant mode has a long-term warming trend, while the second mode is closely associated with the Atlantic tripole pattern (e.g. Seager *et al*., 2000). The tripole pattern is of particular interest since it represents an integration of the mean forcing by the NAO over seasonal timescales, and is thought to feed back into NAO variability, implying predictability at seasonal timescales (e.g. Rodwell and Folland, 2002).

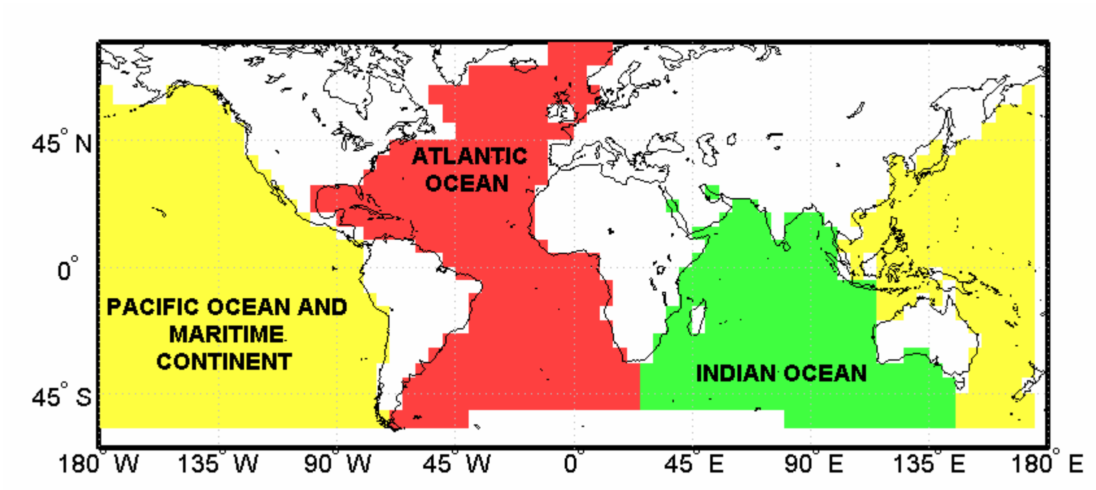

Figure 5.1 Spatial domains of SST Principal Component Analysis (PCA)

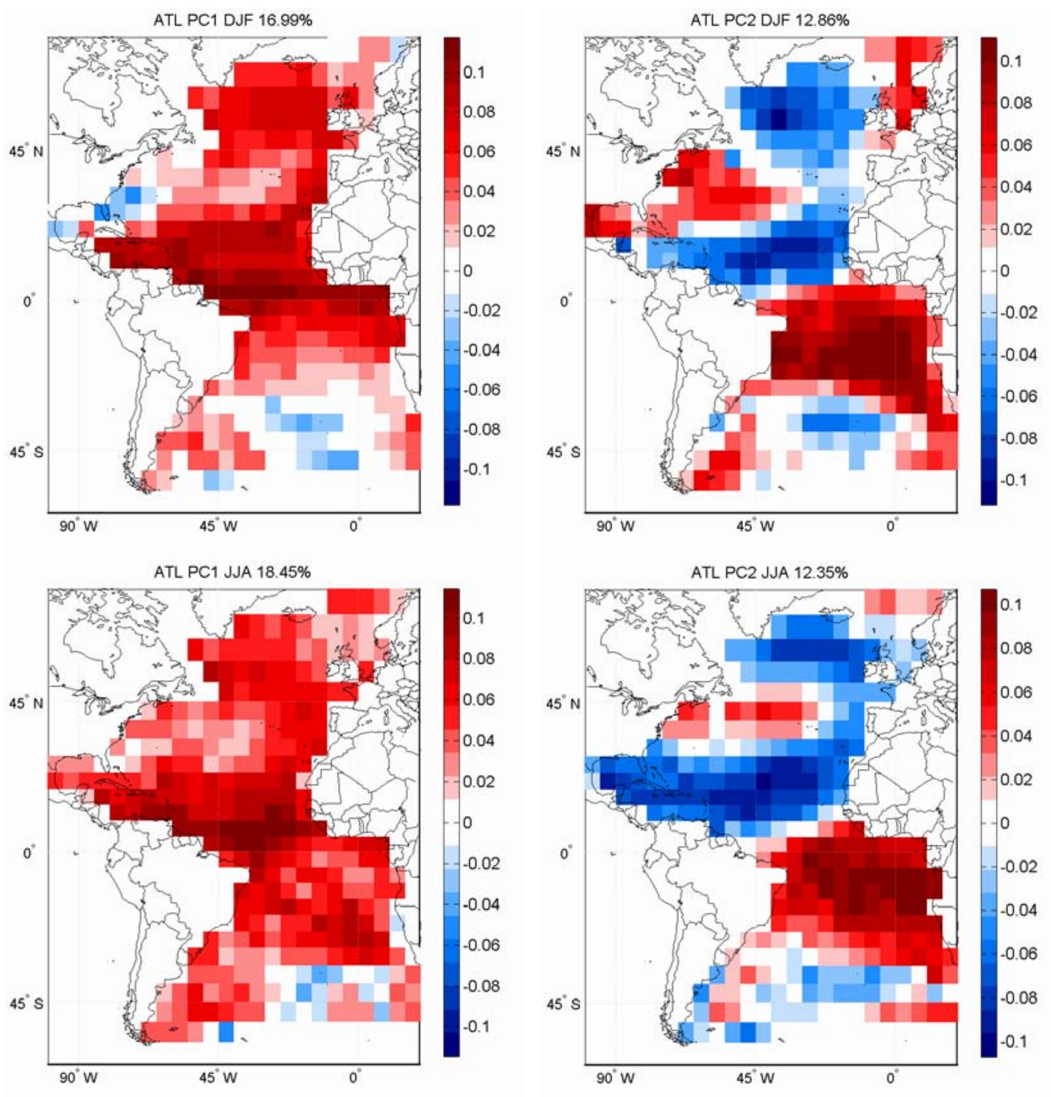

Figure 5.2 PC 1 and 2 of Atlantic SST. DJF (top) and JJA (bottom). PC1 shows a warming trend, and PC2 is highly correlated with the North Atlantic tripole pattern (e.g. Seager *et al.*, 2000).

5.2.2 Stratosphere

5.2.2.1 Background to Stratospheric Predictors

The use of information from the stratosphere to provide potential predictors for seasonal forecasting is a relatively new and untested field. Robock (2001) suggests that improvements in representations of the stratosphere in numerical models are needed to improve dynamical seasonal prediction, and Baldwin (2003) identifies potential predictability at intraseasonal scales in the Northern Hemisphere extratropics. The use of information from the stratosphere is contingent on persistent or predictable anomalies, and interactions between the stratosphere and the troposphere. Recent work has focussed on both, although a comprehensive assessment of the nature and strength of stratospheric coupling with the troposphere is far from complete.

Baldwin and Dunkerton (2001) examine anomalous events in the Northern Hemisphere winter stratospheric circulation, finding observational evidence that these anomalies are often followed by anomalous conditions in the troposphere at timescales of up to 60 days. Additionally, it is found that the strength of the polar vortex is related statistically to the sign of the QBO, implying longer-range predictability. Mechanisms to explain these relationships are not fully developed, although Baldwin *et al.* (2003) succeed in making skilful empirical forecasts of the Northern Annular Mode, defined as the leading mode of extratropical Northern Hemisphere geopotential height variability, at up to a month in advance. Charlton *et al.* (2003) obtain similar results. Subsequent to this work, there has been considerable focus on stratosphere-troposphere coupling in the literature, including Itoh and Harada (2004), who show that the leading tropospheric modes in the Northern Hemisphere (identified here as the Pacific-North American pattern, or PNA, and the NAO) couple with the stratosphere, being significantly correlated over time. Scott and Dritschel (2005) use a simplified model to show that potential vorticity waves tend to propagate downwards through the stratosphere. This has direct implications for the temporal evolution of the polar vortex, and supports the empirical findings of Baldwin and Dunkerton (2001) *inter alia*.

In addition to work on the persistence and tropospheric impact of stratospheric anomalies, it has for some time been well known that upward-propagating anomalies from the troposphere can impact the stratosphere. This is of particular interest in the context of increasing evidence of coupling across the tropopause, since potentially predictable surface phenomena may have an effect on the stratosphere. Although Hamilton (1993) finds that there is no apparent correlation between ENSO and raw measurements of the stratospheric circulation, although composite responses to ENSO warm events are discernable. The problem of separating ENSO and QBO signals in the stratosphere is discussed. Kodera *et al.* (1996) isolate the major stratospheric modes using PCA, and find a strong statistical link with ENSO and the second stratospheric PC. Recently, Brönnimann *et al.* (2004) finds observational evidence from the strong El Niño of 1940-42 for significant impacts on Northern Hemisphere stratospheric heights, and surface temperatures in Europe. Moreover, these results compare strikingly well with model simulations of strong El Niño events. More generally, Calvo Fernandez *et al*. (2006) find a small but significant impact of ENSO on the tropical stratosphere. Scaife *et al*. (2005) use a GCM to identify a possible direct role of the stratosphere in forcing not only NAO variability in the short term, but the longer term positive trend that has been observed since the 1970s. In summary the role of the stratosphere in providing seasonal predictors is potentially large, although much of the theory remains to be clarified.

5.2.2.2 Stratospheric Data and Predictor Indices

Three stratospheric variables are considered, taken from the ERA-40 reanalysis and ECMWF operational analysis: geopotential height, temperature, and potential vorticity. Geopotential height is widely used, in particular to express the Northern Annular Mode (NAM). Haigh *et al.* (2005) examine the effects of stratospheric heating on the troposphere, and find discernable impacts in the meridional extent of the Hadley cells, and of the midlatitude circulation. Hartley *et al.* (1998) find that disturbances to the Northern Hemisphere stratospheric polar vortex result in a redistribution of stratospheric potential vorticity, which results in perturbations in the upper troposphere. A substantial volume of further work (e.g. Ambaum and Hoskins, 2002; Black and McDaniel, 2004; McDaniel and Black, 2005) has explored this relationship further. It is thought that the use of potential vorticity to assess the impact of changes in the stratospheric circulation on the troposphere may be helpful. It should be noted, however, that stratospheric data are of limited reliability in the pre-satellite era (pre-1979). For example Haigh *et al.* (2005) use only post-1979 NCEP data in their analysis. Despite these limitations, it is thought that the omission of the stratosphere from the set of potential predictors on this basis is not justified.

As in the SST predictor set, PCA is used to derive predictor indices from this data. This is a commonly used technique with stratospheric data, for example Baldwin *et al.* (2003) use PCA to define the Arctic Oscillation (AO), and Kodera *et al.* (1996) study the leading modes of Northern Hemisphere stratospheric variability as defined by PCA. The first three principal components of geopotential height, temperature, and potential vorticity are considered, at 150hPa, 100hPa, 50hPa and 30hPa – i.e. approximately from the tropopause into the lower stratosphere. Four different heights are considered since it is known that anomalies in the stratosphere tend to propagate both upwards and downwards over weekly to monthly timescales (e.g. Scott and Dritschel, 2005; Breiteig 2008), and it is thought that the consideration of different heights might allow the inclusion of potentially useful predictors at longer lead times. In particular, Limpasuvan *et al* (2005) find a clear downwards-propagating signal in stratospheric wind and temperature during sudden stratospheric warming (SSW) episodes. Although variability in the extratropical stratosphere is typically on shorter timescales than that of the major SST modes, the use of three month seasons was maintained in order to highlight the more slowly varying modes in the stratosphere. The data were not detrended for the PCA, and in many cases the first component comprises a trend, likely related to the observed cooling in the stratosphere (e.g. Ramaswamy *et al*., 2001). The next most important mode (in some cases this is the most important mode) is highly correlated with timeseries of the AO and the NAM, while the third mode frequently exhibits high correlations with ENSO, comparable to the second mode in Kodera *et al.* (1996). An example provided in Figure 5.3 shows the spatial patterns of the first two leading modes of geopotential height at

150hPa for DJF and JJA. In both cases the leading mode represents a trend, and the NAM pattern is apparent as the second mode in the DJF analysis. The NAM is not apparent in the second JJA pattern, which instead is highly correlated ($r=-0.63$, $p=2\times10^{-6}$, with July Niño 3.4) with ENSO. The choice of only the first three modes of the stratospheric variables was made because typically the variance accounted for by the third mode is already small (of the order of 5%), and little evidence exists to support the physical validity of further modes at seasonal timescales. Given the much lower mass of the stratosphere compared to the troposphere, coupling (i.e. forcing of the troposphere by the stratosphere) is not likely to be accomplished by any but the most coherent and persistent modes. Additionally, given three variables and four levels at which analysis is carried out, the total number of predictors from the stratosphere is 360, given all ten lead times, and it is thought best not to increase this still further. As with the SST predictors, all timeseries are detrended and low-pass filtered using a 20 year window to attempt to isolate interannual variability.

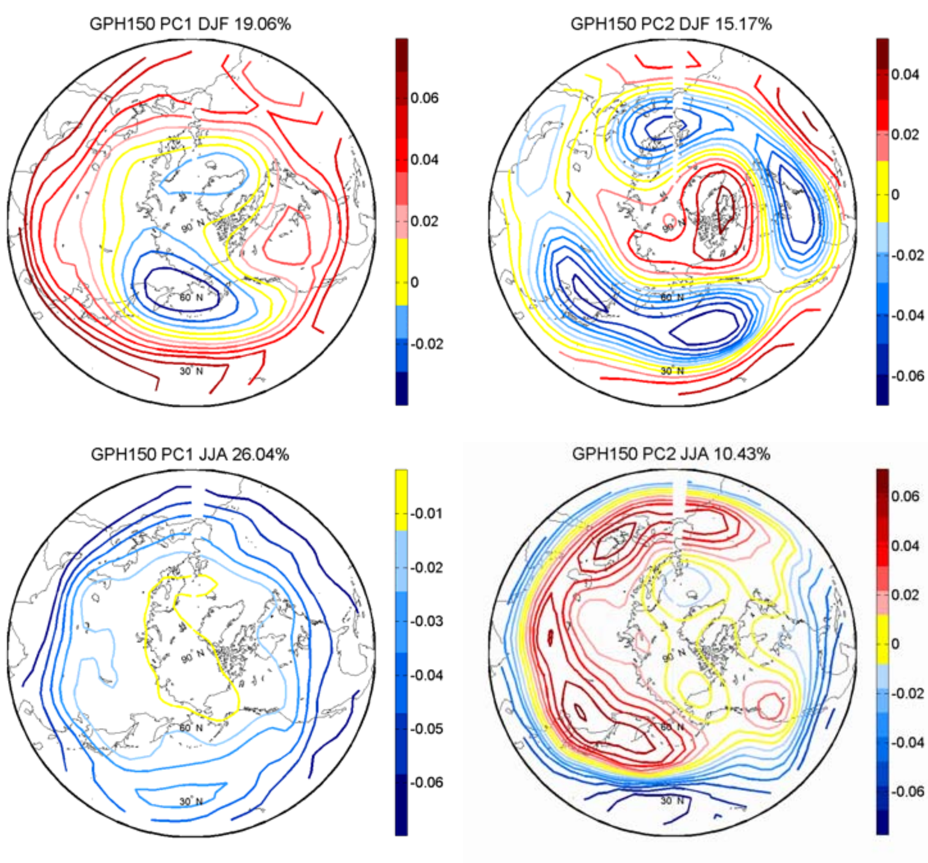

Figure 5.3 PC 1 and 2 of DJF and JJA geopotential height at 150hPa.

5.2.3 Teleconnection Indices

5.2.3.1 Background to Teleconnection Indices

A further series of potential predictor indices representing large-scale variability of the ocean, troposphere and stratosphere are included. They are all thought to describe significant physical processes, and as such are superficially of interest either for their persistence or for their coherent forcing of other boundary processes, such as the NAO influence on the North Atlantic Ocean (e.g. Cayan 1992a, b). In many cases, studies have linked the processes represented by these timeseries with European climate at lead-times suitable for seasonal forecasting (e.g. Fraedrich 1994, Mariotti *et al.* 2005). As well as the summary in Table 5.1, a more detailed description of each index is given in section 5.2.3.2 below.

5.2.3.2 Teleconnection Data and Predictor Indices

Table 5.1 lists the teleconnection indices considered for potential predictors. The Arctic Oscillation (AO) is defined by Thompson and Wallace (1998) as the leading empirical orthogonal function (or principal component) of sea level pressure (SLP) in the Northern Hemisphere extratropics. It is the surface manifestation of the Northern Annular Mode (NAM), which is present throughout the troposphere, and, particularly during the winter months, in the lower stratosphere also. The AO is highly correlated with the North Atlantic Oscillation (NAO), and there is an ongoing discussion in the literature as to whether the two indices describe the same process (e.g. Wallace 2000), or whether they are physically distinct (e.g. Ambaum *et al.* 2001). Cohen *et al*. (2002) identify differences in the seasonal evolution of the SLP anomalies that comprise the major mode of variability in the NH winter. Essentially, two separate categories are described, one which is similar to the NAO in its regional manifestation, and one which is more similar to the AO. Furthermore, it is suggested that the recognition of these separate pathways may offer higher predictability at the seasonal scale than was previously thought possible.

More recently, Feldstein and Franzke (2006) find that composites of SLP and upper atmosphere streamfunction based on NAO and NAM indices are indistinguishable, and cannot be said to be confined to the North Atlantic region, or to be hemispheric modes. The NAO may be interpreted as a local manifestation of the NAM, describing the exchange of atmospheric mass between centres located over the North Atlantic at about 35°N and 65°N, and was first observed by Walker (1924). The NAO index used here is derived from the leading rotated mode of 500hPa height in the Northern Hemisphere north of 20°N. The AO index is based on the same data, but is derived using unrotated PCA. Both datasets are provided by the National Oceanographic and Atmospheric Administration (NOAA) Climate Prediction Centre (CPC). Data can be found at http://www.cpc.noaa.gov/products/precip/CWlink/pna/nao.shtml.

Barnston and Livezey (1987) used rotated principal component analysis (RPCA) to summarize the leading modes of monthly 700hPa heights in the Northern Hemisphere north of 20°N. As well as the NAO, they identified a number of modes which appear to be physically significant. These modes are included as potential predictors. The indices are also provided by the CPC, and full details of their derivation and description can be found in Barnston and Livezey (1987) and on the CPC website (www.cpc.noaa.gov). Numerous studies have identified relationships between these teleconnection patterns and the European climate (e.g. Quadrelli and Wallace 2004; Zvaryaev 2004; Kingston *et al.* 2006;) and although the physical validity of some of the modes has been called into question, they remain useful descriptors of Northern Hemisphere climate variability.

The indices of the Southern Oscillation Index (SOI), and of SLP at Darwin and Tahiti (the two stations from which the SOI is derived, e.g. Ropelewski and Jones, 1987) are included, as are the commonly used regional SST indices of El Niño-Southern Oscillation (ENSO) activity, Niño 1+2, Niño 3, Niño 3.4 and Niño 4, all of which are located in the tropical Pacific, and are designed optimally to describe ENSO evolution at different phases.

Predictor	Description	Dataset	Data provision
AO	Arctic Oscillation		CPC
NAO	North Atlantic Oscillation		CPC
EA	East Atlantic Pattern		CPC
EAWR	East Atlantic/West Russian Pattern		CPC
EPNP	East Pacific/North Pacific Pattern		CPC
PNA	Pacific-North American Pattern		CPC
POL	Polar-Eurasian Pattern		CPC
PT	Pacific Transition Pattern		CPC
SCA	Scandinavia Pattern		CPC
TNH	Tropical-Northern Hemisphere Pattern		CPC
WP	West Pacific Pattern		CPC
SOI	Southern Oscillation Index		CPC
Darwin SLP	Sea Level Pressure at Darwin		CPC
Tahiti SLP	Sea Level pressure at Tahiti		CPC
Nino 1+2	SST anomalies: 0°N-10°S, 90°W-80°W	HadISST 1.1	CPC
Nino 3	SST anomalies: 5°N-5°S, 150°W-90°W	HadISST 1.1	CPC
Nino 3.4	SST anomalies: 5°N-5°S, 170°W-120°W	HadISST 1.1	CPC
Nino 4	SST anomalies: 5°N-5°S, 160°W-150°W	HadISST 1.1	CPC
NH temperature	Northern Hemisphere temperature anomaly	HadCRUT3	CRU
SH temperature	Southern Hemisphere temperature anomaly	HadCRUT3	CRU
	Difference between Hemispheric temperature anomalies	HadCRUT3	CRU
QBO50	Quasi-Biennial Oscillation at 50hPa	1958-1978: CPC Singapore 1979-2005 CPC QBO	CPC
QBO30	Quasi-Biennial Oscillation at 30hPa	1958-1963 FBU 1964-1978: CPC Singapore 1979-2005: CPC QBO	FBU; CPC

Table 5.1 Teleconnection indices, giving the name of the predictor, and describing the nature and source of the data from which the predictor is derived.

Indices of Northern and Southern Hemisphere mean monthly temperature are from the Climatic Research Unit HadCRUT3 dataset (Brohan *et al.*, 2006) which can be obtained at (http://www.cru.uea.ac.uk/cru/data/temperature/).

The quasi-biennial oscillation (QBO) describes an oscillation in the direction of the equatorial stratospheric zonal winds, with a period of approximately 28 months (Baldwin *et al*., 2001). It is thought to play a major role in the variability of the extratropical stratosphere, being linked to the polar vortex and hence, it is thought, surface weather patterns (e.g. Baldwin *et al.*, 2003). Baldwin *et al*, (2001) review research on this phenomenon, identifying its effect on extratropical interannual variability as a crucial area. Indeed further research has been carried out, for example Hampson and Haynes (2006) identify an interaction between the QBO and the extratropical stratosphere in a simplified model, where the phase alignment of the QBO with the annual cycle dictates the manner in which waves propagate into the extratropics.

Timeseries of the QBO used here are taken from measurements at 50hPa (QBO50) and 30hPa (QBO30). The data are of combined indices of radiosonde zonal winds measured at Singapore and provided by the Freie Universitat Berlin (http://strat-www.met.fu-berlin.de/) and the CPC. Since continuous timeseries were not available from one source for the whole period required, data from different sources were combined. Where the data overlap in time, the indices are virtually identical. Table 5.1 provides details of the data sources.

Given 22 timeseries of major atmospheric and oceanic modes, a total of 220 potential predictors are obtained when all lead times are taken into account. The predictors are detrended and filtered as described in 5.2.1.1 above.

5.2.4 Snow cover

5.2.4.1 Snow Cover Background

Northern Hemisphere snow cover has been of interest as a possible predictor of seasonal climate variability for some time. Cohen and Rind (1991) investigated the dynamical effects of snow cover forcing on the overlying atmosphere, and more recent work by Bojariu and Gimeno (2003); Qian and Saunders (2003) and Saunders *et al.* (2003) identifies Eurasian snow cover in particular as a possible predictor for the NAO at timescales of up to several months in advance. Gong *et al.* (2002) identify a modelled AO/NAO response to snow cover anomalies that is consistent with the observed, and Kumar and Yang (2003) identify a marked increase in Northern Hemisphere atmospheric variability in a GCM that is forced with variable snow cover and SSTs. They attribute a significant role in the enhanced variability of the lower troposphere to snow cover. In addition to this work, Gong *et al.* (2004) find that snow extent, depth and albedo all act to some extent, to modify the characteristics of the overlying atmospheric circulation in a GCM. However the physical basis of this relationship is neither straightforward, nor necessarily linear. Saito *et al.* (2004) show that there is a phase change in the link between Eurasian snow cover and the AO in the 1980s, and identify a link between North American snow cover and the NAO. On the basis of considerable statistical evidence, with some support from dynamical experiments, Northern Hemisphere snow cover will be considered as a potential predictor.

5.2.4.2 Snow Cover Data and Predictor Indices

Two datasets were under consideration for deriving predictor indices of snow cover. The observed gridded monthly dataset provided by the Rutgers University Climate Lab (RUCL) (Frei and Robinson 1999; Robinson *et al.* 1993) is one of the most comprehensive products available. However, it only dates back to October 1966, and therefore removes a significant time-slice from the available analysis period, if all potential predictors are to be considered equally. Additionally, there are missing records in the earlier years of the dataset. The ERA-40/ECMWF Operational analysis also provides a snow cover product, covering the period required (1958-2005). Ultimately, for the purposes of developing a prediction system trained on as long a period as possible, it would be more desirable to use the ERA-40 dataset. The two datasets were compared to assess suitability, although it should be noted that there are potential problems with both. Robinson *et al* (1993) note that prior to 1972 when the new AVHRR instrument was introduced, the snow extent – particularly during the autumn – was systematically underestimated, while the ERA-40 data is subject to similar constraints due to the satellite record, and also to the limitations of the precipitation forecast model – which includes snowfall. Martin (2004) finds that for the French Alps region, ERA-40 reproduces the observed snow cover reasonably well, despite significant shortcomings in the snowfall model.

Indices of total continental scale snow cover were obtained from the RUCL, and derived from the gridded ERA-40 snow depth data. The use of total continental or hemispheric snow coverage, as opposed to regional indices or a multivariate index such as a PCA timeseries is adopted here, following the method of Saunders *et al.* (2003). While this may not be the most sensitive index to identify potentially predictable lagged relationships, it is the most concise summary of snow cover, and alternative methods may discount important anomalies which the continental indices would on average be expected to include. Four indices are compared from each dataset, covering Eurasia, North America, North America plus Greenland, and the Northern Hemisphere as a whole. The RUCL indices are obtained from the RUCL website (http://climate.rutgers.edu/snowcover/index.php), and the ERA-40 indices are derived from the N80 gridded snow depth data available from the British Atmospheric Data Centre (BADC, badc.nerc.ac.uk).

There are substantial differences between the datasets, with the ERA-40 coverage biased to larger areas than the RUCL data. This may be due to the different manner in which snow cover is represented in each gridbox. Whereas the RUCL data gives percentage coverage for each gridbox, the ERA-40 snow depth data is binary, giving total or no coverage for each gridbox. In addition, it is recognised that in the early years of the RUCL dataset, snow cover is likely underestimated, as noted above. Figure 5.4 shows comparisons of the datasets on a monthly basis for Eurasia, over the period 1967-2001. The correlations vary as a function of the annual cycle, increasing from January throughout the spring, and disappearing in the summer. They then increase from September, and decay towards the end of the year. In summer, snow cover is much less extensive, and those areas that are snow covered are more likely to be sparse, with respect to percentage gridbox coverage. This is recognised in the RUCL data, but is interpreted as thin but total coverage in the ERA-40 data, possibly leading to the greater percentage bias (not shown) in the summer months. Figure 5.5 shows the same data but for North American snow cover. Correlations are similar throughout the annual cycle, although they are better for December in the North American data. It was decided to use the ERA-40 data based on the requirement for predictors to cover the longest possible period compatible with the rest of the predictor datasets. A total of four indices were derived, covering the major Northern Hemisphere land masses, and giving a total of 40 predictors at all lead times. The data were detrended and filtered as in 5.2.1.1 above.

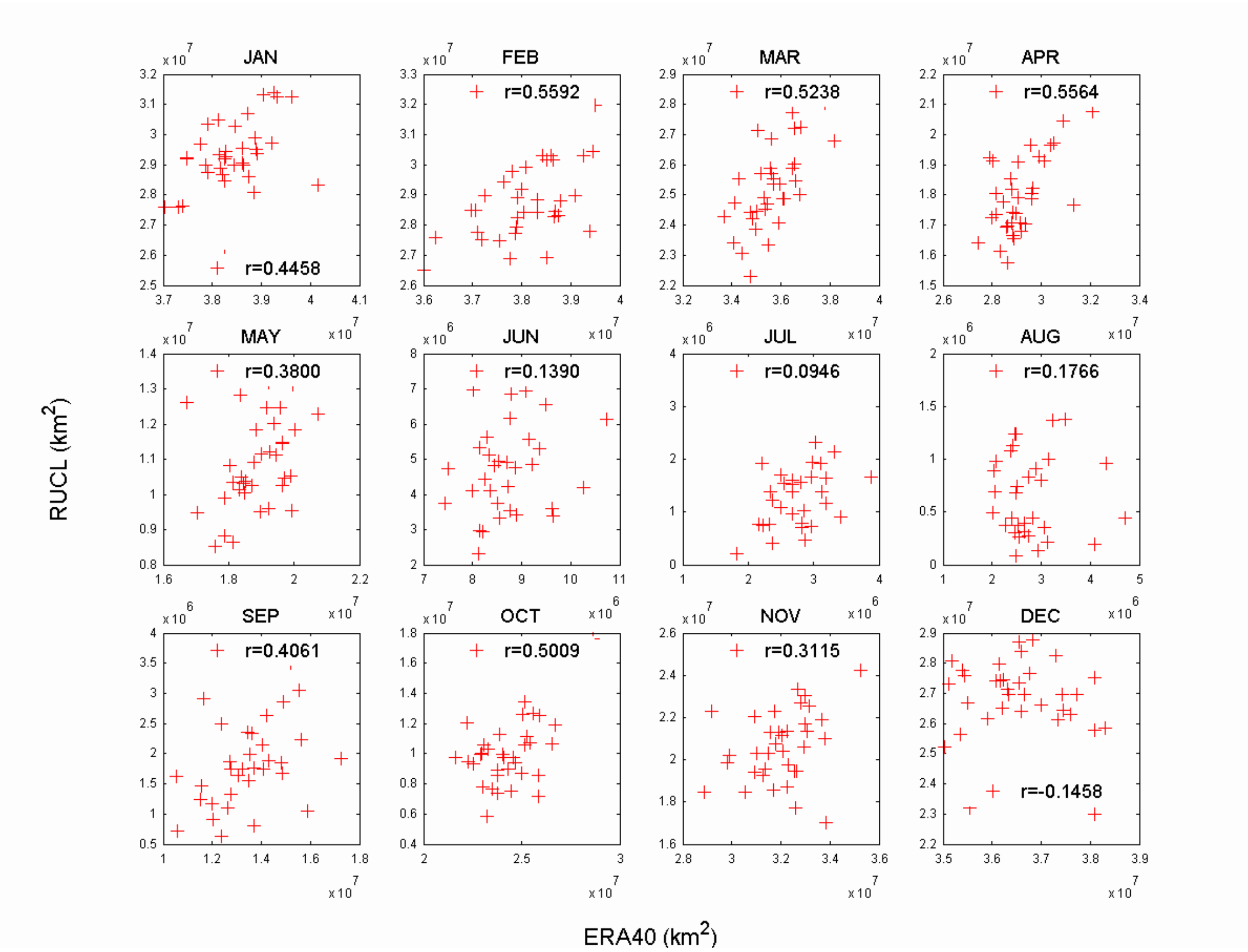

Figure 5.4 Comparison of Eurasian snow-covered area as estimated by the RUCL and ERA-40 datasets. Comparisons are given on a monthly basis, and cover the period 1967-2001. Correlations between the timeseries are shown on the plots for each month

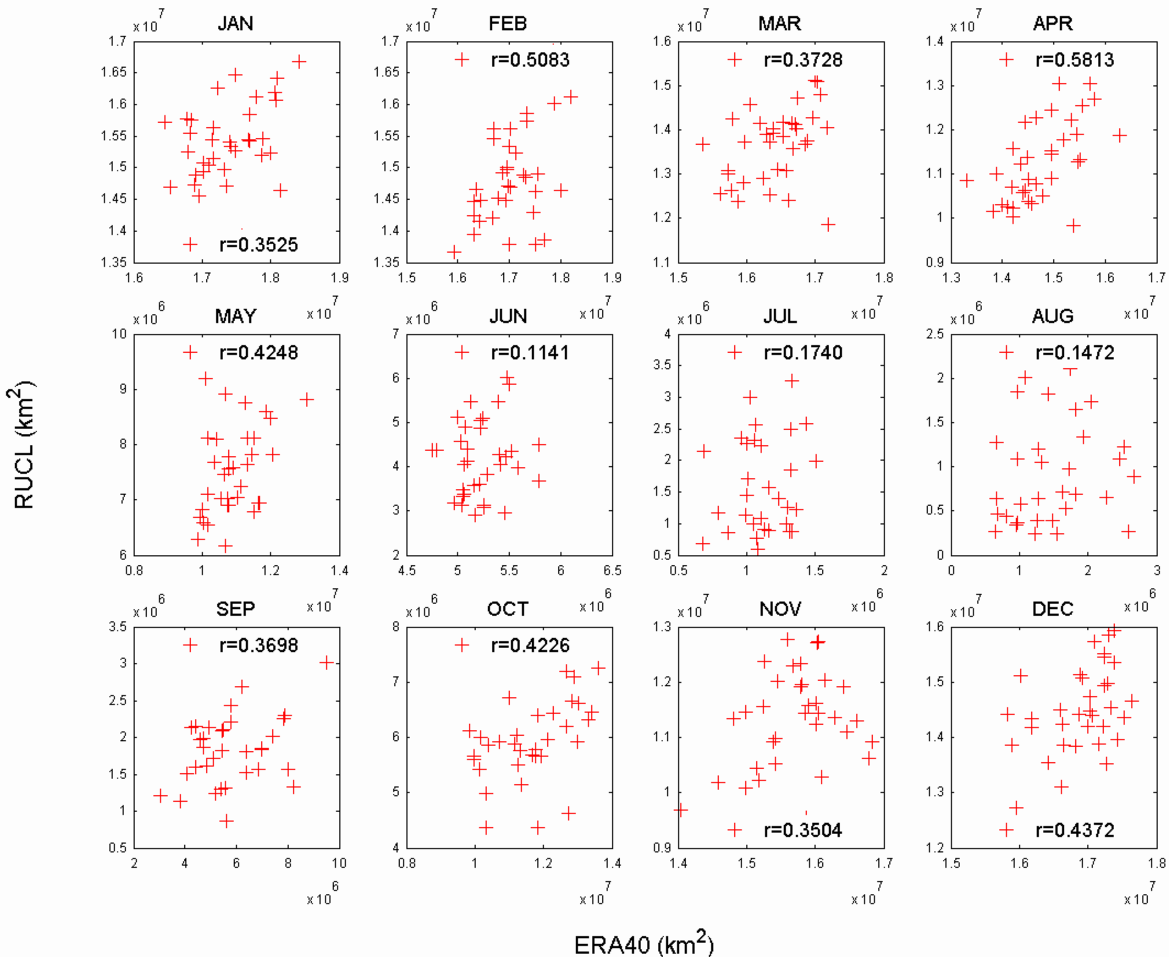

Figure 5.5 Comparison of North American snow-covered area as estimated by the RUCL and ERA-40 datasets. Comparisons are given on a monthly basis, and cover the period 1967-2001. Correlations between the timeseries are shown on the plots for each month

5.2.5 Local SST

5.2.5.1 Background to Local SST Predictors

Recent work has identified useful predictability from local SSTs – that is, SSTs from coastal regions, covering areas orders of magnitude smaller than the large scale PC patterns discussed in 5.2.1 above. Regions as diverse as New Zealand (Zheng and Frederiksen, 2006) and the Western Mediterranean (Lepeaubin *et al*., 2006) have been studied in this context. Hurrell *et al*, (2006) in a review paper on Atlantic climate variability and predictability identify coastal SSTs as warranting further research into their possible role as forcing agents of the local climate, particularly outside the winter season. Zheng and Frederiksen (2006) find evidence for useful predictability of New Zealand summer rainfall based on local SSTs at lead times of several months. Lepeaubin *et al*. (2006) find that anomalously warm SSTs in the Mediterranean are associated with extreme rainfall events in the autumn, when the sea is still warm. The relationships are effectively simultaneous, but given the persistence of even local SST anomalies, predictability may be gained at lead times of a few months based on persistence alone.

5.2.5.2 Local SST Data and Predictor Indices

The same SST dataset (HadISST 1.1) was used to derive local SST predictors as in 5.2.1.1 above. Six coastal regions were selected, based on gridboxes with high spatial correlations. Region 1 (SST1) is derived from SSTs off the north west coast of Scotland. Region 2 (SST2) is the North Sea. Region 3 (SST3) comprises the Baltic Sea, and region 4 (SST4) is the coastal Atlantic from the west coast of Ireland down to the bay of Biscay. Region 5 (SST5) is the Adriatic, and region 6 (SST6) is the western Mediterranean. Because local SSTs are unlikely to force large scale ocean or atmosphere dynamics, any predictability is more likely to result from persistence alone. For this reason, only lead times of up to 5 months are considered, based on an assessment of the e-folding length of the autocorrelation of the SST timeseries. A total of 30 predictors were therefore available, detrended and filtered as in 5.2.1.1.

5.2.6 Solar flux

5.2.6.1 Solar Flux Background and Data

The influence of the solar cycle on climate has been recognised for many years, however to date the specific mechanisms by which it may influence tropospheric dynamics are not fully understood. For example Haigh *et al*. (2005) show that the solar cycle affects the subtropical jets in the stratosphere, whereby an increase in solar output is associated with a poleward shift of the jets. Corresponding to this is a poleward shift of the tropical tropospheric Hadley cells, and the tropospheric midlatitude circulation. Given the long timescale of solar output variability, it may be the case that solar output provides a source of predictability. De la Torre *et al*. (2006) identify a direct link between solar output and NAM variability, with possible implications for tropospheric predictability, at least at the shorter timescales identified by Baldwin and Dunkerton (2001). Typically, the 10.7cm radio flux is used to represent solar output. The data are provided by the National Geophysical Data Centre (http://www.ngdc.noaa.gov/stp/SOLAR/getdata.html).

5.3 Summary

Kushnir *et al*. (2004) discuss the stages by which our understanding of climate predictability develops. Stage one consists of the observation of statistically significant relationships, with stage two providing physical explanations for the observed relationships. Stage three is reached when the observed phenomena are successfully modelled, and the understanding of the dynamics of the relationship can be enhanced, not least as a result of the much larger sample size available. This Chapter presents evidence in the literature for predictability of aspects of the European climate, in particular wind and precipitation. With respect to European seasonal forecasting, much of our understanding is still at the first stage – that is, observational evidence exists, but the nature of the physical mechanisms driving this potential predictability are not yet fully understood. It is felt that this justifies the initial consideration of a large number of potential predictors, given subsequent filtering of these predictors to a more manageable number for the purposes of model development.

In summary: potential predictors are proposed, on the basis that there is empirical or theoretical evidence in the literature to support their inclusion, and each potential predictor is then tested using a field-significance criteria to determine whether it should be included in the next stage of model development. In this way, a wide range of potential predictors can be assessed statistically. As proposed above, there are approximately 850 predictors in total available for the initial predictor selection phase described in Chapter 6. Table 6.1 gives a complete list of the predictors and describes the notation used to refer to them in the rest of this thesis.

It is emphasised that the inclusion of predictors on purely statistical grounds is not a sufficient requirement for a reliable forecasting model. In particular, when a large number of potential predictors are assessed, some will of course pass the significance test purely by chance, which by definition allows a percentile of the most significant results through. For this reason, the model should be viewed as a largely exploratory attempt to improve understanding of potential predictability. Further statistical testing during the model development is detailed in subsequent Chapters.

6. Initial Predictor Selection

This Chapter describes the initial predictor selection problem, and presents possible approaches to solving it, including a discussion of the obstacles to each approach, and the implications of these obstacles for the final prediction models. Section 6.1 comprises the introduction. Section 6.2 introduces the predictor selection problem in model development, with particular reference to seasonal forecasting. Section 6.3 introduces possible approaches to the initial predictor selection problem, and provides detail on the methodology chosen. Results of the intermediate predictor selection stage are then presented in section 6.4, followed by a discussion of the implications for the next stage of model development in section 6.5.

6.1. Introduction

Chapter 5 introduced the indices which describe the major features of climate variability which are, or may be pertinent to the north-west European region, and may offer some predictive skill at seasonal timescales. The process by which potential predictor indices were derived from these features was described. This set of potential predictors will be referred to as the full set. Owing to the large number of predictors in the full set (several times greater than the number of observations under consideration), and the speculative grounds on which they have been considered, the development of a prediction model using a subset selection algorithm directly from the full set is not possible, as it will invariably result in over-fitting. (Burnham and Anderson, 2002). It is therefore necessary to use some intermediate process to reduce the full set to a more manageable set, referred to here as the reduced predictor set. This reduced set can then be used to develop a model, with the final predictors comprising the most skilful subset of the reduced set at each location. The latter method will be described in detail in Chapter 7.

6.2. The predictor selection problem in model development

Predictor selection in the development of regression models, as the name implies, is the process by which a set of predictors is chosen from a larger set, in the interest of describing the maximum amount of 'structure' in the predictand (that is, a signal in the predictand, which is thought to relate to a real physical relationship with the predictor), and in such a way that the resulting model is as parsimonious as possible (that is, as few as possible predictors are used to explain as much variance as possible). Typically, the predictor selection process should constitute a way of fine tuning the input predictors, and should be based on an initial set which are all thought to be causally linked with the predictand. In other words there should be a recognised structure, or signal in the data, before this process commences. This structure should ideally be based on theoretical knowledge of the system under consideration. There are numerous approaches to obtaining a reduced or final predictor set. These will be addressed in section 6.3.1. Firstly, it is necessary to undertake a more fundamental discussion of the rationale behind a statistical approach to this problem.

As discussed in Chapter 5, there is currently little theoretical evidence for skilful seasonal predictors of European climate, in particular for wind and precipitation. The development of a statistical seasonal forecasting model must therefore be exploratory in design, and in application. The nature of the exploratory process by which such a model is defined is highly important, and in the broadest sense can follow one of two approaches – dynamical or statistical. The first option – numerical modelling of the global climate – would take the form of long or ensemble simulations, where the response of the predictands defined in Chapter 4 to each potential forcing agent can be evaluated, and steps towards a mechanistic explanation of any observed relationship can be made. In theory this is the ideal approach, in that the forcing influence of individual factors can be isolated and examined in great detail, for a sample size much larger than that available in the observed record. In fact many such studies have been carried out, and are reviewed in Chapter 3. Principally, the effects of the Atlantic Ocean, and the tropical Pacific are investigated, although some work exists relating stratospheric processed to the European climate (e.g. Baldwin *et al*, 2003). With regard to predictability of the European climate, such studies form the backbone of our current understanding, and support the prevailing view that predictability at seasonal scales is low. This view is also borne out by observational studies, and operational forecasting systems. However, it is not apparent from the literature that a practically exhaustive body of results and theory exists in this area. Furthermore, where experiments to replicate observed relationships have been carried out, it is frequently apparent that there are shortcomings in the model representation of observed climatic processes and relationships. We may also therefore assume that relationships which have yet to be identified in the observed record – which we might principally suspect to be through lack of data – are not guaranteed to be represented correctly in the model. Given that this is the case, further exploratory work using numerical methods must either look for new predictors, or use new models, which better represent climate variability at seasonal scales. In other words, model experiments can be used to further reduce our epistemic uncertainty of the system. However, any model-based attempt at reducing this uncertainty must go hand in hand with an empirical verification of the models. In fact in the atmospheric sciences this relationship between theoretical and empirical understanding may be viewed as an iterative process. The following methods and results are an attempt to contribute to our understanding of climatic relationships from an empirical perspective – the second path suggested above. There are two reasons for the choice of this approach in this case. Firstly, from a practical point of view, the logistical constraints such as access to numerical models, and the expertise required to develop a stand-alone seasonal forecasting system from scratch is beyond the scope of individual PhD theses, and also, a considerable and valuable amount of work has already been done in this area, by large, dedicated groups. Secondly, the shortcomings outlined above suggest that neither empirical nor modelling work is redundant when it comes to the identification of statistical structure in the observed record, where none has previously been identified, or sought.

6.3. *Initial predictor selection methodology*

6.3.1. Alternative approaches to the initial model selection stage

Following on from section 6.2, and given that a statistical scheme is to be used to select the reduced set of predictors, a range of possible techniques are available. These will be outlined with a view to justifying the eventual approach taken. These can be separated broadly into four approaches – firstly, a GLM fit across the whole predictand domain, after e.g. Chandler (2005); model selection at predictand gridbox resolution, using a best subset approach based on all available predictors; the expression of the dominant spatial covariance in the predictor and or predictand field using multivariate techniques; and fourthly, by using a non-parametric technique such as resampling to both reduce the predictor subset with respect to spatial degrees of freedom in the predictand field, and increase spatial coherence in the model selection across gridboxes.

The first approach, in particular developed by Chandler (2005) applies a generalised linear model (GLM) to the multivariate predictand, by fitting functions to account for the autocorrelation in the predictand, geographic location, seasonality, and 'external' forcing factors (such as ENSO, for example). This model can be applied across the whole domain to describe the variability in the predictand.

In the second approach, model selection would include the use of cross-validation to select individual or small subsets of predictors on a gridbox-by-gridbox basis. Superficially, cross-validation is a tool for quantifying the predictive skill of a model, and given circumstances such as these in which a very large initial range of predictors is available, the technique is not suitable since overfitting is a likely result. Although a cross-validation experiment could be designed to minimise this, a further drawback of this approach without an intermediate step is that it does not implicitly encourage spatial coherence between predictand gridboxes with respect to predictor selection, where it might be thought more physically reasonable to suppose that such coherence exists. Another possible model selection approach could entail an information theoretic approach. Burnham and Anderson (2002) have developed this at length in the ecological sciences, and DelSole (2004) has applied these techniques in the climate sciences.

Thirdly – multivariate techniques, such as principal components analysis (PCA), or canonical correlation analysis (CCA) can be used to define the dominant modes of variability in a multivariate climate dataset, and derive both predictors and predictands – for example, Haylock and Goodess (2004) use simultaneous CCA to relate large scale atmospheric predictors to precipitation extremes in Europe. Operational seasonal forecasts using CCA are made by among others, the Climate Prediction Centre of the US National Weather Service.

The fourth possible class of approach is that based on the seminal paper of Livezey and Chen (1983), where the spatial coherence of a predictor-predictand relationship – given a gridded predictand dataset – is tested for statistical 'field significance', based on a re-sampling ('Monte Carlo') simulation, and with a view to determining whether the spatial extent of locally significant correlation exceeds the spatial autocorrelation of the predictor.

Since spatial coherence is implied, the potential to narrow down the set of predictors and offer theoretical justification for their inclusion (either on the basis of published research or further work within this study) is perhaps greater than that offered by the other methods. There is a further advantage of retaining each predictor in its original format, rather than expressing a number of predictors with respect to dominant patterns of covariance, such as is offered by multivariate techniques. In the case of an exploratory study, this is particularly valuable when it comes to proposing theoretical mechanisms for statistically significant relationships. Furthermore, while the approach of Chandler *et al.* *(*2005) may be suitable where there is a clear and dominant set of candidate predictors (in this case simultaneous large-scale circulation characteristics), where this does not exist, it is not apparent that predictors can be applied effectively to the whole spatial domain using a single linear model. Given these circumstances, the process of reducing the full predictor set to a reduced set – based purely on statistical reasoning – was carried out using field significance testing as developed by Livezey and Chen (1983). Further details on this technique will be presented in sections 6.3.1.1 and 6.3.1.2.

6.3.1.1. Predictor selection based on tests for field significance

Field significance tests have been widely used in the climate sciences since the development of this technique by Livezey and Chen (1983). For a full explanation refer to Livezey and Chen (1983), or other works, including for example Wilks, (2005). A shorter explanation will be given here.

The objective of a test for field significance is to identify whether a given predictor has a statistically significant relationship with a climate (or other) predictand variable in a spatial domain.

Fundamentally, the correlation of a single timeseries with a three dimensional (space, time) array of data, and the quantification of the significance of any observed relationships presents two problems. Firstly, it might be expected that given sufficient locations within the spatial array, and a typical threshold of 95% confidence to designate statistical 'significance', a certain number of these locations would be expected to correlate with the predictor purely by chance. The expected level of locally significant chance correlations can be expressed as a binomial function, where given N independent tests (one test for each gridbox or point within the spatial array), we can calculate the expected number of significant tests X (in this case, gridboxes which correlate significantly with the predictor) due to chance, at a given significance level.

$$\Pr\{X = x\} = \binom{N}{x} p^{x} (1-p)^{N-x}, \quad x=0,1,\ldots,N \qquad (6.1)$$

Trivially, we might suppose that an array in which a larger number of locally significant correlations are identified than the threshold level expected from a binomial distribution might be said to display field-significance, in other words the predictor is 'globally significant' for the domain. However if we then consider that this number of locally significant relationships is largely a function of the spatial resolution of the predictand field, and that it may easily be altered by refining or coarsening the spatial resolution, it becomes apparent that some consideration of the spatial degrees of freedom in the predictand data is necessary. That is to say, given that there is an inherent level of spatial autocorrelation within any climate data field, unless this is explicitly accounted for in the gridding of the data, we cannot accept a simple binomial threshold to determine field significance – the 'real' threshold, taking into account spatial autocorrelation, will always be higher. There exists no simple analytical technique to determine the spatial degrees of freedom in a field, and typically, the approach taken to determine this is based on resampling, as introduced by Livezey and Chen (1983). Using this approach, the response of the field to a random variable is tested a large number of times, and an empirical approximation of the degrees of freedom in the field can then be reached. In order for the predictor to be considered as having field significance at a given level, the threshold (which can be expressed as a proportion of the total area covered by the field) set by the resampling experiment must be exceeded. It is this test which will be used to 'filter' out those predictors in the full set which do not display a globally significant relationship with the predictand field at the 95% confidence level. The remaining predictors, for each predictand set, will then comprise the reduced set, and will go on to be subject to further testing at the model development stage.

6.3.1.2. Monte Carlo Methods applied to predictor selection

There are a number of experimental considerations which are required to be met in order to design an effective experiment of this nature. Firstly, and particularly in the case of an exploratory study a set of data should be reserved for a training period, with the remainder being used to test independently the robustness of the model. In this case, a relatively short period (1958-2005) is considered, and in the interest of striking a balance between effectively training the model and quantifying its significance, it is necessary to reserve a period longer than 23 years (half of the available data) for the training set. The period 1958-1995 is used, leaving 10 years of data to test or validate the model. Bearing in mind that this is an exploratory study, and the nature of the system indicates that any predictability will be low, and possibly unstable as a function of time, the main goal of the validation period is not as a precise quantitative indication of model fit, but rather as a qualitative check on the validity of the model. It is therefore thought that given the limitations of the observational record a 10 year validation period is the most workable solution in this case.

Secondly, in order to carry out as true as possible an assessment of the spatial degrees of freedom in the predictand, it is necessary to replicate as closely as possible the characteristics of the predictor timeseries when generating the sample of random 'artificial' timeseries with which to carry out the resampling test. The approach taken here is to use an autoregressive model to replicate the temporal autocorrelation of the predictor timeseries in the artificial series.

6.4. Results

Results for the initial predictor selection stage are presented below. For clarity, Table 6.1, below, explains the notation used to abbreviate the predictor names.

6.4.1. Initial predictor selection for precipitation

This section presents the results of the initial predictor selection stage for the precipitation predictands, focusing on the 90th percentile exceedance counts, with some reference to the 95th percentile counts (which are broadly similar in their response, although generally showing a weaker statistical relationship than the 90th percentile exceedances), and for each of 12 three-month overlapping seasons, as described in Chapter 5. Since a relatively large number of predictors are retained at this stage by the field significance tests introduced in Section 6.3.1.1, only those which display a particularly strong fit with the predictand data, or those which are particularly interesting from a theoretical point of view will be discussed in detail.

It should be noted at this stage that the initial predictor selection phase based on Monte-Carlo testing for field significance fulfils requirements of statistical significance, and spatial coherence for each predictor separately. However, when viewed in the context of multiple field significance tests the issue of false discovery rates – in other words a type 1 error, where the null hypothesis is incorrectly rejected – becomes important. Ventura *et al.* (2004) discuss this issue at length, highlighting possible approaches to minimising or recognising the false discovery rate. In the case of this study, it is clearly important to view all results presented in this section in the light of this problem. In fact, as can be seen from Table 6.2, the percentage of predictors identified as having field significance for each season and each predictand variable is generally small, and may be viewed as the outcome of a series of trials described by a binomial process. Furthermore, given that many of the predictors in the full set are cross-correlated, the entire exercise of selecting the reduced set can be seen as analogous to the selection of a single predictor by Monte-Carlo testing an autocorrelated field. In this way, the initial predictor selection stage is merely a filter, and further testing is required to verify whether these predictors do in fact offer any predictive skill, since, as previously mentioned – if enough potential predictors are considered, a number will be found to be statistically related, commensurate with the stringency of the test. To a certain extent this further testing should account for the problem of the false discovery rate, highlighted by Ventura *et al.* (2004). This issue notwithstanding, the results presented below do highlight some relationships which are either potentially of interest from a theoretical point of view, or corroborate with other research.

Predictor Notation Key

Abbreviation	Description
ATL	Principal component of Atlantic Ocean SST anomalies
PCM	Principal component of Pacific Ocean SST anomalies (including Maritime Continent)
IND	Principal component of Indian Ocean SST anomalies
GLO	Principal component of global SST anomalies
H	Principal component of stratospheric geopotential heights at given pressure level (hPa)
T	Principal component of stratospheric temperature at given pressure level (hPa)
PV	Principal component of stratospheric potential vorticity at given pressure level (hPa)
NAO	North Atlantic Oscillation Index
N12	SST anomalies from Nino 1+2 region
N3	SST anomalies from Nino 3 region
N34	SST anomalies from Nino 3.4 region
N4	SST anomalies from Nino 4 region
SHT	Hemispheric temperature anomaly (Southern Hemisphere)
NHT	Hemispheric temperature anomaly (Northern Hemisphere)
HTD	Difference in Hemispheric temperature anomalies
AOS	Arctic Oscillation (AO)
DAR	MSLP at Darwin, Australia
EAP	East Atlantic pattern (EA)
EAW	East Atlantic/West Russian pattern (EA WR)
EPN	East Pacific/North Pacific pattern (EPNP)
PNA	Pacific North American pattern
POL	Polar Eurasian pattern
PTP	Pacific Transition pattern (PT)
SCA	Scandinavian pattern
SOI	Southern Oscillation Index
TAR	MSLP at Tahiti
TNH	Tropical Northern Hemisphere pattern
WPP	West Pacific pattern (WP)
SDEU	Eurasian snow cover anomaly
SDNA	North American snow cover anomaly
SDNG	North American (including Greenland) snow cover anomaly
SDNH	Northern Hemisphere snow cover anomaly
QBO3	QBO index at 30hPa
QBO5	QBO index at 50hPa
SOL	Solar flux anomaly
SST1	Local SST anomalies (Atlantic Ocean, north west of Scotland)
SST2	Local SST anomalies (North Sea)
SST3	Local SST anomalies (Baltic Sea)
SST4	Local SST anomalies (Atlantic Ocean, west of Ireland to Bay of Biscay)
SST5	Local SST anomalies (Adriatic Sea)
SST6	Local SST anomalies (western Mediterranean Sea)

Predictor names are followed by either P*n*, indicating the principal component number if applicable, or 'Lead *nn*' where *nn* indicated the lead time in months from the middle month of the predictand season

Table 6.1 Each of the predictor groups are assigned a particular abbreviation as shown above. For each predictor either a PC number is given (representing the principal component number) followed by the season from which the timeseries is derived; or a lead time (in months prior to the middle month of the predictand season) is given.

a				b			
Season	Predictors in full set	Predictors in reduced set	Percentage retained in reduced set	Season	Predictors in full set	Predictors in reduced set	Percentage retained in reduced set
JFM	853	49	5.7%	JFM	853	27	3.2%
FMA	852	66	7.7%	FMA	852	82	9.6%
MAM	853	75	8.8%	MAM	853	78	9.1%
AMJ	853	191	22.4%	AMJ	853	164	19.2%
MJJ	852	149	17.5%	MJJ	852	84	9.9%
JJA	852	100	11.7%	JJA	852	71	8.3%
JAS	852	51	6.0%	JAS	852	44	5.2%
ASO	851	49	5.8%	ASO	851	41	4.8%
SON	851	78	9.2%	SON	851	105	12.3%
OND	853	61	7.2%	OND	853	86	10.1%
NDJ	854	59	6.9%	NDJ	854	66	7.7%
DJF	854	33	3.9%	DJF	854	20	2.3%

Table 6.2 Number of predictors retained from the full set after Monte-Carlo testing for field significance for the following predictands: (a) 90th percentile precipitation exceedance counts; (b) 95th percentile exceedance counts, at the 95% confidence level. The aggregate numbers retained show substantial differences between seasons and a strong correlation over the seasonal cycle between the predictands.

6.4.1.1. ENSO

The ENSO family of predictors show broadly similar responses, as might be expected, although there are some notable differences between the different indices, in particular indicating differences in the lead times at which relationships are observed. Figure 6.1 shows the p-values of the field-significance tests for each season of the 90th percentile precipitation exceedance counts with the four Niño SST regions. It is clear that this measure of extreme precipitation responds significantly to all the Niño SST indices over the seasons FMA-MJJ, and at lead times ranging from two to eleven months within the sample. Interestingly, the Niño 3.4 and Niño 4 regions show responses at the longest lead times, where typically Niño 1+2 might be expected to, since the SST anomalies tend to propagate westward in the canonical ENSO variability, and therefore the Niño 1+2 region tends to lead the others. This may be related to the way in which anomalous circulation driven by each ENSO event is propagated into the mid-latitudes and to Europe.

Additionally there is a late autumn/winter response to the Niño 3.4 and Niño 4 indices at shorter lead times of two to four months. Figure 6.2 shows the canonical spatial response of spring/summer 90th percentile precipitation counts to ENSO. In the early spring the response extends from the North Sea coasts of Europe to the east of the domain, in Poland and the Ukraine, and in west-central Europe, south to Switzerland and Austria. Through spring and into summer, the response shifts westward until in MJJ it is located over the Bay of Biscay, with some response in Spain, Portugal and France. As far as precipitation is concerned, it could be argued that this bears some resemblance to the simultaneous NAO response in this region (e.g. Trigo *et al*., 2002). Lead times for the observed responses are wide-ranging as can be seen from Figure 6.1, but in each case the spatial response is broadly consistent with that in Figure 6.2. In advance of further testing of this relationship (see Chapter 7) it is worth noting that an examination of scatter plots of the relationship between ENSO indices and the predictand at each gridbox (not shown) indicate that in general the presence of a small number of outliers at a significant number of locations in the predictand data tend to be associated with El Niño events, and influence the strength of the fit to the extent that it is unlikely that any linear predictability might result from ENSO. That is, if the relationship were to be considered without the outliers, there would not be evidence for a statistically significant relationship. Given the frequency of occurrence of these outliers in the spatial domain, it is likely that they are a real feature of the data, however, the available sample size does not permit a robust estimation of the presence of nonlinear relationships between ENSO and the predictand.

The spatial pattern of the autumn-winter response to Niño 3.4 and Niño 4 is shown in Figure 6.3. Here the OND response to Niño 3.4 is very similar to the predictand response observed in the spring-summer months, consisting of positive anomalies over western and southern Europe. In the NDJ (not shown) and DJF seasons the response becomes a negative one centred over Scandinavia, with only a weak, insignificant positive response over central and western Europe. Similarly to the spring-summer response, an examination of scatterplot matrices (not shown) of the response at single gridboxes indicates that a clear linear signal is not observed. Nonetheless, the persistence of the statistical predictand response to ENSO is of interest, particularly in the light of other research which suggests an ENSO-Europe link, for example Fraedrich and Muller (1992), Fraedrich *et al. (*1992), Brönnimann (2007) and Brönnimann *et al. (*2007).

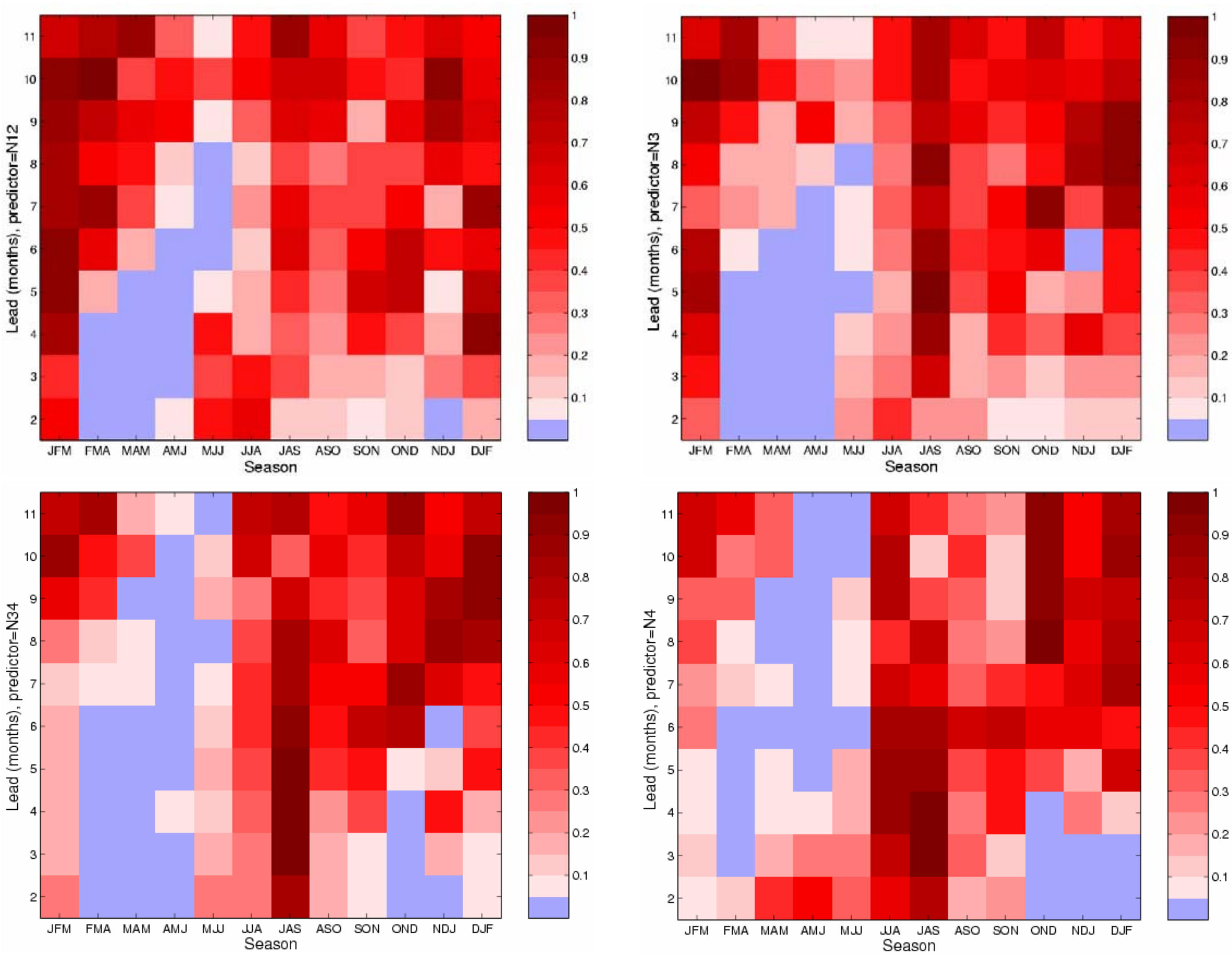

Figure 6.1 P-values indicating field-significance of fit between 90th percentile precipitation and Nino 1+2, Nino 3, Nino 3.4 and Nino 4. Lead times from two to eleven months are shown. Light blue indicates p<=0.05.

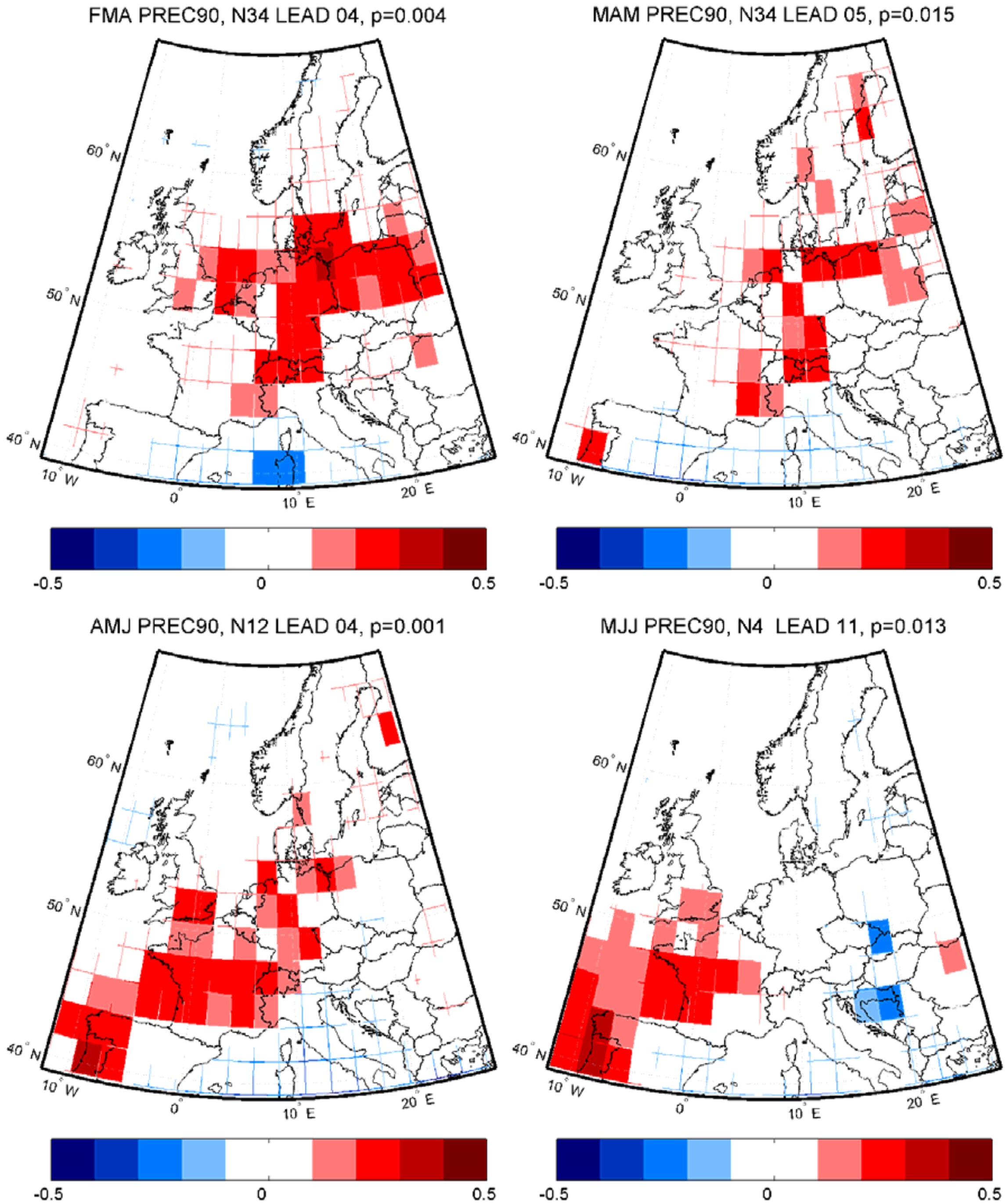

Figure 6.2 Spatial relationship of field-significant responses of 90th percentile precipitation exceedance counts to ENSO SST indices. Predictand season is given in the title of each panel, along with information on the predictor, where N12 is Nino 1+2, N34 is Niño 3.4 and N4 is Niño 4. The response is given as the coefficient of the Poisson model fitted at each gridbox. Filled gridboxes indicate local significance for the model at the 95% confidence level. Empirically calculated p-values from the resampling experiment are also given, where for example a p-value of 0.004 indicates that four of the 1000 random simulations had locally significant responses over a greater spatial extent than the predictor. The lead-time for each predictor is also given, for example where lead=4 for FMA indicates a predictor index from the previous October.

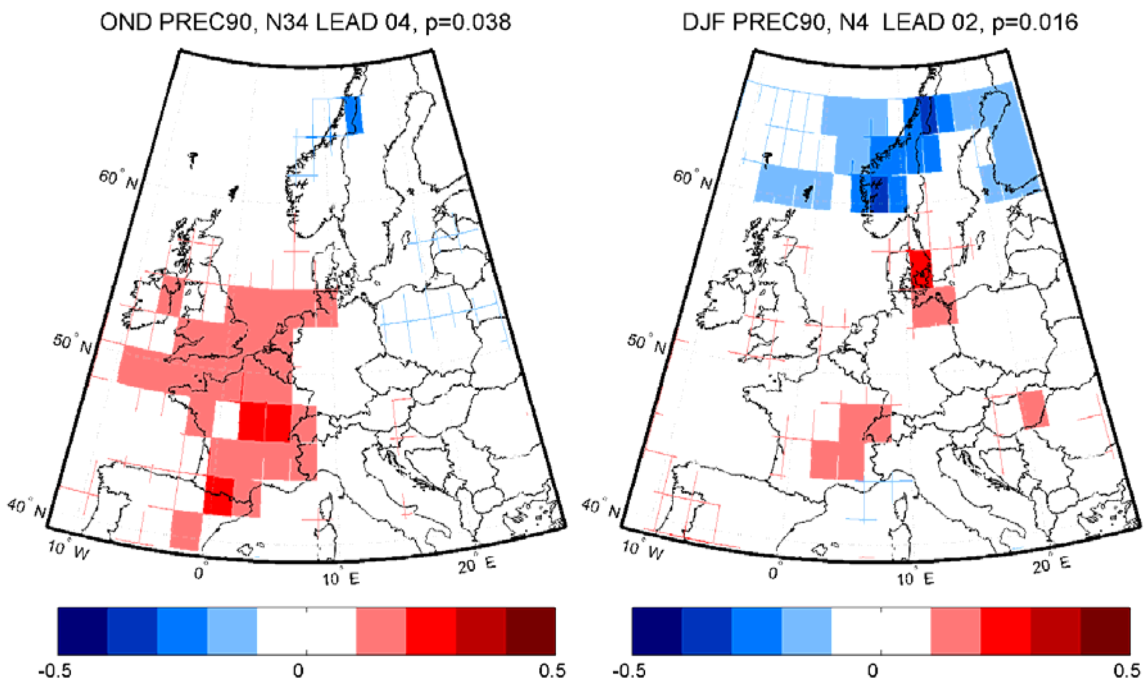

Figure 6.3 As Figure 6.2 but for OND and DJF precipitation exceedance counts, illustrating the marked shift in emphasis from a positive response over western and southern Europe to a negative response over Scandinavia

Of particular interest is Brönnimann *et al. (*2007) who use long records of European precipitation and temperature, and reconstructed ENSO indices to provide evidence for a possible nonlinear relationship between ENSO and Europe, dependent to some extent on slowly varying regimes in the Northern Pacific, and illustrating the likely sensitivity to other climate regimes of signal transmission from the tropics to Europe.

Further ENSO indices including the Southern Oscillation Index (SOI) and its constituent mean sea level pressure timeseries at Tahiti and Darwin, and also PCs of large-scale SST anomalies show a similar relationship, again with notable spatial differences for each index.

6.4.1.2. Large-Scale Atlantic SST Anomalies

As discussed in Chapter 5, the role of large-scale variability in the north and tropical Atlantic in European climate variability has been the topic of much research in recent years. Some success in linking SST variability with the European climate has led to the operational use of Atlantic SSTs as well as stratospheric indices in seasonal forecasts of the winter NAO by for example the United Kingdom Met Office (e.g. Rodwell *et al.*, 1999; Parker *et al*., 2007). Whether, or to what extent this observed skill is applicable to extremes of precipitation is as yet undetermined.

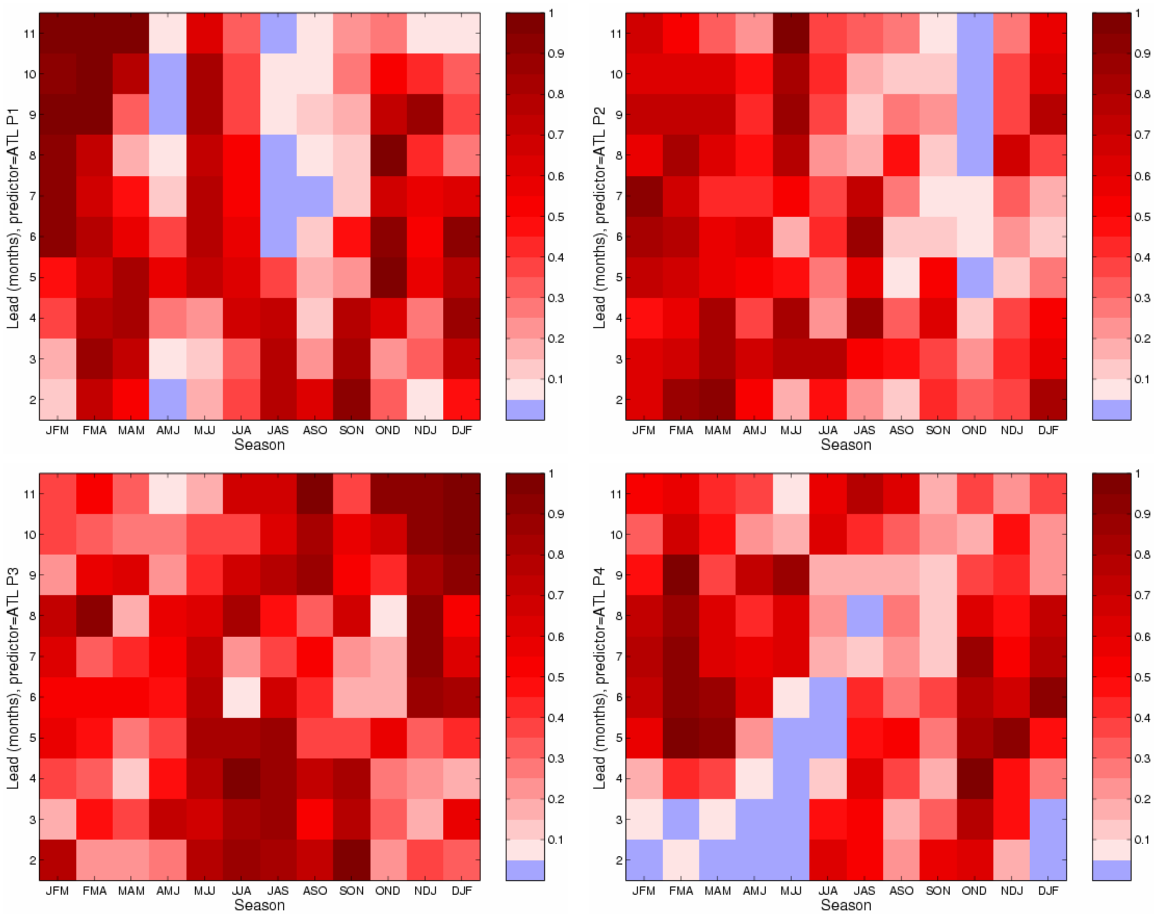

Figure 6.4 P-values indicating field-significance of fit between 90th percentile precipitation and the first four PC indices of Atlantic SST anomalies. Lead times from two to eleven months are shown, where each lead time corresponds to the middle month of a three month season in the predictor data. For example a lead time of two months corresponds to an OND predictor for the JFM predictand. Light blue shading (p<=0.05) indicates field-significance at the 95% confidence level in the predictand response.

From Figure 6.4, it can be seen that when aggregated, Atlantic SST predictors do not produce as many field-significant responses as for example the ENSO predictors. PC3 does not show any significant responses, while PC1 and PC2 show a few responses in spring-summer, and late autumn respectively.

PC4 of Atlantic SST anomalies is a tripole-like pattern in the North Atlantic, with the centre located off the US east coast and surrounded with a sometimes continuous horseshoe pattern of anomalies of the opposite sign. This pattern accounts for only six to seven percent of the overall variability, and from season to season is not stable with respect to its location or the relative strength of the centres of action. Despite this, it is associated with a persistent anomaly in 90^{th} percentile precipitation exceedances over Europe, and – to a lesser degree – with 95^{th} percentile exceedances also (not shown). The response of the 90^{th} percentile exceedance counts begins in the winter (DJF) season, and continues, at between two and 8 month lead times into the summer. These are primarily a response to autumn and winter SST, where the DJF precipitation response is to ASO and SON indices, and typically the late winter through spring precipitation response is to OND to JFM SST. The spatial pattern of responses varies systematically from DJF through to JAS. Initially, the winter response to late summer and autumn SSTs is focussed in the south of the domain, over Spain and the Mediterranean, with a small region over Scandinavia showing anomalies of the opposite sign. This pattern is illustrated in Figure 6.5. There is a weak positive correlation between the SON PC4 index and the December and January NAO indices, which is consistent with the observed precipitation anomalies. Given that the PC4 index is similar in its spatial pattern to the first PC, and that we can assume that the first PC is driven primarily by atmospheric variability associated with the NAO, it is possible that PC4 is associated with a feedback mechanism by which the overlying atmosphere is forced by the Atlantic. However, this relationship is weak, and further testing is required to clarify its potential as a useful predictor at seasonal timescales.

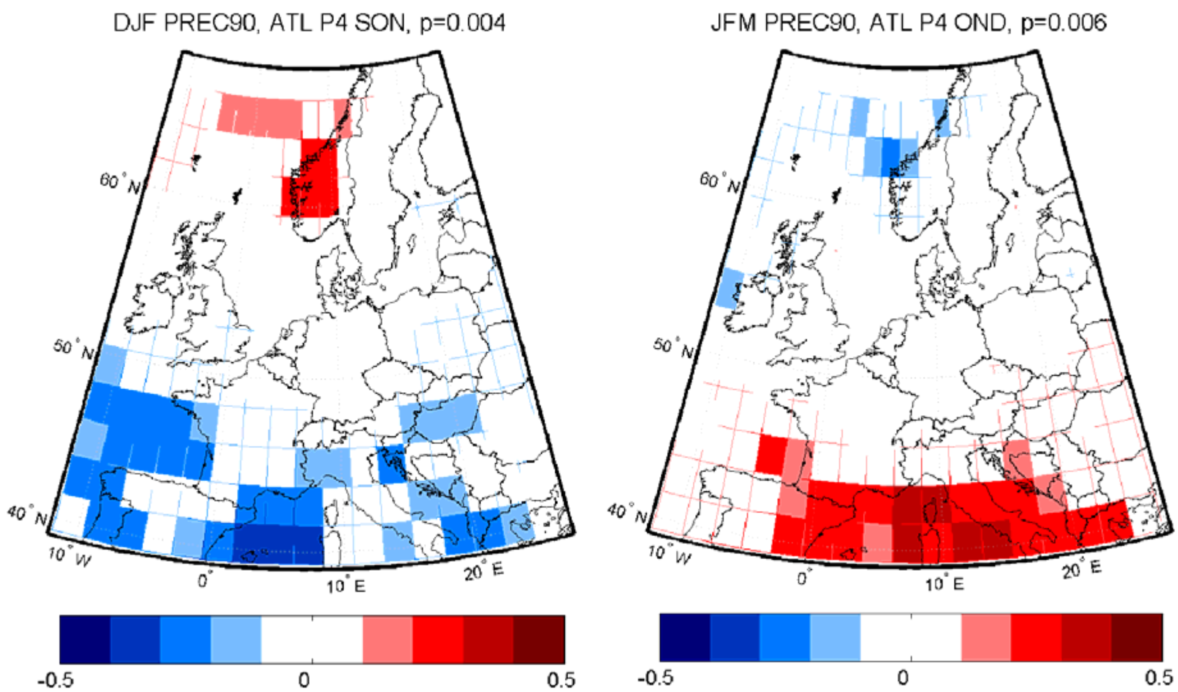

Figure 6.5 As Figure 6.2 but for DJF and JFM precipitation exceedance counts, showing the response in these seasons to the fourth PC of Atlantic SST anomalies in SON and OND respectively. Note that the principal components are derived independently for each season, and so the sign of the main centre of action common to each season can be either positive or negative. The SON and OND predictors are negatively correlated, and therefore the precipitation anomaly is in fact of the same sign in response to the SST pattern in both seasons.

6.4.1.3. Large-Scale Pacific SST Anomalies

In general, the principal components of Pacific SST anomalies elicit a similar response in the predictand to ENSO indices observed in Section 6.4.1.1. A summary of the predictand responses to the first four PCs is given in Figure 6.6. As expected, PC1 has the most frequent response in the predictand, being significant from FMA to MJJ, and at a range of lead times from two to 11 months. Additionally, there is a response in OND and NDJ at shorter lead times of two to three months (JJA to ASO predictor). The larger range of lead times at which precipitation responds to PC1 is to be expected given that PC1 should include elements of the variability from all four Niño SST regions. However, the persistence of PC1 as a statistically significant predictor is somewhat surprising unless it is primarily due to large SST anomalies which persist for more than 12 months – i.e. large ENSO events. The spatial pattern and sign of the precipitation response is similar to that from the ENSO SST regions.

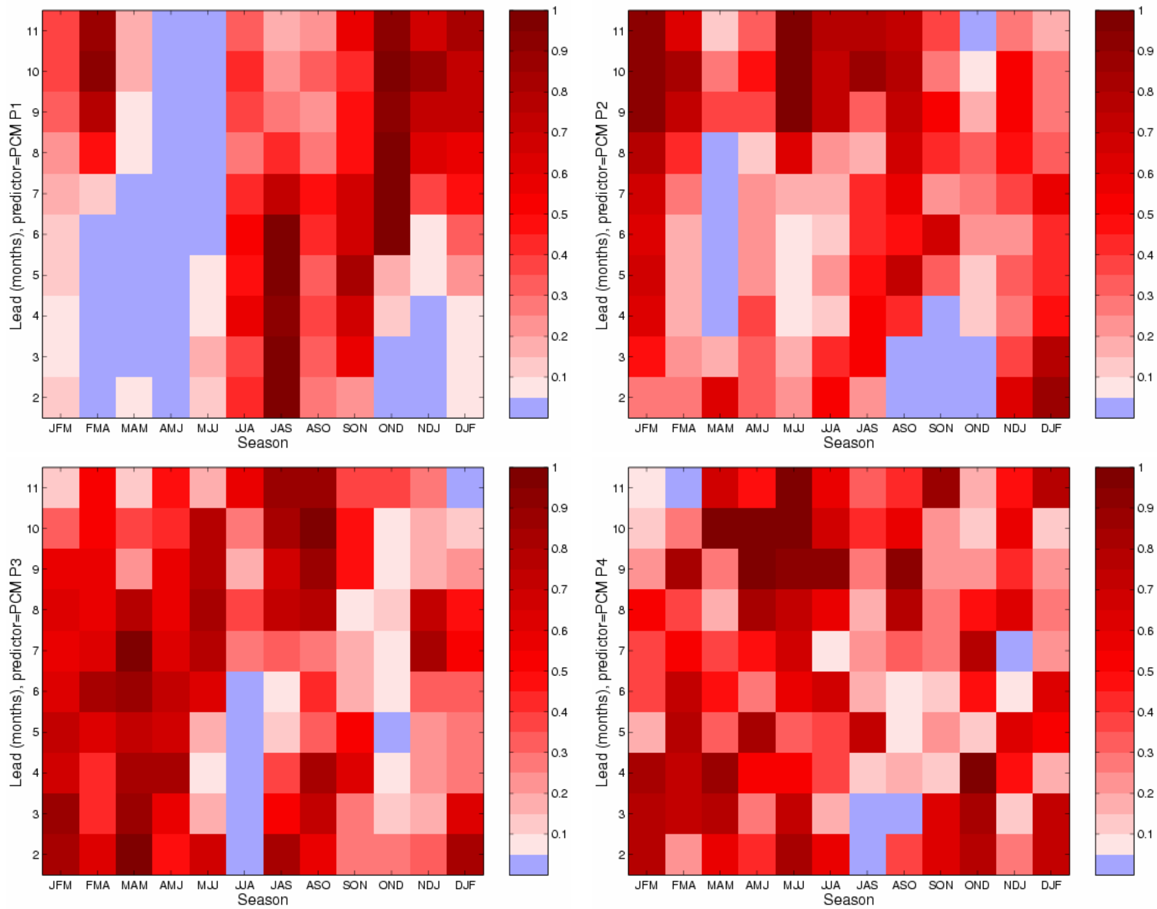

Figure 6.6 P-values indicating field-significance of fit between 90th percentile precipitation and the first four PC indices of Pacific SST anomalies. Lead times from two to eleven months are shown, where each lead time corresponds to the middle month of a three month season in the predictor data. For example a lead time of two months corresponds to an OND predictor for the JFM predictand. Light blue shading (p<=0.05) indicates field-significance at the 95% confidence level in the predictand response.

PC2 has fewer responses, which comprise a MAM response to the preceding summer and autumn SSTs, and an autumn (ASO-OND) response to spring/summer (AMJ-JAS) SSTs. The spring response in MAM comprises a positive anomaly centred over the British Isles, with the ASO predictor having the strongest response. The autumn response is primarily in southern and Europe, and comprises weak and scattered positive anomalies. These are illustrated in Figure 6.7. The second principal component of Pacific SST variability projects onto an ENSO-like pattern, with weak positive anomalies centred on a narrow band along the Equator in the eastern Pacific, and more widespread negative anomalies to the north, south and west. It persistently accounts for 10-12% of the total variability. The ASO PC2 correlation with Niño SST indices is low, comprising a weak negative correlation with the spring ENSO indices, followed by a weak positive correlation when PC2 leads ENSO by two to seven months. The implication here may be that this mode is related to the extratropical ENSO signal in the northern Pacific, which has been proposed as a possible mechanism for the ENSO signal transmission to Europe by Brönnimann *et al. (*2007).

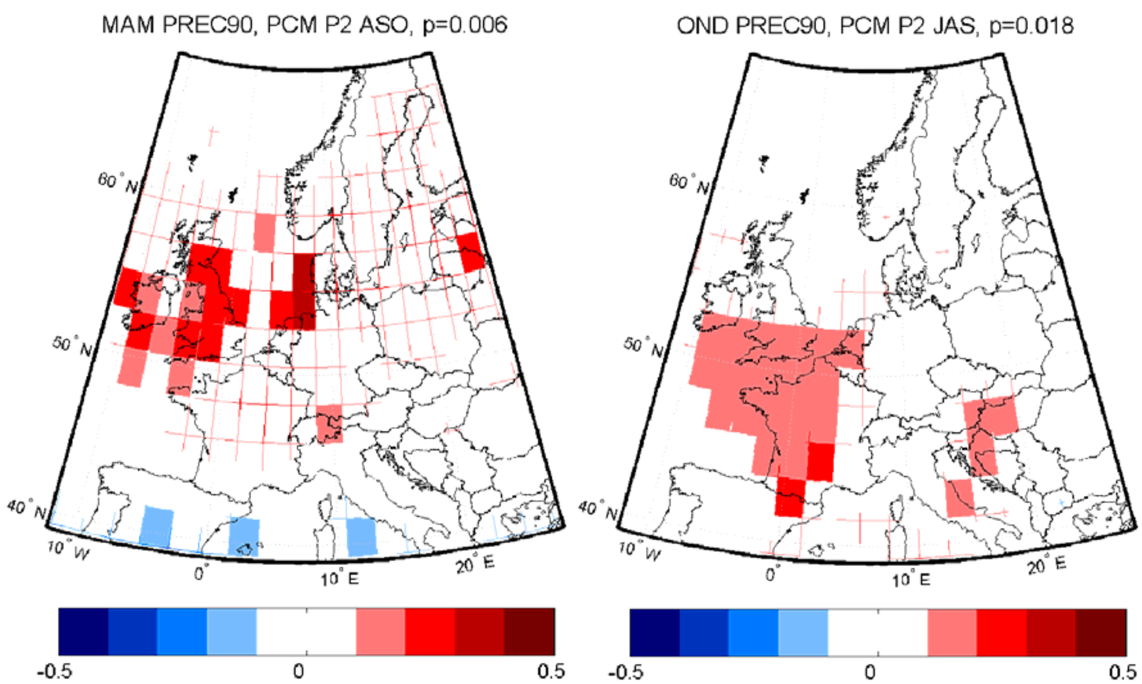

Figure 6.7 As Figure 6.2 but for MAM (left panel) and OND (right panel) precipitation exceedance counts, showing the response to the second PC of Pacific SST variability. The MAM precipitation responds to this index at lead times of between four and eight months (OND to JJA of the preceding year). The OND response is an example of the autumnal response which includes ASO-OND.   The response is spatially variable, being concentrated in southern Europe during ASO, and moving north-west into OND.

6.4.1.4. Large-Scale Indian Ocean SST Anomalies

Indian Ocean PC1 and PC3 are associated with predictand responses from the spring through autumn (FMA to SON), as shown in Figure 6.8. Both of these indices are highly correlated with ENSO variability in the tropical Pacific, and therefore the statistical relationship between Indian Ocean variability and the European climate may be interpreted either as a modulation of the ENSO signal, as a direct effect of Indian Ocean variability, or as a statistical artefact. As with all the potential predictors discussed here, further statistical testing at least is therefore required, before the observed relationships can be confirmed as useful or otherwise. Figure 6.9 gives two examples of the spatial response of precipitation to these indices. In both cases, the response is similar to that observed for ENSO, comprising a coherent positive anomaly across central and western Europe.

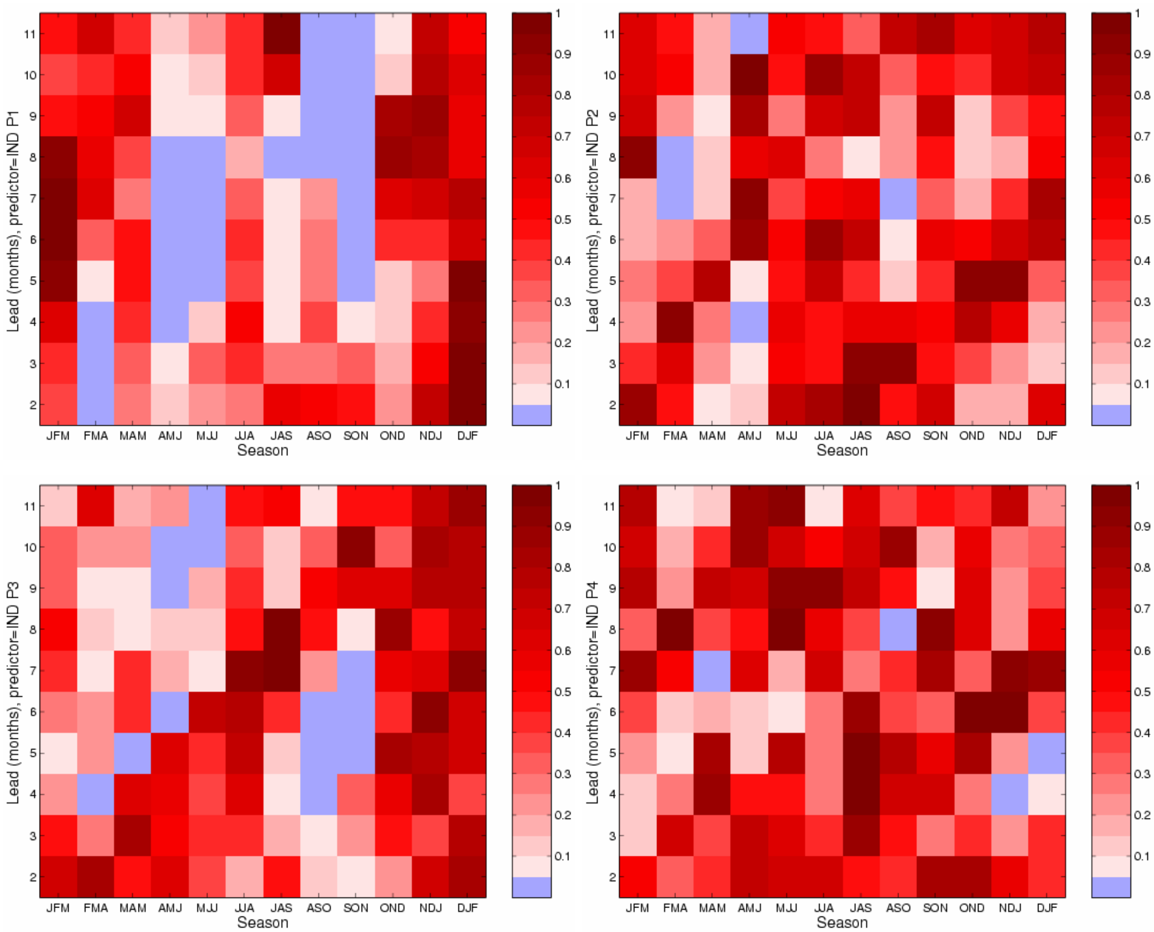

Figure 6.8 P-values indicating field-significance of fit between 90th percentile precipitation and the first four PC indices of Indian Ocean SST anomalies. Lead times from two to eleven months are shown, where each lead time corresponds to the middle month of a three month season in the predictor data. For example a lead time of two months corresponds to an OND predictor for the JFM predictand. Light blue shading (p<=0.05) indicates field-significance at the 95% confidence level in the predictand response.

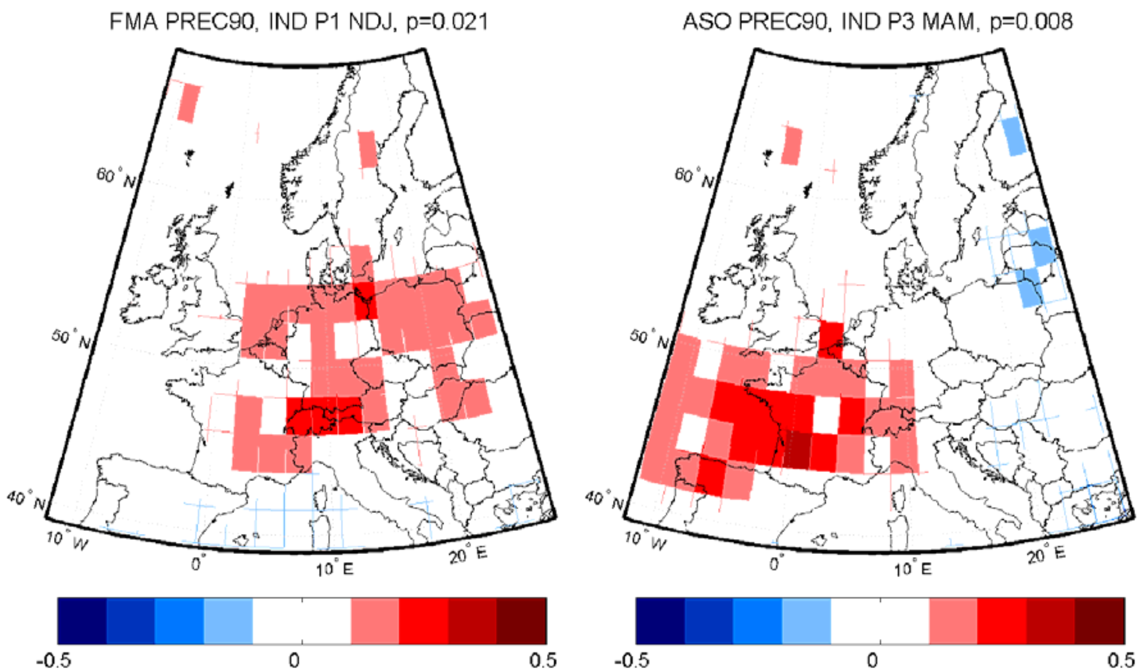

Figure 6.9 As Figure 6.2 but for FMA (left panel) and ASO (right panel) precipitation exceedance counts, showing the response to the first and third PCs respectively of Indian Ocean SST variability. The FMA precipitation responds to preceding NDJ anomalies in a similar fashion to the FMA response to Niño 3.4 SSTs from the preceding October. The ASO response is also similar to the ENSO response in its spatial configuration, although no significant ASO response to ENSO is observed directly.

6.4.1.5. Stratospheric Predictors

Given the large degree of cross correlation between the stratospheric predictors (each predictor is included at 150, 100, 50 and 30 hPa level, and the different variables – geopotential height, potential vorticity and temperature are also cross-correlated), not all of them will be examined in detail.

The inclusion of stratospheric variables as potential predictors is based upon a substantial body of work discussed more extensively in Chapter 5, which aims to explore the possible effect of the stratosphere on the troposphere at lead times of up to several months. It can be seen from the summary plots in Figure 6.10 that a number of these predictors are statistically associated with anomalous precipitation extremes over Europe, at a range of lead times from two to eleven months. In particular, using examples from Figure 6.10, the first and second principal components of 150hPa geopotential height and potential vorticity all have a strong response in European precipitation anomalies particularly in the spring and summer seasons, at a surprising range of lead times – where in many cases significance is observed at up to 12 months preceding.

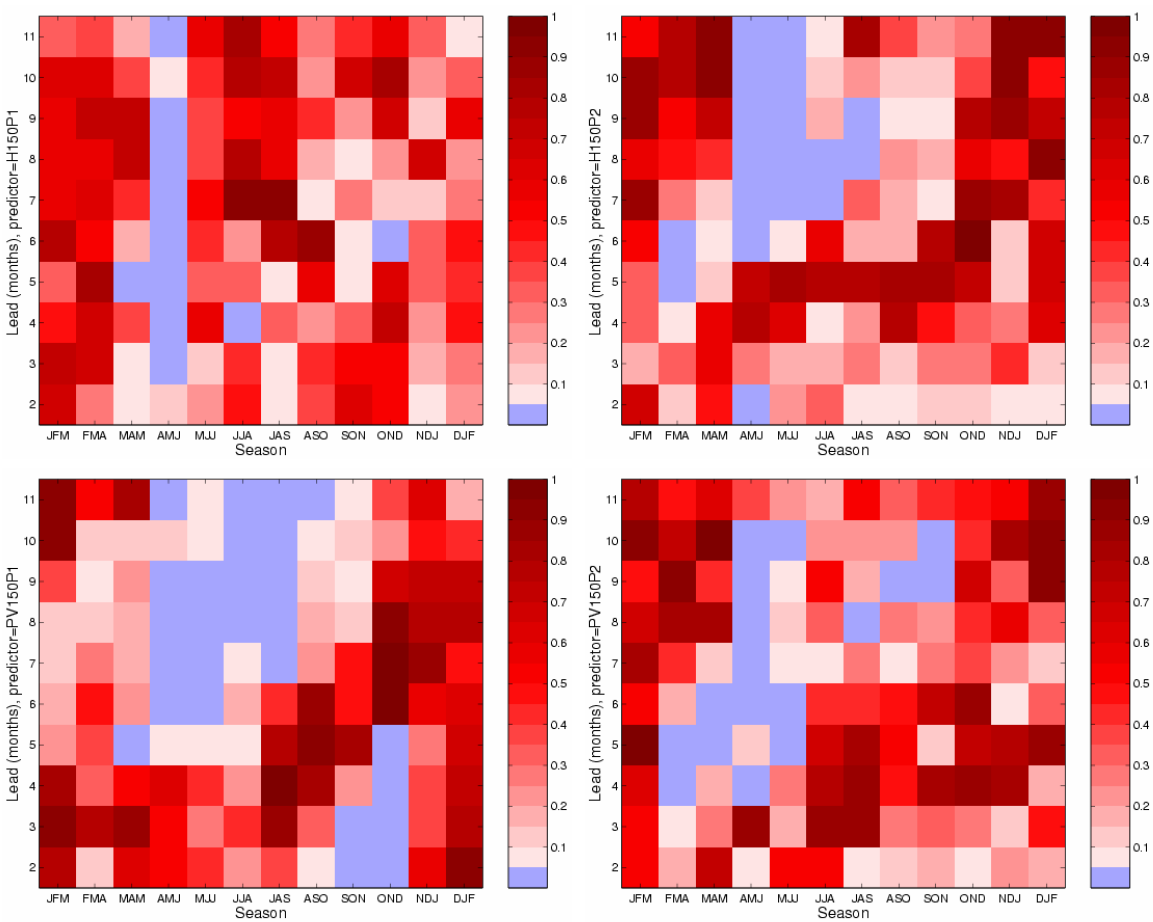

Figure 6.10 P-values indicating field-significance of fit between 90th percentile precipitation and a selection of stratospheric variables, including the first and second principal components of 150hPa geopotential height (top left and right panels), and the first and second principal components of 150hPa potential vorticity (bottom left and right panels). Lead times from two to eleven months are shown, where each lead time corresponds to the middle month of a three month season in the predictor data. For example a lead time of two months corresponds to an OND predictor for the JFM predictand. Light blue shading ($p<=0.05$) indicates field-significance at the 95% confidence level in the predictand response.

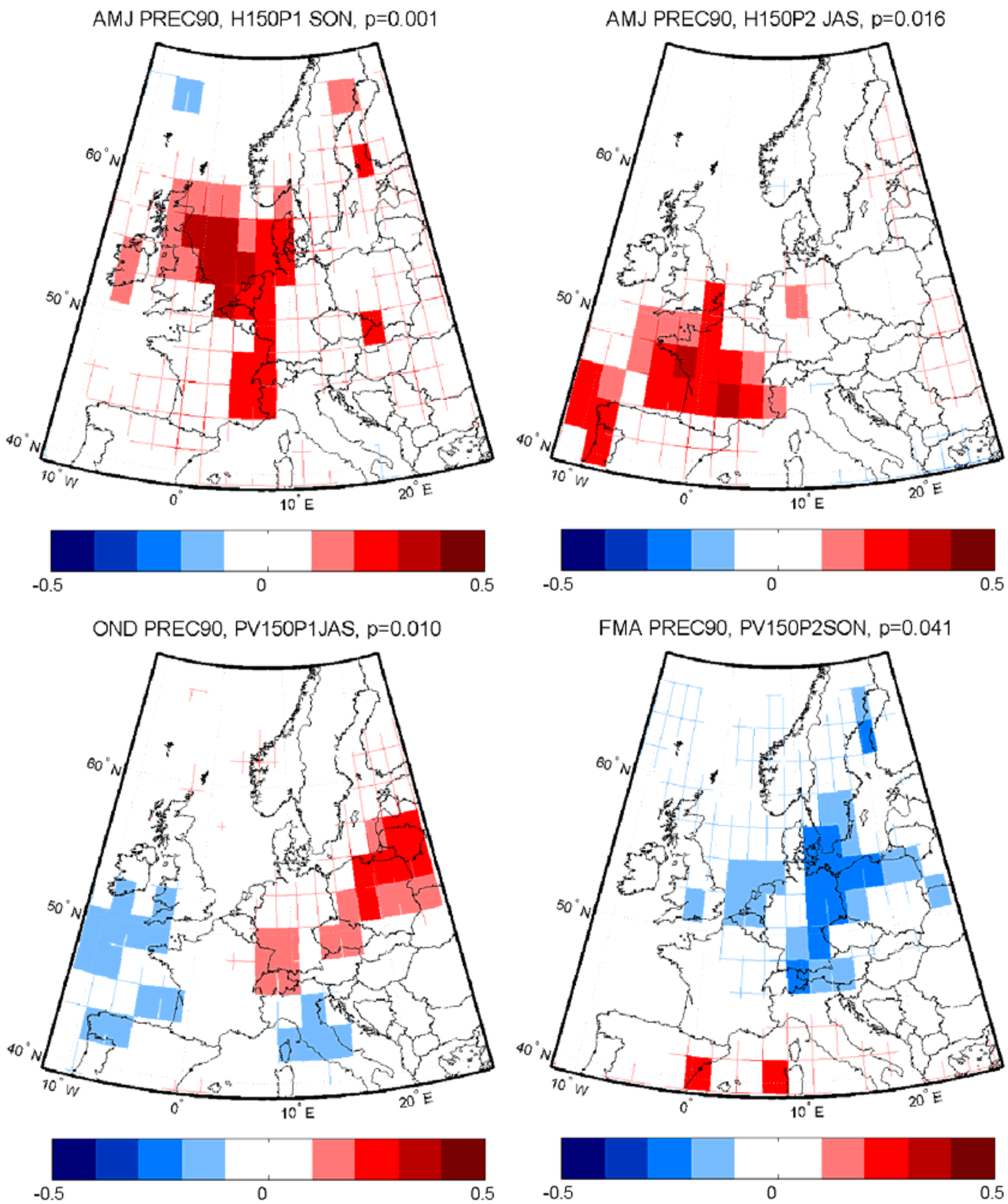

Figure 6.11. The spatial response of 90th percentile precipitation exceedance counts to stratospheric predictors. The top two panels show the AMJ response to the first and second PCs of 150hPa geopotential height from the preceding SON and JAS seasons respectively. The lower two panels show the OND and FMA response to PC1 and PC2 of 150hPa potential vorticity, from the preceding JAS and SON seasons respectively.

The spatial responses to a selection of the field-significant relationships highlighted in Figure 6.10 are shown in Figure 6.11. For example, the spatial response in AMJ to the first two principal components of 150hPa geopotential height, from the preceding autumn and summer is similar in sign although the anomalies occupy a slightly different spatial extent, with the PC1 anomaly centred over the North Sea, and extending down the German/French border, and the PC2 anomaly being centred over France. The SON PC1 and JAS PC2 are uncorrelated. The OND precipitation response to PC1 of 150hPa potential vorticity is illustrated in the lower left panel of Figure 6.11, and the FMA response to PC2 in the lower right panel. While in both cases a field-significant response is observed, the mechanistic explanation for this is as yet unclear, and further investigation is required.

6.4.2. Initial predictor selection for extreme wind events

This section reviews the potential predictors for the wind gust exceedance counts. As for the precipitation predictands, the 90th percentile exceedances will be the main focus, for brevity, and only the predictors where a particularly notable response was observed will be discussed. Similarly, potential predictors identified in this section are based on statistical testing only – and not only is further testing required to assess their utility in a forecasting model, but also a mechanistic explanation of their effect on the predictand is required in order to assess their true potential and reliability as predictors. In general, wind extremes in the ERA-40 dataset show greater spatial autocorrelation than the precipitation data, and the spatial extent of the response to field-significant predictors is therefore relatively greater.

a				b			
Season	Predictors in full set	Predictors in reduced set	Percentage retained in reduced set	Season	Predictors in full set	Predictors in reduced set	Percentage retained in reduced set
JFM	853	26	3%	JFM	853	49	5.7%
FMA	852	27	3.2%	FMA	852	26	3.1%
MAM	853	23	2.7%	MAM	853	11	1.3%
AMJ	853	37	4.3%	AMJ	853	29	3.4%
MJJ	852	51	6%	MJJ	852	58	6.8%
JJA	852	51	6%	JJA	852	36	4.2%
JAS	852	41	4.8%	JAS	852	31	3.6%
ASO	851	57	6.7%	ASO	851	34	4.0%
SON	851	88	10.3%	SON	851	82	9.6%
OND	853	73	8.6%	OND	853	98	11.5%
NDJ	854	64	7.5%	NDJ	854	76	8.9%
DJF	854	52	6.1%	DJF	854	70	8.2%

Table 6.3 Number of predictors retained from the full set after Monte-Carlo testing for field significance for the following predictands: (a) 90th percentile wind gust exceedance counts; (b) 95th percentile exceedance counts, at the 95% confidence level. The aggregate numbers retained show substantial differences between seasons and a strong correlation over the seasonal cycle between the predictands.

6.4.2.1. Large-Scale Atlantic Ocean SST Anomalies

As in the case of the precipitation predictands, the most notable wind predictand response in terms of the persistence of the predictor is to the fourth PC of Atlantic SST. This is during the ASO-SON seasons, at a range of two to 9 months. Figure 6.12 summarizes the seasonality of 90th percentile wind exceedance responses to PCs one, two four and five of Atlantic SST anomalies (PC3 shows no significant responses). The canonical spatial pattern of the autumn response to PC4 is given in the left panel of Figure 6.13. This response differs entirely from the precipitation response, which takes place largely in the winter and spring seasons, comprising a large region of negative anomalies centred over eastern Europe, and bears little resemblance to the NAO like pattern observed in some cases for precipitation. The first PC is associated with a winter response, at a lead time of two months in NDJ and DJF.

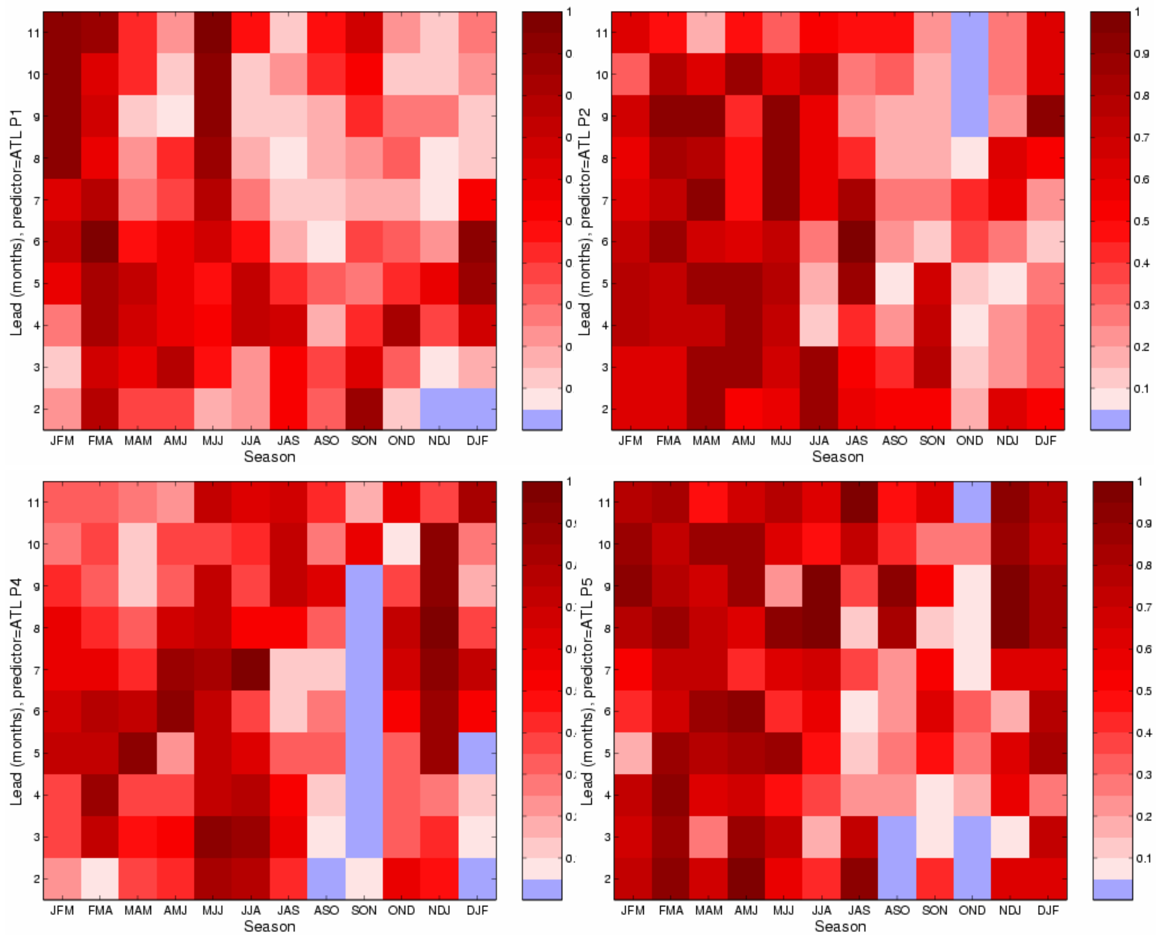

Figure 6.12 P-values indicating field-significance of fit between 90th percentile wind and the first, second fourth and fifth principal components of Atlantic SST variability. Lead times from two to eleven months are shown, where each lead time corresponds to the middle month of a three month season in the predictor data. For example a lead time of two months corresponds to an OND predictor for the JFM predictand. Light blue shading (p<=0.05) indicates field-significance at the 95% confidence level in the predictand response.

This response is more characteristic of the NAO signature, with widespread negative anomalies across northern Europe, and the relationship may bear close resemblance to the predictive pattern used by the UK Met Office in their statistical seasonal prediction scheme for the winter NAO (Rodwell *et al.*, 1999). The spatial configuration is illustrated in the right panel of Figure 6.13.

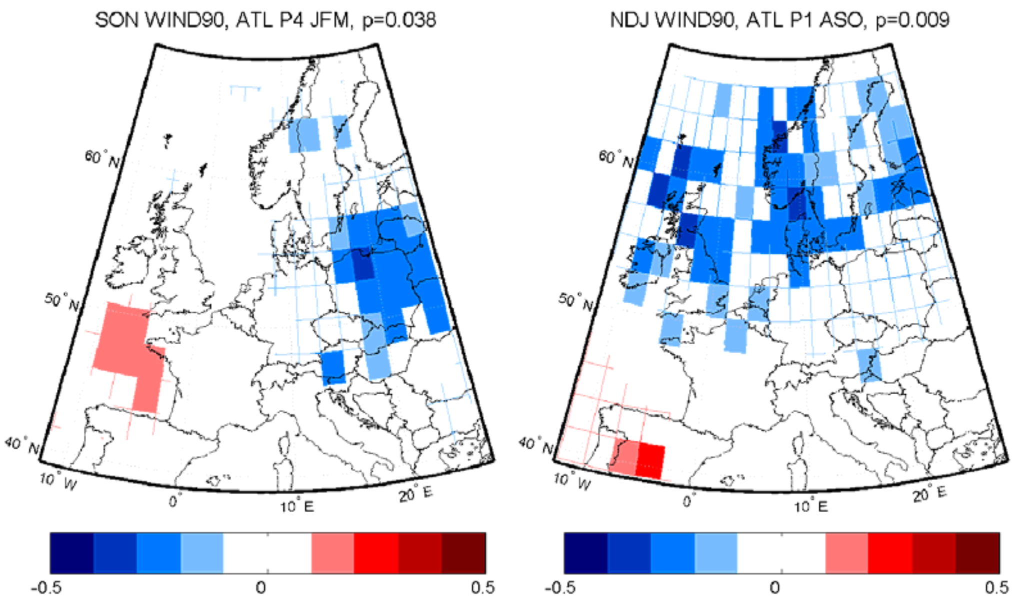

Figure 6.13 Spatial configuration of the 90th percentile wind exceedance count responses in SON (left panel) and NDJ (right panel) to the fourth and first PCs of Atlantic SST respectively.

6.4.2.2. Large-Scale Pacific Ocean SST Anomalies

Interestingly, the ENSO SST region predictors – to which a notable response was observed from the precipitation predictands – do not play such an important role in generating responses from the wind predictands. However, the principal components of Pacific SSTs seem to be more important, indicating that there may still be an ENSO influence on European wind extremes. Additionally, while the ENSO impact on precipitation extremes was observed to be greatest in the spring and summer, the response of wind extremes tends to be concentrated in the autumn and winter, with both Niño 3.4 and PC1 having significant responses in OND and NDJ. Additionally, as can be seen from Figure 6.14, the second, third, fourth (and fifth, not shown) PCs all elicit a number of responses, generally in the autumn and winter, but also, in the case of PC2 and PC3, in the spring and summer respectively. There are a range of different spatial responses to the predictors, with the most frequent being similar to the OND response to Niño 3.4, with a large region of positive anomalies centred off the west coast of France. This is similar in its spatial configuration to the spring/summer precipitation anomalies observed as a function of ENSO. Additionally, particularly in the spring and summer seasons, a dipolar anomaly is observed in response to PC3, consisting of a positive anomaly over Scandinavia and north of Scotland, and a negative anomaly over southern and central Europe. Canonical examples of these patterns are shown in Figure 6.15.

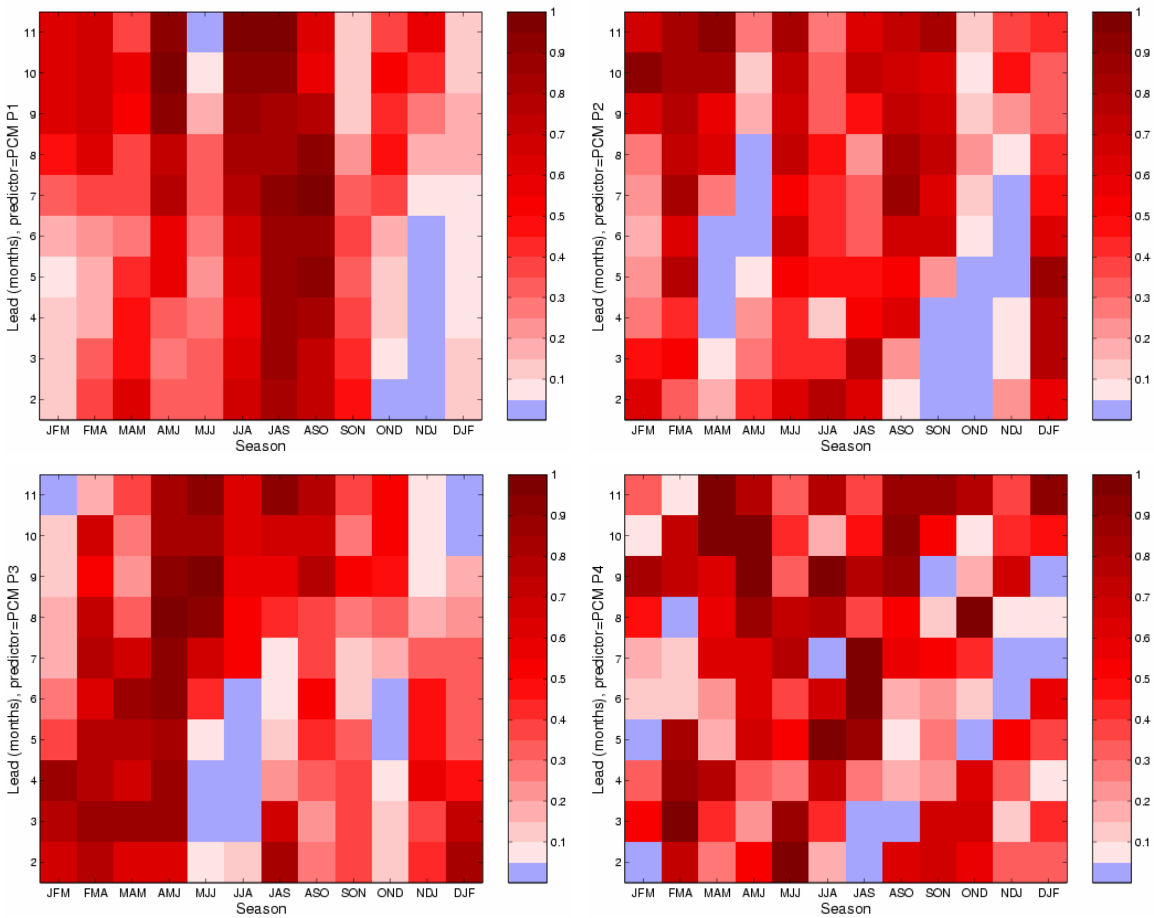

Figure 6.14 P-values indicating field-significance of fit between 90th percentile wind and the first four principal components of Pacific SST variability.

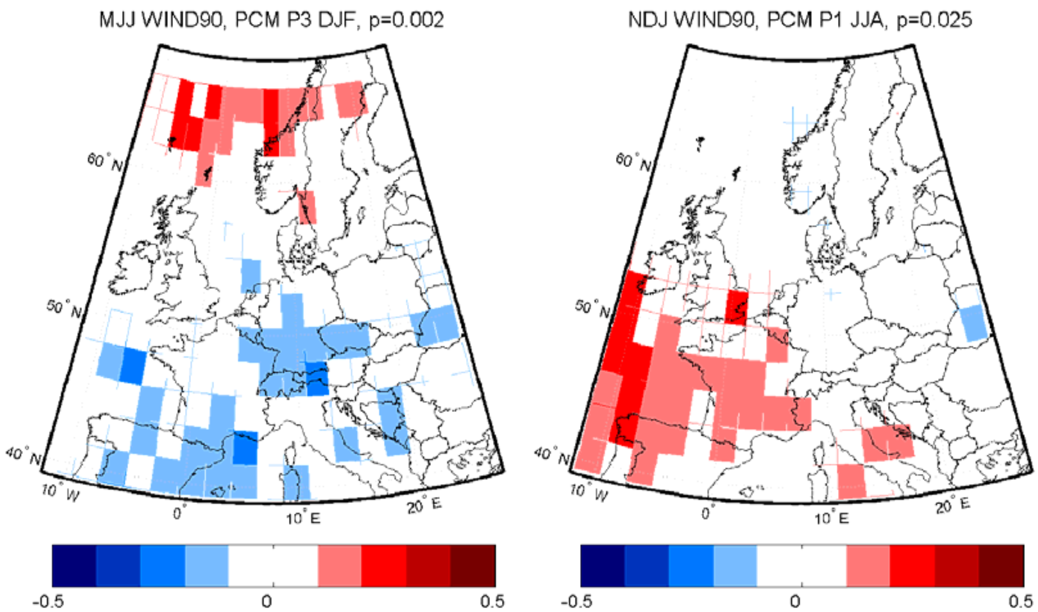

Figure 6.15 Spatial configuration of the 90th percentile wind exceedance count responses in MJJ (left panel) and NDJ (right panel) to the third and first PCs of Pacific SST respectively.

6.4.2.3. Large-Scale Indian Ocean SST Anomalies

The 90th percentile wind exceedance count response to Indian Ocean principal components is dominated by the first component. The major impact of PC1 is observed in the late summer and autumn (ASO-OND), and can be seen at lead times of three to eleven months, and consists of a persistent dipolar anomaly with a positive centre over eastern Europe, and a smaller, more fragmented negative component over Spain and France. Other than the seasonality of the response, there does not seem to be a great deal of overlap with the ENSO-related response, even though this pattern is at least partially associated with ENSO. However it is interesting to note that the response to the Indian Ocean is observed at longer lead times than that to ENSO, and that typically lead times of close to a year are associated with the most widespread geographical response. This is despite the fact that the Indian Ocean typically lags activity in the Pacific by several months. The range of the temporal and spatial responses is shown in Figure 6.16.

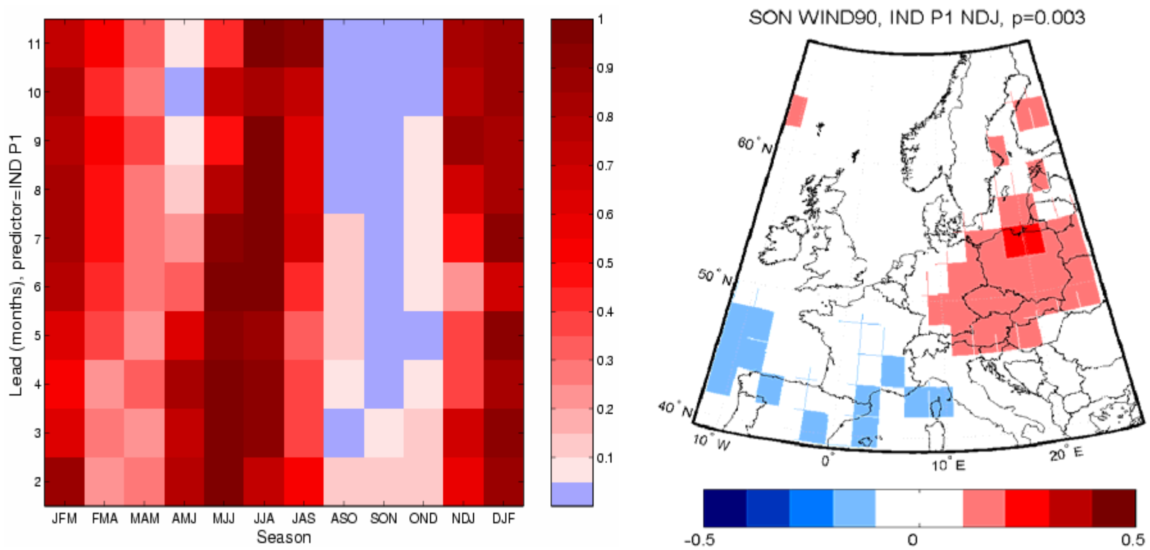

Figure 6.16 Left panel: p-values indicating field-significance of fit between 90th percentile wind and the first principal component of Indian Ocean SST variability. Right panel: Spatial pattern of the anomalous SON response to PC1 of Indian Ocean variability from the preceding NDJ season, showing the canonical dipole associated with this predictor.

6.4.2.4. North Atlantic Oscillation and Arctic Oscillation

A large number of the observed spatial responses of the wind extremes have an NAO or AO-like configuration – that is, a north-south dipole, with centres over northern and southern Europe, that resemble the large-scale circulation response to the NAO over Europe. It may therefore be the case that the predictors identified as statistically significant are projecting onto the NAO or AO, and this is driving the downstream anomalies. Conversely – and this is likely to be the case if the observed relationships are merely statistical artefacts – it may be simply that the NAO related pattern is the most spatially coherent, and any predictor which happens to be correlated with the NAO will be predisposed to be flagged as field-significant as a result. This notwithstanding, the AO and NAO themselves seem to provide some evidence at least of statistical significance as predictors at several months lead. The temporal distribution of significant responses through the seasonal cycle is shown in Figure 6.17. Typically, the responses are at shorter lead times, of up to five months, with a few instances in excess of this. The AO produces more frequent field-significant responses, and spread throughout the annual cycle, while the NAO responses are (perhaps surprisingly) concentrated in the spring and summer. Additionally, significance at 11 months lead is observed in AMJ and MJJ. In the winter seasons (NDJ-FMA), the typical spatial pattern consists of a large positive anomaly over northern Europe. This persists into the spring, for both the AO and NAO, but weakens substantially. In the summer months the response is more variable, in some cases assuming a similar spatial configuration to the winter responses, and in some cases being dominated by a negative anomaly over the south east of the domain. However both of these patterns bear some resemblance to the simultaneous projection of the NAO/AO, and even given the range of lead times at which field-significance is observed, it is interesting to note the rarity of departure from this pattern, indicating that any memory in the system is largely manifest as a positive recurrence of the pattern. Examples of the spatial configuration of the responses are illustrated in Figure 6.18. Given the prominence of summer responses to the NAO in particular this is of particular interest, since the NAO is not particularly dominant in these months. Of course, extremes of wind are unlikely to be of great climatological or meteorological significance during these seasons either.

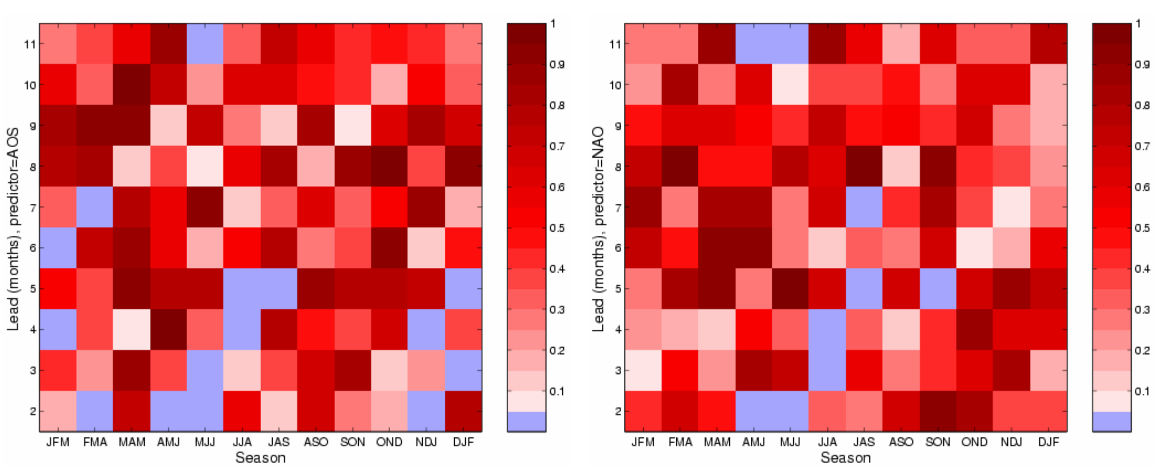

Figure 6.17 P-values indicating field-significance of fit between 90th percentile wind and the Arctic (left panel) and North Atlantic (right panel) Oscillations.

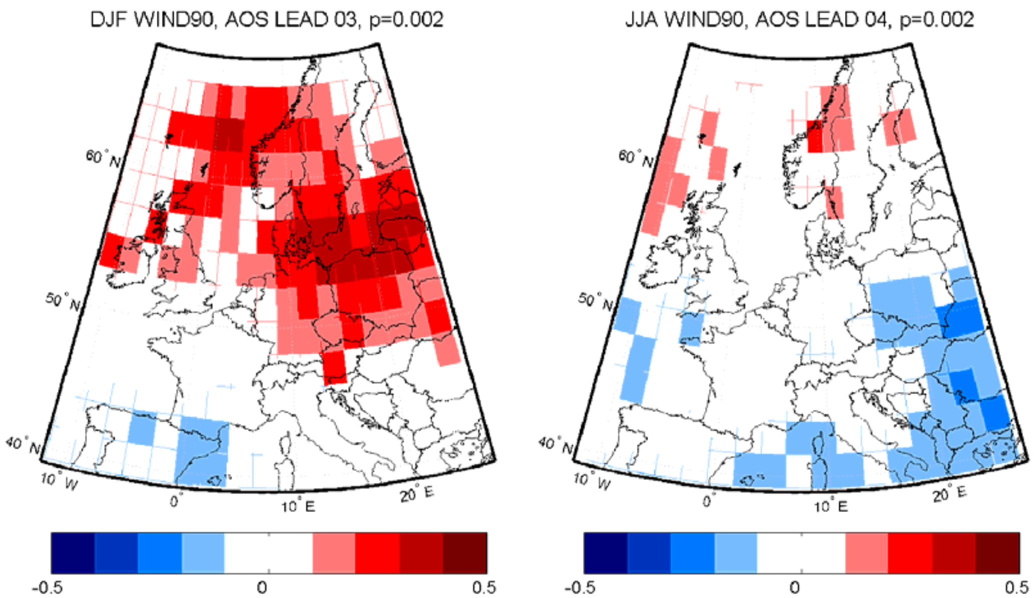

Figure 6.18 Spatial configuration of the 90th percentile wind exceedance count responses in DJF (left panel) and JJA (right panel) to the Arctic Oscillation, representing the canonical response in the winter/spring, and summer respectively.

6.4.2.5. Stratospheric Predictors

As in the case of the precipitation predictands, significant responses are observed in a range of the stratospheric predictors, although unlike for precipitation, the responses are mostly to the temperature and potential vorticity predictors, rather than geopotential height. In many cases there is significant cross-correlation between these predictors. A selection of the more interesting responses is discussed here.

The third principal component of potential vorticity at 30hPa results in a field-significant fit to the predictands at a range of seasons, from MAM to NDJ, at a variety of lead times. These are detailed in Figure 6.19 (top left panel). The remaining panels in Figure 6.19 show a sample of the typical spatial responses to this pattern. As is frequently the case with the wind predictands, an NAO-like pattern is observed, with positive anomalies over northern Europe, and negative anomalies to the south, in response to the JFM predictor. The winter (NDJ) response to the early autumn (ASO) PC index is slightly different, in that the positive anomaly to the north extends further south, and the negative anomaly is spatially more confined to the south west of the domain. This is consistent with the more southerly location of the mid-latitude Atlantic jet during the winter.

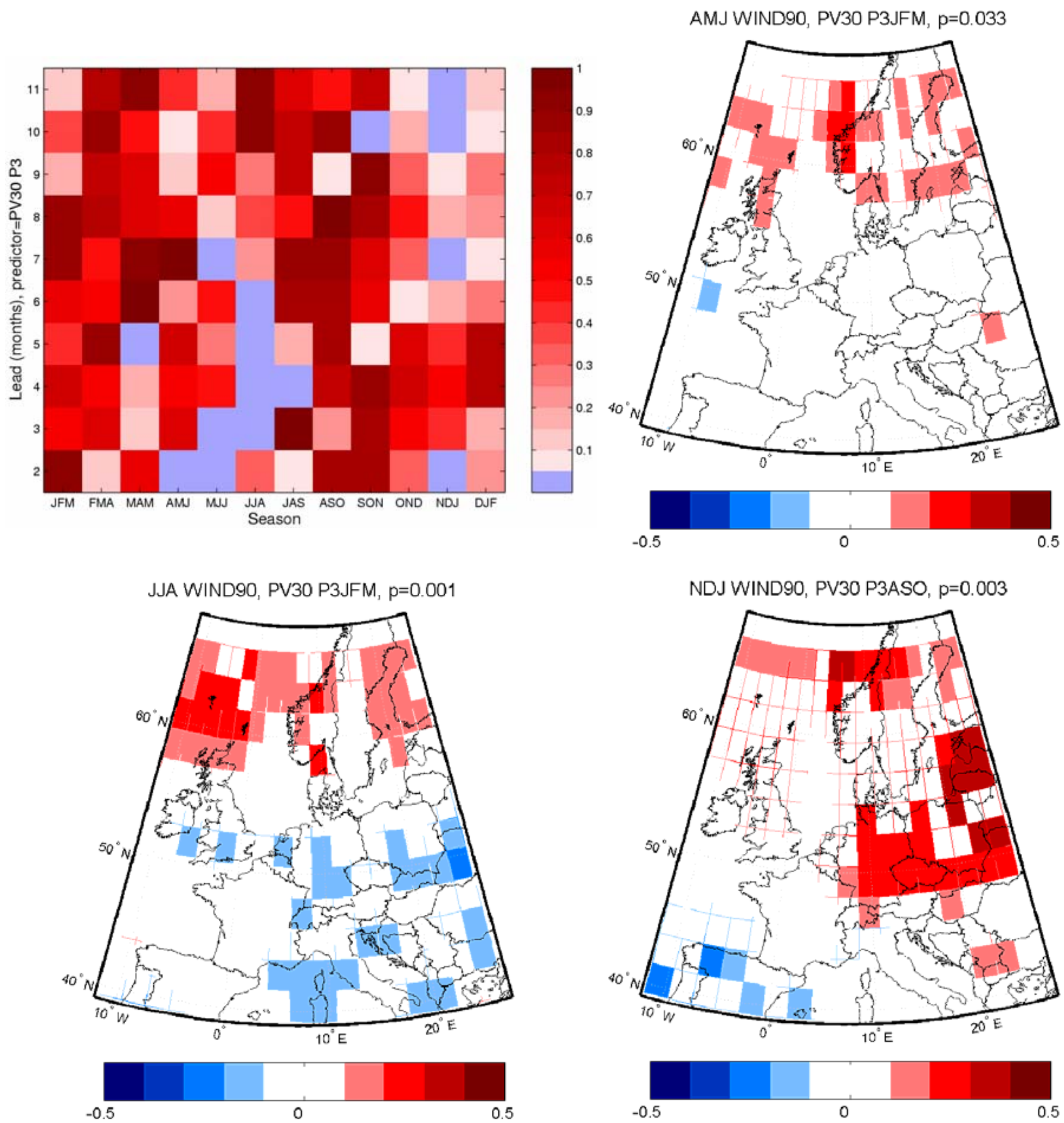

Figure 6.19 P-values (top left panel) of the 90th percentile wind response to PC3 of 30hPa potential vorticity. The remaining panels show the spatial configuration of the response during selected seasons: AMJ (top right panel); JJA (lower left panel) and NDJ (lower right panel).

6.4.2.6. Northern Hemisphere Snow Cover

The potential utility of Northern Hemisphere snow cover as a predictor of the NAO has been examined in a number of studies including Bojariu and Gimeno (2003) and Saunders and Qian (2003). The physical basis for the possible predictive skill is based on the persistence of the snow cover, and its forcing of the overlying atmosphere, which has been shown to occur on a large scale (e.g. Cohen and Rind, 1991; Gong *et al*, 2004). In this case, some response to indices of snow cover over various parts of the Northern Hemisphere, and including the whole Northern Hemisphere is observed. The range of responses is summarized in Figure 6.20. Rather unexpectedly, particularly in light of other research on this topic, Eurasian snow cover does not give the widest range of responses – rather the North American responses are most frequent. Eurasian responses are apparent in late winter/early spring, to snow cover anomalies from the preceding late spring/summer, while responses to the North American snow cover occur in spring through to autumn, at a wide range of lead times. Responses to Northern Hemispheric snow cover anomalies are more concentrated in the autumn and winter.

Typically, the response in the spatial domain as shown in Figure 6.21 is positive – that is, anomalously extensive snow cover is associated with anomalously frequent wind extremes, at a range of lead times. The locations of these anomalous extremes is quite variable, being centred variously over the north, south east, and west of the domain, depending on the season and the lead-time of the predictor. One notable exception to this is the OND response to Northern Hemisphere snow cover at 10 months lead, where the dominant anomaly is negative, over the southern half of the domain. The same predictor five months later gives rise to a strong positive anomaly, although this is centred further to the north.

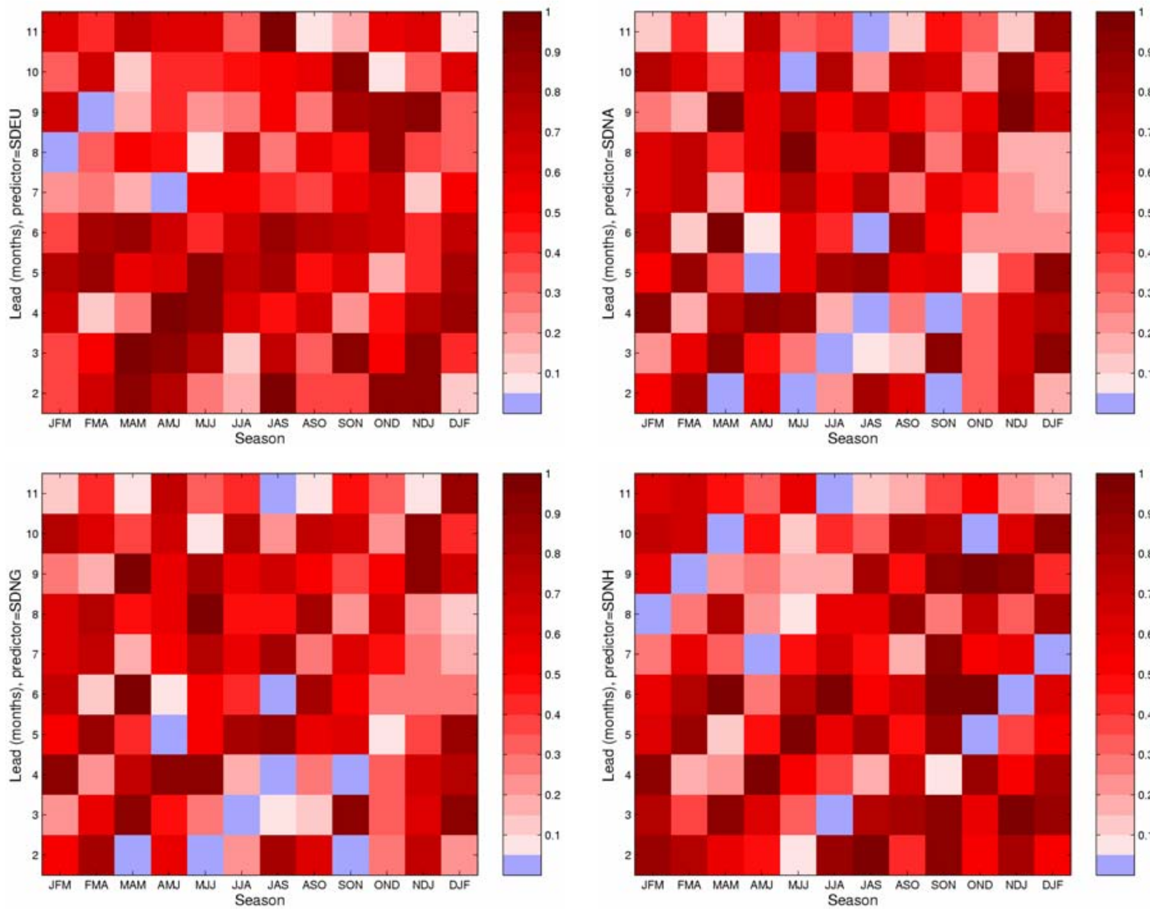

Figure 6.20 P-values indicating field-significance of fit between 90th percentile wind and Northern Hemisphere snow cover: Eurasia (top left panel); North America (top right panel); North America including Greenland (lower left panel) and Northern Hemisphere (lower right panel).

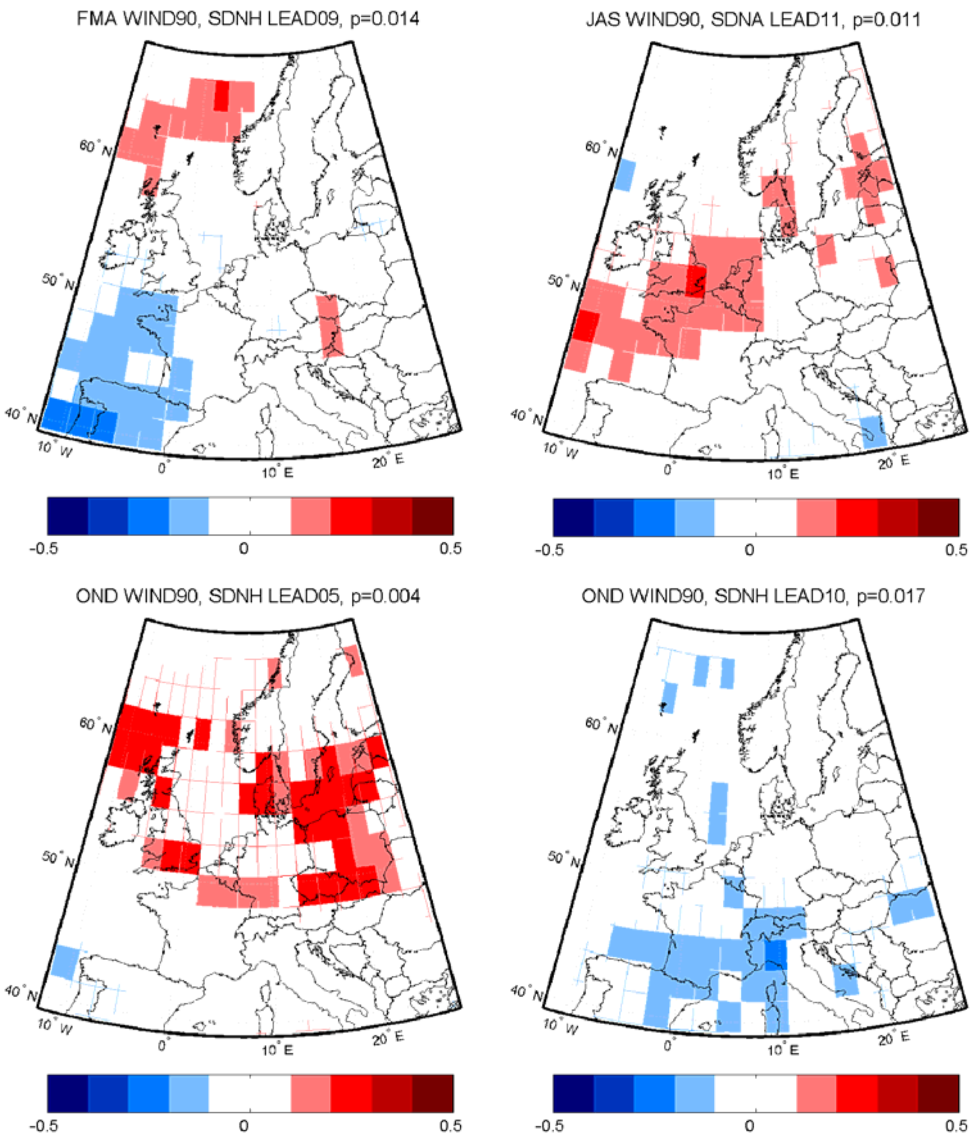

Figure 6.21 Spatial configuration of the 90th percentile wind exceedance count responses to indices of Northern Hemisphere snow cover.

6.5. Discussion and implications for model development

This Chapter has outlined the problems faced when carrying out a largely exploratory analysis of potential predictors for a seasonal forecasting model for Europe, using statistical methods. Initially, the main problem is one of narrowing down a large set of potential predictors to a reduced set that is practical for model design. This was carried out using Monte-Carlo resampling to determine field significant responses in the predictand domain. In the process of doing this, it is instructive to consider the role of each potential predictor individually, with respect to its statistical significance, and how it might contribute to the model.

Section 6.4 presents the results, focussing on those predictors which display the strongest or most striking fit with the predictand data. For brevity, the 90^{th} percentile counts of precipitation and wind exceedances were considered in detail, and generally it is the case that the 95^{th} percentile exceedance counts for both wind and precipitation are similar to the 90^{th}, although in most cases show a weaker fit.

Overall, for the 90^{th} (95^{th}) percentile precipitation exceedance counts, 9.4% (8.5%) of the full initial predictor set passed the field significance testing (with 95% confidence). For the 90^{th} and 95^{th} percentile wind exceedances, the total is lower, at 5.8% and 5.9% respectively. Although these are all above the (naively) expected 5% which would pass by chance, we must consider cross-correlation in the predictors, which would act to reduce the degrees of freedom. However whether the net effect of this would be to inflate or reduce the chance of a type I error is unclear.

An additional factor that should be considered in the interpretation of field-significant predictor-predictand relationships is the spatial coherence of the predictand response. It is frequently observed in the field-significant responses that there are small numbers of (‘scattered’) locally significant gridboxes separate from the main (‘coherent’) response region. There is a significant likelihood that these will have arisen by chance, and therefore contribute to an overestimation of the field-significance. While it might be possible to automate the minimisation of this effect by using some sort of nearest-neighbour algorithm to exclude these isolated responses, it should also be recognised that at this stage the field-significance testing is primarily a filter on the full set of predictors, with further model selection and testing in place as presented in the following chapters. It may therefore be of questionable benefit to exclude predictors for gridboxes where at this stage there is at least some evidence for a field-significant response over the domain.

This approach notwithstanding, it is possible to gain some appreciation of how many such scattered locally significant gridboxes might appear by chance in the general case, using the binomial distribution. If we assume that each gridbox is entirely independent from the others, then at the 95% confidence level the total of 252 gridboxes would be expected to result in up to 18 locally significant responses purely by chance. We know this not to be the case, since the spatial scale of the processes is generally larger than the gridboxes (hence the need for Monte Carlo simulations of the predictand response). If we assume – given a region of coherent (as opposed to scattered) field-significant responses – that the remainder of the gridboxes outside this region are subject to displaying a ‘scattered’ locally significant fit with the predictor as a function of the binomial distribution then we would expect the number of locally significant outliers to follow the pattern illustrated in Figure 6.22. For example, if a spatially coherent region of 150 gridboxes shows a field-significant response, then up to 12 gridboxes outside this domain might be expected to show scattered local significance. However, the gridboxes outside this region are likely also to have some spatial coherence in their lack of fit with the predictor, so taking this into consideration, it might be expected that the observed number of scattered locally significant gridboxes – given a field-significant response is somewhat lower than the cases given in Figure 6.22. While this might be sufficient to boost the rate of field-significant responses, given the further testing carried out on each predictor, it is thought that this effect is likely to be largely negated during the model selection phase.

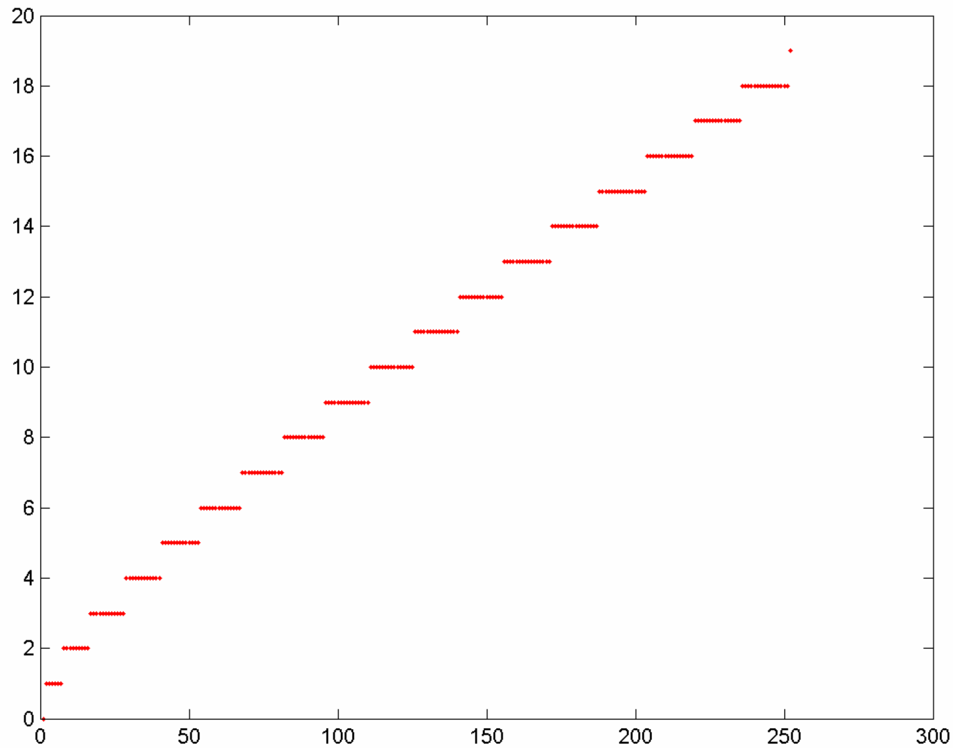

Figure 6.22. The number of locally significant gridbox responses expected by chance (shown on the Y-axis) at the 95% confidence level, as a function of the number of gridboxes (shown on the X-axis), and assuming independence between gridboxes.

What can be ascertained at this stage is that many of the predictors are associated with a response in the predictands that is supported by other literature on the subject – although none of it specifically on precipitation and wind extremes. It is encouraging to see some correspondence although further testing is required before any of the reduced set of predictors obtained by this process can be incorporated into a model where the skill can be reliably quantified.

7 Model Selection

Having obtained a reduced predictor set by retaining predictors which display a significant fit with the predictands at 95% confidence, using a field-significance test, it is then possible to select the optimum model at each gridbox using the reduced predictor set. The procedure by which the reduced set is obtained is described in chapter 6. This chapter describes the process of selecting prediction models at each gridbox, and presents the results over the model training period of 1958-1995. Section 7.1 introduces the chapter, section 7.2 describes in detail the method by which the models are derived, section 7.3 presents the results of the analysis, and section 7.4 concludes the chapter.

7.1 Introduction

Chapter 6 presented an exploratory study of the linear statistical significance of the relationships between a wide range of potential predictor indices describing large scale ocean, atmosphere and land surface variability, and the predictands. This chapter seeks to extend this work to an assessment of potential predictive skill derived from an all-subsets model selection exercise using the reduced predictor set. The assumptions of linearity and stationarity over the model training period still hold in the methodology. Since the skill of the models is tested over an independent time period separately to the results shown here, any discussion of model skill in this chapter should be interpreted strictly in the context of the statistical fit shown over the training period. Furthermore, given the large number of predictors available from the reduced predictor set, and the complexity of summarising all the observed model selections, only a sample of the apparently more important predictors and models will be discussed in detail here – although skill for all models will be presented in graphical format. This approach is taken partly due to the issue of overfitting, which will be illustrated in chapter 8, and which will allow a more focussed and meaningful discussion of predictability, beyond the constraints of assumptions of linearity and nonstationarity as required by the available data.

7.2 Method

As with the reduced predictor set selection, the model development uses Poisson regression – a member of the family of general linear models (GLMs) to fit the best model from all possible predictor subsets based on the model training period of 1958-1995 as in chapter 6. Further detail on the methodology and its suitability to this problem is given in chapter 3. As in chapter 6, the emphasis must still be on the exploratory nature of this work, since there is little evidence for a linear physical relationship between the predictors and predictands – although as demonstrated in chapter 6, there is substantial evidence that statistical relationships (albeit weak) exit. The aim of the work presented in this chapter is therefore to attempt to optimise the predictors locally at each gridbox, and separately by predictand and season, so that a best fit model is available for each location and season. However, each of these models should be treated as an optimum case of linear predictability given the exploratory nature of the work, and the relatively short period over which data are widely available for this work. In light of this, one of the principal goals of this work is to develop models that are not over fitted to the available data, such that further testing on independent data does not result in significant degradation of any observed relationships – or in the case of significant degradation, that the observed relationships can be explored further on the basis that they may be either nonlinear or nonstationary. In addition, the final models will be constrained to have a maximum of two predictors. Given the limitations of the data with respect to the period of observation, and the lack of clear physical mechanisms this is thought to be prudent.

Given these considerations, the models are selected as follows. For each gridbox within the domain the locally significant predictors (significance level 0.05) from the reduced set of predictors are first filtered for cross-correlation with other predictors. However, since there are further checks for multicolinearity, a very high threshold (significance level $1x10^{-5}$) is used at this stage. This step is simply a filter to reduce processing time. When two or more predictors correlate at this level, the one with greater field significance over the whole domain is retained. For the remaining predictors, all possible subsets with maximum two predictors are tested. All-subset selection is chosen as it offers the most comprehensive assessment of the 'best' model, given all other constraints. For example, backward selection (e.g. Jagger and Elsner, 2002) starts with all available predictors, and removes the least important until a pre-defined set of criteria relating to improvement in fit, and model parsimony are met. Since the constraint of having two predictors is thought to be of primary importance here, backward (or forward) selection algorithms are not preferable to the all subsets method Drobot (2003) highlights the suitability of this method on the grounds that the user can pre-define the required selection criteria, rather than having a machine algorithm take the decision. For each subset, given that the predictors are not cross-correlated with significance less than 0.05, a leave-one-out cross validation is carried out on the model, over the training period. That is, for *n* observations the model is fitted *n* times using the same predictor set, and leaving out one observation each time. This missing observation is then predicted using the fitted model. A number of considerations are necessary at this stage. Firstly, the predictor data are standardised, and a leave-one-out cross validation scheme in this case will tend towards reproducing the left out observation by default as a result (von Storch and Zwiers, 1999). Therefore the predictor data are re-standardised at each stage of the cross-validation, to eliminate any possibility of reconstructing the missing observation. Since only one observation is left out, the effect of re-standardisation on the remaining observations is minimal. For the purposes of this analysis, it is assumed that serial correlation of the predictands from year to year is minimal, and will not have a significant effect on cross-validation performance, therefore a buffer of observations between the training and validation data at this stage is not required.

For each model, the predicted values from the cross-validation step are used to estimate the fit of the model within the training period using the mean absolute error (MAE) of the predicted values relative to the observed predictands. The final model selected at each gridbox is that with the lowest MAE of all the subsets tested. MAE is used as the selection criteria following Willmott and Matsuura (2005), who identify this statistic as preferable to, for example, the root-mean-square error (RMSE), which is sensitive to the variance of the distribution of error magnitudes. The MAE is also more meaningful in this case since it relates directly to the magnitude of the errors on timeseries of percentile exceedance counts.

7.3 Results

This section starts by summarising the number of predictors retained from the field-significance testing phase detailed in chapter 6. Then for each predictand, results are summarised and illustrated by season with respect to the model fit. Although model skill within the training period is shown for all models, there are too many to discuss each in detail, instead, for each season, the impact of some of the most important predictors are described in more detail.

Figure 7.1 shows the number of predictors retained in the reduced subset predictor set by gridbox for 90th percentile precipitation exceedance counts. As well as marked regional variation across the domain, there is also variation by season, with winter and summer having fewer predictors and spring and autumn having large areas of the domain with more than 30 predictors from which to derive models. Gridboxes in white are those where no significant predictors are found. It should be noted that the spatial distribution of predictor density suggests that the method of selection might be susceptible to edge effects, that is predictors which have local significance towards the centre of the domain are more likely to pass the field-significance criteria, since those which have local significance towards the edge of the domain are more likely to be under-represented in terms of numbers of significant gridboxes. However, the domain is large enough to include more than just Western Europe (the study region), and the computational intensity of the predictor selection is too great to include a region substantially larger than the one given. Additionally, the gridboxes with high significant predictor counts are not exclusively located in the centre of the domain, suggesting that the edge-effect is not crucial in this case. Figure 7.2 shows the reduced predictor set for 95th percentile exceedance precipitation. The spatial and seasonal distributions are very similar to the 90th percentile predictands, as expected, although the spatial extent of the high predictor counts across most of the domain in spring is smaller. Table 6.1 and 6.2 in Chapter 6 provide the total numbers of retained predictors for precipitation and wind respectively, and it can be seen that for 95th percentile precipitation

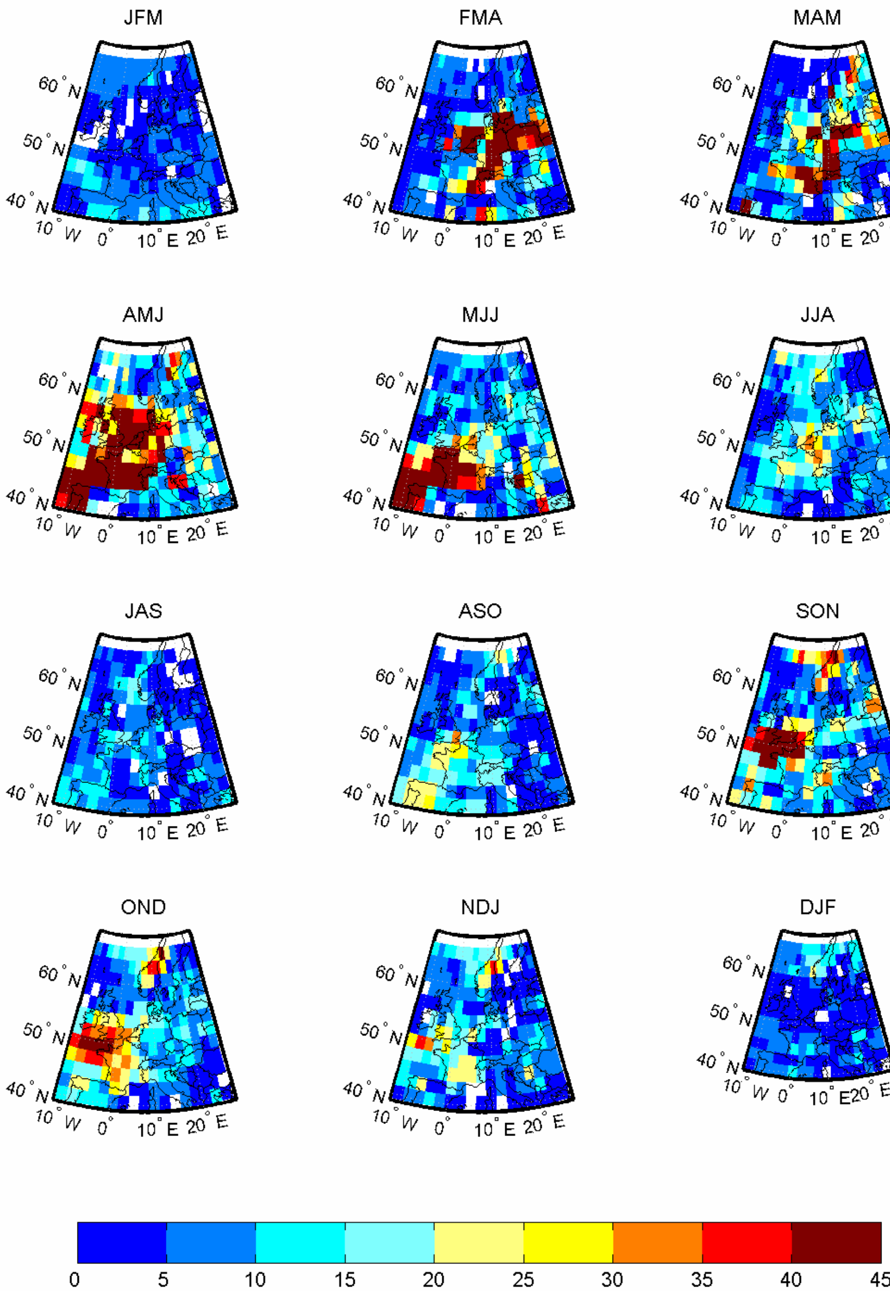

Figure 7.1 Number of predictors retained in the reduced predictor set for 90th percentile precipitation exceedance counts, by season.

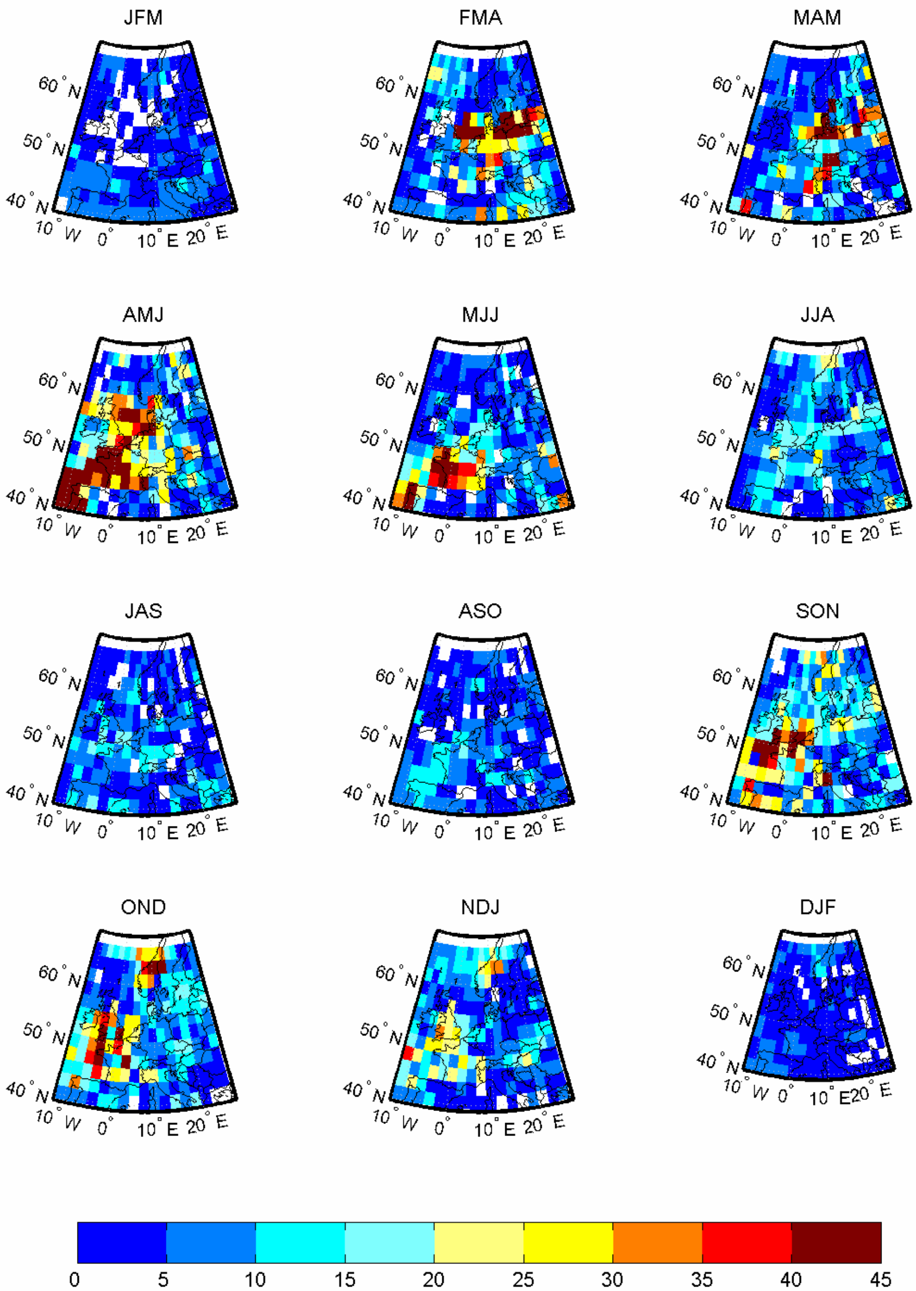

Figure 7.2 Number of predictors retained in the reduced predictor set for 95th percentile precipitation exceedance counts, by season.

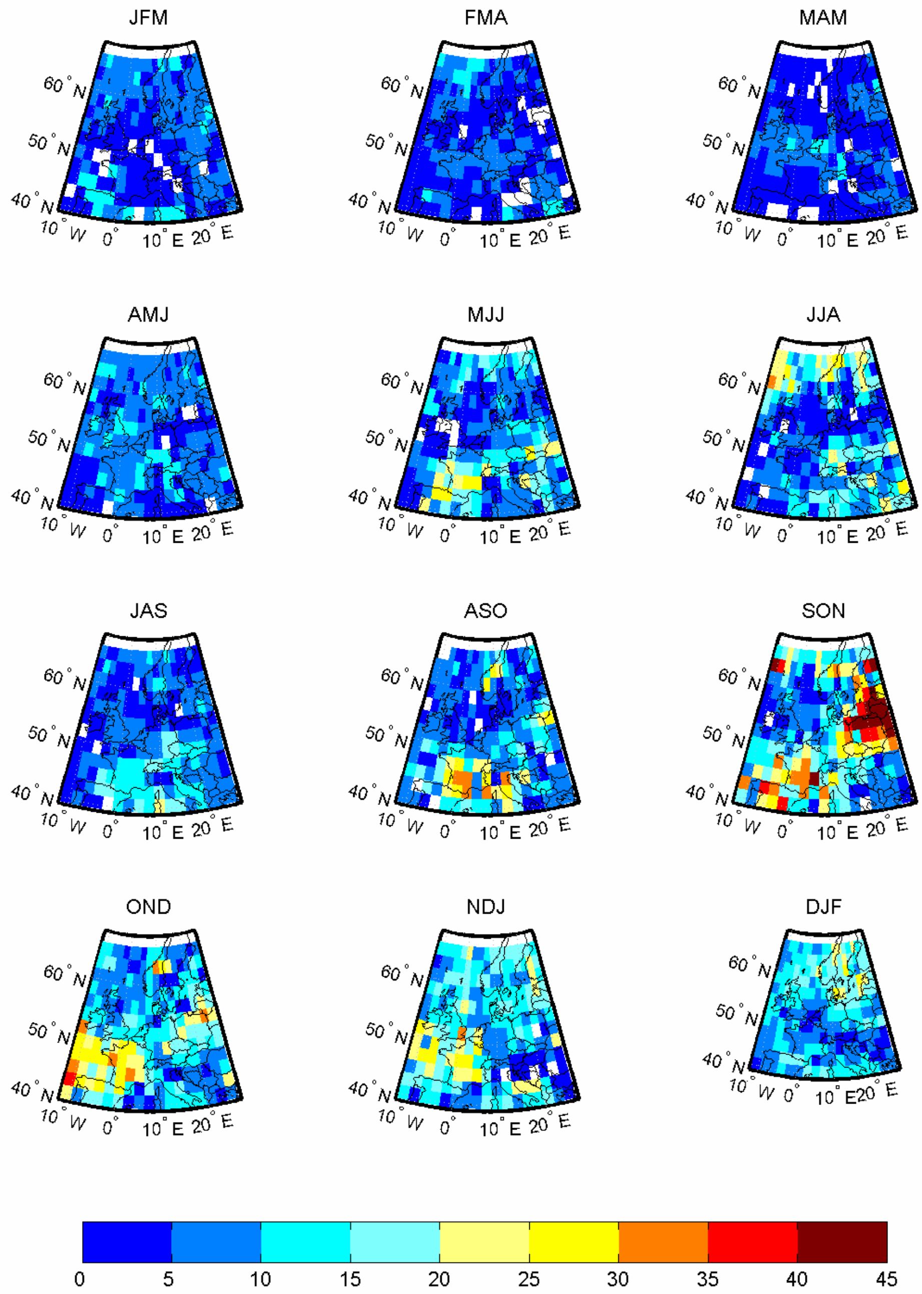

Figure 7.3 Number of predictors retained in the reduced predictor set for 90th percentile wind exceedance counts, by season.

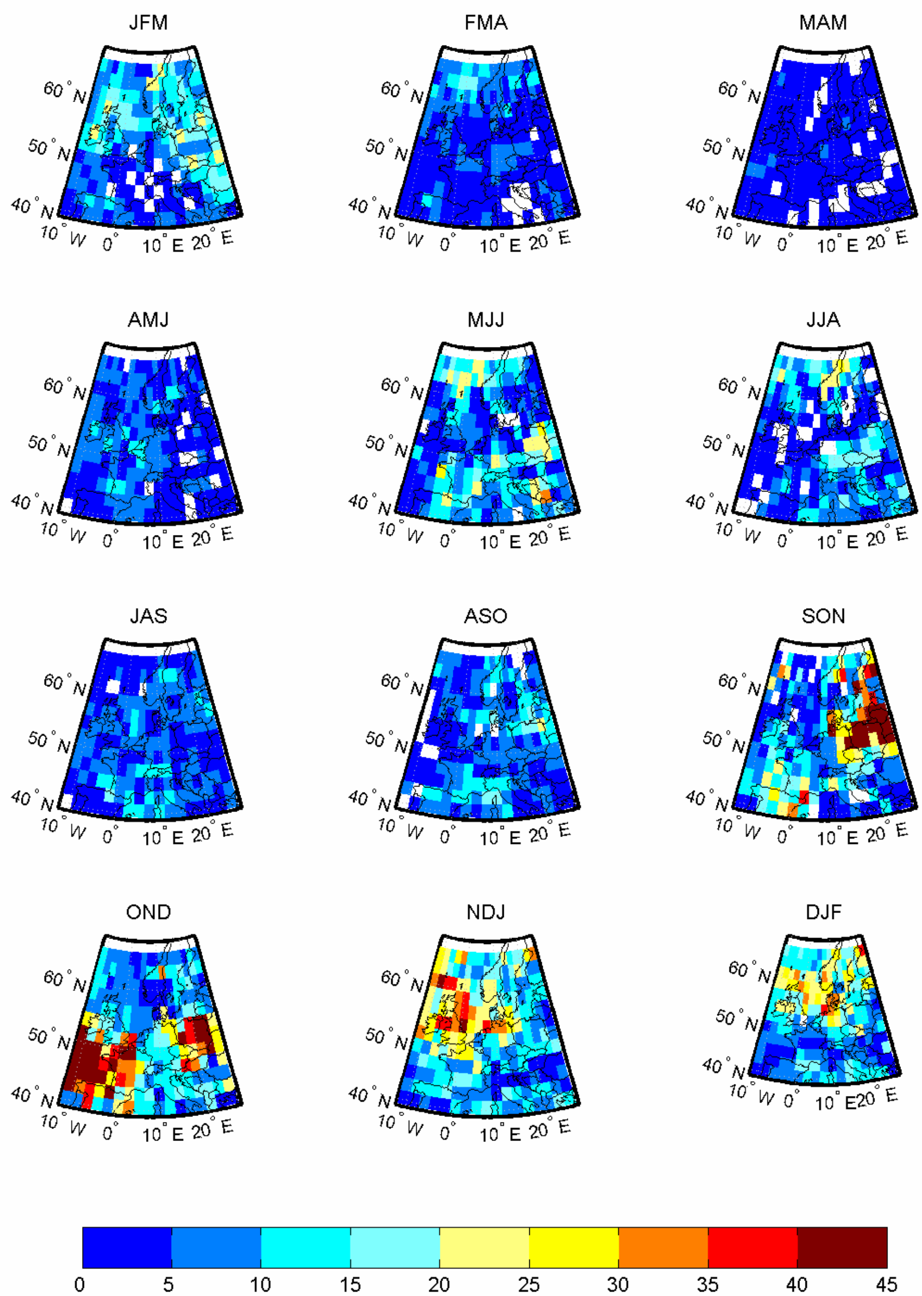

Figure 7.4 Number of predictors retained in the reduced predictor set for 95th percentile wind exceedance counts, by season.

exceedance counts there are overall fewer predictors retained in the spring-summer period, but in the autumn there are more. Figure 7.3 illustrates the same results for 90th percentile wind exceedance counts. In this case, there is a smoother annual cycle of significant predictor counts, with the late winter and spring typically having more gridboxes with low counts, and late summer through to early winter having more predictors. However, compared to precipitation, there are fewer predictors overall. In the autumn peak of predictor numbers, the spatial distribution of field-significant predictors bears some similarity with the precipitation predictors, with a high concentration of predictors in the south west of the domain. The pattern for the 95th percentile wind exceedance counts is similar spatially, and over the annual cycle. This is illustrated in Figure 7.4.

7.3.1 Model selection for 90th percentile precipitation exceedance counts

Figure 7.5 shows the R^2 values by location for each season of the 90th percentile precipitation exceedance models, and Figure 7.6 shows the MAE for the same models. Both diagnostics are shown, since R^2 gives an indication of the closeness of the fit with respect to the first difference of the observed and predicted timeseries at each gridbox, and the MAE gives some idea of whether there is any bias or substantial difference in the sharpness (variance) of the forecast compared to the observed series. Over the 12 seasons, a similar pattern emerges when compared with the predictor density counts in Figure 7.1 – that is, spring and autumn tend to have the largest concentrations of models with relatively high R^2 (and correspondingly low MAE) values, and the spatial distribution of the more skillful models also bears some resemblance. R^2 values describe the amount of variance explained by each model, and the values shown here range from zero to over 0.5 (50%). The latter is an unexpectedly high level of skill for European seasonal forecasting, and typically it might be assumed that the models are somewhat overfitted. This will be investigated further in chapter 8. In general, the distribution of skill displays some spatial smoothness, with homogenous regions of higher or lower skill beyond the gridbox scale, although by no means could the whole domain be described as displaying a smooth variation in skill. This is likely due to a number of reasons. Firstly, skill levels over the training period are generally marginal, and also the frequently large number of predictors retained up to the all-subsets selection stage described in section 7.2 means that there is potential volatility in the model selection, particularly given the low skill levels. Note that these factors apply to the models selected for all predictands, to a greater or lesser extent.

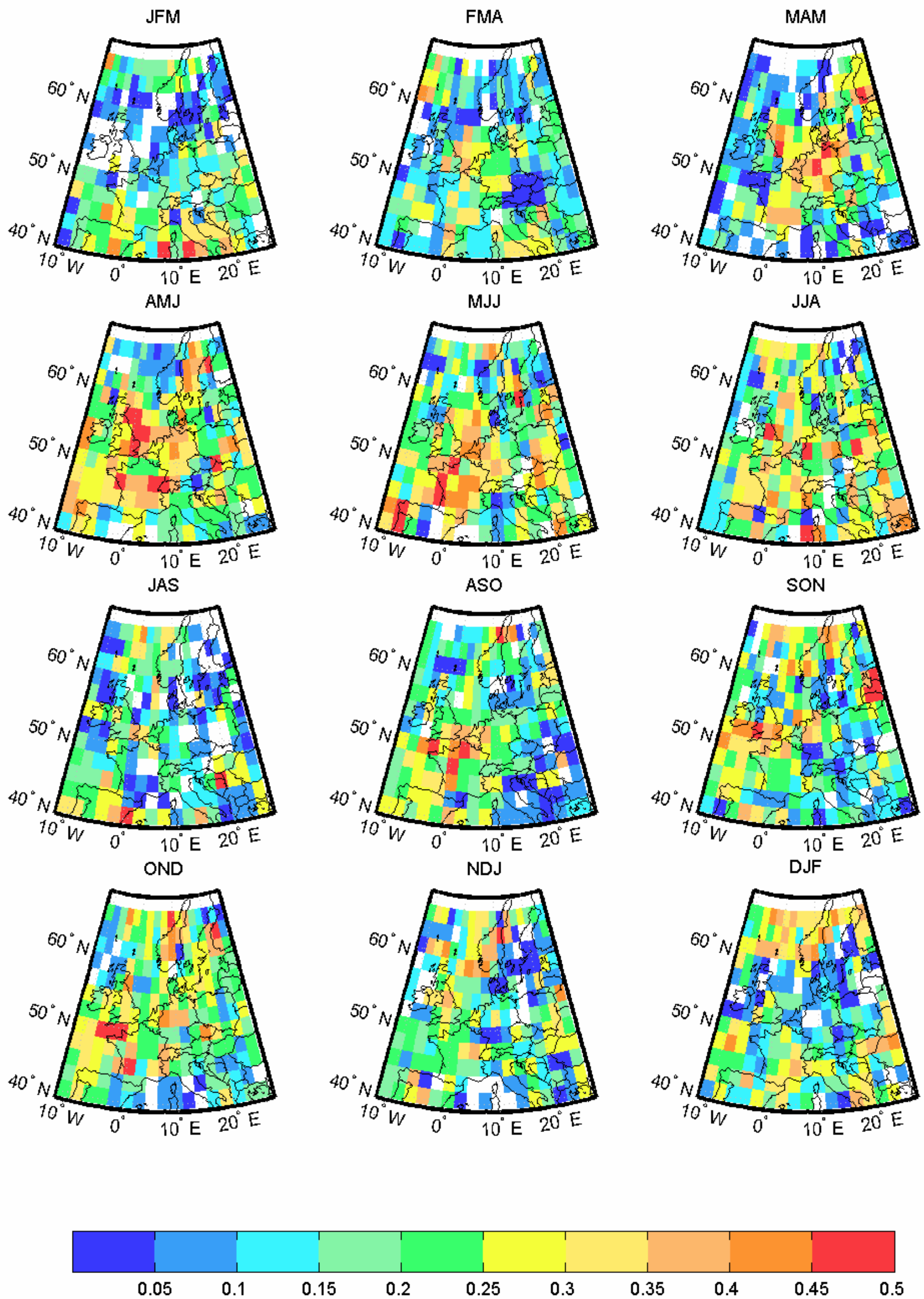

Figure 7.5 R^2 by season for 90th percentile precipitation exceedance counts. R^2 values are for the final model selected for each gridbox, based on the lowest cross-validation mean absolute error (MAE). Gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria.

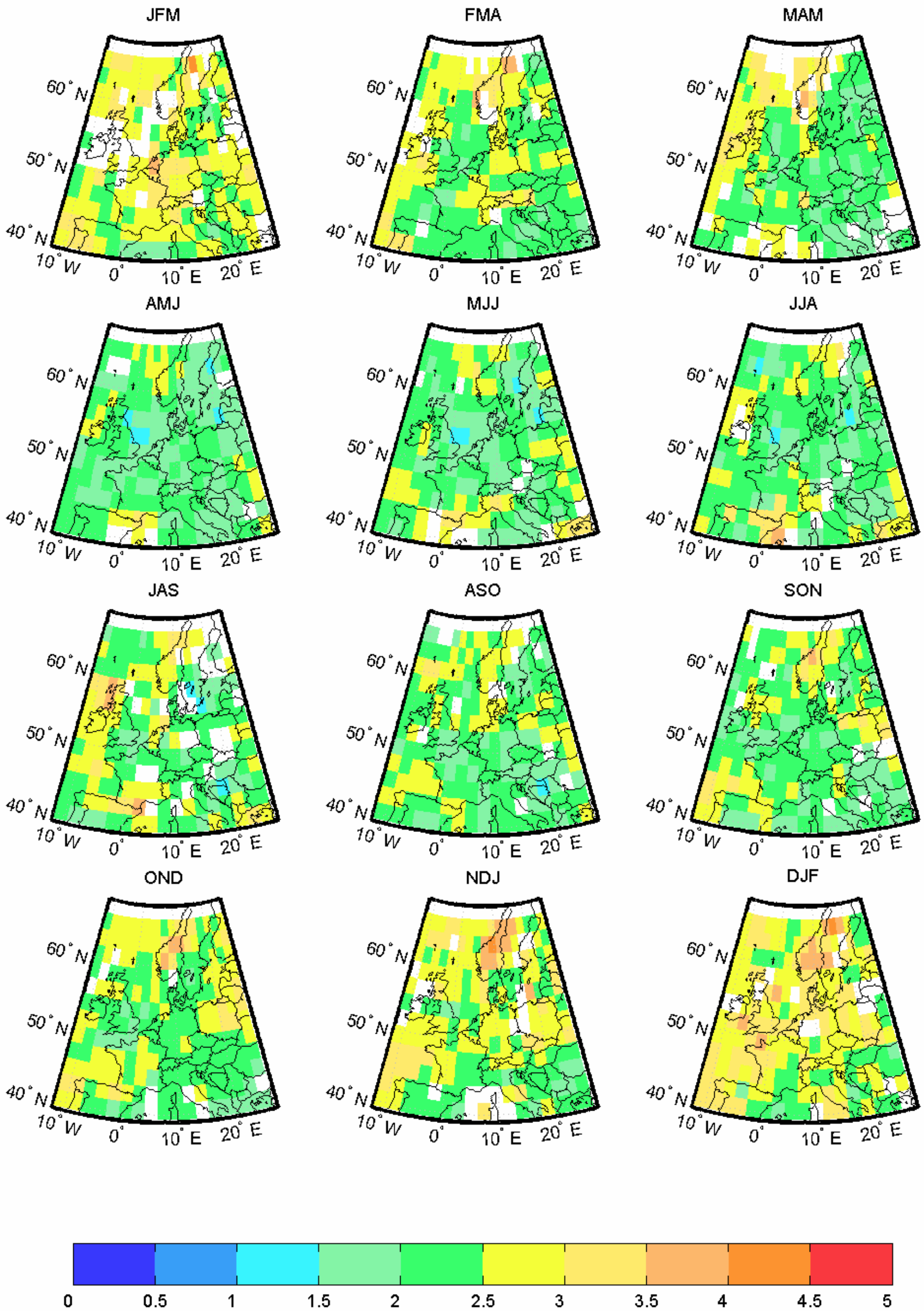

Figure 7.6 Mean Absolute Error (MAE) by season for 90th percentile precipitation exceedance counts. Values are for the final model selected for each gridbox, having the lowest cross-validation mean absolute error (MAE) of all predictor subsets. As Figure 7.5, gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria.

90th Percentile Precipitation

	JFM		**FMA**		**MAM**	
	SST4 Lead 03	54	EPN Lead 04	39	T150 P1 NDJ	35
	SOL Lead 06	38	SOL Lead 07	38	PV30 P2 ASO	25
	ATL P4 OND	33	H30 P1 NDJ	34	H100 P1 DJF	21
	TAR Lead 06	29	PCM P4 FMA	24	T30 P2 JAS	20
	SST3 Lead 04	20	DAR Lead 06	23	DAR Lead 07	18
Total		339		350		337
	AMJ		**MJJ**		**JJA**	
	H100 P1 SON	25	ATL P4 DJF	39	PCM P3 DJF	31
	H50 P1 NDJ	23	H50 P1 JFM	30	SCA Lead 05	29
	H100 P1 DJF	20	T30 P3 JAS	27	PV150 P1 SON	28
	IND P3 MJJ	17	PV150 P1 SON	23	PV150 P1 JAS	24
	EAP Lead 07	15	T30 P3 FMA	21	EAW Lead 02	21
Total		405		415		418
	JAS		**ASO**		**SON**	
	PV150 P1 JAS	45	PV150 P1 ASO	33	H50 P1 DJF	56
	PCM P5 MAM	30	PNA Lead 03	32	PV150 P2 OND	22
	PV100 P2 AMJ	28	PV30 P3 OND	24	EAW Lead 07	21
	SCA Lead 06	25	PV150 P2 OND	24	PV150 P1 JJA	21
	EAW Lead 07	19	SST4 Lead 02	22	QBO3 Lead 05	20
Total		324		392		406
	OND		**NDJ**		**DJF**	
	SOI Lead 02	41	EAP Lead 03	39	SST4 Lead 02	64
	PV30 P3 OND	40	H150 P3 JJA	33	TAR Lead 05	52
	ATL P2 OND	38	SOI Lead 03	31	SST3 Lead 05	32
	SHT Lead 10	20	AOS Lead 04	25	TNH Lead 11	24
	PV150 P1 JAS	18	AOS Lead 02	20	SST5 Lead 03	20
Total		406		370		376

Table 7.1 Top five predictors across whole domain for each season for the 90th percentile precipitation exceedance counts models. Totals given are for all predictors. Given a domain comprising 252 gridboxes, this includes a mix of gridboxes having models with one or two predictors, and some gridboxes with no viable model.

Table 7.1 shows the top five predictors for the whole domain, for each season of the 90th percentile precipitation models. Typically, the most commonly selected predictors are present in 20-50 gridboxes within the domain. However, each season for each predictand typically has a large number of predictors, and not all of them can be discussed in detail here. Several predictors are of particular interest in this sample. Some of these are illustrated in Figure 7.7 and are described in some more detail below. Local SST anomalies play a role in models in the winter season – this is at odds with some of the literature on this subject. For example, Hurrell *et al* (2006) postulate that local SST anomalies might be important for predictability outside the winter season, relating in particular to the frequency and intensity of convective precipitation events. For SST region 4 (from the west coast of Ireland to the Bay of Biscay) November anomalies are widely selected as explaining variability in the DJF and JFM seasons. The models using this predictor relate to gridboxes located over the Bay of Biscay itself, and also around the Iberian Peninsula. There is a corresponding response of the opposite sign to the north of the domain, with the whole comprising a dipolar response reminiscent of an NAO-like pattern. It is certainly possible that the local SST anomalies are partly a response to NAO forcing, although the correlation between the local SST and the NAO during the preceding months is weak – as might be expected during the late summer. However, it is interesting to note that climatologically, SSTs in these regions are at their warmest in the autumn, and they may be most inclined to influence the atmosphere as a result – possibly through persistence, rather than dynamical processes. SST region 3 (Baltic Sea) similarly explains variability in DJF and JFM, this time due to September and October anomalies respectively.

Indices associated with ENSO result in some skill from the Autumn (OND) through to spring (MAM), with the atmospheric indices having more importance than the oceanic ones. In OND and NDJ the September SOI is negatively related to precipitation over the English Channel region, and has a positive fit with precipitation over Scandinavia. From DJF through to MAM, indices derived from MSLP at Tahiti and Darwin are more frequently selected, with lead times up to seven months from the mid-month of the predictand season – i.e. August and September MSLP indices. From DJF through MAM, the responses are more widely spread through Western Europe, with negative responses to the Tahiti indices, and positive responses to the Darwin indices. Typically the Tahiti responses are more spatially coherent, and stronger. It is interesting to note that Jia et al (2008) find a link between the AO and tropical Pacific SSTs, with a three month lead time. Moreover, they find that the western tropical Pacific tends to exert relatively more influence on the North Atlantic, while the eastern tropical Pacific affects the north Pacific, and PNA pattern to a greater extent. The greater significance of the responses to Tahiti pressure anomalies may be connected with this – although unlike Jia et al (2008), in general atmospheric, rather than oceanic influences from the tropical Pacific are found to be more important in this study.

Some stratospheric indices are also selected frequently. Notably, indices associated with geopotential height have some skill from DJF through to MJJ, typically at two to three months lead time. There does not appear to be a single preferred metric in this case – that is to say, of the four pressure levels and three PC modes considered, none is dominant, although the first PC tends to be more commonly selected. Note that the dominant mode numbers are somewhat interchangeable from one season to the next, and at different pressure levels.

Stratospheric temperature and potential vorticity are also frequently selected, particularly in the summer months. This is of interest since variability in summer extreme precipitation is closely associated with variability in the position and strength of the Atlantic jet stream – for example, the UK floods of summer 2007 were associated with an anomalous southward displacement of the jet-stream. The causes of this displacement are uncertain, but there is speculation that they may be related to the La Nina event developing at the time (for example see www.walker-institute.ac.uk/news/summer_2007.pdf). Other notable responses include the NAO in JFM, which has skill as a predictor at an 11 month lead time across a large portion of central and Western Europe, from the English Channel through Poland and the Ukraine. Recent work by Parker *et al*. (2007) identifies potential predictors for the NAO including May North Atlantic SSTs, which are used operationally in UK Met Office seasonal forecasts of the NAO, with appreciable levels of skill above climatology. Some of the other Northern Hemisphere teleconnection indices are also selected over large regions of the domain, including the Scandinavian (SCA) pattern, which shows a positive fit to the predictand over parts of northern Europe in summer seasons, from AMJ to JAS, at lead times of several months, but typically involving the March SCA index.

In general, the spatial distribution of predictors selected in the final models is rather scattered. This is likely due to both the potentially large number of cross-correlated predictors, so that models in neighbouring gridboxes might have different, but highly correlated predictors, and the relatively weak fit of the predictors to the predictand data. An additional factor may be the spatial distribution of the rainfall extremes themselves.

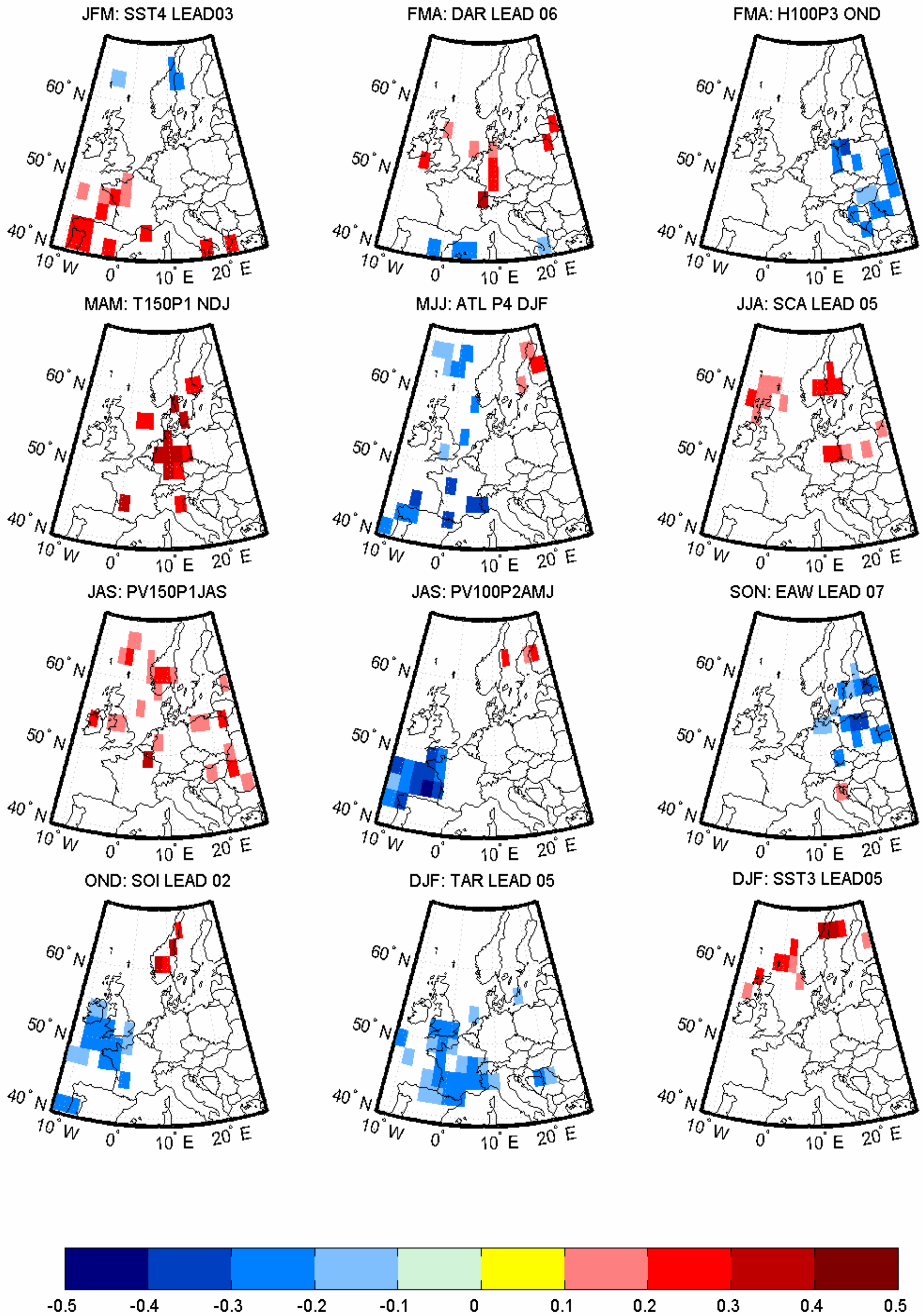

Figure 7.7 Spatial configuration for a sample of the most commonly selected predictors for 90th percentile precipitation exceedance counts. For each map, the title gives the predictand season, followed by the predictor, and the gridbox shading illustrates the coefficient for that predictor in the model.

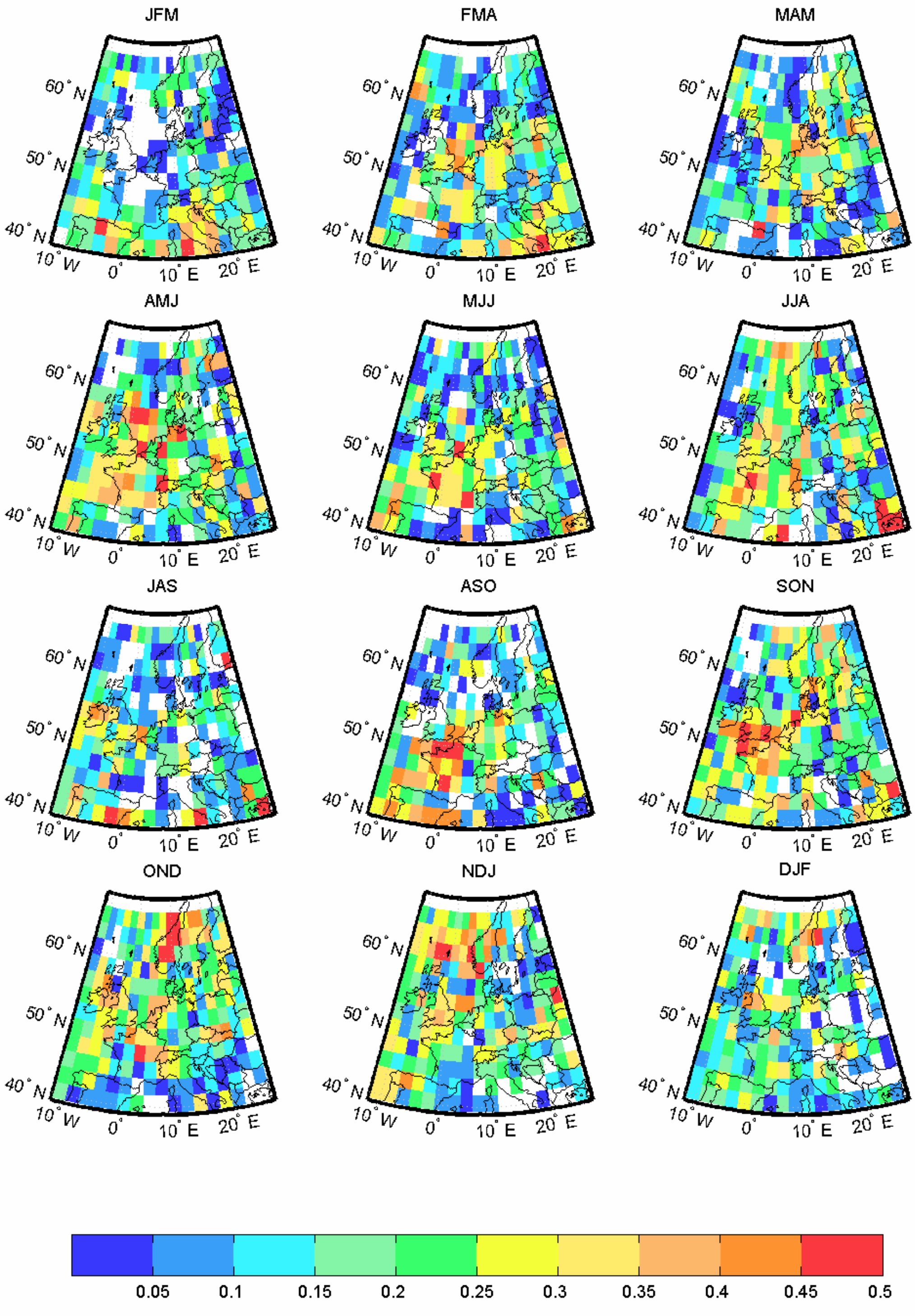

Figure 7.8 R^2 by season for 95th percentile precipitation exceedance counts. R^2 values are for the final model selected for each gridbox, based on the lowest cross-validation mean absolute error (MAE). Gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria.

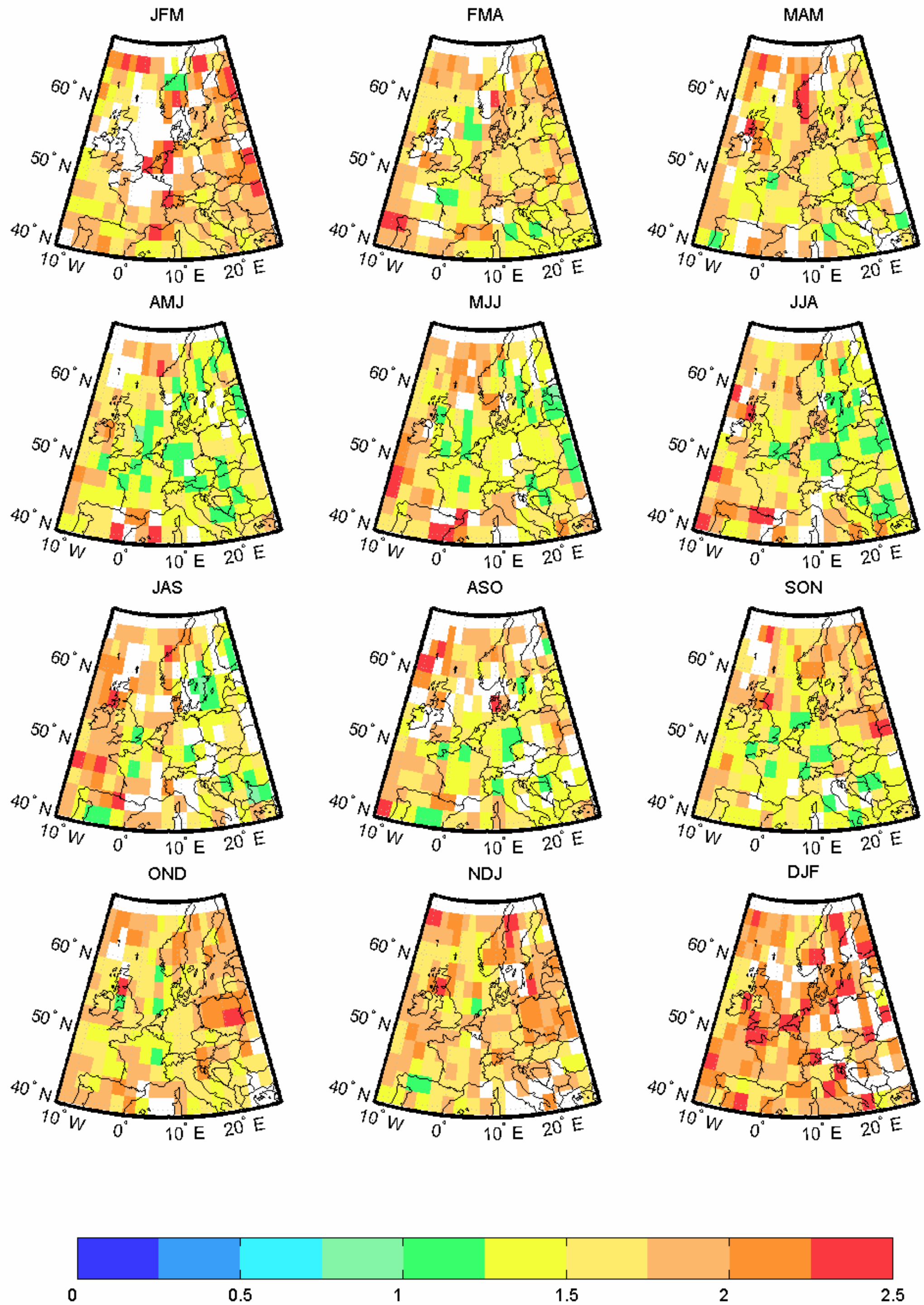

Figure 7.9 Mean Absolute Error (MAE) by season for 95th percentile precipitation exceedance counts. Values are for the final model selected for each gridbox, having the lowest cross-validation mean absolute error (MAE) of all predictor subsets. As Figure 7.5, gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria. Note that the scale is different to that for the 90th percentile MAE values in Figure 7.6, as appropriate for the smaller absolute values of the 95th percentile exceedance counts.

7.3.2 Model selection for 95th percentile precipitation exceedance counts

Figure 7.8 shows the R^2 values for the models fitted to the 95th percentile predictand by season. There is less of a pronounced seasonal cycle than for the 90th percentile models, with the DJF, JFM, and JAS seasons showing the least skill, and having the greatest number of gridboxes without any viable model. Generally, there is a larger number of gridboxes where no skill is identified. This is also evident from the totals provided in Table 7.2. From Figure 7.9 it can be seen that similar to the 90th percentile models, the MAE for the models is smallest in the late spring and summer seasons, having a more noticeable seasonal cycle than the R^2 metric. The peak MAE values tend to occur from autumn through to late winter, corresponding with the lower R^2 values. The JFM season contains a large region in the centre of the domain where no viable models are found.

95th Percentile Precipitation

	JFM		**FMA**		**MAM**	
	SST4 Lead 03	38	SOL Lead 07	35	H50 P1 DJF	26
	SOL Lead 06	36	SOI Lead 06	31	IND P5 NDJ	19
	T50 P3 MJJ	23	PCM P4 FMA	29	SOL Lead 07	17
	ATL P4 OND	20	H150 P3 SON	25	T30 P2 JAS	16
	EAP Lead 03	19	EPN Lead 04	18	H100 P3 SON	15
Total		283		355		314
	AMJ		**MJJ**		**JJA**	
	H100 P1 NDJ	37	T30 P3 JAS	28	IND P5 NDJ	30
	EAW Lead 06	20	ATL P4 DJF	27	PV150 P1 SON	28
	HTD Lead 03	17	H30 P1 JFM	25	PCM P3 DJF	21
	H30 P1 JFM	16	T30 P3 FMA	21	EAW Lead 06	19
	SCA Lead 02	12	PV150 P1 SON	18	SCA Lead 05	17
Total		373		342		372
	JAS		**ASO**		**SON**	
	PV150 P1 ASO	33	T150 P1 DJF	31	H150 P3 NDJ	33
	PV100 P2 AMJ	25	PNA Lead 03	27	SHT Lead 10	31
	PCM P5 MAM	23	PV30 P3 OND	20	T150 P1 DJF	29
	SCA Lead 06	22	PV150 P2 OND	19	DAR Lead 04	26
	DAR Lead 09	17	PV150 P3 NDJ	19	DAR Lead 07	20
Total		308		309		399
	OND		**NDJ**		**DJF**	
	SOI Lead 02	35	H150 P3 JJA	27	SST4 Lead 02	65
	ATL P2 OND	25	SOI Lead 03	25	TAR Lead 05	43
	T30 P3 AMJ	24	POL Lead 07	20	TNH Lead 11	29
	PV30 P3 OND	21	EAP Lead 03	18	SST1 Lead 02	27
	ATL P2 JFM	18	SDNH Lead 06	17	ATL P4 SON	25
Total		386		376		326

Table 7.2 Top five predictors across whole domain for each season for the 95th percentile precipitation exceedance counts models. Totals given are for all predictors. Given a domain comprising 252 gridboxes, this includes a mix of gridboxes having models with one or two predictors, and some gridboxes with no viable model.

This follows from the small number of field-significant predictors selected for this region in JFM, and can be seen in Figure 7.2. Note that the scale for the MAE values in Figure 7.9 is half that of Figure 7.6 (for the 90th percentile models), to correspond with the lower exceedance counts, and overall the MAE is higher relative to the absolute exceedance counts for the 95th percentile precipitation than it is for the 90th percentile. It is possible that this apparently lower predictive skill within the model training period is due to some inherently less predictable property of more extreme precipitation. However, it must also be considered that the statistical properties of the timeseries of exceedance counts lend themselves to lower skill metrics, and larger relative errors in the forecast.

Table 7.2 shows the top five predictors by numbers of gridboxes for which they are selected for the 95th percentile precipitation exceedance counts. There are considerable similarities with the predictors selected for the 90th percentile precipitation models, as might be expected. In particular, the local SST predictors in the DJF and JFM seasons are the same, with a similar spatial manifestation and despite the low skill exhibited for these seasons. Also, the SOI predictors are similar, and have a considerable influence in a number of seasons, although they become more important earlier in the autumn seasons than they do for the 90th percentile predictand. A sample of the most spatially important predictands is shown in Figure 7.10. Another predictor to feature substantially in both the 90th and 95th percentile precipitation predictands in JFM is the fourth PC of Atlantic SST for the preceding OND season, which is selected across a large portion of the Mediterranean region. The Solar flux index from the preceding September and October also features in the JFM, FMA and MAM seasons for both precipitation predictands, showing a dipolar fit with one pole centred over south-eastern Europe and a more scattered response across the north of the domain. This is one of several predictors where the spatial response is broadly aligned to the canonical NAO/AO pattern (and the major mode of large-scale precipitation variability) across Europe. The spatial pattern of predictive skill is very patchy into the Spring and Summer, and the main region of coherent skill from AMJ to JJA is located over France and the Bay of Biscay. Although there is little skill here in JAS, this area of greater skill reappears in the ASO and SON seasons. There is no single dominant predictor or family of correlated predictors driving this response, however some of the following predictors do contribute. Indian Ocean SST predictors feature in a number of seasons, including the fifth PC during NDJ, which is selected in MAM, over a large region to the north of the domain, and JJA, where the response is over France and the UK. The February index of Scandinavian teleconnection pattern features in the JJA and JAS seasons, where it is selected over an arc stretching from eastern Europe across to the North Sea and northern Scotland. Throughout the summer and autumn, a number of stratospheric predictors are selected across large parts of the domain, including the third PC of stratospheric temperature during MAM and AMJ. These responses are illustrated in Figure 7.10.

In the Autumn and early winter, indices associated with the SOI, typically at lead times of two to four months become more prominent. In ASO and SON, the June Darwin pressure indices are selected over the southern and south-western domain, and in OND and NDJ the September SOI is associated with a negative response over western France and the southern UK, and a positive response over Scandinavia. That is to say, negative SOI (El Nino) events are associated with more frequent precipitation extremes over France, and fewer events over Scandinavia. The NDJ season responds to a number of predictors besides the SOI, and including the Northern Hemisphere snow extent in June – which is rather earlier than might be expected given the absolute extent of the snow cover at this time of year. There are also notable responses to some of the Northern Hemisphere teleconnection patterns, including the May Polar Eurasian pattern, the September East Atlantic pattern, and the Tropical Northern Hemisphere (TNH) pattern from the preceding January. There is also a response to the AO, at two and four month lead times, although these do not follow closely the canonical pattern for this mode, comprising small and scattered mostly positive responses, over Scandinavia and the Baltic states, with a small negative response over the Mediterranean from the four month (September) AO index. The TNH index (this time from the preceding February) becomes increasingly prominent as a predictor in DJF. As well as this, a number of local SST predictors seem to be important at lead times of two to five months. Two other notable predictors in this season are the July Tahiti MSLP index, which is associated with a coherent negative response in precipitation events over south western Europe, and the Western Pacific pattern, also in July, which gives rise to a large region of positive responses, stretching from northern France to the Baltic.

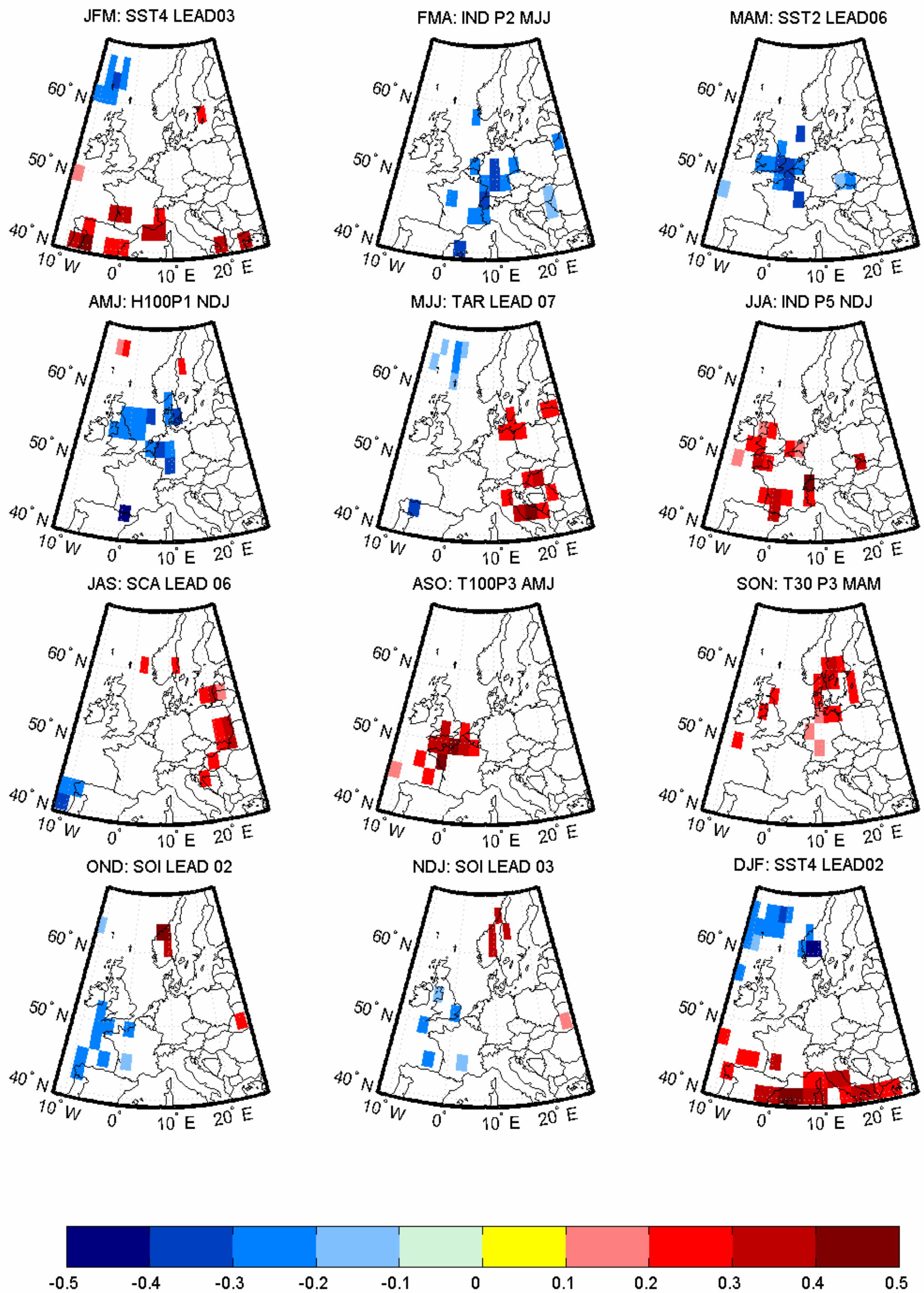

Figure 7.10 Spatial configuration for a sample of the most commonly selected predictors for 95th percentile precipitation exceedance counts. For each map, the title gives the predictand season, followed by the predictor, and the gridbox shading illustrates the coefficient for that predictor in the model.

7.3.3 Model selection for 90th percentile wind exceedance counts

Compared to the precipitation predictands, the model selection process for the 90th percentile wind exceedance counts results in spatially smoother levels of skill. There is also a more pronounced annual cycle in the values of R^2. These details are illustrated in Figure 7.11. The R^2 values for JFM are generally low in the south west of the domain, and increase diagonally to levels in excess of 0.4 in the north east, with some areas where no significant model is identified. Throughout FMA and MAM R^2 levels are generally low across the domain, typically with values in the 0.1 to 0.2 range. From AMJ, they tend to increase in patches throughout the summer, before increasing more notably in SON through to DJF. In these last four seasons, R^2 frequently attains levels in excess of 0.4, and much of the domain has levels greater than 0.25. In particular, the north of the domain in the winter seasons of NDJ and DJF exhibits the highest levels of skill, and these seasons have very few gridboxes where no viable model is selected. Through comparison with Figure 7.3, showing the predictor density for 90th percentile wind exceedance counts, it can be seen that the areas if highest skill in NDJ and DJF correspond with gridboxes having a large number of predictors retained in the reduced subset – generally in excess of 15 predictors – although the regions with the highest number of retained predictors towards the south of the domain, do not have such high skill levels. Figure 7.12 illustrates the corresponding MAE values for these models. A clear annual cycle is observed, although it is at odds with the annual cycle in R^2, having generally lower values in the summer, and higher values in the winter. From a purely statistical point of view, this indicates that although the fit with respect to the first difference of the observed and predicted timeseries – in other words the interannual variability – may be better in the winter, overall the errors are larger. However, when it is considered that in general the interannual variability of these predictand indices in the summer is lower than for the winter seasons, it becomes plausible that this may be a valid climatological feature of the observed skill. Indeed, the absolute values of MAE are almost universally lower across the whole domain and in every season, when compared to the values for the precipitation predictors – seldom exceeding values of 2.5.

Table 7.3 shows the top five predictors ranked by the number of gridboxes in which they are selected, for the 90th percentile wind predictand. One feature that is apparent compared to the 90th percentile precipitation models, is that overall the number of predictors is greater from the late summer (JAS) through to JFM, with the exception of ASO. This corresponds with a smaller number of gridboxes for which no statistically significant model is found in these seasons. The spatial patterns of selected predictors are show in Figure 7.13 and Figure 7.14. A larger number of predictors are illustrated than for precipitation, since a greater number have a spatially coherent relationship with the predictand, and it is useful to illustrate some of the more persistent relationships.

Indices associated with ENSO are common from OND through to JFM, but are very infrequently selected outside these seasons. Local SST anomalies are much less prevalent than for the precipitation predictands. This follows reasonably since local SSTs are unlikely to affect the dynamics of the large-scale circulation – which drives wind extremes – beyond having some effect on levels of evaporation, and hence possibly affecting precipitation extremes. However, the exception to this case is the apparently widespread fit between Baltic SSTs in August, and wind extremes over the north west of the domain in DJF and to a lesser extent in JFM. This is illustrated in Figure 7.14 for DJF.

The set of Northern Hemisphere teleconnection indices (after Barnston and Livezey, 1987) feature prominently, including the NAO/AO. This is present (mostly as the AO) in the top five predictors in six seasons throughout the year, and in almost all other seasons to some extent, at lead times of up to seven months. Interestingly, the response is almost invariably positive in the north of the domain, and negative in the south, and seldom does the response comprise both of these dipoles in the same season for the same lead time. The fact that the response is positive implies that the AO displays some persistence at timescales of several months and although this process does not have an established physical explanation, the observed link here with wind extremes at relatively long lead times is notable. The East Atlantic Pattern (EA – or EAP as given in the predictor nomenclature here) is important from OND through to AMJ, with the September EAP being selected from OND through to JFM, and then for FMA, MAM and AMJ the October, February and June indices are selected, respectively. The spatial pattern of this response is at its most widespread in the JFM season, and is shown in Figure 7.13 for JFM, comprising a positive response, across the north of the domain. The OND to DJF responses are similar, but the predictor is selected for fewer gridboxes. The FMA response to October forcing consists of a large region of strongly negative responses to the west of the domain, from the Bay of Biscay to the Atlantic west of

Scotland. This is replicated to a much lesser extent in MAM and AMJ, where the main feature is a weak negative response to the west, and a slightly stronger positive response in the south east of the domain, over the Mediterranean. The other notable instance where NH teleconnection patterns are important is MAM, where the Western Pacific pattern (WPP) is also associated with a large positive response in the north of the domain – illustrated in Figure 7.13.

90^{th} Percentile Wind

	JFM		**FMA**		**MAM**	
JFM	TAR Lead 06	56	IND P4 ASO	57	WPP Lead 02	75
	EAP Lead 05	40	SOL Lead 05	50	EAP Lead 02	35
	AOS Lead 04	36	PCM P5 AMJ	27	T150 P1 NDJ	20
	SST3 Lead 06	35	SDNH Lead 09	22	QBO3 Lead 11	17
	SOL Lead 06	33	T150 P3 ASO	20	T30 P2 JAS	17
Total		341		304		306
	AMJ		**MJJ**		**JJA**	
	AOS Lead 02	32	T30 P3 FMA	55	EAW Lead 02	37
	SST6 Lead 04	31	T100 P3 FMA	34	T30 P3 MAM	31
	H50 P1 JFM	30	T30 P3 JJA	32	AOS Lead 04	29
	SDEU Lead 07	26	NAO Lead 02	30	PCM P3 JFM	27
	SDNG Lead 05	25	PCM P3 DJF	26	POL Lead 09	24
Total		359		413		383
	JAS		**ASO**		**SON**	
	AOS Lead 05	37	H50 P1 DJF	39	SHT Lead 10	40
	NAO Lead 07	37	T150 P1 NDJ	34	SHT Lead 04	25
	SDNG Lead 11	35	ATL P4 MJJ	32	H30 P1 DJF	24
	PV30 P2 ASO	26	ATL P5 AMJ	28	PV150 P1 JJA	23
	PCM P4 AMJ	25	SHT Lead 10	26	H100 P1 DJF	23
Total		372		371		417
	OND		**NDJ**		**DJF**	
	PCM P5 JFM	76	PCM P5 JJA	54	SST3 Lead 05	64
	SDNH Lead 05	45	H150 P3 JAS	46	TAR Lead 05	64
	T50 P2 JFM	29	T50 P3 JAS	45	AOS Lead 03	50
	EAP Lead 02	28	EAP Lead 03	37	EAP Lead 04	44
	TAR Lead 02	27	T150 P2 ASO	31	T50 P3 SON	25
Total		440		423		462

Table 7.3 Top five predictors across whole domain for each season for the 90^{th} percentile wind exceedance counts models. Totals given are for all predictors. Given a domain comprising 252 gridboxes, this includes a mix of gridboxes having models with one or two predictors, and some gridboxes with no viable model.

Northern Hemisphere snow cover is much more important for this predictand than for the precipitation predictands. This is possibly related to the observed link between snow cover and the NAO/AO at lead times of several months, as documented in e.g. Cohen *et al.* (2001), which in turn may be more closely associated with wind extremes as observed here. Varying levels of fit between the snow cover predictors and 90^{th} percentile wind is present throughout the year, at a range of lead times from two to 11 months. During AMJ, the Eurasian snow cover index from the preceding October is selected as a predictor across portions of the southwest of the domain, having a positive fit to the predictand, as shown in Figure 7.13. In addition to this in AMJ the index describing North American and Greenland snow cover during December displays a positive fit with the predictand over central and eastern Europe, extending as far north as Scandinavia (not shown). The spatial pattern shown in response to this predictor does not clearly conform to the canonical NAO response, being centred somewhat too far to the north of the domain to be entirely a response to wind anomalies in southern Europe associated with a negative NAO (Trigo *et al.*, 2002). Other seasons with a notable response to snow cover indices include JAS, where positive North American and Greenland snow cover anomalies from the preceding September are associated with positive anomalies in the predictand from the Atlantic south of Ireland across to northern France and Belgium. A positive response to May North American and Northern Hemisphere snow cover is observed respectively in SON and OND, comprising a positive response across the domain centred on about 55N. The OND response is shown in Figure 7.14. Much of the literature on snow cover as a predictor for the NAO, or European climate focuses on the response to summer snow cover anomalies (e.g. Saunders *et al*., 2003), so the apparent importance of an autumn response to late spring snow cover is of interest – at least statistically speaking.

Pacific Ocean SST PC indices also feature strongly in the final model selection process, generally at relatively long lead times. These indices are present in models throughout the year, but predominantly in OND and NDJ, and to a lesser extent from DJF through to JAS. The first PC – relating to ENSO activity – does not feature in any of the models despite showing a significant response in OND and NDJ at the field-significance testing stage. In general, the European response to the Pacific Ocean here appears to be constrained to variability in the northern extratropical Pacific. The OND response to the fifth PC from the preceding JFM season is shown in Figure 7.14. It comprises a large region of positive responses off the Atlantic coast of Norway, and a smaller region of negative responses off the northwest coast of Portugal. The positive response to this predictor is similar although weaker in NDJ, and the negative response is no longer present. The fifth PC pattern in the Pacific in JFM has its dominant centre of correlation weightings over the North Pacific, to the west of Japan, with a negative horseshoe-shaped region surrounding it, and further coherent regions of covariance in the tropics, which do not appear to correspond with the spatial manifestation of ENSO variability. This pattern accounts for 5% of the overall variance in JFM. This pattern exhibits a weak (r = -0.32) negative correlation with the DJF PNA pattern, which implies that it is driven by the PNA, similar to the Atlantic tripole response to the NAO/AO.

Stratospheric indices also account for a considerable proportion of the predictors selected for the final models – in particular the stratospheric temperature indices, which are particularly prevalent in the summer and autumn, from MJJ to SON.

Other predictors of interest include the Southern Hemisphere temperature index, for which the November anomaly shows a positive fit with the predictand during SON over central Europe, and the May anomaly has a positive fit with the predictand over the North Sea and Scandinavia, as shown in Figure 7.14. Indian Ocean PC indices are also associated with a significant response, concentrated in FMA and MAM. In FMA, there is a negative response to the fourth PC from the preceding ASO season across the southern part of the domain. In MAM there is a response to the second PC in both AMJ and DJF. The AMJ response is negative, and spread across central and western Europe, while the DJF response is positive, and concentrated in the south of the domain, over the Balkan states and the Adriatic.

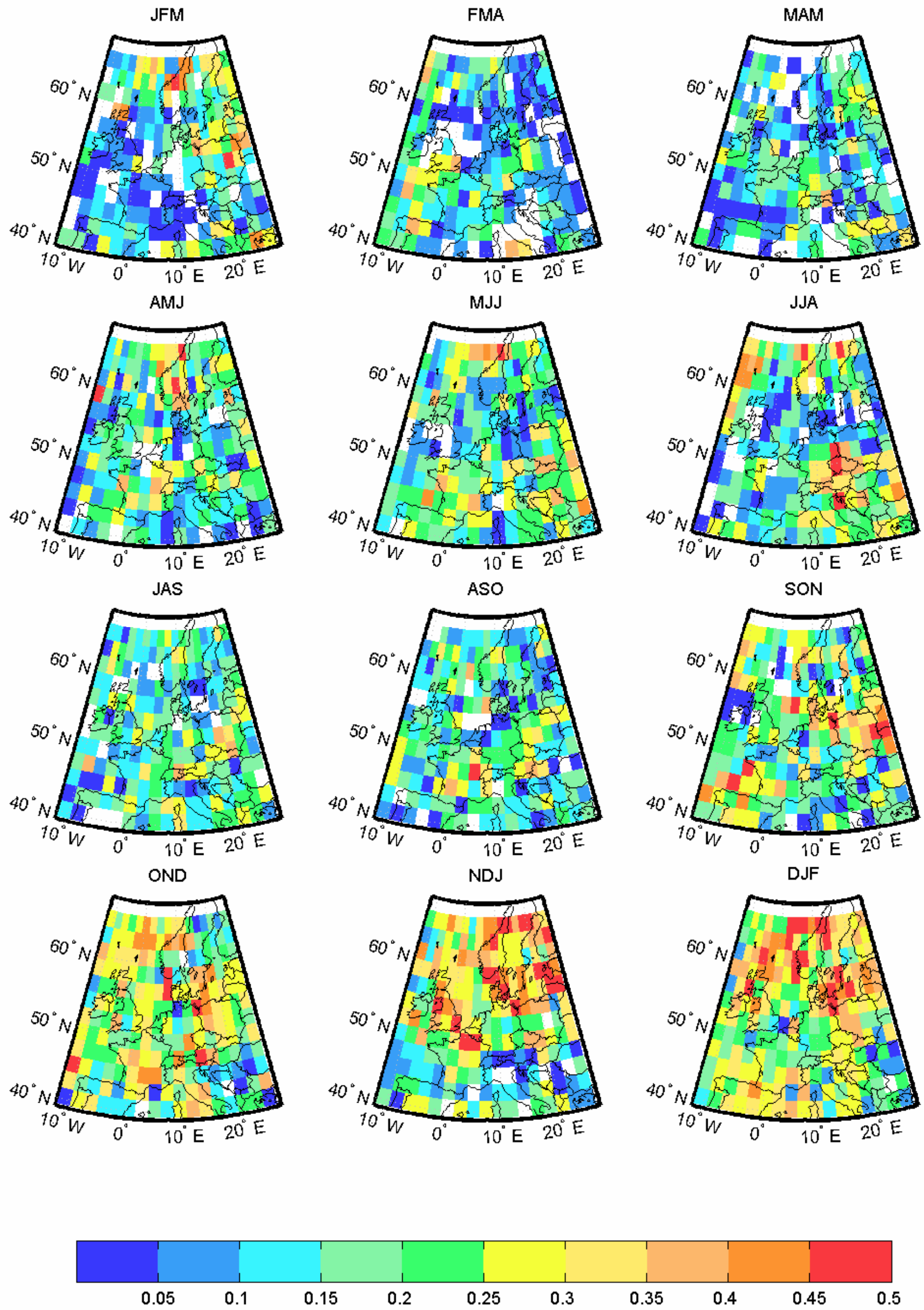

Figure 7.11 R^2 by season for 90th percentile wind exceedance counts. R^2 values are for the final model selected for each gridbox, based on the lowest cross-validation mean absolute error (MAE). Gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria.

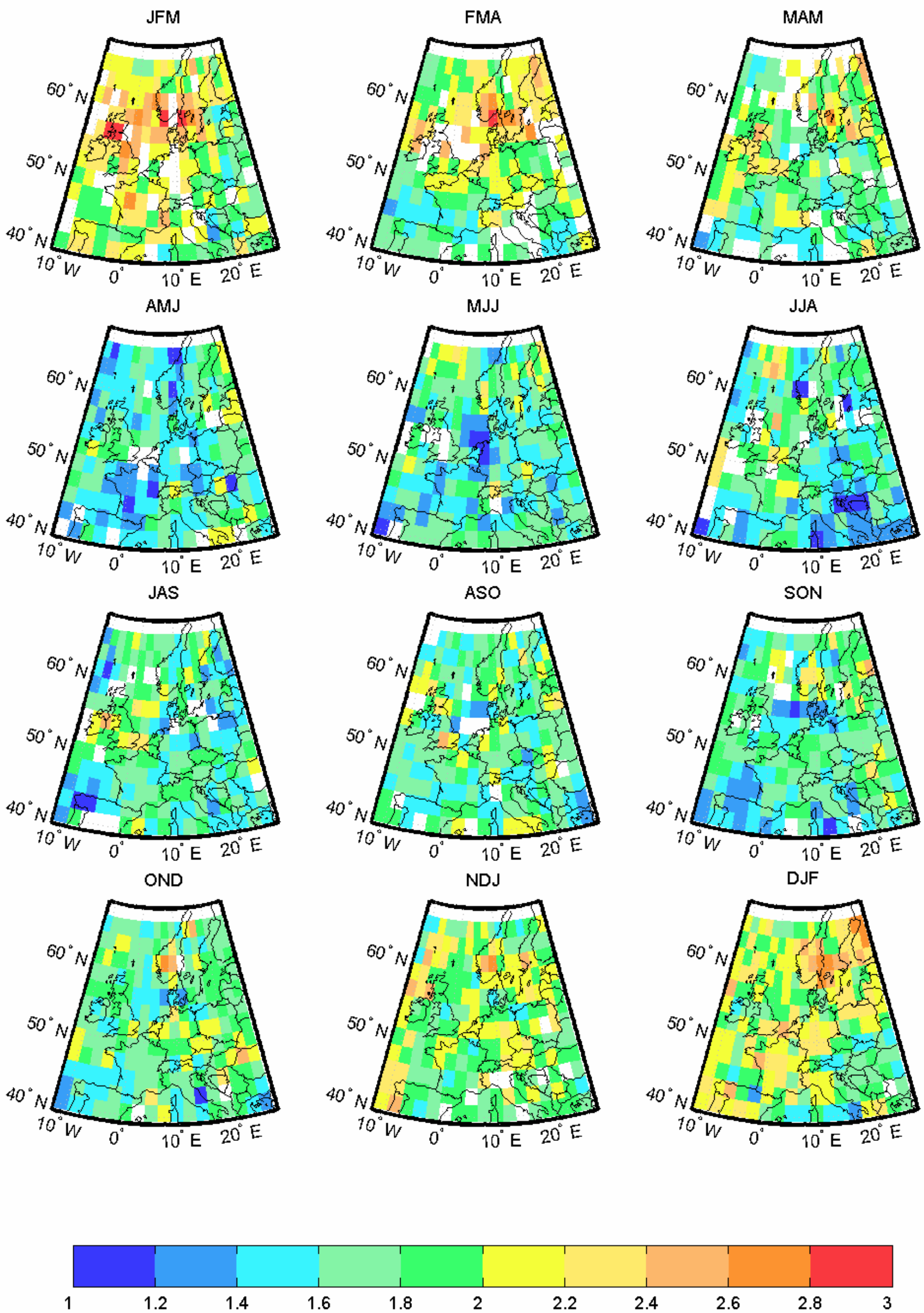

Figure 7.12 Mean Absolute Error (MAE) by season for 90th percentile wind exceedance counts. Values are for the final model selected for each gridbox, having the lowest cross-validation mean absolute error (MAE) of all predictor subsets. As Figure 7.5, gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria. Note that the scale is different to that for the precipitaion MAE.

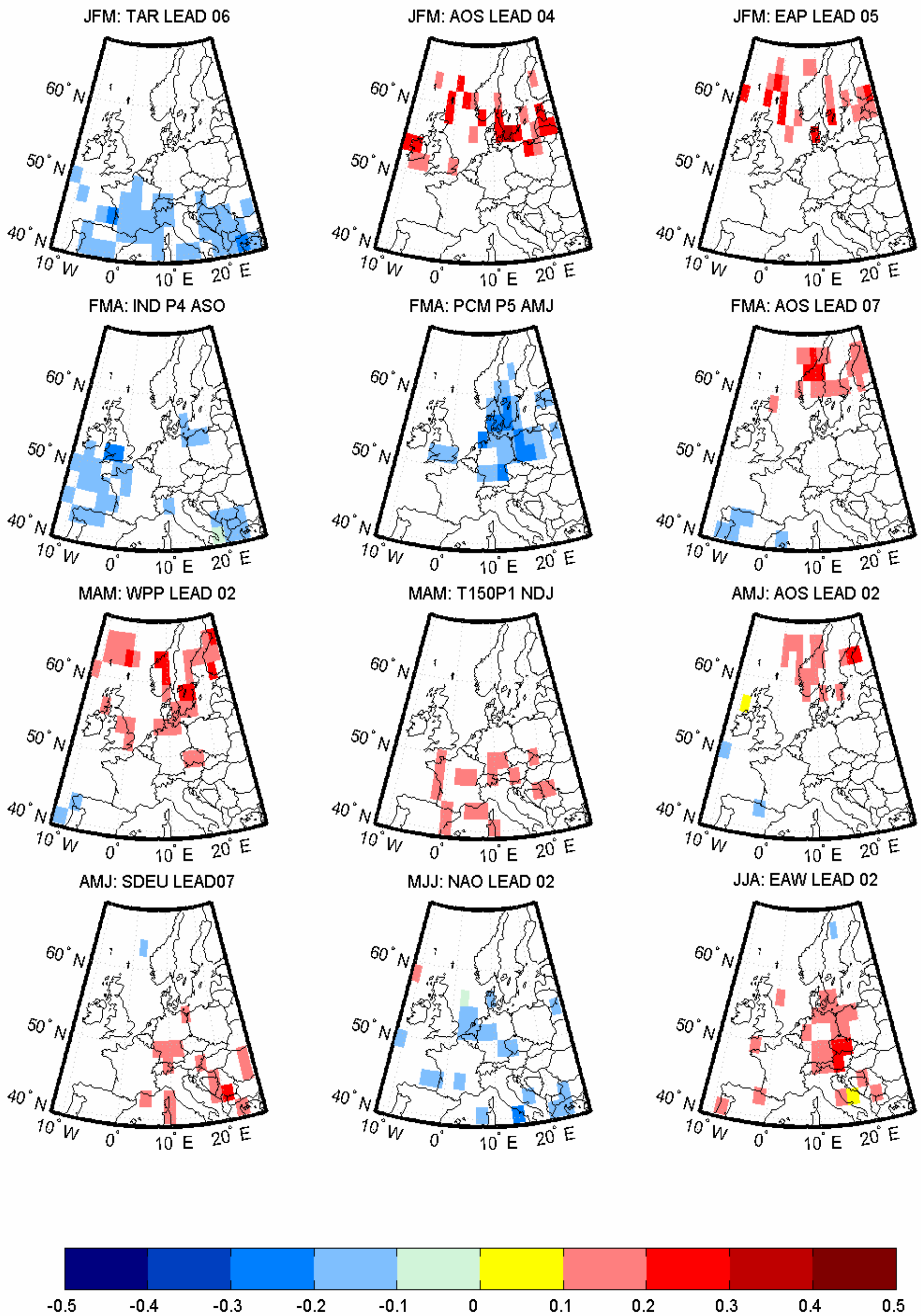

Figure 7.13 Spatial configuration for a sample of the most commonly selected predictors for 90th percentile wind exceedance counts, from JFM to JJA. For each map, the title gives the predictand season, followed by the predictor, and the gridbox shading illustrates the coefficient for that predictor in the model.

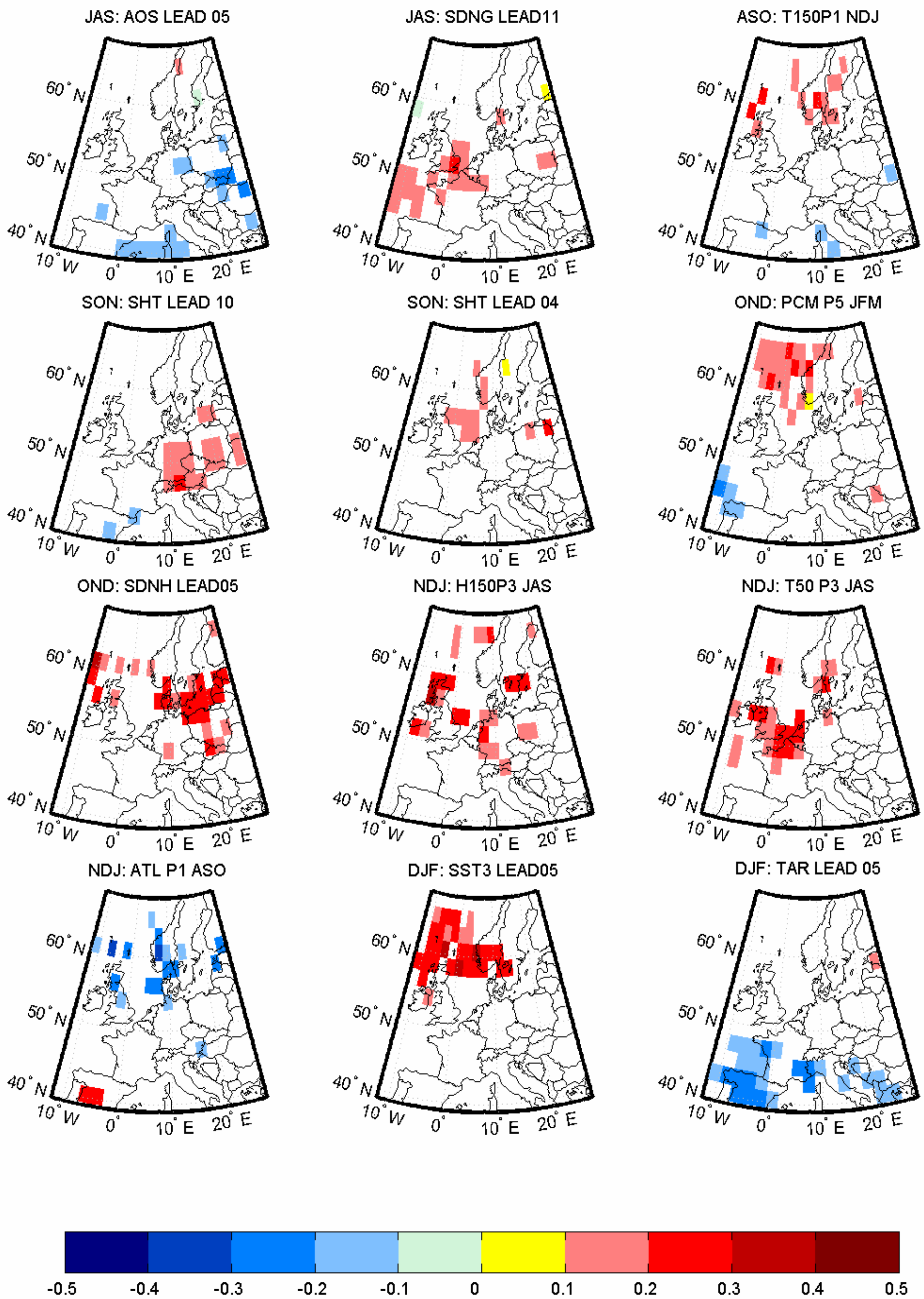

Figure 7.14 Spatial configuration for a sample of the most commonly selected predictors for 90th percentile wind exceedance counts, from JAS to DJF. For each map, the title gives the predictand season, followed by the predictor, and the gridbox shading illustrates the coefficient for that predictor in the model.

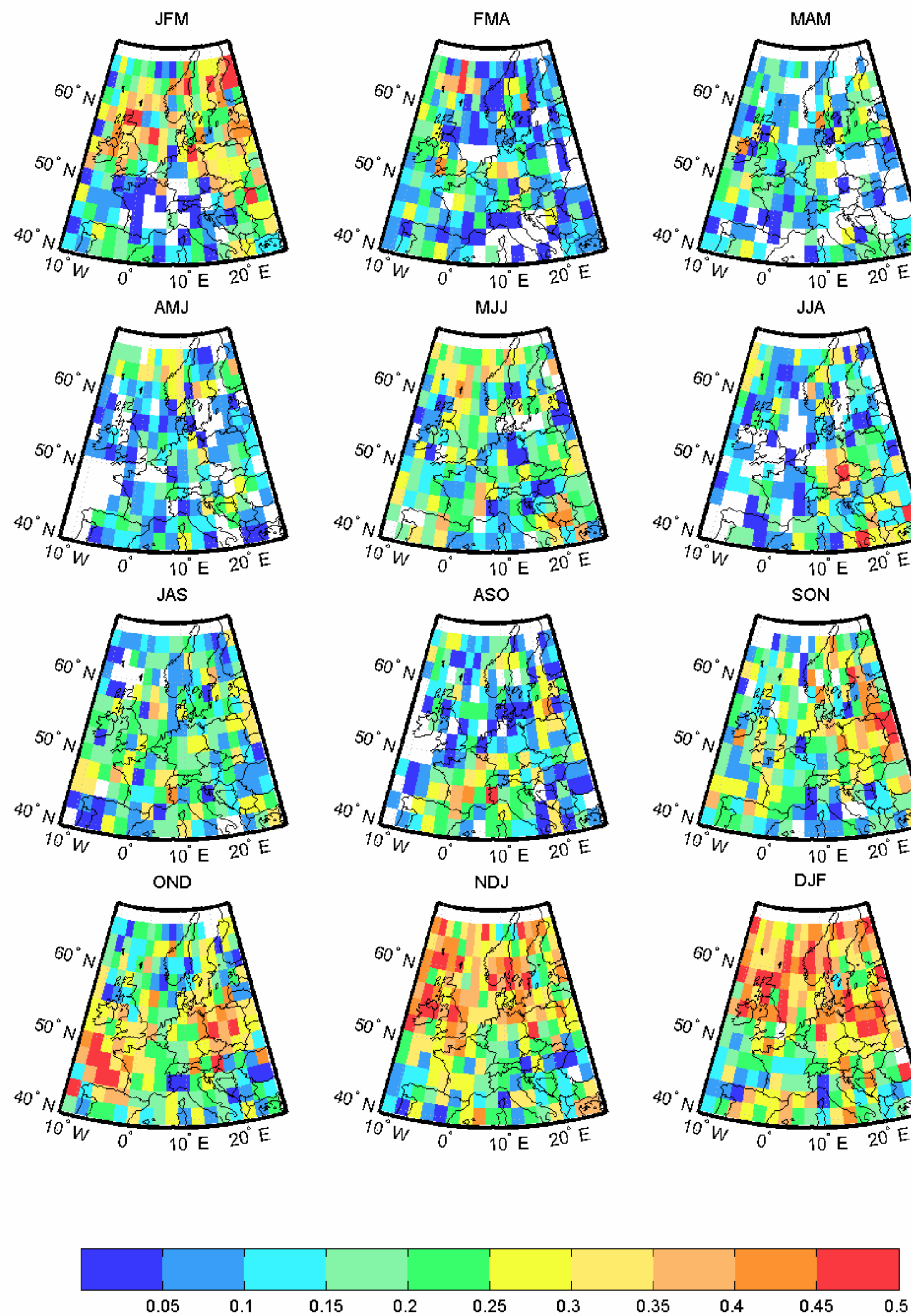

Figure 7.15 R^2 by season for 95th percentile wind exceedance counts. R^2 values are for the final model selected for each gridbox, based on the lowest cross-validation mean absolute error (MAE). Gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria.

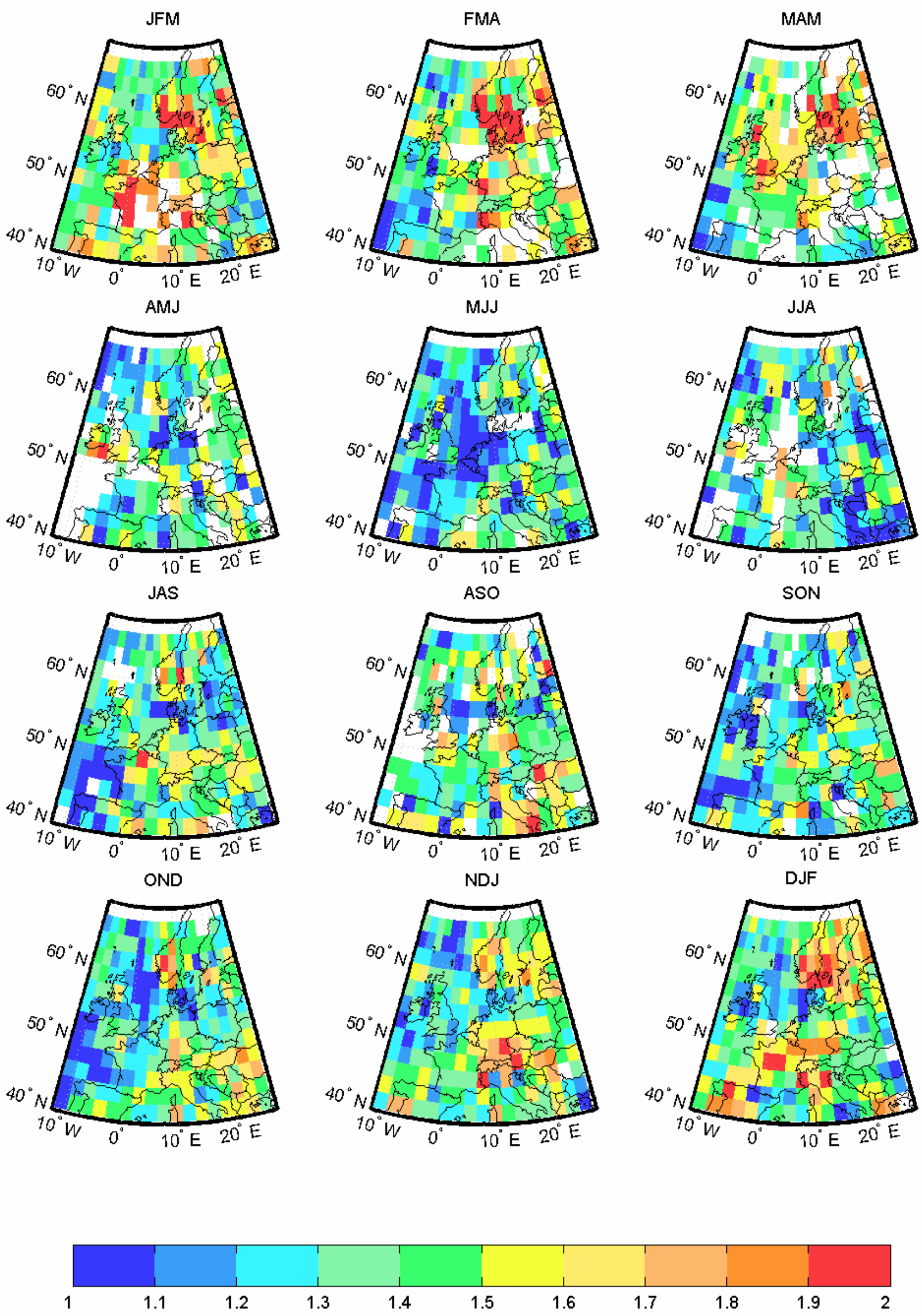

Figure 7.16 Mean Absolute Error (MAE) by season for 95th percentile wind exceedance counts. Values are for the final model selected for each gridbox, having the lowest cross-validation mean absolute error (MAE) of all predictor subsets. As Figure 7.5, gridboxes in white are those where no statistically significant model exists – either because no field-significant predictands are locally significant, or weakly significant predictors do not pass the cross-validation criteria. Note that the scale is different to that for the 90th percentile MAE values in Figure 7.12, as appropriate for the smaller absolute values of the 95th percentile exceedance counts.

7.3.4 Model selection for 95th percentile wind exceedance counts

Model selection for the 95th percentile wind exceedance predictand has much in common with the 90th percentile wind predictand, as might be expected. Interestingly, as illustrated in Figure 7.3 and Figure 7.4, the 95th percentile predictand generally has a higher count of retained predictors per gridbox than the 90th percentile predictand, although a similar spatial pattern and seasonal cycle is observed. The model fit for the final models selected at each gridbox also follow a similar spatial and seasonal pattern to the 90th percentile predictand, R^2 values generally being highest from SON to DJF as illustrated in Figure 7.15. The associated MAE has the same seasonal cycle as the 90th percentile predictand, with minima in the summer seasons, counter to the R^2 metric. This pattern is illustrated in Figure 7.16.

Table 7.4 shows the top 5 predictands by gridbox count for this predictand, by season. Many of the most commonly selected predictors are identical to those selected for the 90th percentile wind predictand, and some are similar in that they refer to the same index, but lead or lag by one or two seasons. Statistically we can therefore say that the high correlation between the 90th and 95th percentile seasonal exceedance counts for wind has resulted in very similar model selection, with similar levels of skill, at least in the training phase of the model development. It is not unreasonable to suppose that the large-scale variability which affects the likelihood of extreme wind events is also very similar for these two predictand indices.

Where predictors are common to the 90th and 95th percentile wind predictands, the spatial configuration of predictor selection is also similar, as expected. Figure 7.17 and Figure 7.18 illustrate a sample of the commonly selected predictors for 95th percentile wind exceedance counts. ENSO indices are primarily atmospheric, and as is the case with the 90th percentile wind predictand, are concentrated in the autumn and winter seasons, this time being present from OND until FMA – slightly later in the winter than the 90th percentile predictand. The response to these indices is largely concentrated in the southern portion of the domain, with a small number of gridboxes in the north of the domain responding in the opposite sign. Other teleconnection indices, including the East Atlantic (EA) and West Pacific (WP) patterns are selected frequently. The EA pattern figures prominently as a predictor from NDJ through to

MAM, at lead times of two to five months. The JFM response to the September EA pattern is shown in Figure 7.17, and comprises a large proportion of the north and north west of the

95^{th} Percentile Wind

	JFM		FMA		MAM	
	T50 P3 SON	51	SOL Lead 04	46	WPP Lead 02	65
	EAP Lead 05	47	EAP Lead 03	30	T150 P1 NDJ	38
	TAR Lead 06	35	PCM P5 AMJ	27	GLO P5 NDJ	28
	AOS Lead 04	25	SOI Lead 07	26	IND P4 JAS	28
	T150 P3 OND	25	EAW Lead 07	26	EAP Lead 02	22
Total		410		297		252
	AMJ		**MJJ**		**JJA**	
	NAO Lead 11	33	T30 P3 FMA	45	AOS Lead 04	34
	SDNH Lead 07	31	AOS Lead 02	28	EAW Lead 02	32
	IND P3 OND	28	T30 P3 JJA	27	PCM P3 DJF	29
	SDNA Lead 06	20	T150 P1 SON	25	H150 P1 JFM	28
	POL Lead 10	20	PNA Lead 05	21	PV30 P3 FMA	24
Total		256		403		303
	JAS		**ASO**		**SON**	
	AOS Lead 05	47	PV30 P3 OND	41	PV150 P1 JJA	32
	SDNG Lead 11	33	IND P1 SON	38	H30 P1 NDJ	31
	WPP Lead 10	31	IND P1 MJJ	32	SHT Lead 04	28
	NAO Lead 07	28	H50 P1 NDJ	26	H100 P1 DJF	24
	PCM P4 MAM	22	T150 P1 DJF	21	PCM P2 MJJ	19
Total		357		309		405
	OND		**NDJ**		**DJF**	
	PCM P5 JFM	59	H150 P3 JAS	53	TAR Lead 05	46
	TAR Lead 02	41	PV30 P3 ASO	40	PV50 P3 JFM	46
	ATL P2 OND	39	ATL P1 OND	36	T50 P3 SON	26
	PCM P4 AMJ	26	EAP Lead 03	25	T100 P1 AMJ	26
	ATL P5 OND	23	PV50 P3 FMA	25	EAP Lead 04	24
Total		428		450		462

Table 7.4 Top five predictors across whole domain for each season for the 95^{th} percentile wind exceedance counts models. Totals given are for all predictors. Given a domain comprising 252 gridboxes, this includes a mix of gridboxes having models with one or two predictors, and some gridboxes with no viable model.

domain, in which a positive response is observed. The West Pacific pattern is selected in the FMA, MAM and JAS seasons, and most prominently in MAM, as shown in Figure 7.17, at a lead time of two months. Again, the response is concentrated in the north of the domain, and is positive, with a much smaller area of negative responses in the south west of the domain. The NAO and AO are also selected and similar to the 90^{th} percentile wind predictand, the AOS is the more frequently selected of the two. On this point, it is interesting to note the findings of Hu and Huang (2006), who relate indices of the NAO to antecedent North Atlantic SST anomalies, and find that the statistical significance of the relationships vary markedly, to the extent that the standard NAO index, derived from MSLP in the Azores and Iceland (as used in this study) does not respond to leading SST anomalies, and a regional index of the NAO derived from regional 500hPa height anomalies does respond, although they note that this response is conditioned largely on a small number of years which are not anomalous in the station index. Although this study related primarily to the summer months, it is nevertheless important to note the sensitivity of such outcomes to small variations in both the predictand and predictor indices, and in particular, the role of outliers may mask or enhance levels of potential predictability. Hu and Huang (2006) also note the potential importance of tropical and North Pacific SST anomalies in affecting NAO predictability – which may well relate to the observed statistical fit between the Pacific teleconnection indices illustrated here and for the 90th percentile wind predictand.

Snow cover indices also feature prominently, with the strongest responses in the late spring and summer, and also to a lesser extent in the late autumn and early winter. Figure 7.17 shows the AMJ response to the preceding October Northern Hemisphere snow cover extent, where a large region of southern Europe – coincident with the area of Southern Europe which responds negatively to the NAO – exhibits a positive response. To a much smaller extent, there is a negative response in the north of the domain, centred over northern Scotland. The OND response to Northern Hemisphere snow cover extent from the preceding June, illustrated in Figure 7.18 is positive, and centred in the north of the domain, over the southern Baltic Sea region.

Large-scale SST variability also seems to drive some response in the 95th percentile wind predictand. Notably, the fifth PC of AMJ Pacific SST anomalies elicits a large negative response across the centre of the domain in the following AMJ season as shown in Figure 7.17. This compares with a similar FMA response to the same predictor in the 90th percentile wind predictand. The fifth PC of global NDJ SST anomalies is selected as a predictor across the southern Baltic Sea, into Scandinavia and across to the west of Ireland for MAM, and is associated with a strong positive response. This pattern is centred in the north western Pacific, and is similar to some of the Pacific PCs discussed in 7.3.3, above.

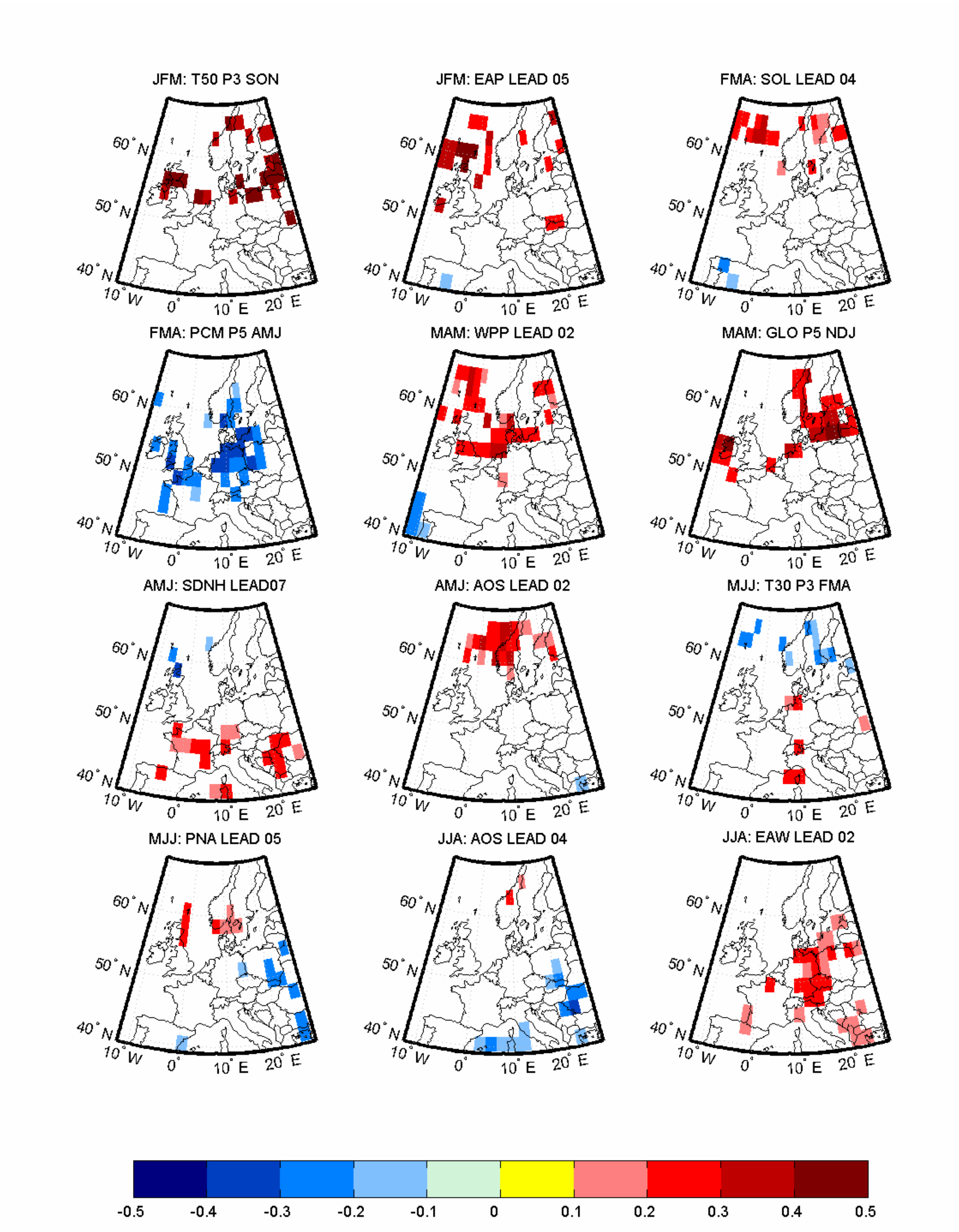

Figure 7.17 Spatial configuration for a sample of the most commonly selected predictors for 95th percentile wind exceedance counts, from JFM to JJA. For each map, the title gives the predictand season, followed by the predictor, and the gridbox shading illustrates the coefficient for that predictor in the model.

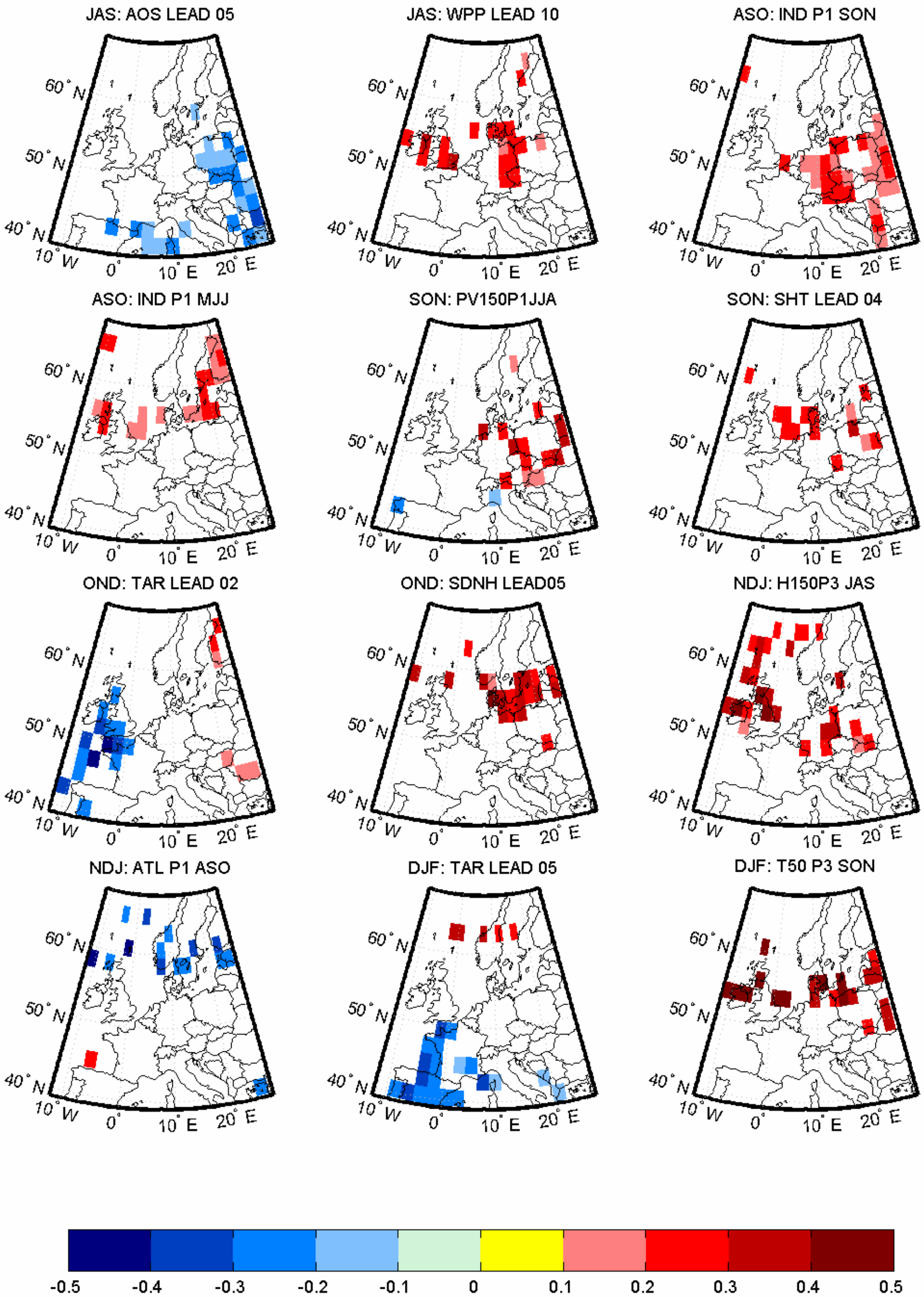

Figure 7.18 Spatial configuration for a sample of the most commonly selected predictors for 95th percentile wind exceedance counts, from JAS to DJF. For each map, the title gives the predictand season, followed by the predictor, and the gridbox shading illustrates the coefficient for that predictor in the model.

7.4 Summary

Based on a set of predictors identified as having field-significance over the predictor domain, a prediction model was then developed for each gridbox, season and predictand variable, using cross-validation over the model training period of 1959-1995. The final model at each gridbox is the one which minimises the MAE of the cross-validated predictions.

Skill levels are presented for each model in graphical format, comprising the R^2 values and MAE values separately, to indicate the extent of the fit, and the mean magnitude of the forecast error. R^2 values range from zero to 0.5 (0% to 50% of variance explained). The upper end of the range implies useful predictive skill, beyond what is commonly thought to be the case for the European climate – at least for precipitation and wind predictands. For example, Saunders and Qian (2002) identify levels of skill in the region of R^2~0.5 - 0.6 as being useable for predictions of the NAO. This apparent skill implies the need for further testing on a further independent validation period. Generally, skill levels are higher for the 90th percentile predictands than for the 95th percentile predictands, and the models for the wind predictands tend to be associated with more spatially coherent patterns of predictors, as illustrated. This likely relates to the larger spatial scale of wind extremes as opposed to precipitation extremes. One possible alteration to the approach taken here would be to restrict the number of predictors retained based on isolating one common' predictor from a set of similar predictors, and thus increasing the spatial coherence of the models with respect to individual predictors – although this might marginally lower the skill for individual gridboxes, it might aid in the interpretation of results, and in the development of a fully operational prediction scheme. There is a notable seasonal cycle in skill levels for both precipitation and wind predictands, with precipitation skill peaking in the late spring and summer, and remaining high to a lesser extent in the autumn and early winter. Skill for the wind predictands is highest in autumn and winter.

Notable predictors – i.e. those which are frequently selected – include those associated with atmospheric indices of ENSO, northern hemisphere teleconnection patterns – in particular the AO (and to a lesser extent the NAO) – and also the EA pattern. Northern Hemisphere snow cover also features, particularly for the wind predictands, and this may well relate to the apparent influence of snow cover on the AO and NAO. In general, large-scale SST predictors are not a dominant influence, counter to the prevailing view that these offer the greatest predictive skill at seasonal timescales. However some SST patterns do feature – perhaps most notably those associated with variability in the north Pacific, and to a lesser extent in the north Atlantic. North Pacific indices tend to be associated with atmospheric teleconnection patterns centred in the north Pacific, such as the PNA and EPNP patterns. Local SST anomalies are identified as useful predictors for extreme precipitation events over the training period, and in some cases for extreme wind events also. For precipitation, at least, this is consistent with the view that warm SSTs allow the overlying atmosphere to retain more moisture – and the responses observed here largely correspond to that, although some anomalies in the predictands associated with local SST anomalies seem to be related to larger-scale dynamic effects.

While the results presented here are of interest within the model training data, and allows insight into a wide range of atmospheric, oceanic and land-based phenomena which may be useful seasonal predictors of extremes in Europe, without further testing, and a full assessment of the physical links between the predictors and predictands it is not advisable or possible to implement a fully operational prediction scheme based on these results. Primary considerations in the treatment of these results must include the possibility that the models (having been trained on a relatively short time period) are not calibrated correctly to account for periods outside the one considered, or indeed that they are over-fitted within the training period. Furthermore, it must not be assumed that the relationships identified here are stationary in time or space, or that they are linear, even when identified using linear techniques. The next chapter assesses the model fit outside the training period, and discusses the further validation of these models.

8 Model Testing

The previous chapter addressed results of the model selection exercise, based on an all subset selection method, using cross-validation on predictors with a field-significant response in the predictand data. This chapter presents the results of further model testing on an independent sample of data, in order to validate the levels of skill apparent in these models. Section 8.1 introduces the chapter and sets out the aims; section 8.2 describes the method, following on closely from Chapter 7; section 8.3 presents the results of the model testing and section 8.4 concludes the chapter.

8.1 Introduction

The fit of the models developed and presented in Chapter 7 was assessed by considering R^2 values and the mean absolute error (MAE) between the predictand and the predicted timeseries at each gridbox. Although many of the models did not result in what might be considered useful skill, a large number of them did – having R^2 values of up to c.0.5, corresponding to 50% of interannual variance in the predictand being described by the selected model. Although not every model at every gridbox could reasonably be discussed in depth, the apparently more important and more widely selected predictors driving this response were discussed in some detail. However, as emphasised previously, this is an exploratory study, on predictands which are thought likely to have low levels of predictability at seasonal timescales, which furthermore may not be linear or stationary over the observational period considered here. In particular, the potential for nonlinear or nonstationary relationships means that despite statistically significant cross-validation skill identified in the model fitting component, it does not follow that the same levels of skill will persist outside the model training period. For this reason, it is essential to carry out further testing, on an independent sample of data in order to further quantify the levels of skill apparent in these models (e.g. Lloyd-Hughes and Saunders 2002). Another concern relating to the model fitting process is that given the large number of predictors – despite the condition of field significance imposed on each predictor, and the constraint of two predictors at most – there is potential for the models to be overfitted to the predictand data.

This chapter presents the results of further testing of the models developed in Chapter 7, over a ten year period from 1996 to 2005, referred to here as the validation period. Results are presented for each predictand variable, first as an overview of the broad changes in skill, including a brief discussion of each season, a summary of the predictors which retain some useful skill, and a graphical illustration of changes in skill by season. Secondly, a selection of key regions and seasons where predictive skill is retained throughout the validation period are discussed in more detail, including an illustration of the model parameters, and a discussion of the potential mechanisms driving the apparent predictability.

8.2 Method

Having developed models trained on the predictand data from 1959 to 1995, the models are then tested on an independent period from 1996 to 2005. For each gridbox, the training model coefficients are applied using Poisson regression, and the predicted values are compared with the observed in the same manner as the training period – that is, using the R^2 and the MAE values.

In determining whether potentially useful skill has persisted into the validation period, it is necessary to apply some threshold of skill to the validation period fit, given that the training period fit shows useful skill. On the latter condition, for the purposes of this study given the context of inherently low (at best) seasonal predictability in the midlatitudes, a threshold of R^2=0.2 is set for the determination of predictive skill in the training period – that is, the predictive model explains 20% of interannual variation in the predictand. It is thought that this level of fit – if it were supported by theory – would provide at least a marginal improvement over climatology.

Superficially, it should follow that R^2=0.2 would then be a suitable threshold for the validation period skill. However, since a significant degradation of skill is expected in many cases, a slightly relaxed threshold of R^2=0.15 is used for the validation data. It is important to note that this should not be interpreted as a criterion for validating a model with a deemed level of skill for operational purposes, but rather as a slightly relaxed filter to allow a more coherent picture of where and how skill might persist in the validation period.

In summary, models are considered to be of potential interest if $R^2 \geq 0.2$ during the training period and $R^2 \geq 0.15$ during the validation period, assuming of course that there is a positive fit between the validation data.

8.3 Results

Results illustrating the change in R^2 and MAE values from the training period to the independent validation period are presented for each season, and each predictand variable. Changes in skill are discussed in summary for each predictand. In the great majority of cases, the model skill degrades significantly in the validation period, indicating that the relationships identified over the period 1959-1995 are either over-fitted, or the observed relationships are nonstationary in time, or possibly nonlinear. This breakdown in skill is described for each season and predictand, and cases where skill is maintained into the validation period are examined in more detail.

8.3.1 90th Percentile Precipitation Model Validation: Overview

Figure 8.1 to Figure 8.12 illustrate the change in R^2 and MAE for each season from JFM to DJF, and for each model. The top left panel in each figure shows the R^2 values over the training period (1959-1995), and the top right panel shows the corresponding values for the validation period of 1996-2005. The solid gridboxes are those where potentially useful skill persists through the validation period as defined in 8.2, above, and the outline gridboxes are those where skill is insufficient. The lower panels in each figure show the same information for the MAE values. The same colour scale is used as for Chapter 7, for ease of comparison.

There follows a brief summary of each season, relating to the change in skill in the validation period.

- JFM: Figure 8.1 shows the change in skill for JFM. Almost without exception, R^2 values reduce, and MAE values increase across the domain, indicating that the fit obtained in the model training process does not apply in the validation period in most cases. There are a small number of cases where R^2 values indicate useful skill in both the training and validation periods, although the MAE values are higher. These tend to be concentrated over the Mediterranean region. It should be noted for these and for following results that given the relatively short length of the validation period, there is a substantial chance of spuriously skillful results from the validation period. This is illustrated by the instances of high values of R^2 in the validation period, where the training period skill was lower.

- FMA: Figure 8.2 illustrates the degradation in skill for FMA. A small number of gridboxes in southern France and northern Spain, and one over northern Poland retain some skill.

- MAM: Figure 8.3 illustrates the degradation in skill for MAM. Three gridboxes situated over Denmark, Lithuania and the Ukraine respectively retain some skill.

- AMJ: Figure 8.4 illustrates the degradation in skill for AMJ. A number of gridboxes over the Bay of Biscay and southern France retain some skill, as well as six gridboxes surrounding the British Isles. These models will be discussed in more detail in 8.3.5.

- MJJ: Figure 8.5 illustrates the degradation in skill for MJJ. Some scattered gridboxes (situated over Sweden, Poland, Greece, England and the Atlantic) retain some skill.

- JJA: Figure 8.6 illustrates the degradation in skill for JJA. Some scattered gridboxes – mainly over northern Europe retain skill, although R^2 values are generally below 0.25 in the validation period. One gridbox over southern Italy has $R^2>0.45$, with a low MAE value.

- JAS: Figure 8.7 illustrates the degradation in skill for JAS. Four gridboxes located across southern Europe retain some skill. Most predictors for these models are associated with indices of stratospheric temperature and potential vorticity from the preceding winter.

- ASO: Figure 8.8 illustrates the degradation in skill for ASO. Only two gridboxes – located over Norway and to the north of Scotland – retain skill.

- SON: Figure 8.9 illustrates the degradation in skill for SON. A number of gridboxes retain skill. Four of these are located over southern Europe, one to the south west of England, and the remainder in the north of the domain over Scandinavia and the Atlantic.

- OND: Figure 8.10 illustrates the degradation in skill for OND. A number of predictors – primarily to the west of the domain – retain some skill. In particular, these are located over the Bay of Biscay and Spain, and over the Atlantic to the north of Scotland.

- NDJ: Figure 8.11 illustrates the degradation in skill for NDJ. Several models over Romania and Bulgaria retain skill, as do two over the English Channel region, where validation period R^2 levels are in excess of 0.45. Predictors for these latter models include the SOI from the preceding September, and the third PC of 30hPa potential vorticity from the preceding DJF season.

- DJF: Figure 8.12 illustrates the degradation in skill for DJF. The main region where some skill is retained is located off the west coast of Scandinavia, where the dominant predictors include local SST anomalies, and the fifth PC of Indian Ocean SSTs from the preceding MJJ season.

Table 8.1 shows the predictors retained, by season, for the 90th percentile precipitation models. For each season, they are ranked by the frequency of occurrence and parameter (β). Throughout the year, the stratospheric predictors are the most important – geopotential height in particular, although in the summer seasons stratospheric temperature and potential vorticity feature more frequently. Local SST indices also feature in the winter and spring seasons, and indices of tropical Pacific MSLP also result in potentially useful skill in the winter and spring seasons, for a small number of gridboxes. Generally, no predictors are dominant to the extent that large regions of coherent skill may be attributed with any confidence.

Overall, for 90th percentile precipitation exceedance counts, very little skill is retained from the training period to the validation period, indicating that the potential predictive skill identified in Chapter 6 and Chapter 7 is at best nonstationary. Some models do retain their skill in the validation period, however, given the small numbers that do, there is a significant likelihood that this is artificial skill – in other words the skill is retained in the validation period due to chance. Without confirmation of a physical basis for the relationship, the results must still be treated with caution. Of most interest from the results presented here, is the persistence of skill into the validation period in AMJ, situated over the Bay of Biscay, and regions surrounding the UK. These results will be discussed in more detail in 8.3.5

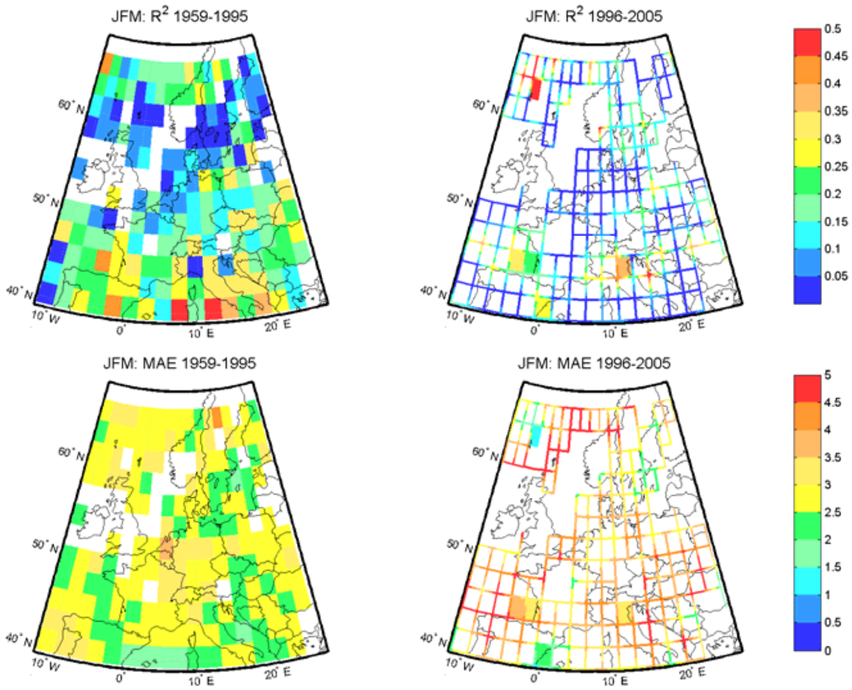

Figure 8.1. Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JFM 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

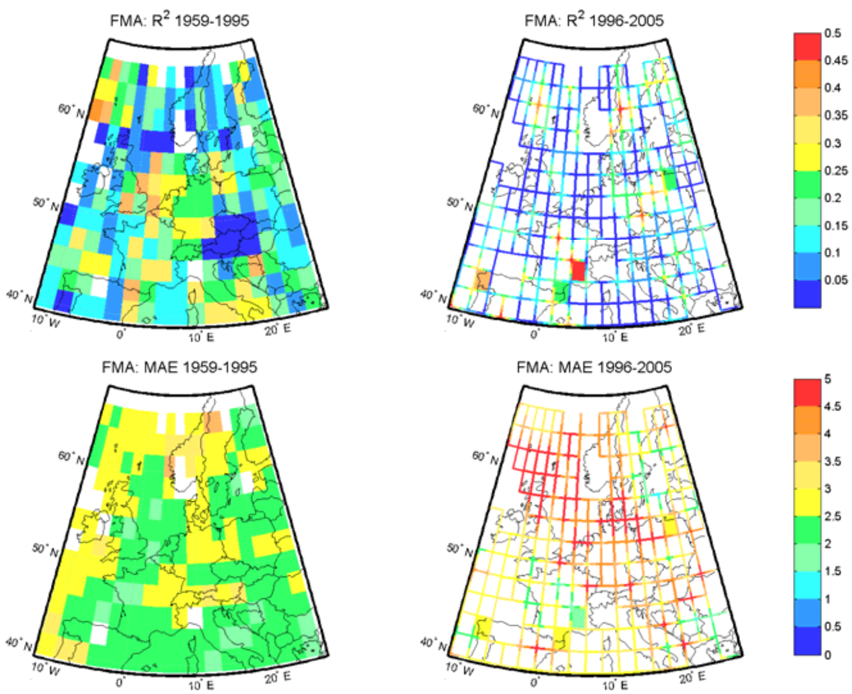

Figure 8.2 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for FMA 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

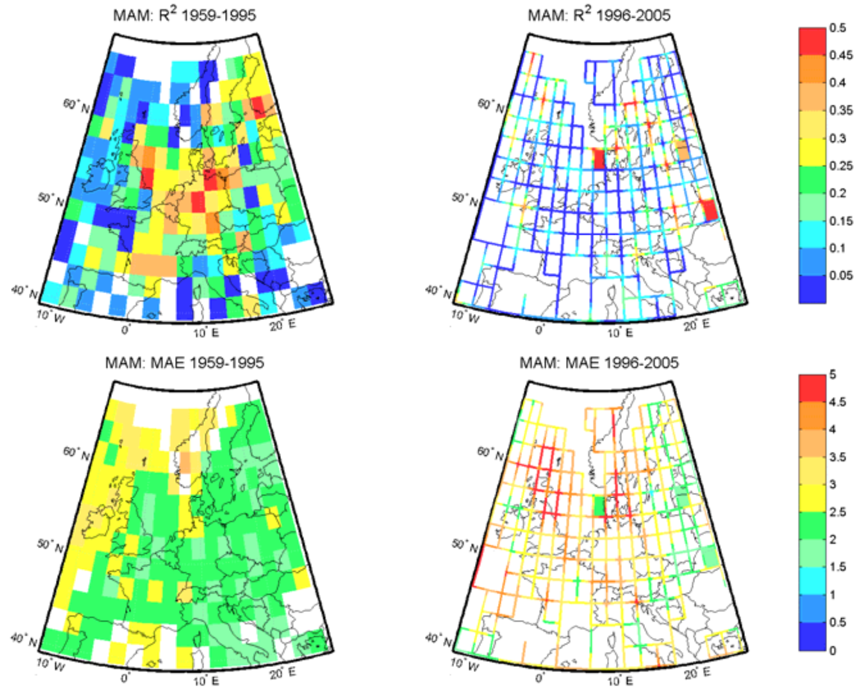

Figure 8.3 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MAM 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

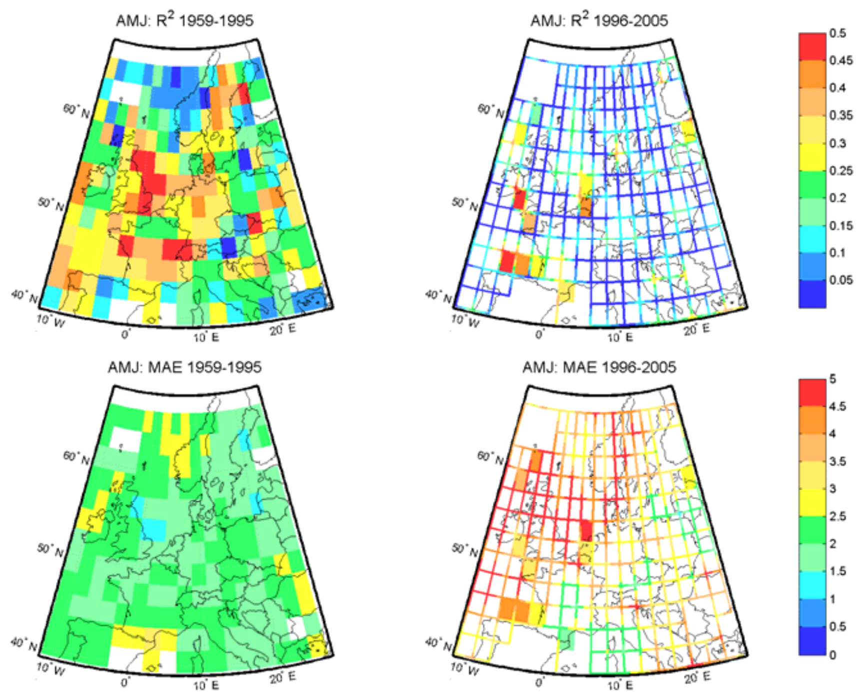

Figure 8.4 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for AMJ 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

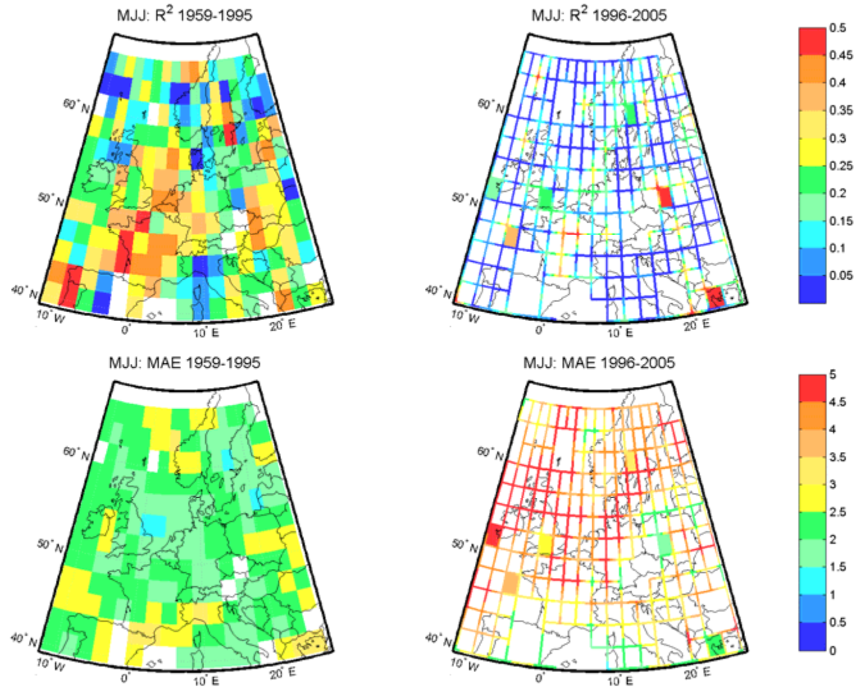

Figure 8.5 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MJJ 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

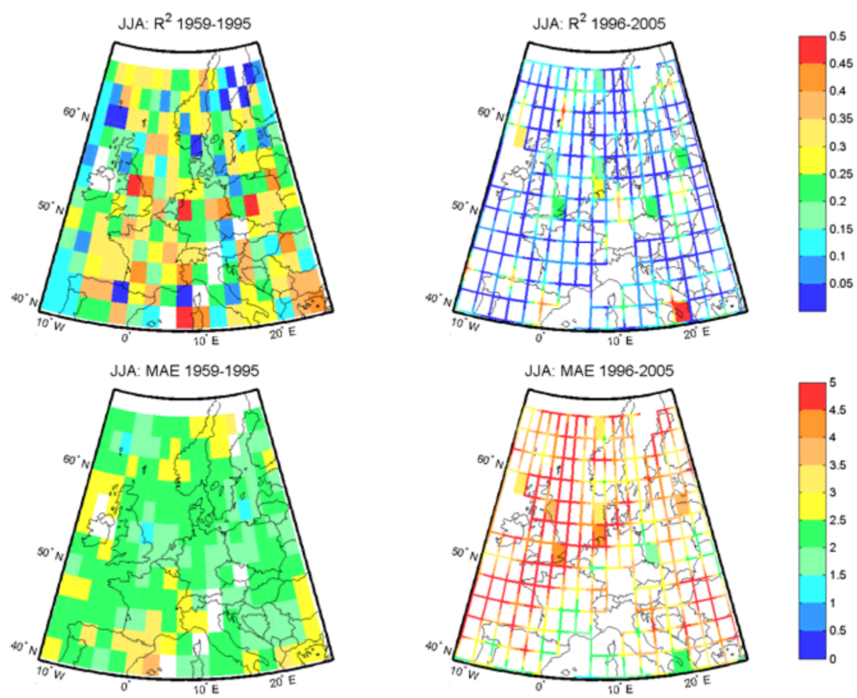

Figure 8.6 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JJA 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

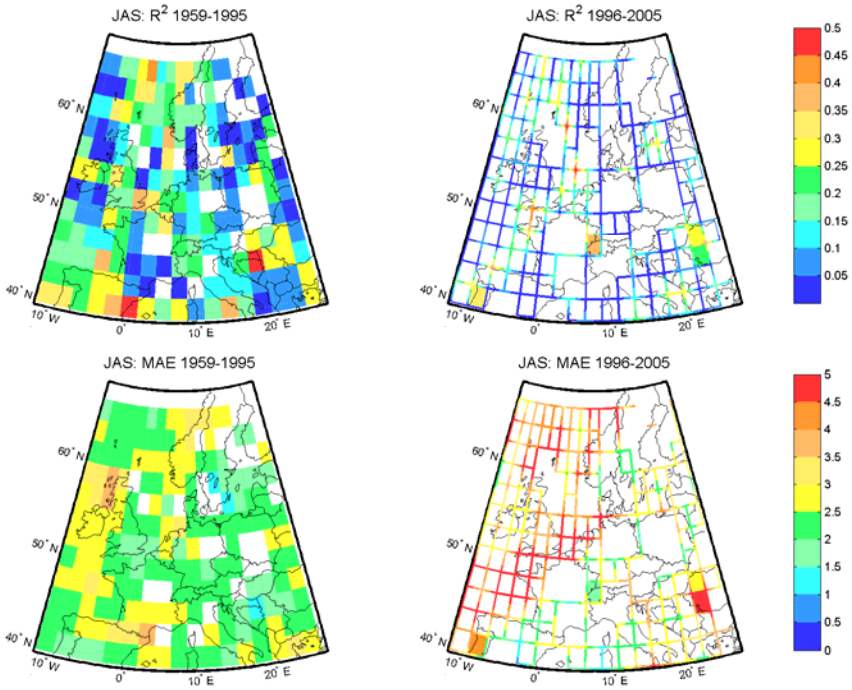

Figure 8.7 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JAS 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

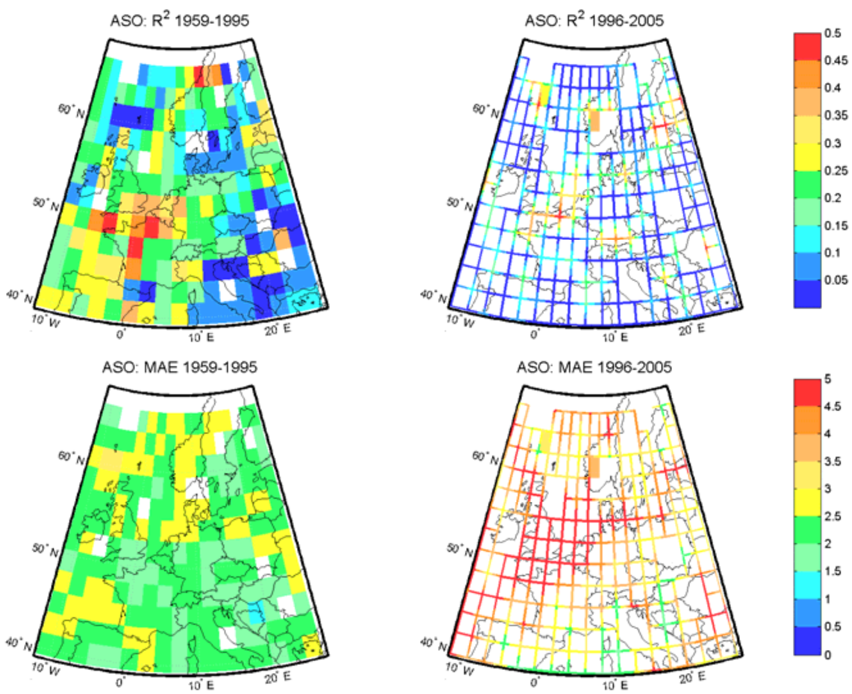

Figure 8.8 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for ASO 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

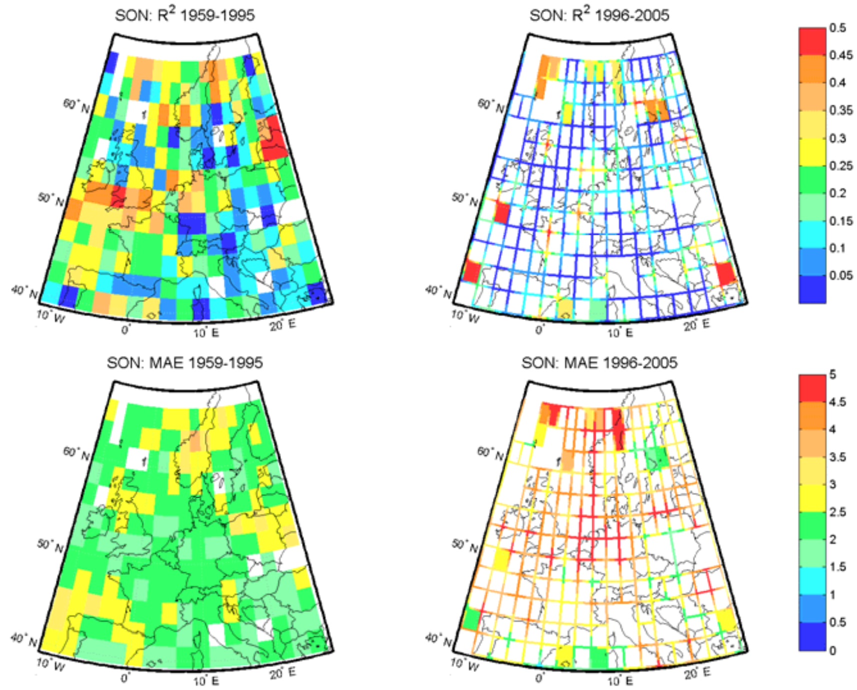

Figure 8.9 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for SON 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

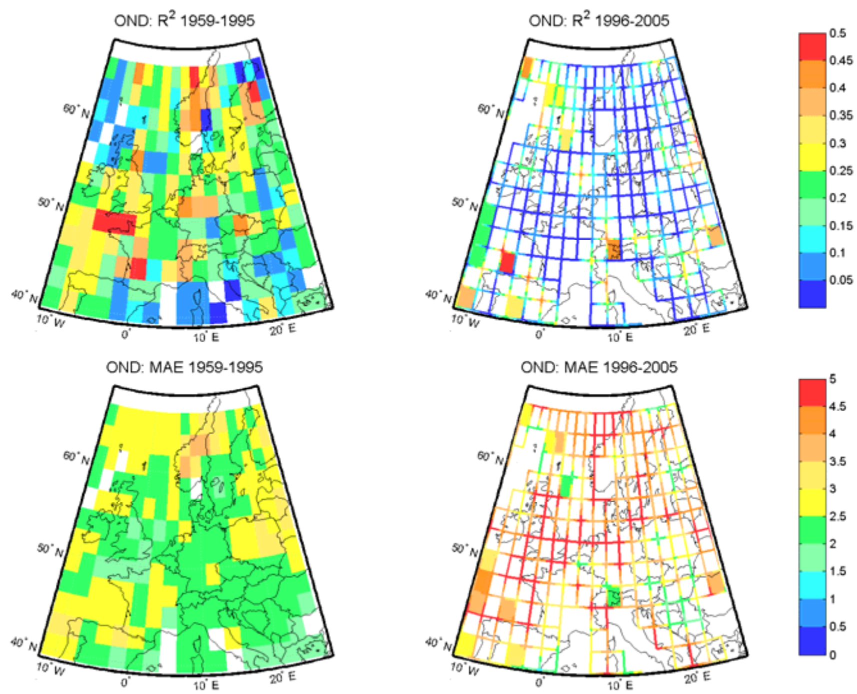

Figure 8.10 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for OND 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

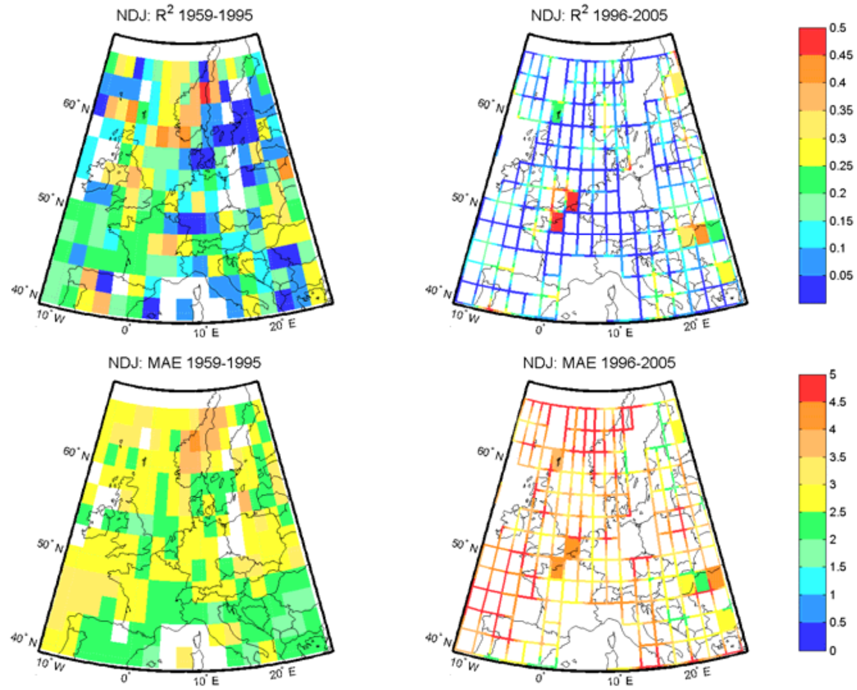

Figure 8.11 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for NDJ 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

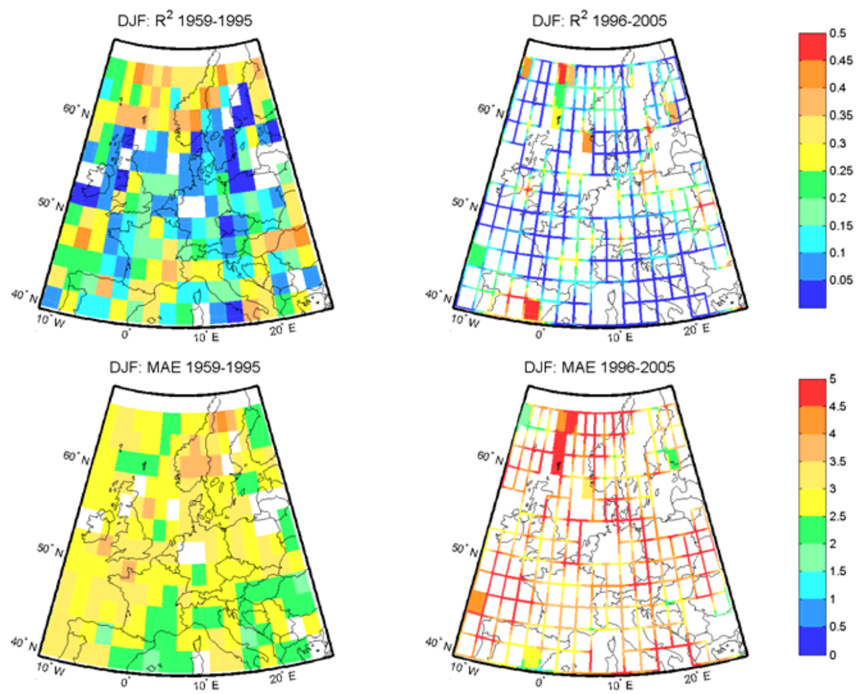

Figure 8.12 Change in R^2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for DJF 90th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

Season	Predictor	Frequency	β
JFM	H30 P1 OND	3	0.343
	SST4 Lead03	2	0.034
	NAO Lead 03	1	0.231
	T30 P2 SON	1	0.204
	SST3 Lead04	1	0.199
	H30 P1 JFM	1	0.193
	DAR Lead 05	1	- 0.189
FMA	H150P3 SON	2	0.223
	H30 P1 NDJ	2	- 0.229
	DAR Lead 10	1	0.239
	PV30 P2ASO	1	0.222
	PNA Lead 07	1	- 0.170
	AOS Lead 07	1	- 0.245
MAM	PCM P2 OND	1	0.243
	H50 P1 NDJ	1	0.216
	DAR Lead 07	1	0.192
	T100P2 MJJ	1	- 0.194
	TAR Lead 06	1	- 0.225
	ATL P4 DJF	1	- 0.280
AMJ	DAR Lead 09	2	0.187
	H100P3 SON	2	- 0.017
	EAP Lead 07	2	- 0.124
	SST3 Lead05	2	- 0.247
	PV100P3JAS	1	0.236
	GLO P2 AMJ	1	0.225
	PV30 P3AMJ	1	0.217
	T150P1 NDJ	1	0.214
	H150P1 JJA	1	0.200
	IND P2 NDJ	1	0.175
	GLO P5 NDJ	1	0.164
	T150P3 JJA	1	- 0.154
	T30 P3 JJA	1	- 0.216
	TAR Lead 03	1	- 0.226
	QBO3 Lead09	1	- 0.231
	ATL P1 MJJ	1	- 0.245
	H100P1 SON	1	- 0.278
	ATL P4 JFM	1	- 0.399
MJJ	T150P2 ASO	1	0.347
	NAO Lead 11	1	0.302
	H50 P3 SON	1	0.267
	PV150P3FMA	1	0.231
	PV30 P1JFM	1	0.168
	IND P5 DJF	1	- 0.117
	T150P1 JJA	1	- 0.145
	T30 P1 JJA	1	- 0.168
	ATL P4 DJF	1	- 0.175
	EAP Lead 08	1	- 0.216
	T30 P3 JAS	1	- 0.225
JJA	PV100P3JAS	2	0.241
	SCA Lead 05	2	0.207
	H50 P3 OND	2	- 0.241
	PV150P1SON	2	- 0.253
	EAW Lead 06	1	0.470
	GLO P4 MAM	1	0.290
	H50 P1 JFM	1	0.200
	PCM P5 MAM	1	0.183
	T100P2 DJF	1	0.175
	H150P1 JFM	1	0.169
	PCM P5 SON	1	0.155
	EPN Lead 11	1	- 0.137
	POL Lead 04	1	- 0.248
JAS	T150P1 NDJ	3	0.088
	PV150P1JAS	1	0.248
	H100P1 FMA	1	0.187
	PV150P2OND	1	- 0.213
	PV30 P2JAS	1	- 0.291
	IND P1 OND	1	- 0.300
ASO	IND P2 DJF	1	0.215
	PV150P1ASO	1	0.203
	PV150P3AMJ	1	- 0.204
SON	PCM P5 JFM	7	0.103
	H50 P1 DJF	5	0.071
	PV150P2NDJ	2	- 0.224
	H150P3 NDJ	2	- 0.253
	H50 P1 NDJ	1	0.274
	EAW Lead 09	1	0.252
	PV100P1MJJ	1	0.221
	T30 P1 OND	1	0.167
	DAR Lead 04	1	0.165
	NAO Lead 05	1	- 0.126
	SST1 Lead03	1	- 0.147
	PCM P2 JJA	1	- 0.153
	PNA Lead 07	1	- 0.176
	EAW Lead 07	1	- 0.178
	H150P3 AMJ	1	- 0.214
	GLO P3 JJA	1	- 0.222
	GLO P3 JFM	1	- 0.230
	H100P1 DJF	1	- 0.266
OND	PCM P5 JFM	3	0.146
	EAP Lead 11	3	0.061
	PV150P1JAS	2	- 0.022
	H150P3 JJA	2	- 0.198
	SOI Lead 02	2	- 0.218
	H150P1 MAM	1	0.220
	SHT Lead 10	1	0.166
	T50 P2 JFM	1	0.165
	PCM P2 OND	1	0.147
	H30 P3 MJJ	1	0.129
	AOS Lead 02	1	0.113
	EAP Lead 09	1	- 0.175
	GLO P5 JFM	1	- 0.209
	EAP Lead 02	1	- 0.213
	SST6 Lead02	1	- 0.232
	T30 P2 OND	1	- 0.252
NDJ	IND P4 JJA	4	- 0.117
	PV30 P3DJF	2	0.140
	PV30 P2MAM	2	0.042
	QBO5 Lead02	1	0.216
	GLO P4 DJF	1	0.198
	T30 P3 FMA	1	0.192
	AOS Lead 04	1	0.142
	SOI Lead 03	1	- 0.166
	PV50 P1JAS	1	- 0.197
	PV100P1MAM	1	- 0.209
DJF	T50 P3 JAS	3	0.176
	EAP Lead 04	3	- 0.055
	IND P5 MJJ	3	- 0.200
	SST3 Lead05	2	0.217
	SST4 Lead02	2	0.045
	SST3 Lead03	1	0.262
	SST2 Lead06	1	0.162
	IND P4 JJA	1	0.106
	PV100P3SON	1	- 0.196
	TAR Lead 04	1	- 0.224

Table 8.1. Predictor frequency for 90th percentile precipitation exceedance models, where skill is retained in the validation period. For each season, all predictors which feature in skillful models are included, together with their frequency of occurrence, and the mean value of the model coefficient (β).

8.3.2 95th Percentile Precipitation Model Validation: Overview

Figure 8.13 to Figure 8.24 illustrate the change in R^2 and MAE for each season from JFM to DJF, and for each model. The top left panel in each figure shows the R^2 values over the training period (1959-1995), and the top right panel shows the corresponding values for the validation period of 1996-2005. The lower panels in each figure show the same information for the MAE values. The same colour scale is used as for Chapter 7, for ease of comparison. There follows a brief summary of each season, relating to the change in skill in the validation period.

- JFM: Figure 8.13 illustrates the degradation in skill for JFM. Three gridboxes to the west of France, and one in the north west of the domain retain some skill, although R^2 levels do not exceed 0.35, and MAE values are generally in excess of two.

- FMA: Figure 8.14 illustrates the degradation in skill for FMA. Two gridboxes retain skill – one located over southern France, with R^2>0.45 and one over the western Ukraine, with R^2=0.2.

- MAM: Figure 8.15 illustrates the degradation in skill for MAM. Three gridboxes retain skill – two over Denmark, with R^2>0.3, and one over southern France (the same gridbox as that in FMA) with a lower R^2 of 0.2.

- AMJ: Figure 8.16 illustrates the degradation in skill for AMJ. Similar to the 90th percentile precipitation models, this appears to be the most skillful season, and a relatively large number of models retain skill into the validation period. The majority of these are located over of near the Irish Sea and Wales, with others situated over France, Spain, Germany, Scandinavia and Estonia. For the cluster of significant models centred on the Irish Sea, there do not appear to be any dominant predictors. These results will be discussed further in 8.3.6.

- MJJ: Figure 8.17 illustrates the degradation in skill for MJJ. Fewer skillful models are retained in MJJ, with three located to the west of Norway, southern Sweden and Slovakia respectively. R^2 values do not exceed 0.3.

- JJA: Figure 8.18 illustrates the degradation in skill for JJA. Eight models retain skill in the validation period here, located mostly in the North and Baltic Seas, and also southern England and France, and the Atlantic south of Ireland. In some instances R^2 values are in excess of 0.45.

- JAS: Figure 8.19 illustrates the degradation in skill for JAS. Four gridboxes retain some skill. These are located over Ireland, France, Poland and Romania – little spatial coherence is shown.

- ASO: Figure 8.20 illustrates the degradation in skill for ASO. Three models retain skill, located over the Atlantic west of Norway, and southern Germany.

- SON: Figure 8.21 illustrates the degradation in skill for SON. Some skill is retained for models located to the north west of Spain, over northern England and to the north of Scotland, and over France and the Czech Republic. R^2 values vary from 0.2 to >0.45.

- OND: Figure 8.22 illustrates the degradation in skill for OND. Five models retain skill, located over the Atlantic west of Norway, the North Sea, central France and the Balearic Islands respectively.

- NDJ: Figure 8.23 illustrates the degradation in skill for NDJ. Eight models retain some skill. These are located around the edges of the domain, and include the western Ukraine, where R^2 values are greater than 0.25, and the west coast of Norway, where R^2=0.6. The predictors for this latter model are May Northern Hemisphere snow cover anomalies, and the August Scandinavia teleconnection pattern (SCA).

- DJF: Figure 8.24 illustrates the degradation in skill for DJF. Only two models are retained, located over southern Norway (R^2=0.34) and the Czech Republic (R^2=0.21) respectively.

Table 8.2 shows the predictors retained, by season, for the 95th percentile precipitation models. For each season, they are ranked by the frequency of occurrence and parameter (β). There are fewer predictors here than for the 90th percentile precipitation models, although there are some similarities, in that stratospheric height indices feature prominently. SOI related indices do not feature as frequently, although four gridboxes show potentially useful skill from spring and summer Darwin MSLP during SON. R^2 values for models featuring these predictors range from 0.20 to 0.42. Additionally, some skill is due to Niño 3.4 and Niño 4 indices, unlike the 90th percentile precipitation models.

In summary, fewer of the 95th percentile precipitation models retain potentially useful skill into the validation period than the 90th percentile precipitation models, and there are no large coherent areas of skill that persist into the validation period. A small number of the models do retain skill, and of particular interest is the region over the UK in AMJ, where a coherent region of skill persists through the validation period. This will be examined in more detail in 8.3.6.

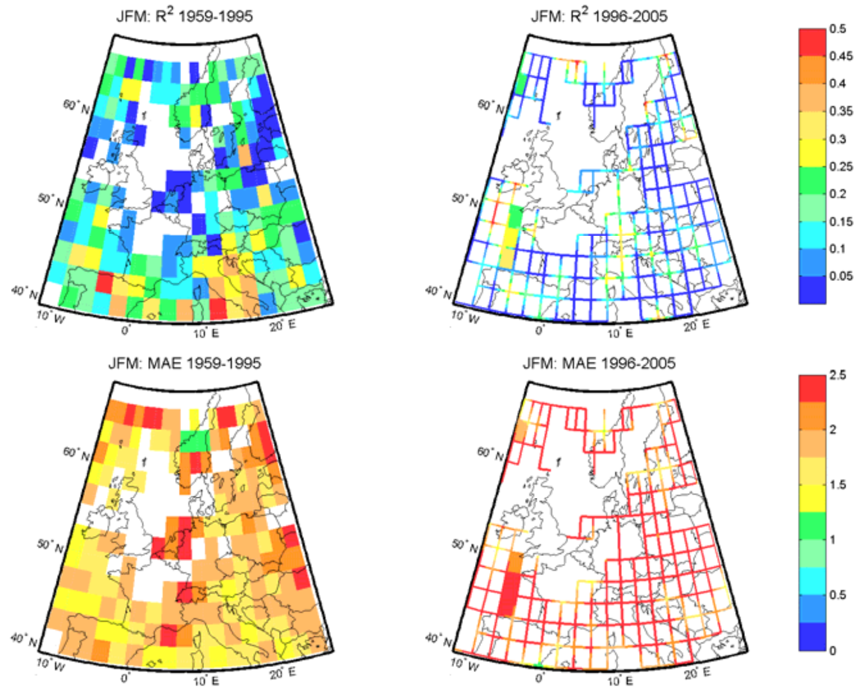

Figure 8.13. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JFM 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

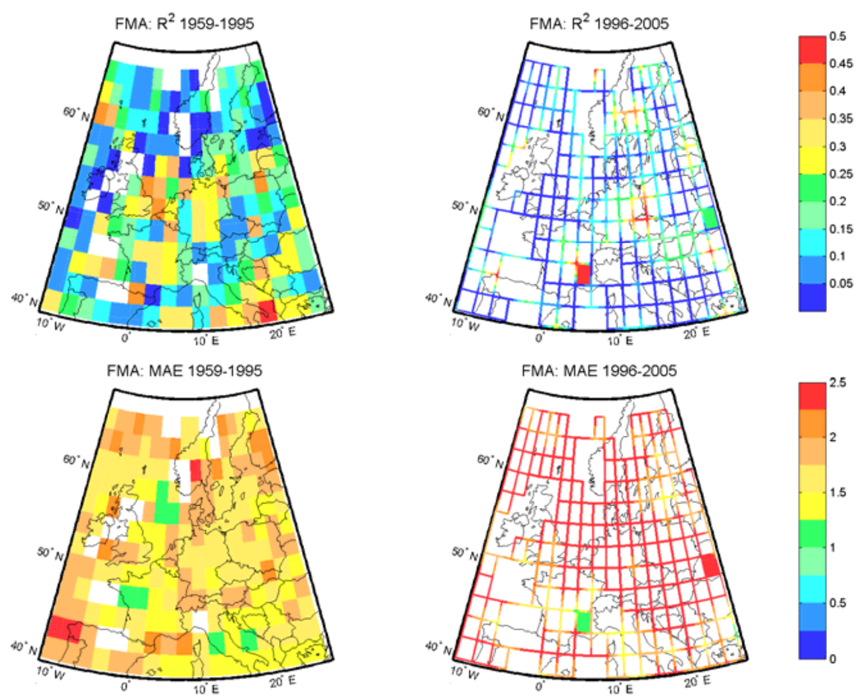

Figure 8.14. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for FMA 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

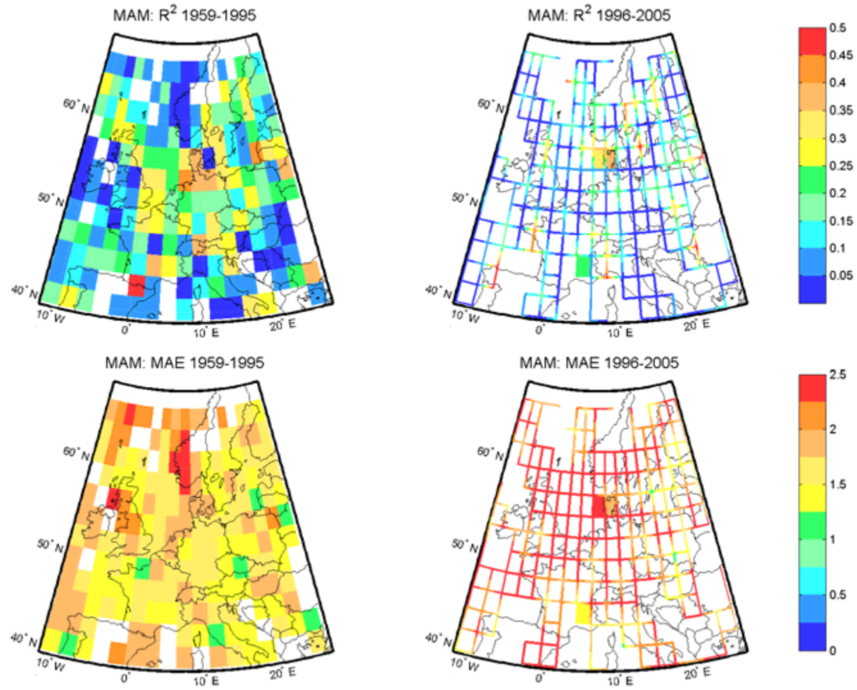

Figure 8.15. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MAM 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

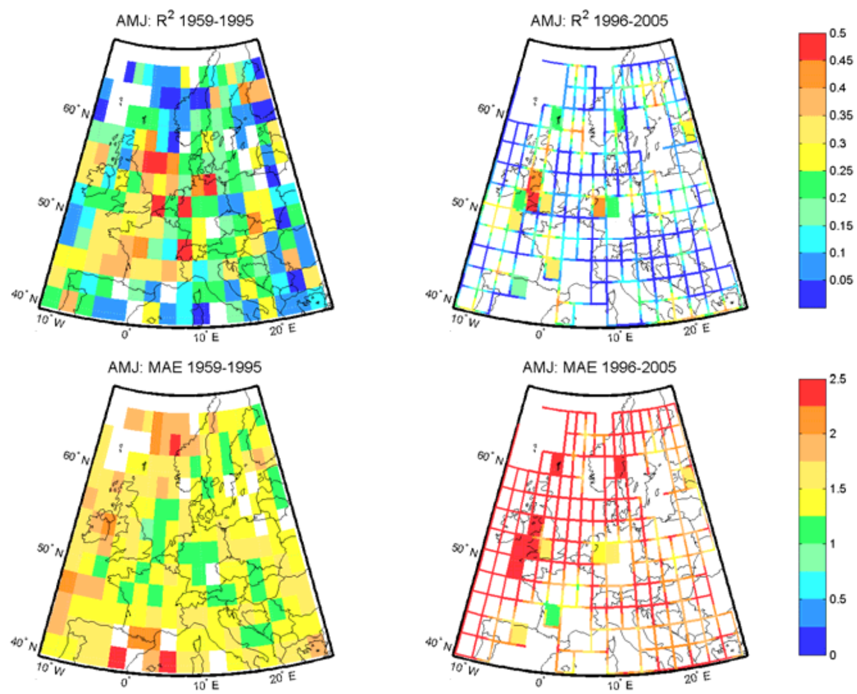

Figure 8.16. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for AMJ 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

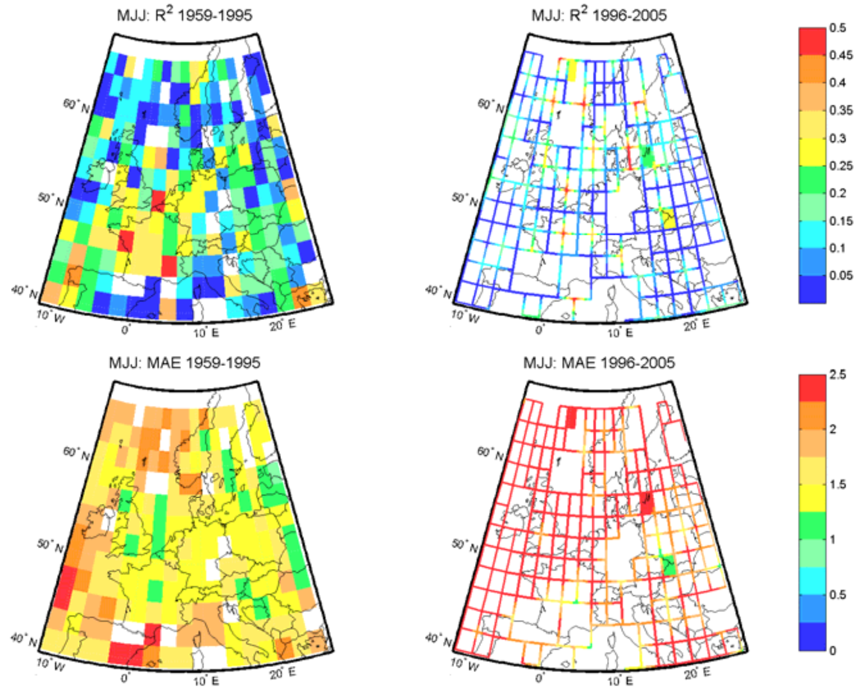

Figure 8.17. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MJJ 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

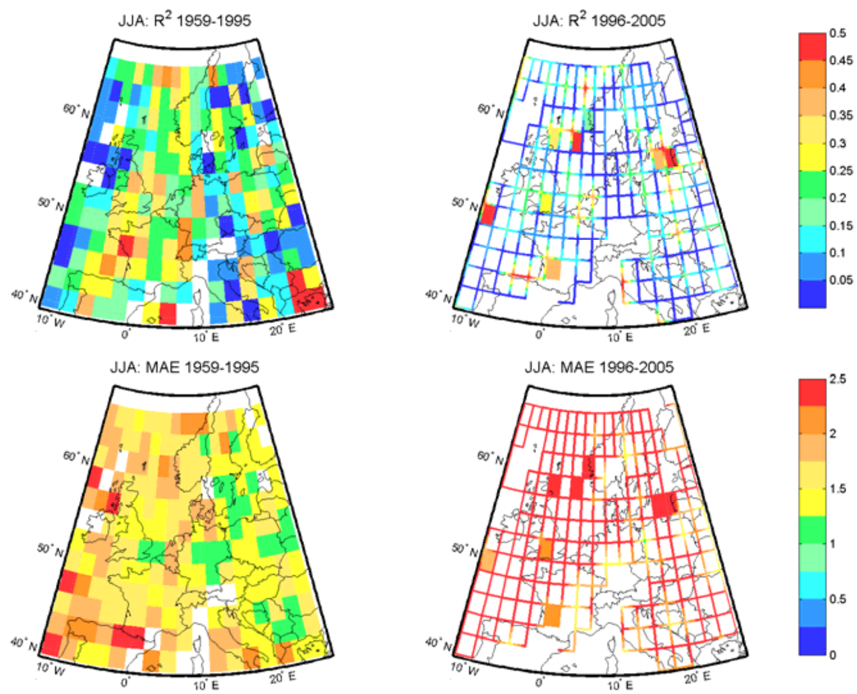

Figure 8.18. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JJA 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

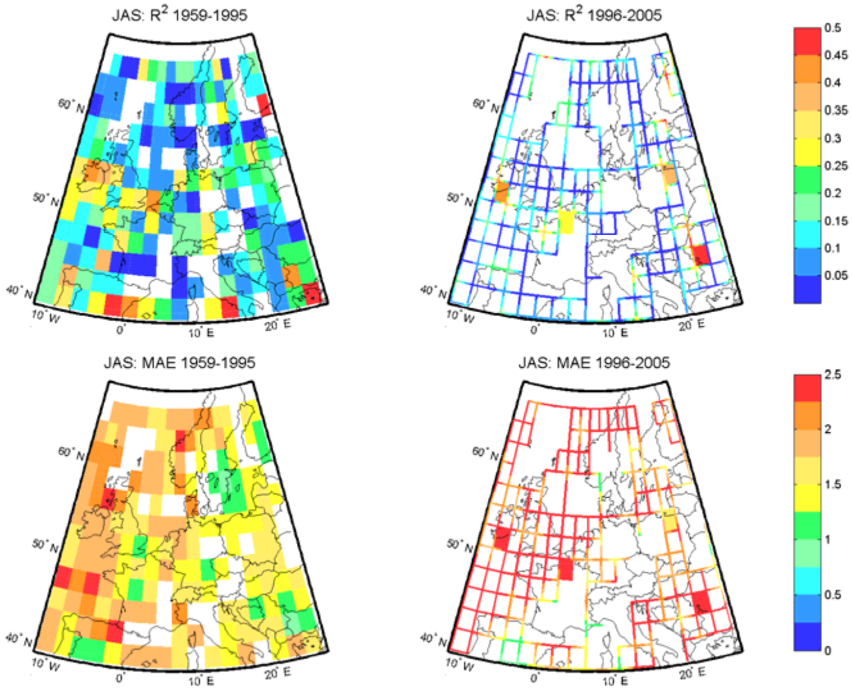

Figure 8.19. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JAS 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

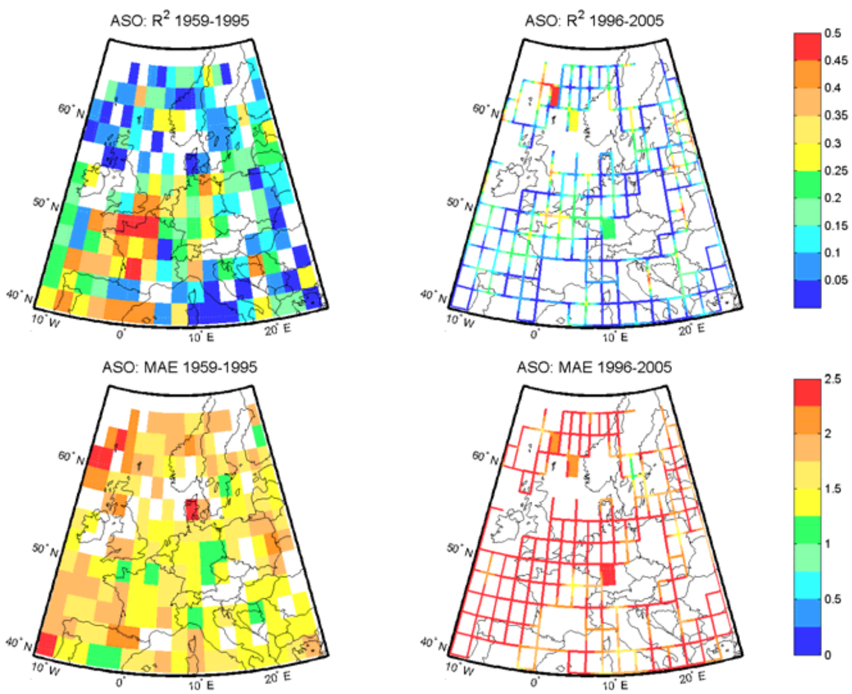

Figure 8.20. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for ASO 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

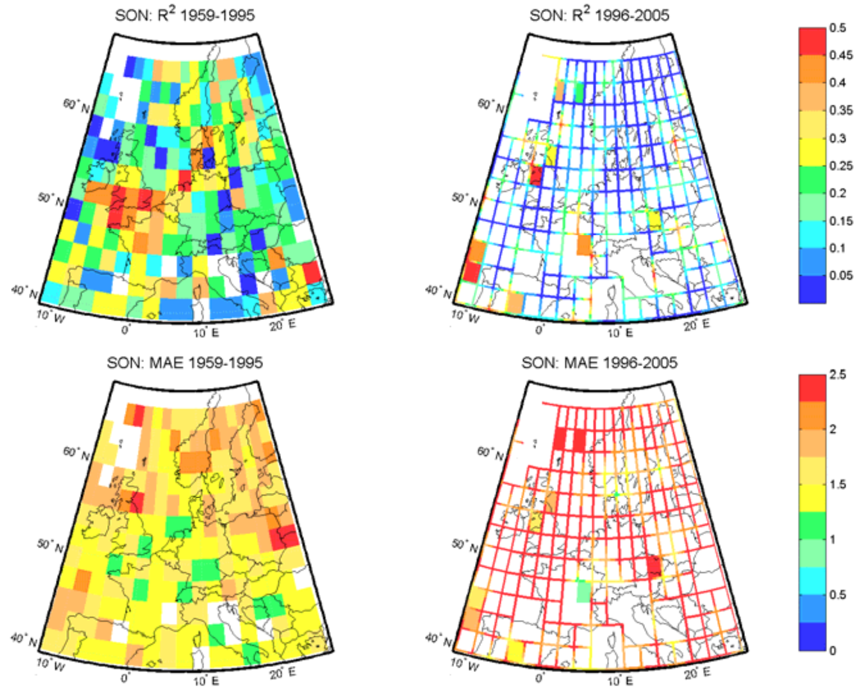

Figure 8.21. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for SON 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

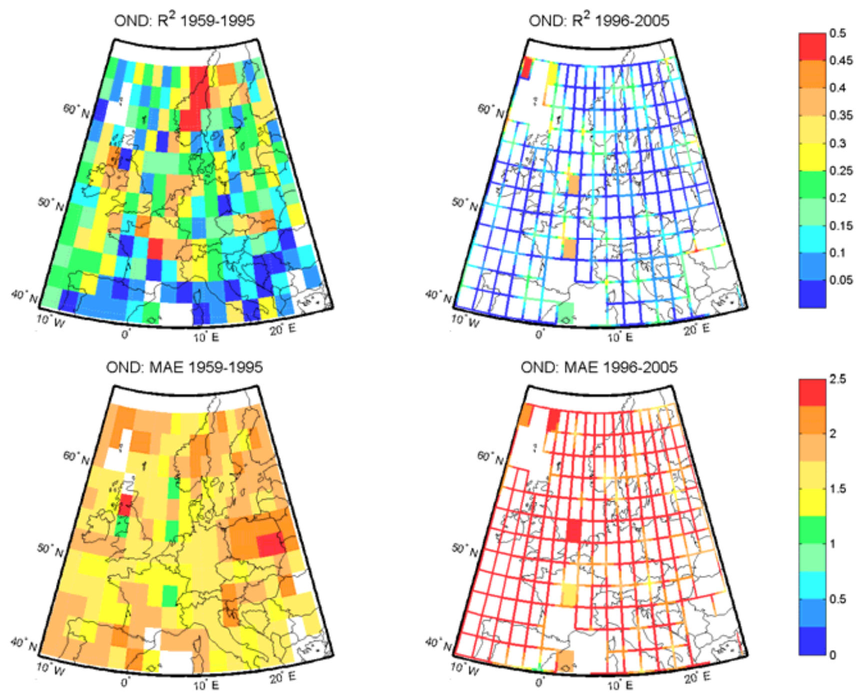

Figure 8.22. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for OND 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

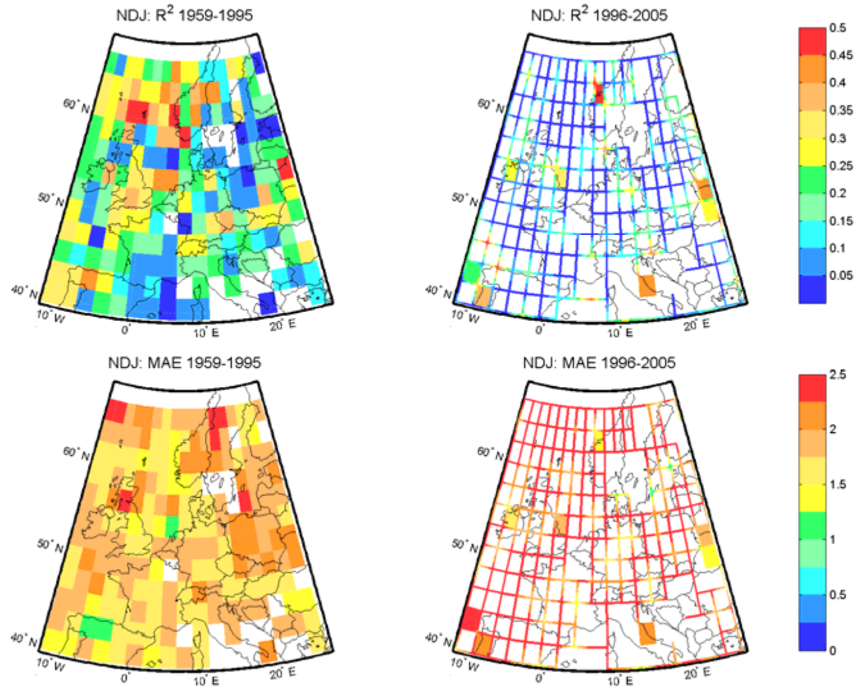

Figure 8.23. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for NDJ 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

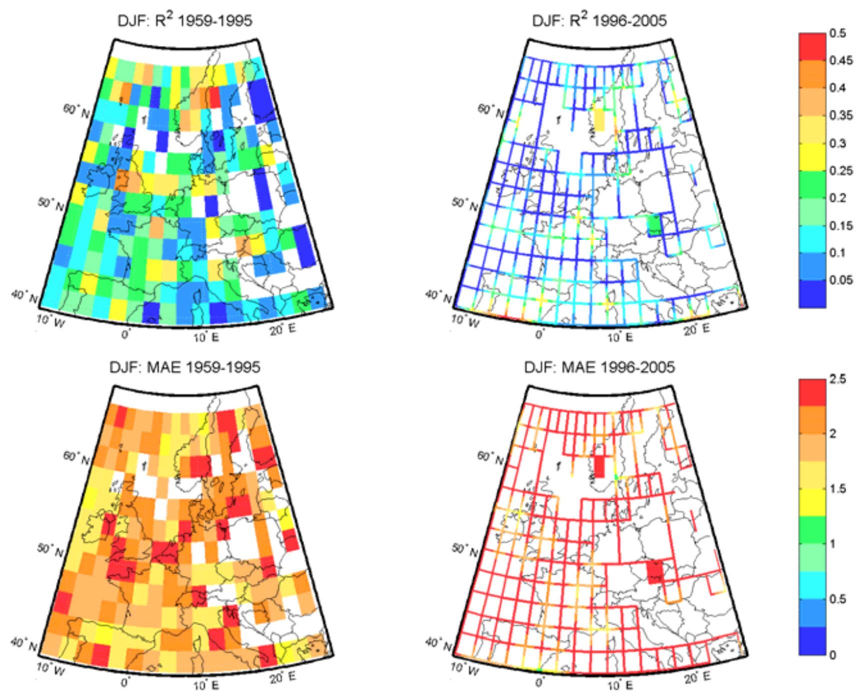

Figure 8.24. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for DJF 95th percentile precipitation exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

Season	Predictor	Frequency	β
JFM	PNA Lead 06	2	-0.223
	SST3 Lead06	1	0.256
	PV150P3ASO	1	-0.175
	EAP Lead 05	1	-0.185
	T100P2 MAM	1	-0.199
	SST4 Lead02	1	-0.228
	T150P3 MJJ	1	-0.246
FMA	T150P3 ASO	1	0.300
	DAR Lead 08	1	0.222
	EPN Lead 04	1	-0.252
	IND P2 MJJ	1	-0.268
MAM	DAR Lead 07	2	0.245
	IND P2 JJA	2	-0.255
	H50 P1 DJF	1	0.292
	GLO P2 SON	1	-0.265
AMJ	H100P1 NDJ	3	-0.255
	PNA Lead 09	2	-0.002
	H50 P2 JFM	2	-0.259
	N4 Lead 10	1	0.326
	SDNH Lead10	1	0.322
	EAW Lead 06	1	0.286
	T100P2 SON	1	0.275
	H100P2 ASO	1	0.269
	H30 P1 JFM	1	0.219
	PV150P2MJJ	1	0.219
	EAW Lead 11	1	0.217
	SDEU Lead06	1	0.213
	SHT Lead 06	1	-0.152
	TAR Lead 04	1	-0.168
	ATL P1 JJA	1	-0.178
	H50 P1 OND	1	-0.183
	EAP Lead 07	1	-0.250
	SST2 Lead03	1	-0.352
	PV50 P3AMJ	1	-0.455
MJJ	GLO P4 DJF	1	0.391
	H50 P1 JFM	1	0.243
	EAW Lead 05	1	0.187
	TAR Lead 07	1	-0.188
	T50 P3 JAS	1	-0.342
JJA	H30 P1 DJF	2	-0.223
	PV150P1SON	2	-0.494
	IND P5 NDJ	1	0.304
	AOS Lead 05	1	0.290
	SHT Lead 09	1	0.231
	PV100P3JAS	1	0.186
	SDEU Lead03	1	0.169
	IND P4 JJA	1	-0.128
	AOS Lead 04	1	-0.205
	T100P2 NDJ	1	-0.223
	PV30 P1DJF	1	-0.272
	SDNG Lead02	1	-0.302
	T100P1 JFM	1	-0.328
	ATL P4 NDJ	1	-0.363

Season	Predictor	Frequency	β
JAS	PV150P1ASO	2	0.239
	SDNG Lead11	1	0.326
	T150P1 DJF	1	0.311
	H100P1 MAM	1	0.297
	ATL P1 JAS	1	0.272
	H50 P1 MAM	1	-0.207
	PV30 P2JAS	1	-0.287
ASO	PV150P3AMJ	2	-0.232
	PV30 P3OND	1	0.398
	IND P3 MAM	1	0.137
	T100P3 FMA	1	-0.210
SON	H150P3 NDJ	3	-0.303
	T30 P3 MAM	2	0.222
	DAR Lead 04	2	0.207
	DAR Lead 07	2	0.203
	H100P3 NDJ	1	0.246
	H50 P2 JFM	1	0.227
	H30 P1 DJF	1	0.215
	T50 P1 JFM	1	0.199
	IND P3 MAM	1	0.188
	T100P2 JJA	1	0.187
	T30 P3 AMJ	1	-0.263
	T150P1 DJF	1	-0.278
	T100P3 AMJ	1	-0.336
OND	PCM P2 JAS	1	0.330
	N34 Lead 02	1	0.267
	SDNG Lead06	1	0.246
	SDNH Lead05	1	0.218
	H100P1 JFM	1	0.217
	GLO P2 JAS	1	0.194
	AOS Lead 02	1	0.176
	PCM P5 JFM	1	0.143
	H30 P3 JJA	1	-0.239
	T150P1 JAS	1	-0.332
NDJ	GLO P4 JFM	2	0.282
	PV100P2JFM	2	0.232
	SDNH Lead06	1	0.423
	PV50 P3FMA	1	0.339
	T30 P1 DJF	1	0.296
	OBO5 Lead02	1	0.204
	SOI Lead 03	1	0.193
	PV30 P3NDJ	1	-0.217
	PV50 P1ASO	1	-0.226
	TNH Lead 10	1	-0.235
	NHT Lead 10	1	-0.240
	EAP Lead 03	1	-0.279
	SCA Lead 03	1	-0.306
DJF	AOS Lead 03	1	0.466
	SDNH Lead07	1	0.350
	PV50 P3MAM	1	0.222
	SST4 Lead02	1	-0.487

Table 8.2. Predictor frequency for 95th percentile precipitation exceedance models, where skill is retained in the validation period. For each season, all predictors which feature in skillful models are included, together with their frequency of occurrence, and the mean value of the model coefficient (β).

8.3.3 90th Percentile Wind Model Validation: Overview

Validation of the 90th percentile wind exceedance models is presented in the same format as for the precipitation models. First a summary of the breakdown in skill is presented by season, and particular features of interest are noted.

- JFM: Figure 8.25 illustrates the degradation in skill for JFM. A small cluster of models located over northern Germany and Poland retain skill into the validation period, with R^2>0.35. The common predictor here is the third PC of 50hPa temperature from the preceding SON season. Other predictors for these models include PCs of Pacific Ocean SST also from the preceding autumn and summer.

- FMA: Figure 8.26 illustrates the degradation in skill for FMA. Four gridboxes in the extreme south of the domain retain some skill, with R^2 values in excess of 0.25.

- MAM: Figure 8.27 illustrates the degradation in skill for MAM. Similar to FMA, five models in the south of the domain retain some skill, with R^2 values ranging from 0.33 to 0.46.

- AMJ: Figure 8.28 illustrates the degradation in skill for AMJ. Five models retain some skill in AMJ. Three are located over Scandinavia, one over southern Germany, and one to the south of Ireland. The Scandinavian models show most skill (R^2>0.3), and are associated with the AO, or stratospheric patterns from the preceding winter.

- MJJ: Figure 8.29 illustrates the degradation in skill for MJJ. A number of models retain relatively high levels of skill (R^2>0.45). These are split into two meridionally constrained regions – one stretching from the Atlantic across southern Scandinavia, and the other from the Bay of Biscay across to Romania.

- JJA: Figure 8.30 illustrates the degradation in skill for JJA. Five models retain potentially useful skill. These are located across Scandinavia and Germany, with one in Romania.

- JAS: Figure 8.31 illustrates the degradation in skill for JAS. A coherent region of skillful models is located over eastern Poland and the Ukraine, with R^2 values ranging from 0.25 to 0.55. North American snow cover anomalies at lead times up to the previous autumn, and stratospheric height anomalies from the preceding winter feature as the most common predictors here.

- ASO: Figure 8.32 illustrates the degradation in skill for ASO. A number of gridboxes centred over Austria and Poland retain some skill in the validation period. R^2 values range from 0.19 to 0.62. Predictors for these models include Southern Hemisphere temperature from the preceding October, and stratospheric potential vorticity anomalies form the preceding winter seasons.

- SON: Figure 8.33 illustrates the degradation in skill for SON. The region of coherent skill identified in ASO over eastern Europe persists into SON, although it shifts slightly to the north and west. The preceding October Southern Hemisphere temperature anomaly is still the dominant predictor for these models, resulting in R^2 values from 0.36 to 0.55.

- OND: Figure 8.34 illustrates the degradation in skill for OND. A relatively large number of models (19) retain skill into the validation period during OND. The bulk of these are situated over northern Scotland, the North Sea and to the west of Norway, with other regions in the Bay of Biscay and over southern Germany and Austria. The latter region shows the highest skill levels ($R^2 \approx 0.7$), and the predictors for this region are the fifth PC of Atlantic SSTs from the preceding OND season, and the September SCA pattern. The region to the west of Norway is skilfully predicted by the March SCA pattern and the fifth PC of Pacific Ocean SSTs from the preceding January, among others.

- NDJ: Figure 8.35 illustrates the degradation in skill for NDJ. A coherent region of skillful models is centred over the Benelux countries. R^2 values are in excess of 0.35. The dominant predictors for this region are the second PC of JAS 50hPa temperature, and the second PC of MJJ global SST anomalies.

- DJF: Figure 8.36 illustrates the degradation in skill for DJF. This is the season where most skill is retained into the validation period. An extensive region stretching from the North Sea across Denmark to the Baltic States exhibits high levels of skill (R^2 ranges from 0.25 to 0.65). Another coherent level of skill is located to the south, stretching from southern France across to Romania. Typically, for all these models the MAE values are substantially higher then during the training period. These results will be discussed in more detail in 8.3.7.

Table 8.3 shows the predictors retained, by season, for the 90th percentile wind models. For each season, they are ranked by the frequency of occurrence and parameter (β). There are considerably more predictors and models for this predictand than for the precipitation predictands, and two seasons (OND and DJF) stand out as having a large number of models which retain skill in the validation period. The most frequently selected group of predictors is that describing stratospheric temperature variability. These are particularly prevalent in the autumn and winter seasons. In DJF the third PC of 50hPa temperature during the preceding SON season contributes to potentially useful skill in nine gridbox models, and the JAS index of this predictor features in three further models. Generally indices of snow cover, the AO and NAO and Indian Ocean PCs contribute more to potentially useful skill than they do for the precipitation models, and indices associated with the SOI and local SST contribute less.

Skill levels for the 90th percentile wind exceedance models reduce significantly in the validation period, in a similar fashion to the precipitation models. However, there are some seasons and regions where skill appears to persist, and spatially, more coherence is apparent. In particular, the autumn and winter seasons show potentially useful levels of skill for large regions throughout the validation period. Further detail on the most notable areas where skill is retained is discussed further in 8.3.7

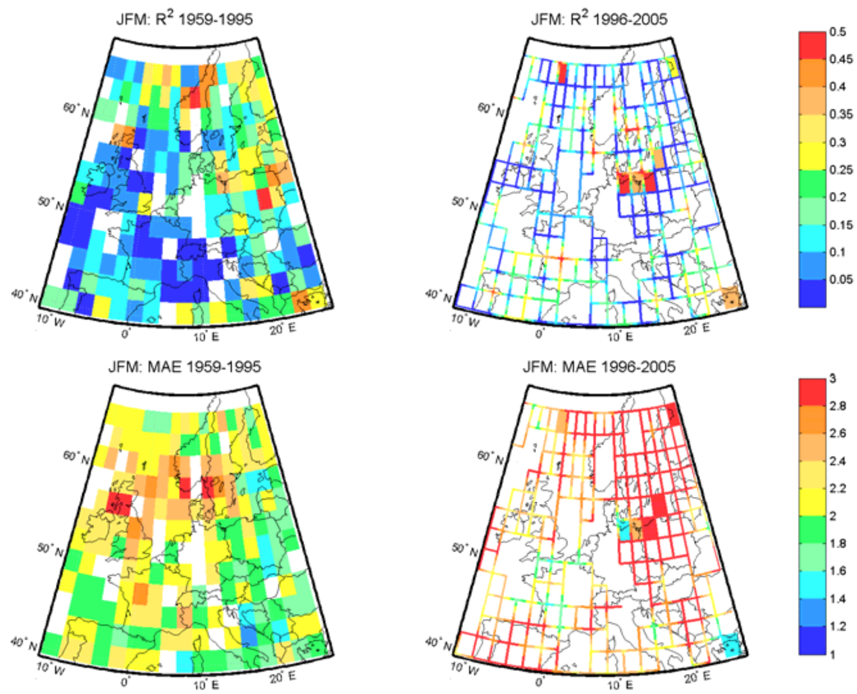

Figure 8.25. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JFM 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

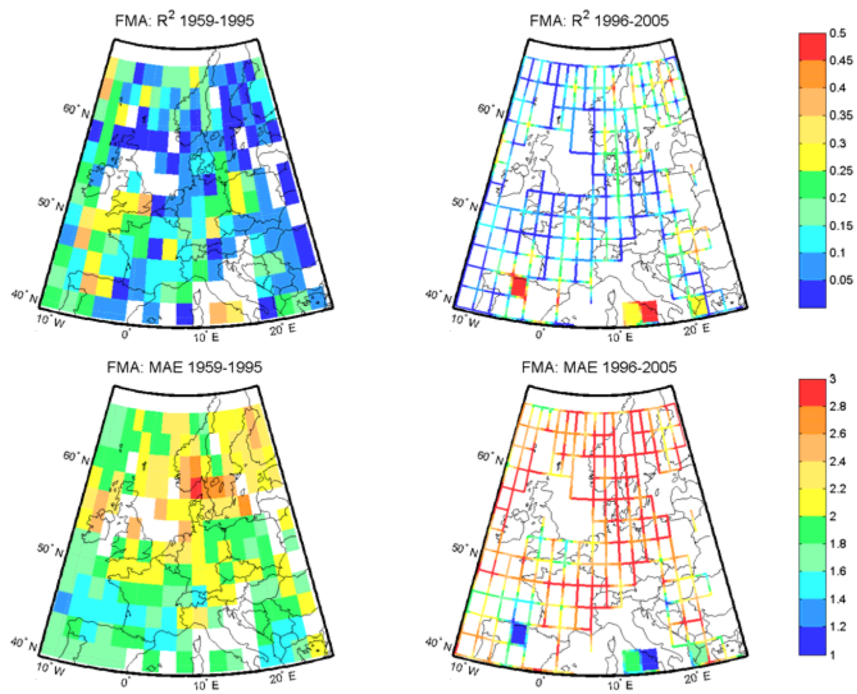

Figure 8.26. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for FMA 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

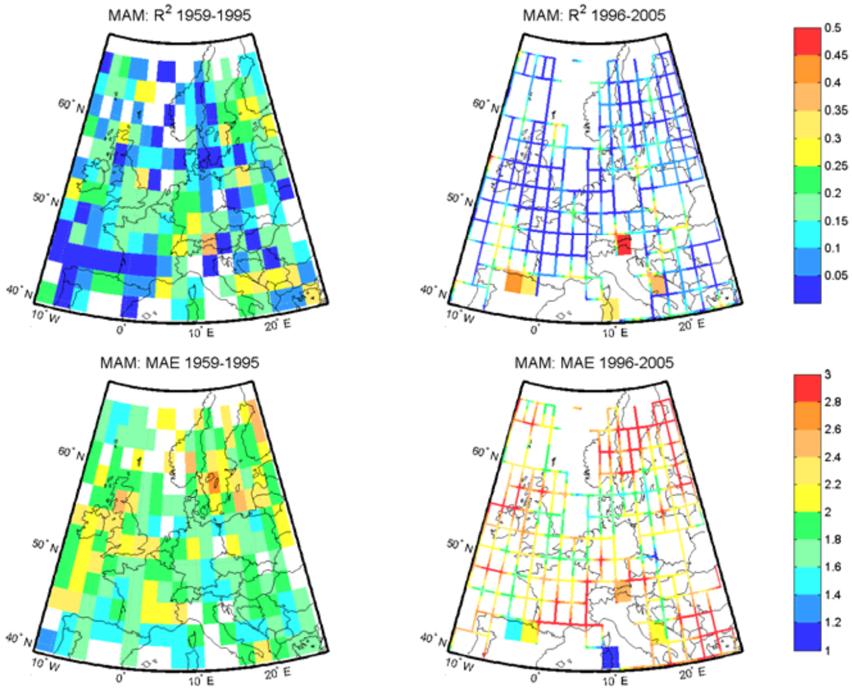

Figure 8.27. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MAM 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

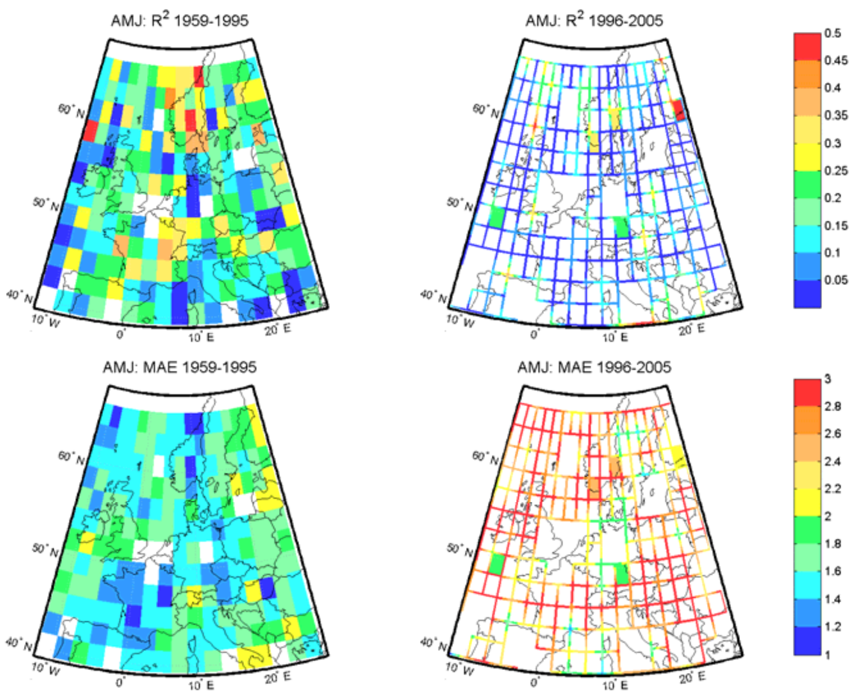

Figure 8.28. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for AMJ 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

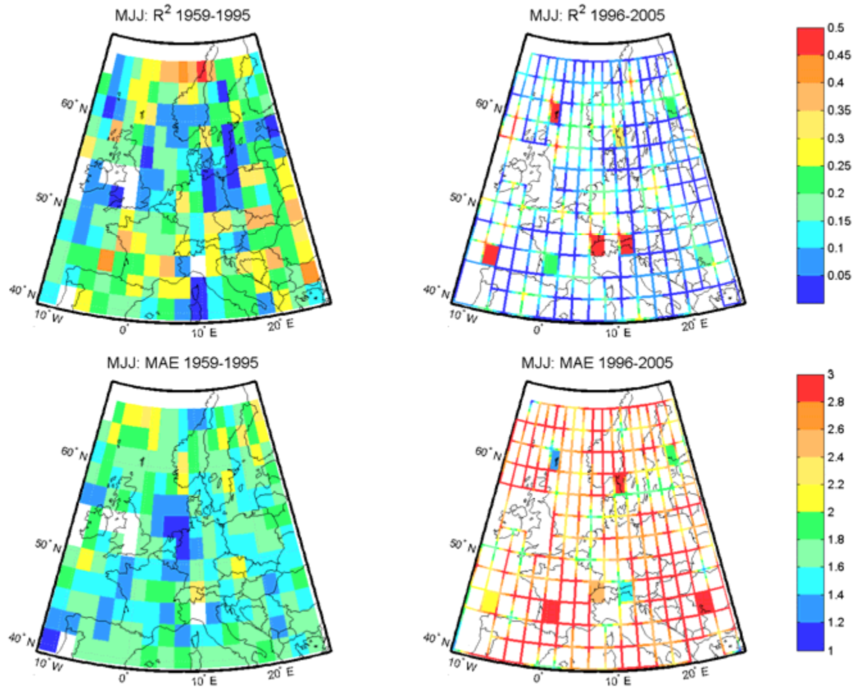

Figure 8.29. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MJJ 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

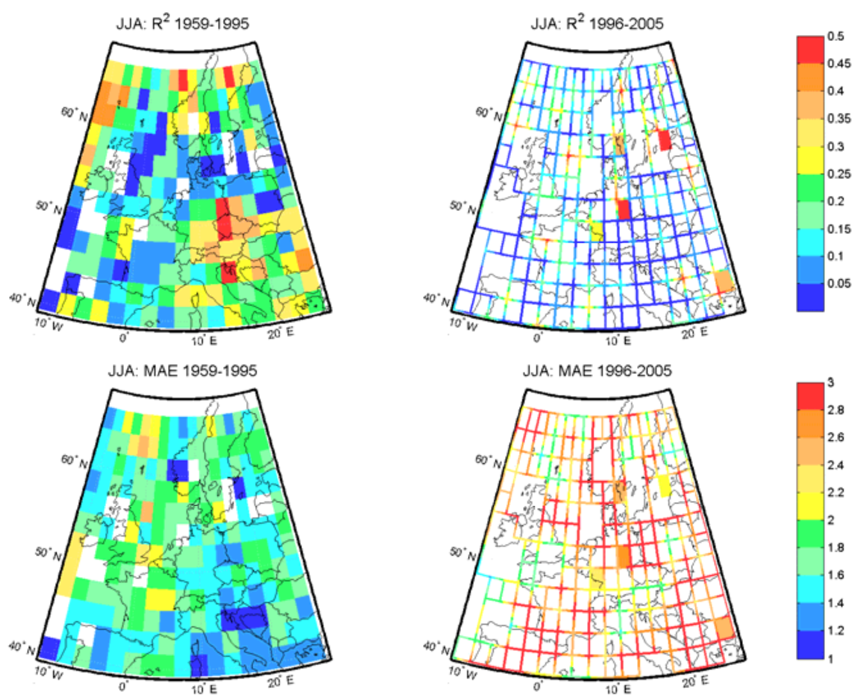

Figure 8.30. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JJA 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

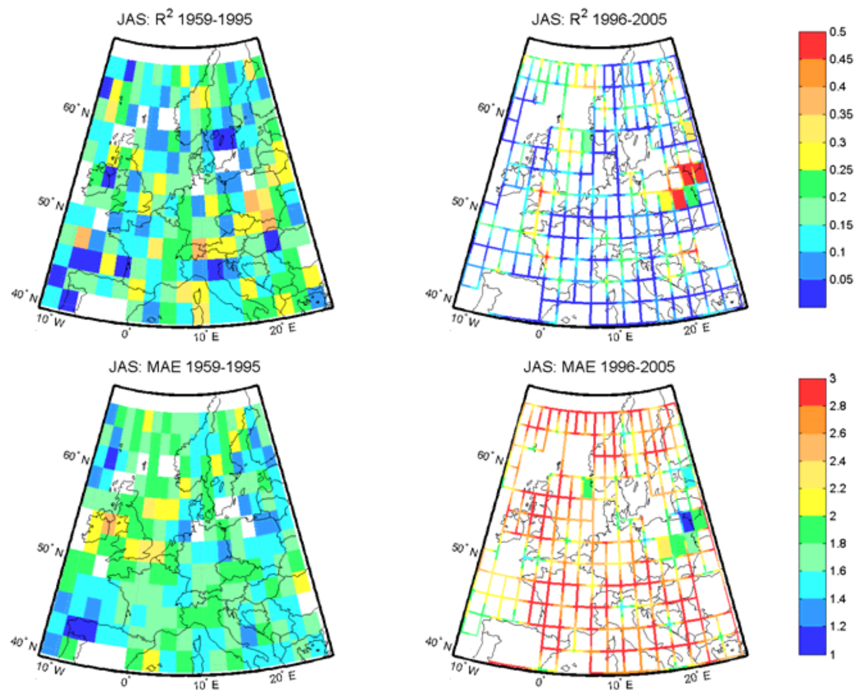

Figure 8.31. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JAS 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

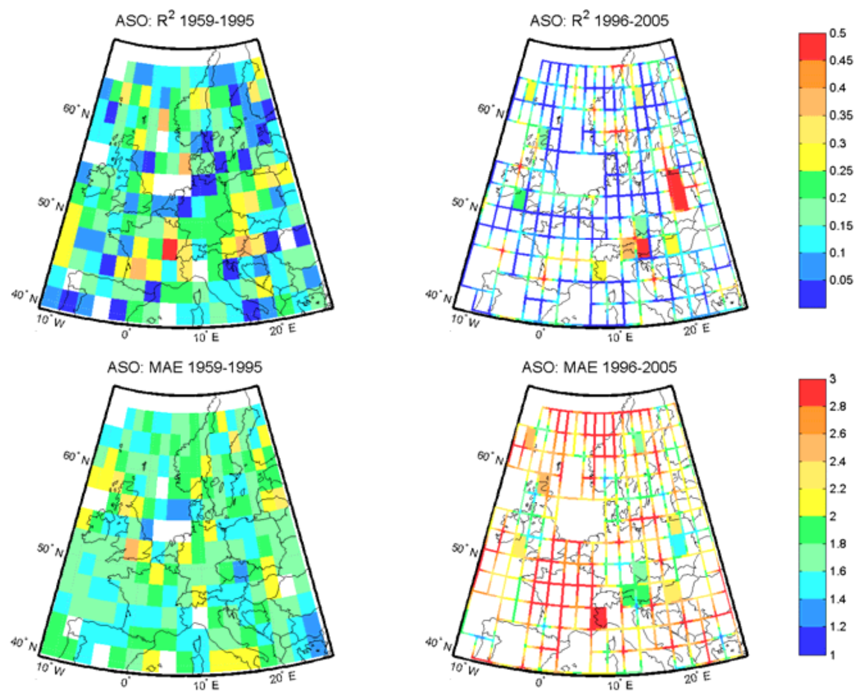

Figure 8.32. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for ASO 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

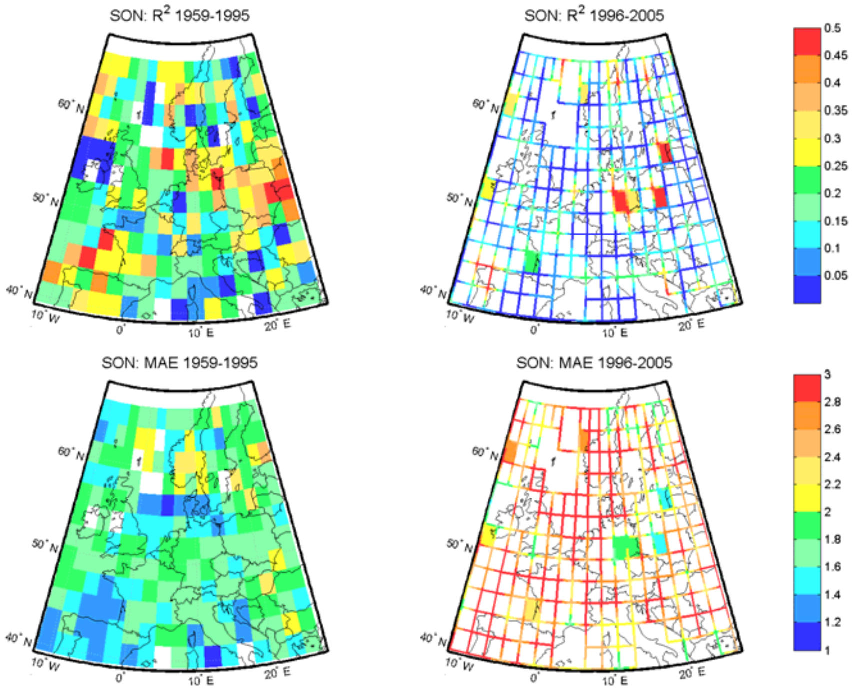

Figure 8.33. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for SON 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

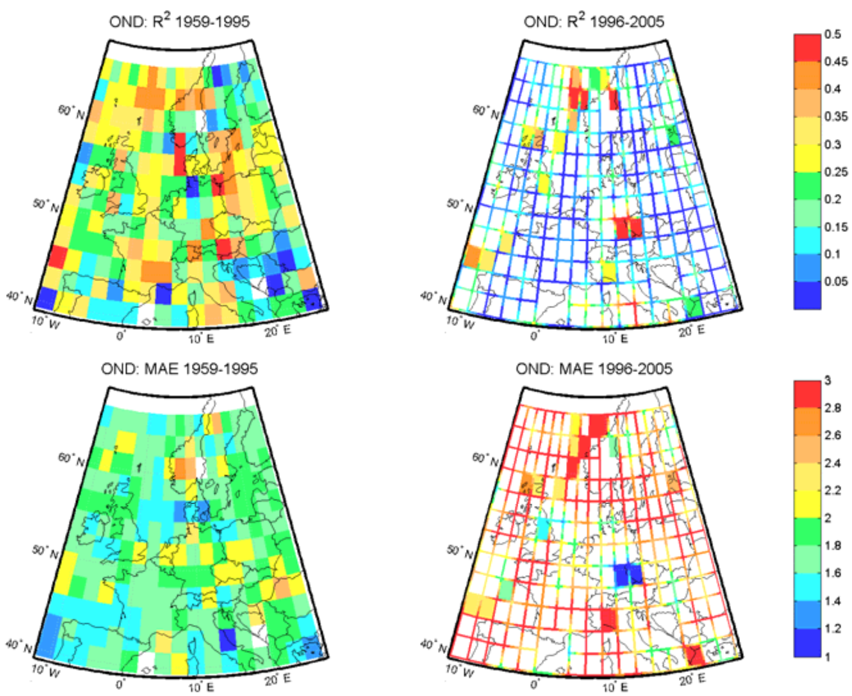

Figure 8.34. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for OND 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

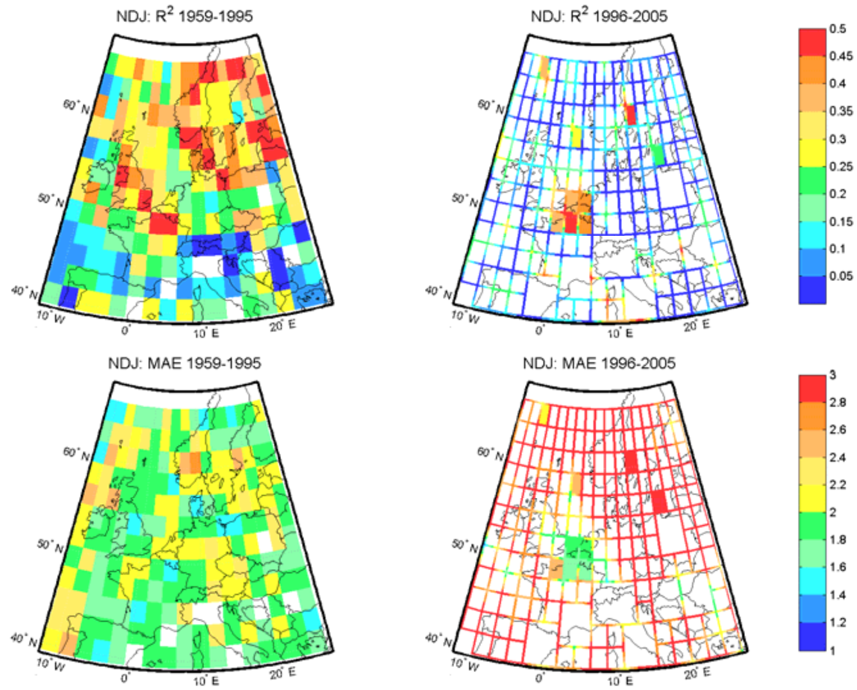

Figure 8.35. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for NDJ 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

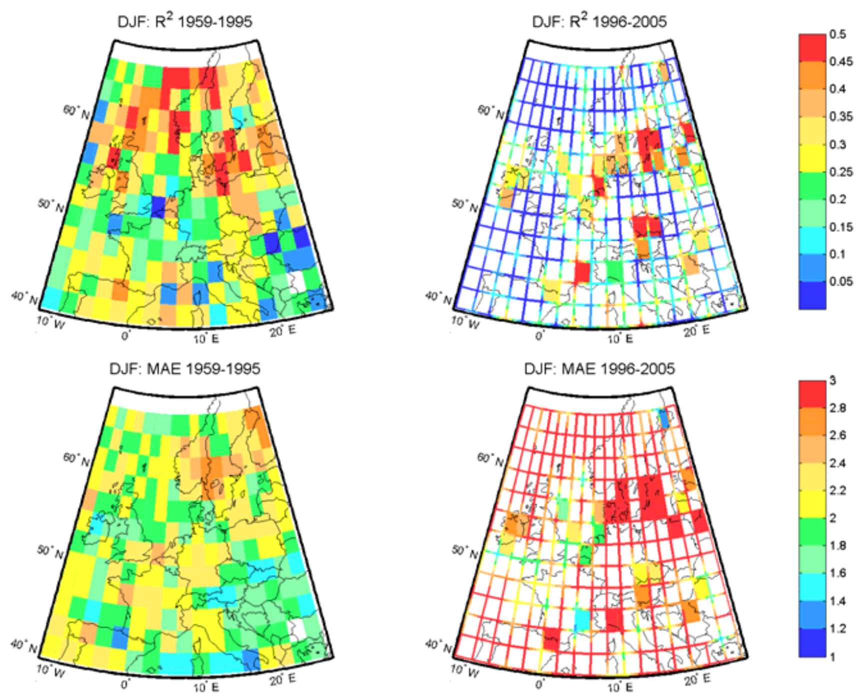

Figure 8.36. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for DJF 90th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

Season	Predictor	Frequency		β
JFM	T50 P3 SON	4		0.253
	T100P2 MAM	1		0.212
	SST3 Lead04	1		0.204
	EAP Lead 05	1		0.183
	POL Lead 04	1		0.174
	TAR Lead 06	1	-	0.150
	IND P5 MJJ	1	-	0.175
	PCM P4 OND	1	-	0.187
	PCM P4 JAS	1	-	0.216
	EAP Lead 03	1	-	0.229
FMA	SOI Lead 07	2	-	0.131
	IND P4 ASO	2	-	0.154
	T150P1 NDJ	1		0.171
	T30 P3 NDJ	1	-	0.130
	AOS Lead 07	1	-	0.191
	H30 P1 NDJ	1	-	0.203
MAM	SDNH Lead10	2	-	0.165
	IND P2 DJF	1		0.189
	PCM P2 SON	1		0.184
	T150P1 NDJ	1		0.152
	EPN Lead 02	1		0.112
	EAP Lead 02	1		0.109
	IND P2 AMJ	1	-	0.172
	IND P3 SON	1	-	0.236
AMJ	SST6 Lead04	2	-	0.135
	QBO5 Lead09	1		0.175
	IND P5 OND	1		0.169
	AOS Lead 02	1		0.160
	T30 P3 NDJ	1		0.147
	SDEU Lead07	1		0.102
	SDNG Lead05	1		0.086
	H50 P1 JFM	1	-	0.180
	NAO Lead 11	1	-	0.183
MJJ	N4 Lead 11	1		0.216
	GLO P4 DJF	1		0.213
	T150P3 SON	1		0.206
	PNA Lead 05	1		0.158
	NAO Lead 11	1		0.157
	PV100P2FMA	1		0.140
	T150P2 ASO	1		0.137
	SDNA Lead10	1		0.113
	T30 P3 FMA	1	-	0.120
	T30 P3 ASO	1	-	0.131
	H150P3 MJJ	1	-	0.151
	H30 P3 NDJ	1	-	0.159
	PV30 P3SON	1	-	0.167
	PCM P3 DJF	1	-	0.190
	PV150P1OND	1	-	0.194
	PV50 P2SON	1	-	0.252
JJA	T50 P3 ASO	2	-	0.126
	PV150P2NDJ	2	-	0.175
	POL Lead 09	1		0.161
	T30 P3 MAM	1		0.131
	AOS Lead 05	1		0.100
	WPP Lead 05	1	-	0.140
	GLO P4 NDJ	1	-	0.149
	AOS Lead 04	1	-	0.244
JAS	H150P3 NDJ	4		0.211
	PV150P2NDJ	3	-	0.168
	SDNG Lead11	2		0.155
	SDNG Lead06	1		0.154
	WPP Lead 10	1		0.127
	AOS Lead 05	1	-	0.146
	PCM P5 JFM	1	-	0.229
ASO	SHT Lead 10	4		0.167
	PV150P2DJF	3		0.193
	PCM P4 AMJ	2	-	0.154
	IND P2 ASO	2	-	0.165
	T150P1 NDJ	1		0.178
	H150P3 DJF	1		0.157
	PV100P2JFM	1		0.154
	PNA Lead 06	1		0.144
	PTP Lead 11	1		0.120
	H50 P1 DJF	1	-	0.106
	ATL P4 MJJ	1	-	0.109
	T30 P2 DJF	1	-	0.114
	ATL P5 AMJ	1	-	0.162
	PV150P3AMJ	1	-	0.175
SON	SHT Lead 10	4		0.162
	T100P3 JFM	2		0.183
	SST1 Lead02	2	-	0.166
	T30 P2 DJF	2	-	0.184
	H100P1 DJF	1		0.175
	PV30 P3OND	1		0.166
	T100P2 DJF	1		0.157
	T50 P1 AMJ	1		0.151
	SHT Lead 04	1		0.132
	NAO Lead 05	1	-	0.154
OND	SCA Lead 08	7		0.194
	PCM P5 JFM	7		0.063
	GLO P2 MJJ	3	-	0.190
	EAP Lead 09	2		0.223
	SCA Lead 02	2	-	0.186
	T50 P2 JFM	2	-	0.191
	ATL P5 OND	2	-	0.224
	SDNH Lead05	1		0.249
	SST5 Lead02	1		0.186
	ATL P2 NDJ	1		0.186
	SHT Lead 10	1		0.172
	IND P5 MAM	1		0.166
	IND P5 AMJ	1		0.157
	SHT Lead 11	1		0.155
	EAP Lead 02	1		0.155
	SHT Lead 05	1		0.137
	H30 P1 DJF	1	-	0.097
	PV100P1AMJ	1	-	0.105
	PV50 P3FMA	1	-	0.165
	SDNH Lead10	1	-	0.180
NDJ	T50 P3 JAS	6		0.207
	GLO P2 MJJ	3	-	0.192
	AOS Lead 02	1		0.237
	EAP Lead 03	1		0.188
	PCM P2 MAM	1		0.184
	SST3 Lead04	1		0.152
	AOS Lead 04	1		0.145
	ATL P1 ASO	1	-	0.138
	TAR Lead 04	1	-	0.150
	T30 P2 JFM	1	-	0.189
	PV50 P3JFM	1	-	0.213
DJF	T50 P3 SON	9		0.249
	T100P2 MAM	7		0.274
	H150P3 JAS	6		0.237
	SST5 Lead03	4		0.190
	T50 P3 JAS	3		0.243
	AOS Lead 03	3		0.209
	TAR Lead 05	3	-	0.054
	SST3 Lead05	2		0.215
	POL Lead 03	2		0.194
	EAP Lead 04	2		0.185
	SST1 Lead02	2	-	0.008
	TNH Lead 11	2	-	0.200
	DAR Lead 04	1		0.290
	AOS Lead 05	1		0.235
	IND P4 JJA	1		0.226
	H100P2 SON	1		0.190
	SST2 Lead06	1		0.179
	EPN Lead 06	1		0.170
	IND P5 JAS	1	-	0.140
	PV30 P2SON	1	-	0.243
	POL Lead 08	1	-	0.251
	T150P2 ASO	1	-	0.276

Table 8.3. Predictor frequency for 90th percentile wind exceedance models, where skill is retained in the validation period. For each season, all predictors which feature in skillful models are included, together with their frequency of occurrence, and the mean value of the model coefficient (β).

8.3.4 95th Percentile Wind Model Validation: Overview

Results for the 95th percentile wind exceedance models are presented in the same format as the other predictands shown above. First a summary of the breakdown in skill is presented by season, and particular features of interest are noted.

- JFM: Figure 8.37 illustrates the degradation in skill for JFM. The same region of validation period skill identified in DJF for the 90th percentile wind models is present in JFM for the 95th percentile wind models. There is a slight shift to the east, particularly for the southern component, and the R^2 values are of a similar magnitude, as are most of the MAE values relative to the predictand climatology. These results will be discussed in more detail in 8.3.8.

- FMA: Figure 8.38 illustrates the degradation in skill for FMA. The pattern of skillful models in JFM is not present in FMA, and only three gridboxes show potentially useful skill. These are located over Norway, southern France and Spain.

- MAM: Figure 8.39 illustrates the degradation in skill for MAM. Five models retain some skill in the validation period. These are scattered across the domain, with no spatial coherence, and generally low R^2 values.

- AMJ: Figure 8.40 illustrates the degradation in skill for AMJ. Four gridboxes retain skill, located over eastern Germany and the Balkan states. R^2 values range from 0.21 to 0.53.

- MJJ: Figure 8.41 illustrates the degradation in skill for MJJ. Five models retain some skill, although these are scattered across the domain, and skill levels are relatively low.

- JJA: Figure 8.42 illustrates the degradation in skill for JJA. No models retain potentially useful skill into the validation period.

- JAS: Figure 8.43 illustrates the degradation in skill for JAS. Some skill is retained for models located over Belgium, western France and Spain, with R^2 values between 0.22 and 0.49. One model off the west coast of France has R^2=0.49 for the validation period, and an MAE of 1.1. The predictors for this model are North American and Greenland snow cover and the SCA teleconnection pattern, both from the preceding September.

- ASO: Figure 8.44 illustrates the degradation in skill for ASO. Ten models retain skill. These are mostly scattered across eastern and central Europe, with some skill over the Mediterranean.

- SON: Figure 8.45 illustrates the degradation in skill for SON. Five models retain some skill, located over Scandinavia and Eastern Europe. R^2 values range from 0.20 to 0.47.

- OND: Figure 8.46 illustrates the degradation in skill for OND. Seven models retain some skill during the validation period. Three of these are located over the UK, with R^2 values ranging from 0.43 to 0.55. the dominant predictor for these models is the third PC of JAS 50hPa temperature.

- NDJ: Figure 8.47 illustrates the degradation in skill for NDJ. A similar pattern of persistent skill over the UK exists in NDJ, with R^2 levels ranging from 0.40 to 0.80. The third PC of JAS 50hPa temperature, and 150hPa geopotential height are the dominant predictors.

- DJF: Figure 8.48 illustrates the degradation in skill for DJF. The pattern observed in the 90^{th} percentile wind exceedance models for DJF, and the 95^{th} percentile models for JFM is present here, although to a substantially lesser extent, particularly in the southern region. The Baltic region shows relatively high levels of potential skill, with R^2 values ranging from 0.28 to 0.55. Predictors for these models include the third PC of SON 50hPa temperature, the early autumn Arctic Oscillation, and other stratospheric indices.

Table 8.4 shows the predictors retained, by season, for the 95th percentile wind models. For each season, they are ranked by the frequency of occurrence and parameter (β). Slightly fewer predictors contribute to potentially useful skill here than for the 90th percentile wind models, although a greater number are retained than for either of the precipitation predictands. JFM has the largest number of potentially useful models, and as for the 90th percentile wind models, the third PC of SON 50hPa temperature features relatively prominently, in eight gridboxes. Interestingly, the first PC of this variable during SON also features. While these modes are orthogonal, it may be that they describe respectively the intensity and location of anomalies, and both contribute towards potential predictability. Stratospheric temperature predictors also feature in other seasons throughout the year, but are particularly important in the autumn and winter. The AO is also relatively important, and to a lesser extent the NAO. Other Northern Hemisphere teleconnection patterns also feature, most notably the EA pattern during the winter seasons. Indian Ocean SST predictors – generally from the preceding summer and autumn – contribute to potentially useful skill in a small number of locations during the spring, and also to some extent in the ASO and NDJ seasons, where SST anomalies with lead times of up to a year appear to be significant. Snow cover indices contribute to a slightly larger number of models than for the 90th percentile wind models, and North American (including Greenland) snow cover anomalies from the preceding August feature in three models during the JAS season. Since snow cover anomalies are typically related to NAO variability during the winter, it is not clear what causes this relationship, particularly since it is driven by August snow anomalies, which are likely to be less widespread in absolute terms.

Validation results for the 95th percentile wind exceedance models are generally similar to those for the 90th percentile wind models. Overall, skill levels decrease dramatically beyond the training period, although some regions do show coherent skillful responses to the models throughout the validation period, particularly in the autumn and winter seasons. Further detail on the most notable areas where skill is retained is discussed further in 8.3.8.

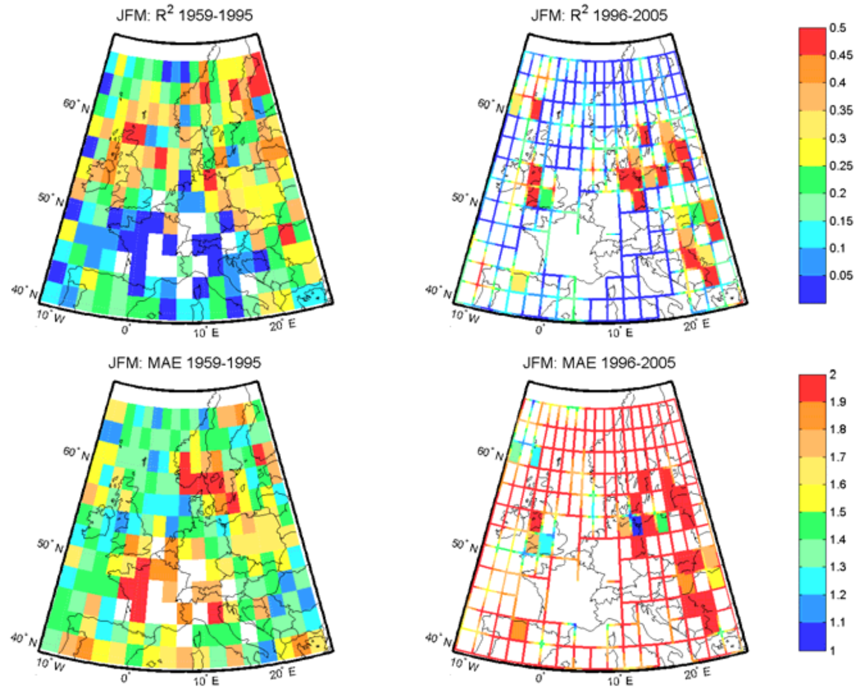

Figure 8.37. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JFM 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

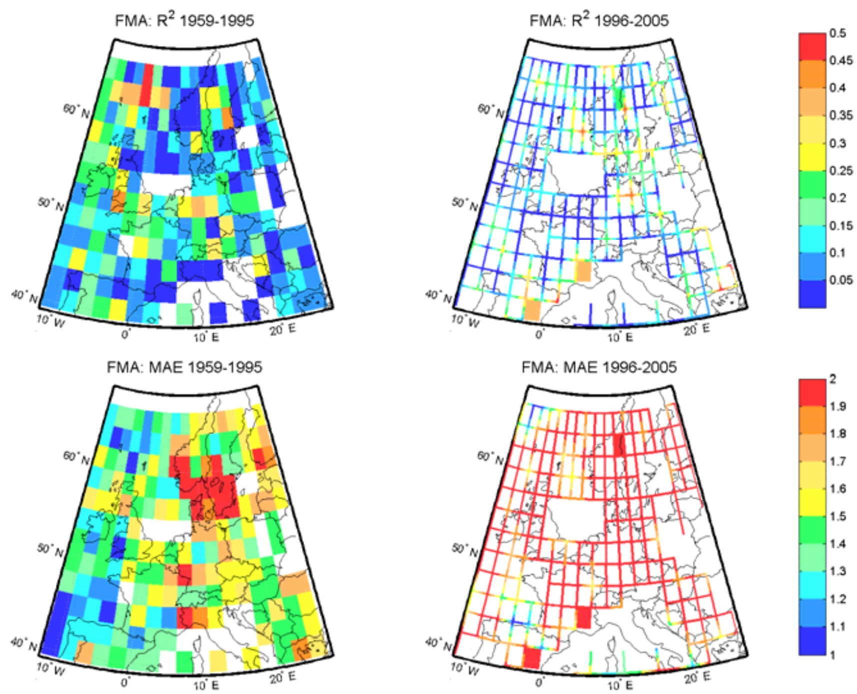

Figure 8.38. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for FMA 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

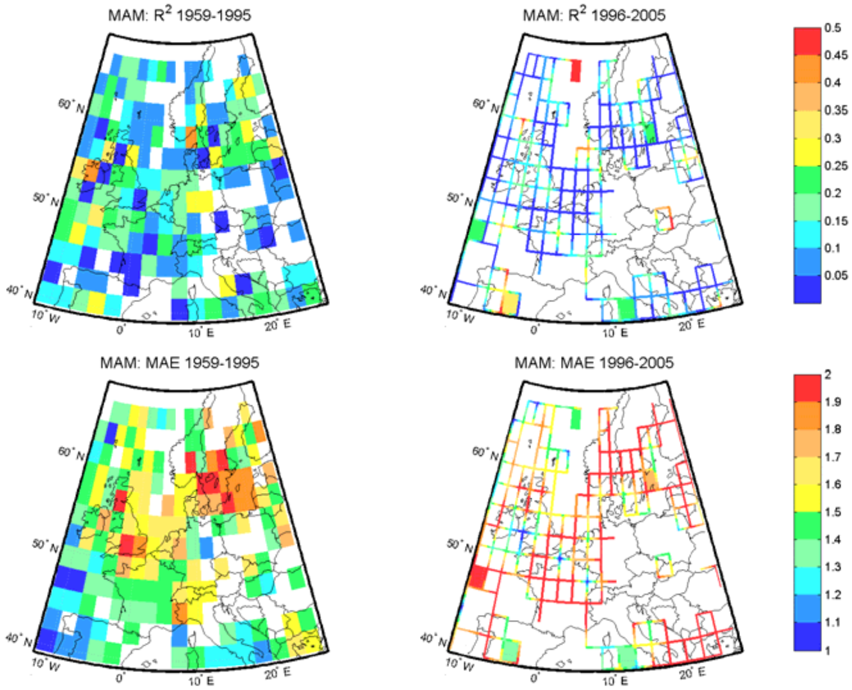

Figure 8.39. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MAM 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

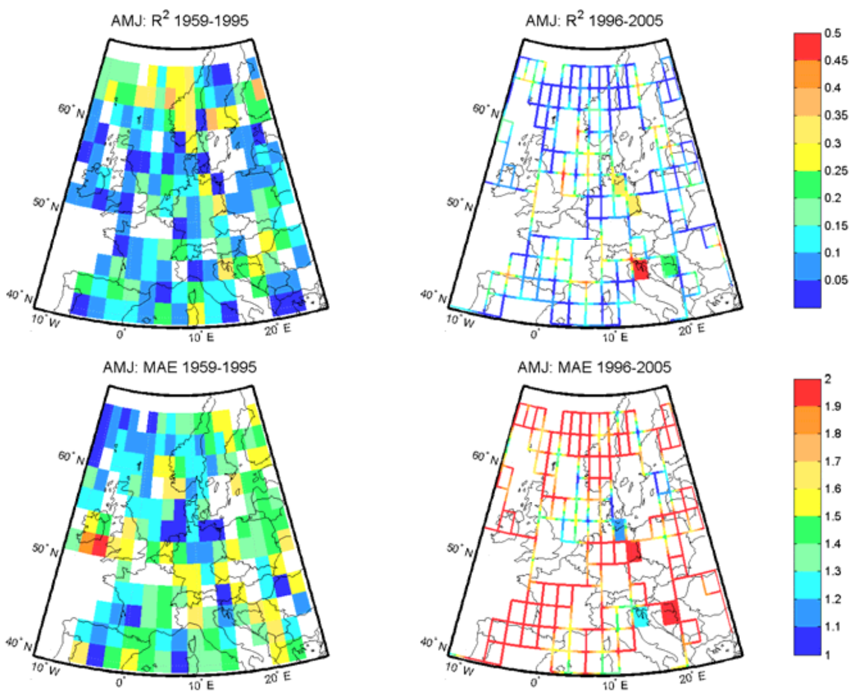

Figure 8.40. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for AMJ 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

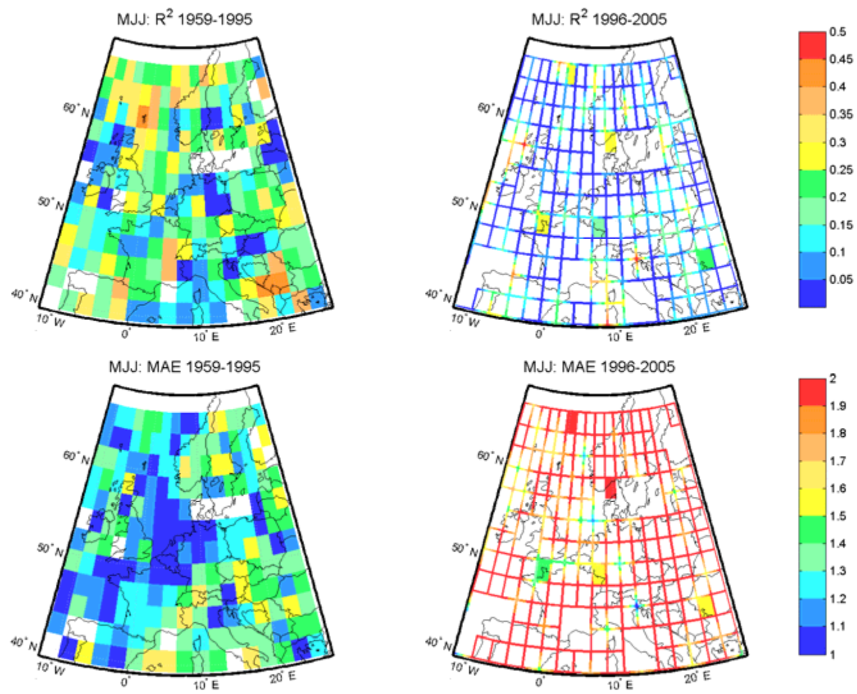

Figure 8.41. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for MJJ 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

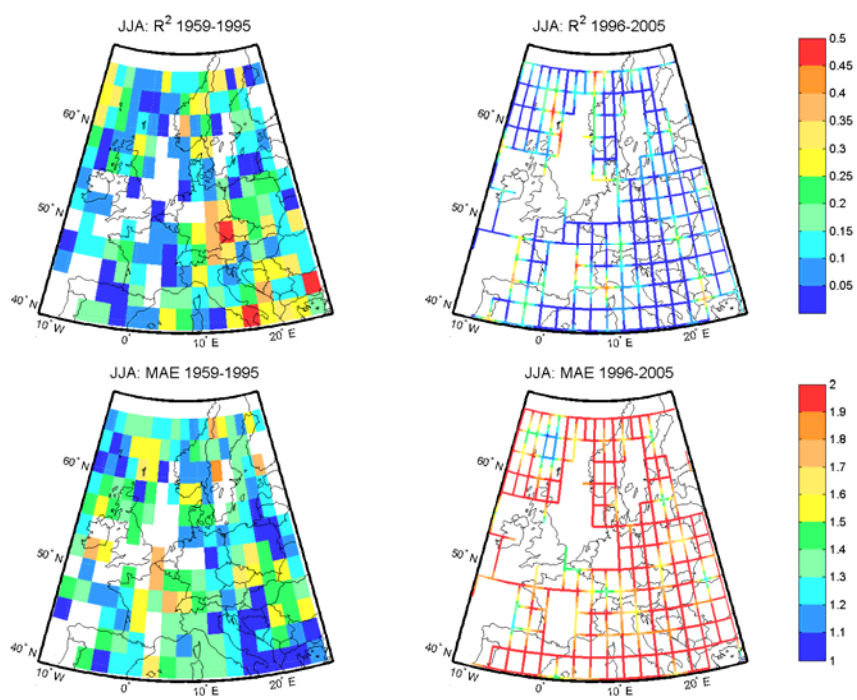

Figure 8.42. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JJA 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

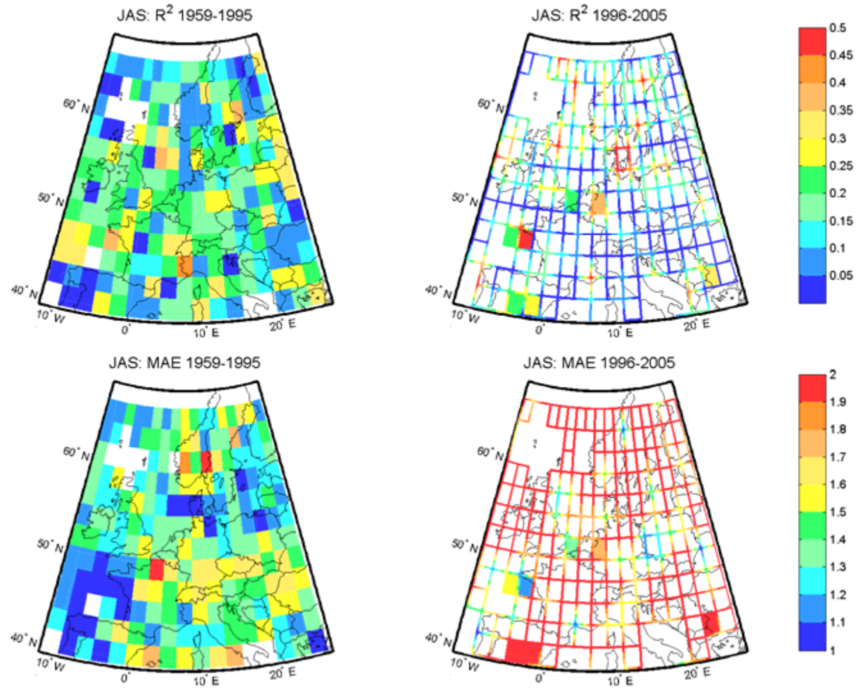

Figure 8.43. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for JAS 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

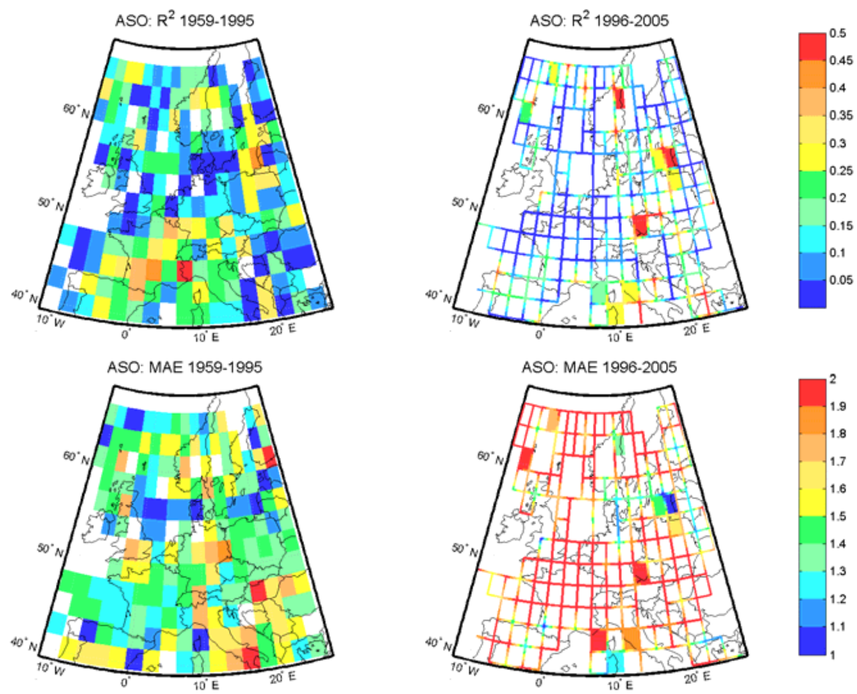

Figure 8.44. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for ASO 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

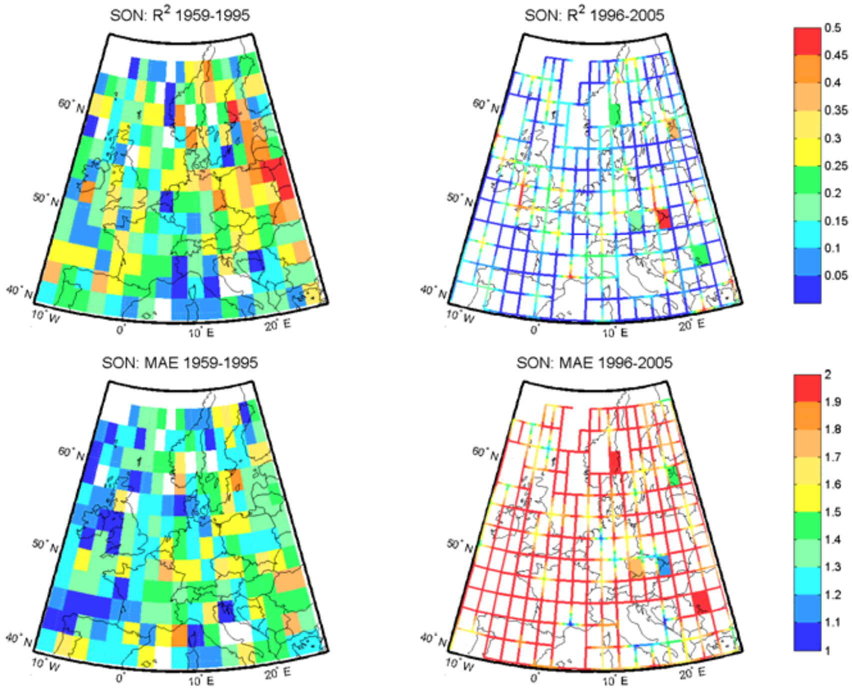

Figure 8.45. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for SON 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

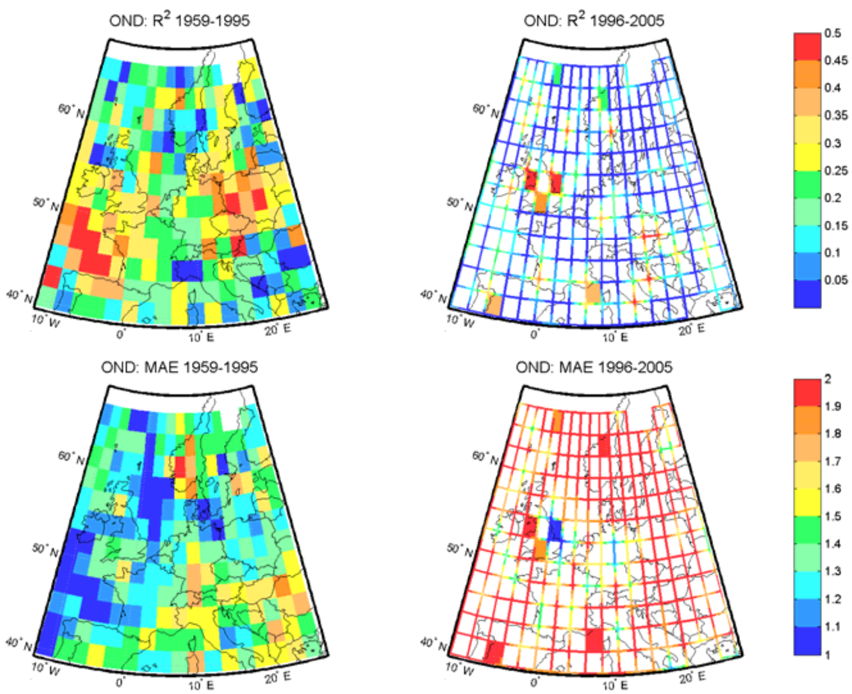

Figure 8.46. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for OND 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

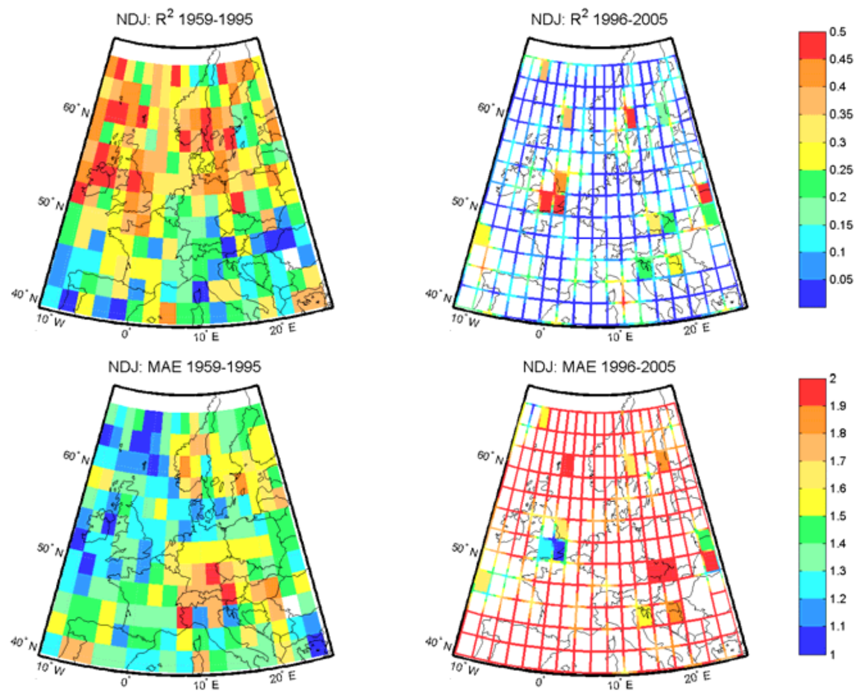

Figure 8.47. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for NDJ 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

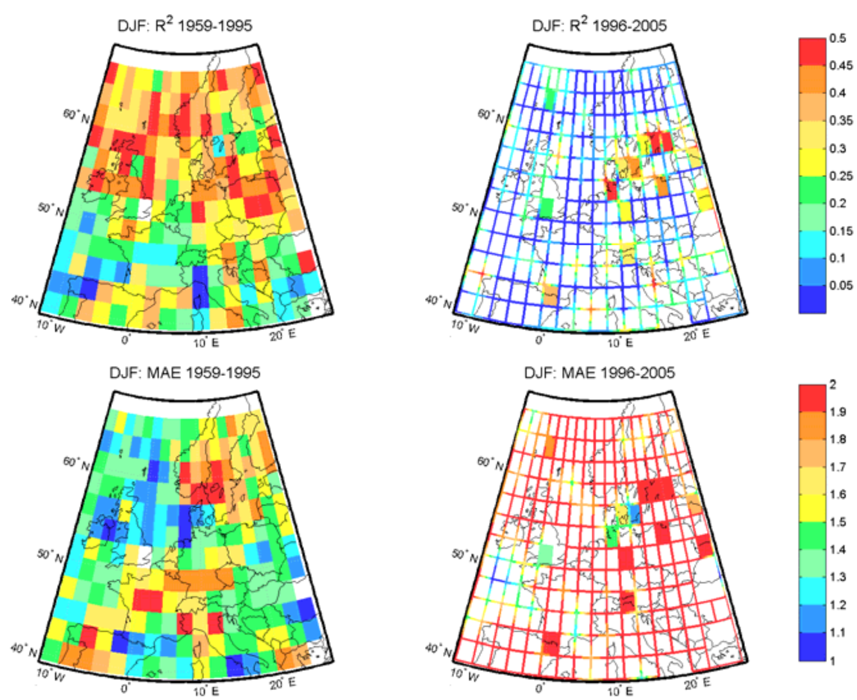

Figure 8.48. Change in R2 and MAE values from model training period (1959-1995) to model validation period (1996-2005) for DJF 95th percentile wind exceedance models. Solid gridboxes in the validation plots indicate models where useful skill persists beyond the training period.

Season	Predictor	Frequency		β
JFM	T50 P3 SON	8		0.447
	EAP Lead 05	4		0.351
	T50 P1 SON	4		0.241
	AOS Lead 06	3		0.296
	H100P2 SON	3		0.295
	ATL P3 JAS	3		0.181
	T50 P1 JJA	3	-	0.218
	AOS Lead 04	2		0.227
	T150P3 OND	2	-	0.237
	T150P2 AMJ	2	-	0.237
	EAP Lead 03	2	-	0.274
	SDNH Lead08	1		0.366
	SST2 Lead06	1		0.341
	T100P2 MAM	1		0.266
	IND P4 JJA	1		0.256
	POL Lead 04	1		0.235
	PV30 P3JFM	1		0.217
	IND P4 ASO	1	-	0.144
	T30 P3 FMA	1	-	0.204
	T50 P1 ASO	1	-	0.236
	TAR Lead 06	1	-	0.238
	T100P1 MJJ	1	-	0.245
	PCM P4 OND	1	-	0.263
	T50 P3 OND	1	-	0.380
FMA	AOS Lead 07	2		0.076
	TAR Lead 07	1	-	0.141
	SDNH Lead09	1	-	0.178
	IND P2 MJJ	1	-	0.200
MAM	AOS Lead 08	3	-	0.101
	IND P4 JAS	2		0.036
	GLO P5 NDJ	1		0.330
	SDNH Lead10	1		0.224
	EPN Lead 02	1		0.177
	PV50 P3OND	1	-	0.231
AMJ	SDNG Lead05	2		0.221
	NAO Lead 11	2	-	0.182
	IND P3 OND	2	-	0.196
	SDNH Lead07	1		0.220
	POL Lead 10	1	-	0.175
MJJ	T100P3 JFM	1		0.238
	TNH Lead 03	1		0.232
	SDNA Lead10	1		0.218
	PNA Lead 05	1		0.192
	T150P2 ASO	1		0.183
	NAO Lead 11	1		0.141
	PV100P3JAS	1		0.134
	POL Lead 08	1	-	0.163
	H30 P3 NDJ	1	-	0.233
	AOS Lead 11	1	-	0.276
JAS	SDNG Lead11	3		0.205
	PNA Lead 02	2		0.336
	PV30 P2ASO	2		0.268
	SCA Lead 11	2	-	0.224
	SHT Lead 11	1		0.236
	PCM P4 MAM	1	-	0.211
	NAO Lead 07	1	-	0.226
ASO	T30 P2 DJF	3	-	0.245
	IND P1 MJJ	2		0.262
	IND P1 SON	2		0.233
	ATL P5 AMJ	2		0.204
	ATL P2 FMA	2	-	0.225
	PV150P3AMJ	2	-	0.264
	NHT Lead 04	1		0.228
	T30 P1 SON	1		0.227
	T50 P1 MAM	1	-	0.149
	T100P3 FMA	1	-	0.196
	SST2 Lead03	1	-	0.211
	H100P1 DJF	1	-	0.213
SON	SHT Lead 04	2		0.243
	ATL P4 NDJ	2	-	0.152
	AOS Lead 09	1		0.283
	HTD Lead 04	1		0.215
	SHT Lead 10	1		0.148
	EAW Lead 09	1	-	0.183
	DAR Lead 03	1	-	0.240
	SST1 Lead05	1	-	0.320
OND	T50 P3 JAS	3		0.320
	PCM P5 JFM	2		0.238
	TAR Lead 02	2	-	0.248
	PV100P2OND	1		0.321
	PCM P4 AMJ	1		0.270
	H30 P1 NDJ	1		0.209
	ATL P2 OND	1		0.200
	EAW Lead 08	1		0.178
	SDNH Lead10	1	-	0.153
	GLO P4 MAM	1	-	0.289
NDJ	T50 P3 JAS	3		0.311
	PV30 P2MAM	3		0.310
	IND P5 NDJ	3		0.102
	H150P3 JAS	2		0.280
	EAP Lead 03	2		0.240
	SST3 Lead04	1		0.399
	PV30 P3ASO	1		0.344
	IND P5 AMJ	1		0.310
	SDNH Lead06	1		0.285
	ATL P3 JAS	1		0.259
	AOS Lead 04	1		0.237
	PCM P5 JJA	1		0.224
	SST5 Lead03	1		0.193
	SDEU Lead07	1		0.185
	H50 P2 JJA	1		0.171
	GLO P2 MJJ	1	-	0.216
	T30 P2 JFM	1	-	0.261
	T30 P3 FMA	1	-	0.347
DJF	T50 P3 SON	5		0.362
	T100P2 MAM	2		0.458
	T50 P3 JAS	2		0.410
	H150P3 JAS	2		0.406
	EAP Lead 04	2		0.383
	H100P2 SON	2		0.371
	PV50 P3AMJ	1		0.450
	AOS Lead 03	1		0.351
	SST3 Lead06	1		0.337
	PV30 P2MAM	1		0.281
	SDNH Lead07	1		0.273
	EAW Lead 11	1		0.249
	AOS Lead 05	1		0.234
	TNH Lead 11	1	-	0.138
	ATL P2 MJJ	1	-	0.157
	ATL P1 SON	1	-	0.267
	EAP Lead 02	1	-	0.289
	PV50 P3JFM	1	-	0.443

Table 8.4. Predictor frequency for 95th percentile wind exceedance models, where skill is retained in the validation period. For each season, all predictors which feature in skillful models are included, together with their frequency of occurrence, and the mean value of the model coefficient (β).

8.3.5 90th Percentile Precipitation Model Validation: Skillful Models

The most noteworthy case of potentially useful skill that persists into the validation period for the 90th percentile precipitation models is during AMJ, including a number of the gridbox models in Western Europe, from the Bay of Biscay and the Pyrenees, north to the Irish Sea and the Netherlands. A sample of these models is described in Table 8.5, and the exceedance count predictions illustrated in Figure 8.49.

The Irish Sea location model uses indices of stratospheric potential vorticity and temperature from the preceding summer. This is a very long lead time for stratospheric predictors, and there is little evidence to support the persistence of such features. Both predictors are highly correlated with other stratospheric indices into the autumn and winter, and it is possible that either persistence or dynamical evolution of the summer anomalies is pertinent to the troposphere during the following AMJ.

The Bay of Biscay predictors (for both models) are February Tahiti MSLP values, August Darwin MSLP values, and December Baltic SST anomalies. The Baltic SST anomalies are unlikely to have a direct influence on the predictand – it is more likely that they are the result of some large-scale dynamical process which affects the Bay of Biscay also, although this index is not correlated with any of the Bay of Biscay SST indices (SST region 4). The SOI indices selected here are weakly correlated with NAO indices the following spring.

Despite the apparent spatial concentration of skill that persists into the validation period, the models and predictors associated with this skill do not lend themselves to easy interpretation – there being a large number of different predictors, which do not appear to be related between all the models, and with some of them having long lead times which do not plausibly signify a direct link between that predictor and the subsequent precipitation extremes.

For most of these models, while the R^2 values appear significant, the MAE values are all higher during the validation period. Figure 8.49 illustrates a consistent negative bias to the forecasts, partly due to an apparent slight upward trend in the precipitation extremes in these gridboxes.

Location	Irish Sea	Bay of Biscay (W)	English Channel	Bay of Biscay (E)	Netherlands
Predictor1	PV100 P3 JAS	TAR LEAD 03	GLO P2 AMJ	DAR LEAD 09	H100 P1 SON
Predictor2	T150 P3 JJA	SST3 LEAD 05	PV30 P3 AMJ	SST3 LEAD 05	IND P2 NDJ
β_0	1.9811	1.8058	1.7886	1.7387	1.7759
β_1	0.2361	-0.2255	0.225	0.1808	-0.278
β_2	-0.1535	-0.211	0.2171	-0.282	0.1748
R^2 training	0.231	0.293	0.3537	0.3053	0.2655
R^2 validation	0.5037	0.4634	0.3667	0.4355	0.4001
MAE training	2.3239	1.7293	1.9952	1.7771	1.8438
MAE validation	3.2226	4.2155	3.4529	4.3597	3.1056

Table 8.5. Selected 90th percentile precipitation model specifications and skill for the AMJ season. Five gridbox locations are included where potentially useful skill persists into the validation period. Gridbox locations are given in the first row, and can be compared with the top right panel of Figure 8.4. Predictor names, together with the predictor season or lead time, and the associated coefficients complete the model specifications, and R^2 and MAE values from the training and validation periods are included for comparison.

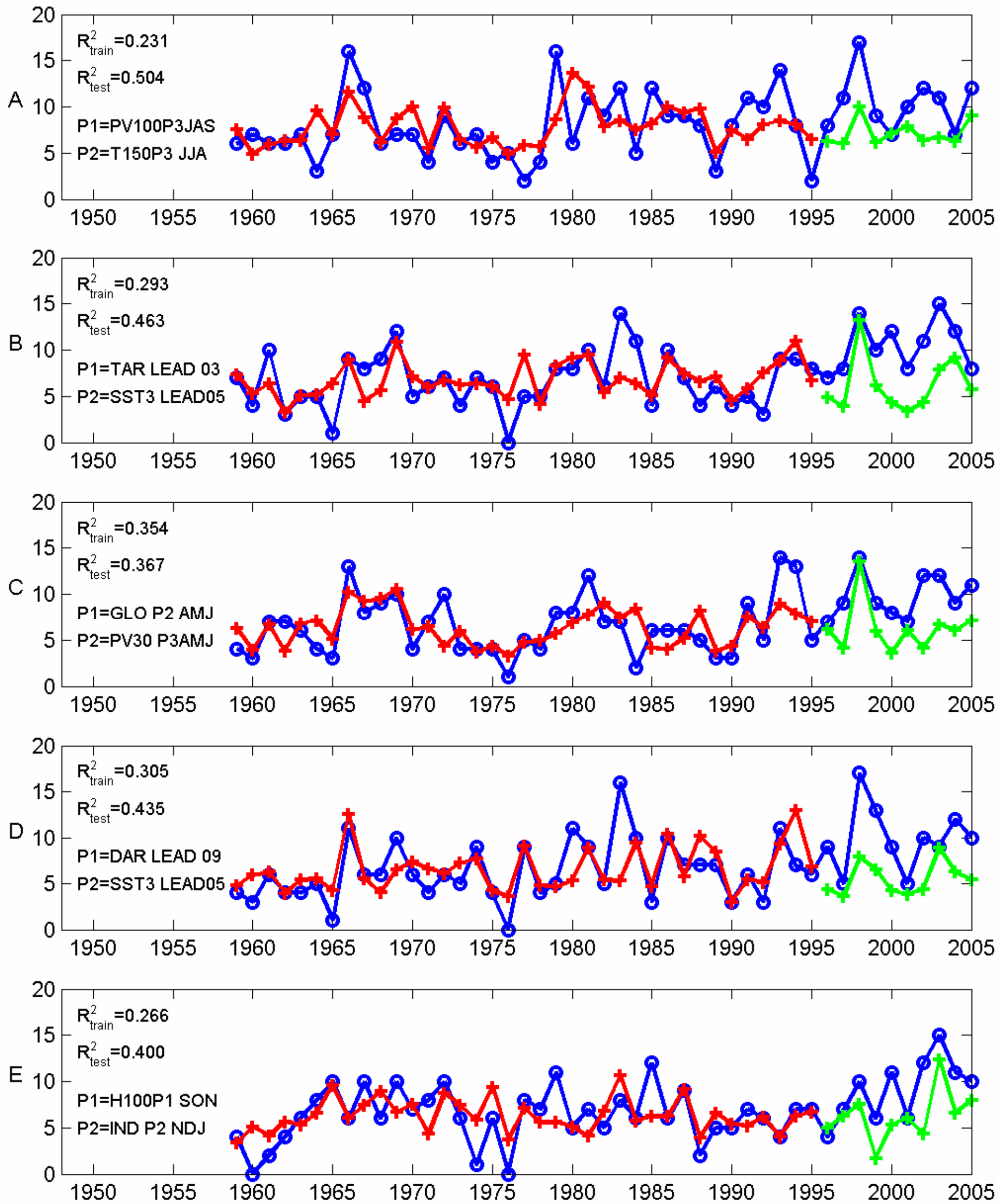

Figure 8.49. Observed and predicted AMJ 90th percentile precipitation exceedance counts for selected locations across western Europe, where potentially useful predictive skill persists into the validation period. Blue lines show the observed values, red shows the values predicted during the training period, and green shows the validation period predictions. R^2 values during the training and validation period are shown for each gridbox, as are the predictors for each model. Locations are as follows: A – Irish Sea; B – Bay of Biscay (W); C – English Channel; D – Bay of Biscay (E); E - Netherlands

8.3.6 95th Percentile Precipitation Model Validation: Skillful Models

For the 95th percentile precipitation models, AMJ also shows the largest region of coherent skill that persists throughout the validation period, comprising a group of gridboxes over the south and west UK, as well as others in France, Benelux and Scandinavia (see Figure 8.16). Table 8.6 shows the parameters for five selected models, located over the UK and surrounds, including indicators of models skill during the training and validation periods. Figure 8.50 shows the corresponding observed and predicted timeseries. As with the 90th percentile models discussed in 8.3.5, the fit in the training and validation periods is quite good, with respect to the R^2 metric. However the MAE increases notably during the validation period – and as is the case with the 90th percentile models, this appears to be due to a positive tendency in the exceedance counts from 1995 onwards.

The predictors selected for these models superficially bear little resemblance to those selected for the 90th percentile models – however there are some correlations between the indices. In particular, the SON stratospheric temperature index in the Irish Sea model is highly correlated with simultaneous ENSO indices (both SST and MSLP indices), including the Niño 4 region index at 10 months lead time, selected for the central England gridbox. More generally, this is a feature of a number of the stratospheric indices – they seem to lag ENSO/SOI indices by some months, in some cases with remarkably strong correlations. This response has been noted by Kodera *et al.* (1996), Broennimann (2004) and Broennimann *et al*. (2007).

Together with the AMJ response of 90th percentile precipitation, these comprise the most notable potentially useful predictability in the precipitation predictands on the basis of the analyses carried out. A wide range of predictors are identified as having potential skill, with SOI and stratospheric predictors featuring most strongly. Further work is required on the mechanisms which might link these processes before seasonal forecasts can be made with confidence.

Location	Irish Sea	Scilly Isles	North Wales	South Wales	Central England
Predictor1	T100 P2 SON	SDEU LEAD 06	H100 P1 NDJ	EAW LEAD 11	N4 LEAD 10
Predictor2	PNA LEAD 09	PV50 P3 AMJ	EAW LEAD 06	ATL P1 JJA	SST2 LEAD 03
β_0	1.3026	1.3424	1.4108	1.1922	1.1699
β_1	0.2753	0.2131	-0.203	0.2173	0.3257
β_2	-0.213	-0.455	0.2862	-0.178	-0.352
R^2 training	0.2098	0.3334	0.2697	0.2256	0.3145
R^2 validation	0.2341	0.3326	0.4245	0.7807	0.3212
MAE training	1.7509	1.2451	1.7342	1.3918	1.5811
MAE validation	2.3446	2.9268	3.3819	3.6703	1.7275

Table 8.6. Selected 95th percentile precipitation model specifications and skill for the AMJ season. Five gridbox locations are included where potentially useful skill persists into the validation period. Gridbox locations are given in the first row, and can be compared with the top right panel of Figure 8.16. Predictor names, together with the predictor season or lead time, and the associated coefficients complete the model specifications, and R^2 and MAE values from the training and validation periods are included for comparison.

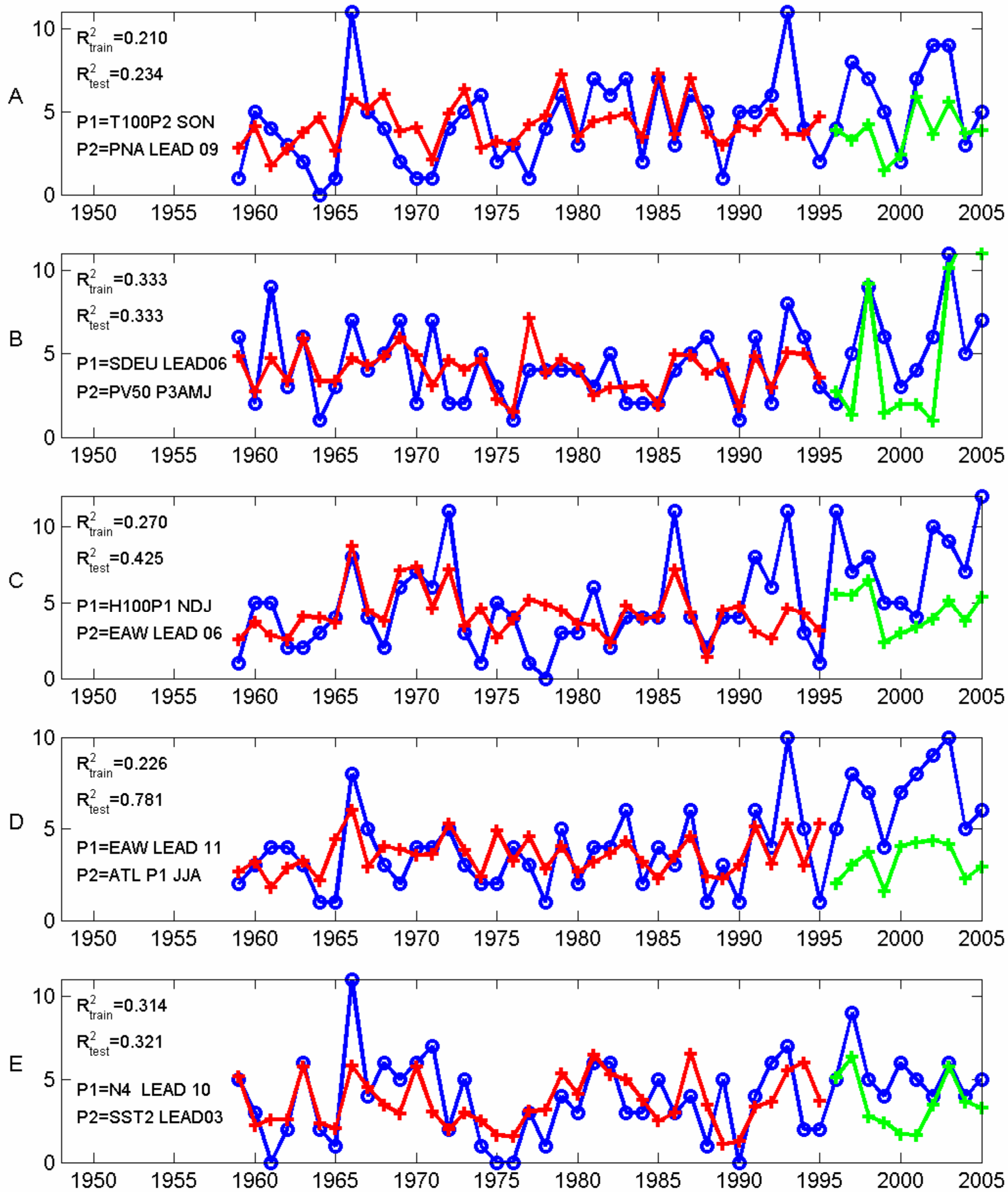

Figure 8.50. Observed and predicted AMJ 95th percentile precipitation exceedance counts for selected locations across the UK, where potentially useful predictive skill persists into the validation period. Blue lines show the observed values, red shows the values predicted during the training period, and green shows the validation period predictions. R^2 values during the training and validation period are shown for each gridbox, as are the predictors for each model. Locations are as follows: A – Irish Sea; B – Scilly Isles; C – North Wales; D – South Wales; E – South-Central England

8.3.7 90th Percentile Wind Model Validation: Skillful Models

A number of responses stand out for the 90th percentile wind models. Those discussed here are the NDJ response over the Benelux countries, and the DJF response over the Baltic, and southern and central Europe.

The NDJ response over the Benelux countries is summarised in Table 8.7, which describes the model parameters and skill for five selected gridboxes showing persistent skill, Figure 8.52 which illustrates the observed and predicted timeseries for the same gridboxes, and Figure 8.35 which illustrates the spatial pattern of the response.

The model fit in the validation period is substantially higher than that for the precipitation models, with MAE values that are similar or in some cases lower than during the training period. The predictor set is more constrained, with every model in this case having the third PC of JAS 50hPa temperature as a predictor, and a number of the models sharing the 2nd PC of MJJ global SST anomalies as a predictor. JFM stratospheric temperature, and Tahiti MSLP from the preceding August comprise the rest of the predictors.

The JAS 50hPa temperature predictor is shown in Figure 8.51, and consists of a dipole with the positive loadings over Canada and Greenland, and the negative loadings over Alaska and eastern Siberia. This pattern is also related to North American and Greenland snow cover anomalies during October and November, as well as the September EA pattern.

It is not clear how this pattern might physically relate to NDJ wind extremes – in particular, does it propagate downwards into the troposphere, and if so over what period of time? Also, does it cause the snow cover anomalies over North America, or is it driven by these anomalies? The nature of the lagged relationship surely indicates the former. However, this mode is only associated with 3% of variance in this field. While is certainly appears to be a coherent pattern, it must also be questioned whether the variance can be measured accurately, and whether such a predictor can be regarded as useful. It is worth noting that the first PC of this field during JAS is associated with a strong trend, and accounts for 66% of the total variability. If the detrending stage is carried out prior to the PCA the third PC accounts for 5% of the observed variability – a somewhat greater amount, but still possibly too small to justify its association with the observed predictability. This is certainly an area that warrants further investigation.

The other notable predictor during NDJ is the 2nd PC of MJJ global SST anomalies. This pattern describes the MJJ variability of ENSO, and is positively correlated to the SOI indices, whereas the first PC is positively correlated with the Nino regional SST anomalies. It also seems to be related to tropical Atlantic variability, more so than the first PC of Pacific variability. The SOI link seems to correspond with the observed pattern in this study that atmospheric variability in the tropical Pacific seems to be more important as a predictor for the European climate than variability in the SSTs – in other words, some independent component of the atmospheric variability is of interest here. This possibly accounts for the fact that generally the responses to Tahiti MSLP are stronger than those for Darwin, and a further analysis of tropical Pacific MSLP might yield a more optimal predictor for the European climate.

Location	France (W)	Belgium (W)	France (E)	Netherlands	Belgium (S)
Predictor1	T50 P3 JAS	T50 P3 JAS	T50 P3 JAS	T50 P3 JAS	T50 P3 JAS
Predictor2	TAR LEAD 04	T30 P2 JFM	GLO P2 MJJ	GLO P2 MJJ	GLO P2 MJJ
β_0	1.8381	1.8203	1.8177	1.777	1.863
β_1	0.2321	0.2335	0.2581	0.2141	0.1622
β_2	-0.1497	-0.1893	-0.1729	-0.1932	-0.2085
R^2 training	0.312	0.318	0.4637	0.2893	0.4995
R^2 validation	0.3544	0.3739	0.6053	0.4409	0.4091
MAE training	2.0459	2.017	1.8199	2.0439	1.4916
MAE validation	2.4244	1.9256	1.6567	1.9488	1.6141

Table 8.7. Selected 90th percentile wind model specifications and skill for the NDJ season. Five gridbox locations are included where potentially useful skill persists into the validation period. Gridbox locations are given in the first row, and can be compared with the top right panel of Figure 8.35. Predictor names, together with the predictor season or lead time, and the associated coefficients complete the model specifications, and R^2 and MAE values from the training and validation periods are included for comparison.

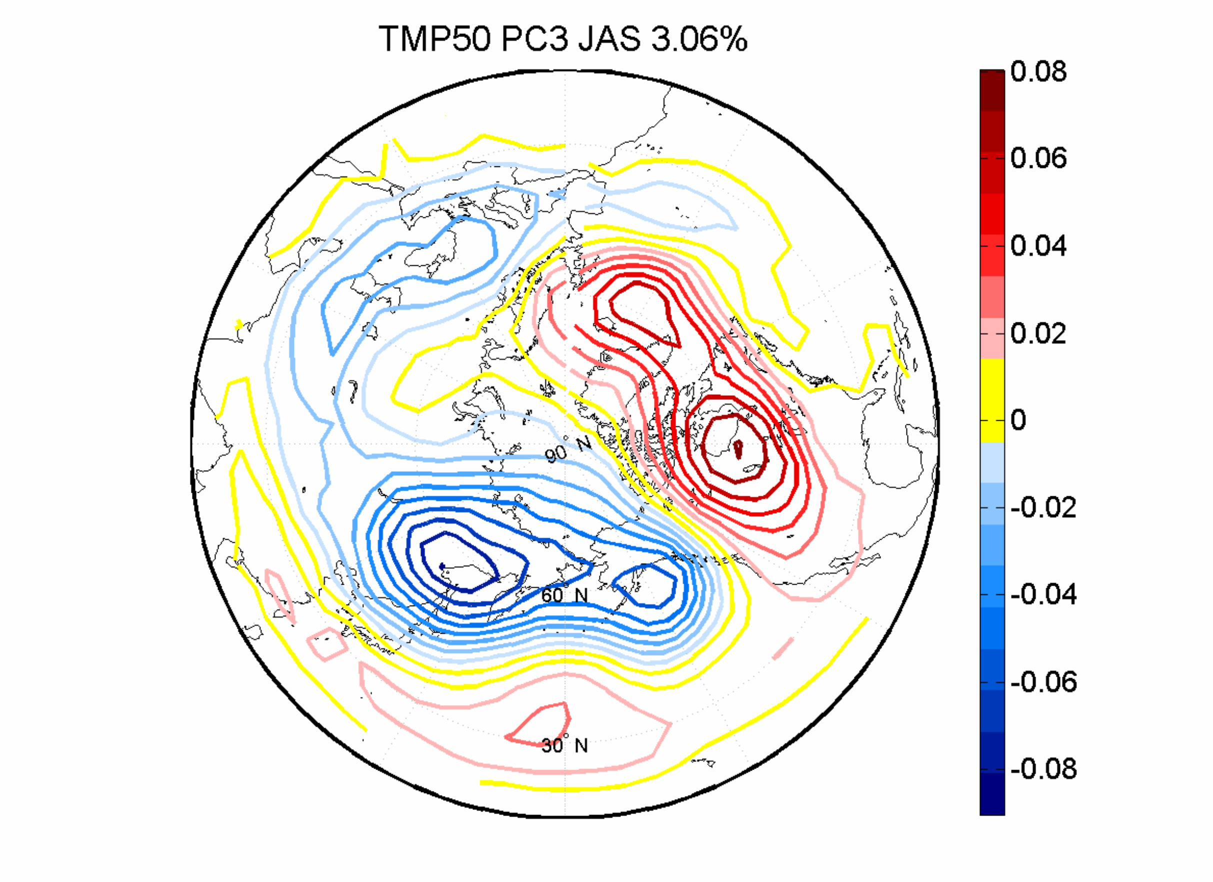

Figure 8.51. The third principal component of JAS 50hPa temperature anomalies. This predictor shows potentially useful skill in predicting 90th percentile wind exceedances over regions of western Europe during the following NDJ season. The pattern comprises positive loadings over Canada and Greenland, and negative loadings over Alaska and eastern Siberia, and accounts for 3.06% of variability in this field.

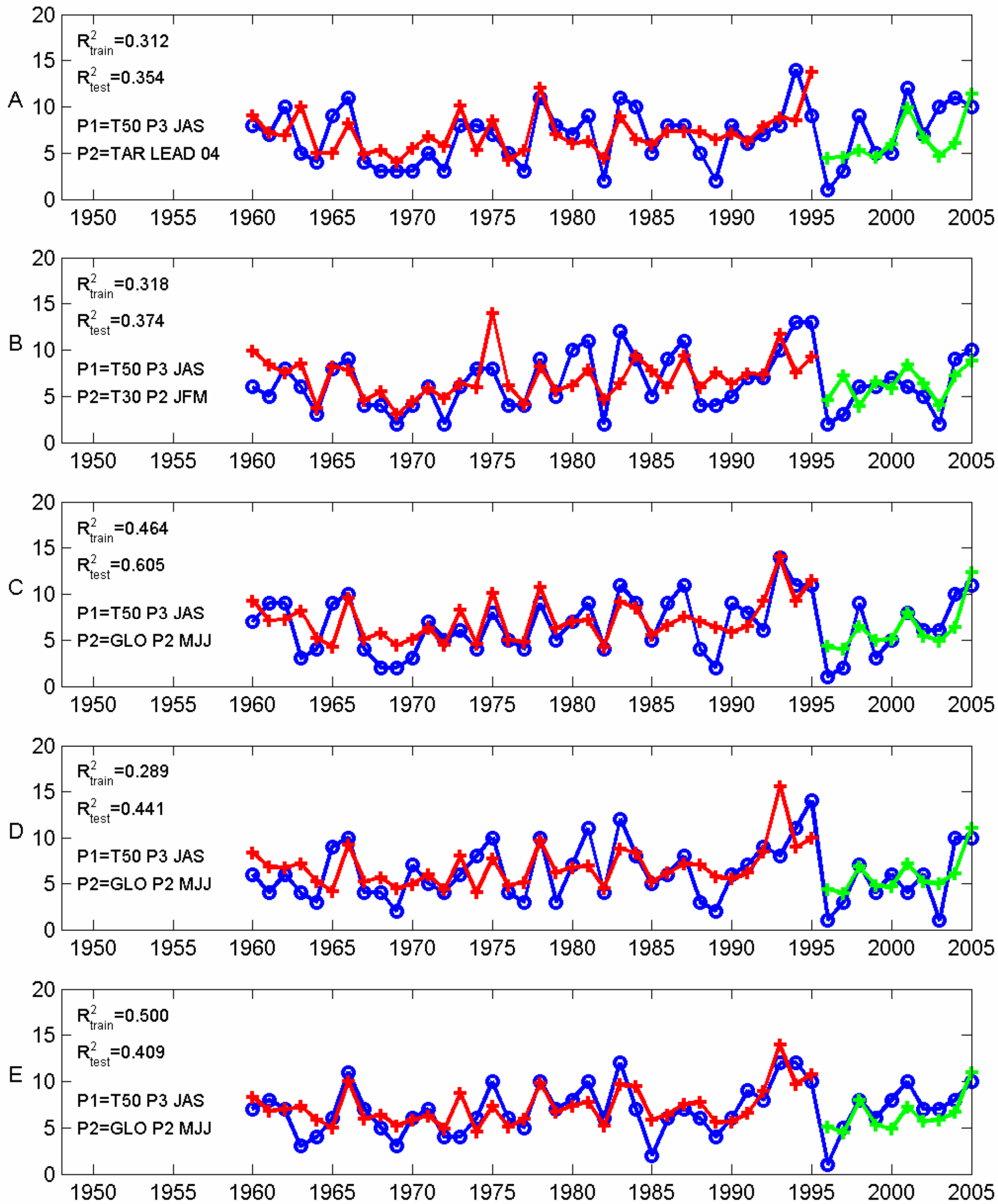

Figure 8.52. Observed and predicted NDJ 90th percentile wind exceedance counts for selected locations across the Benelux region, where potentially useful predictive skill persists into the validation period. Blue lines show the observed values, red shows the values predicted during the training period, and green shows the validation period predictions. R^2 values during the training and validation period are shown for each gridbox, as are the predictors for each model. Locations are as follows: A – France (W); B – Belgium (W); C – France (E); D – Netherlands; E – Belgium (S)

The DJF season has the largest coherent regions of persistent skill for the 90th percentile wind models (and indeed for all the predictand variables). The spatial pattern is illustrated in Figure 8.36, and comprises a region of skill stretching from the North Sea to the Baltic states (referred to here as the northern region), and a region from the Pyrenees to the Czech Republic and the Balkans (referred to here as the southern region).

Selected models for the northern region are described in Table 8.8, and the corresponding observed and predicted timeseries are illustrated in Figure 8.53. Generally the interannual variability in the extremes is captured well throughout the validation period, although the magnitude of the variance is sometimes low compared to the observations. The T50 P3 JAS predictor features in one of the models over the Netherlands, as for the NDJ models, and other stratospheric predictors related to this, such as the equivalent index during SON also feature. The 2nd PC of 100hPa temperature is also frequently selected. This is a quasi-annular pattern, with positive loadings over the North Pole, and negative loadings in the midlatitudes, which are almost continuous but interrupted over North America. This pattern is weakly correlated with the AO and NAO, and this relationship is strongest when it leads the AO by several months, into the late summer. This lead time is not sufficient to directly account for the observed predictability.

The southern region shows similarly strong fitting compared to the northern region, as illustrated in Figure 8.55, although the predictor set is somewhat different. Local SST anomalies feature more prominently – in particular the SST5 region, which comprises the Adriatic, and is a predictor for the Czech Republic models, as shown in Table 8.9. Stratospheric temperature predictors from the preceding summer and autumn also feature prominently.

90th percentile wind models show the most skill during the winter seasons, and it appears that the dominant features driving this predictability are related to stratospheric temperature, at lead times of up to eight months, and to a lesser extent indices associated with tropical Pacific atmospheric variability. Further work is required to uncover physical mechanisms supporting this potential predictability.

Location	Netherlands	Denmark	SE Sweden	Baltic Sea	Estonia
Predictor1	T50 P3 JAS	T100 P2 MAM	T100 P2 MAM	T100 P2 MAM	TAR LEAD 05
Predictor2	DAR LEAD 04	H100 P2 SON	H150 P3 JAS	T50 P3 SON	H150 P3 JAS
β_0	1.6224	1.7156	1.8085	1.6901	1.8666
β_1	0.2579	0.2685	0.2901	0.3299	0.181
β_2	0.2899	0.1904	0.3013	0.3109	0.25
R^2 training	0.3607	0.345	0.4677	0.4109	0.3666
R^2 validation	0.3161	0.3656	0.4706	0.4832	0.5349
MAE training	2.23	1.967	2.1694	2.2999	2.17
MAE validation	1.9571	2.8654	3.5632	3.9322	2.6362

Table 8.8. Selected 90th percentile wind model specifications and skill for the DJF season over northern Europe. Five gridbox locations are included where potentially useful skill persists into the validation period. Gridbox locations are given in the first row, and can be compared with the top right panel of Figure 8.36. Locations chosen are a sample of those that retain skill in the north of the domain. Predictor names, together with the predictor season or lead time, and the associated coefficients complete the model specifications, and R^2 and MAE values from the training and validation periods are included for comparison.

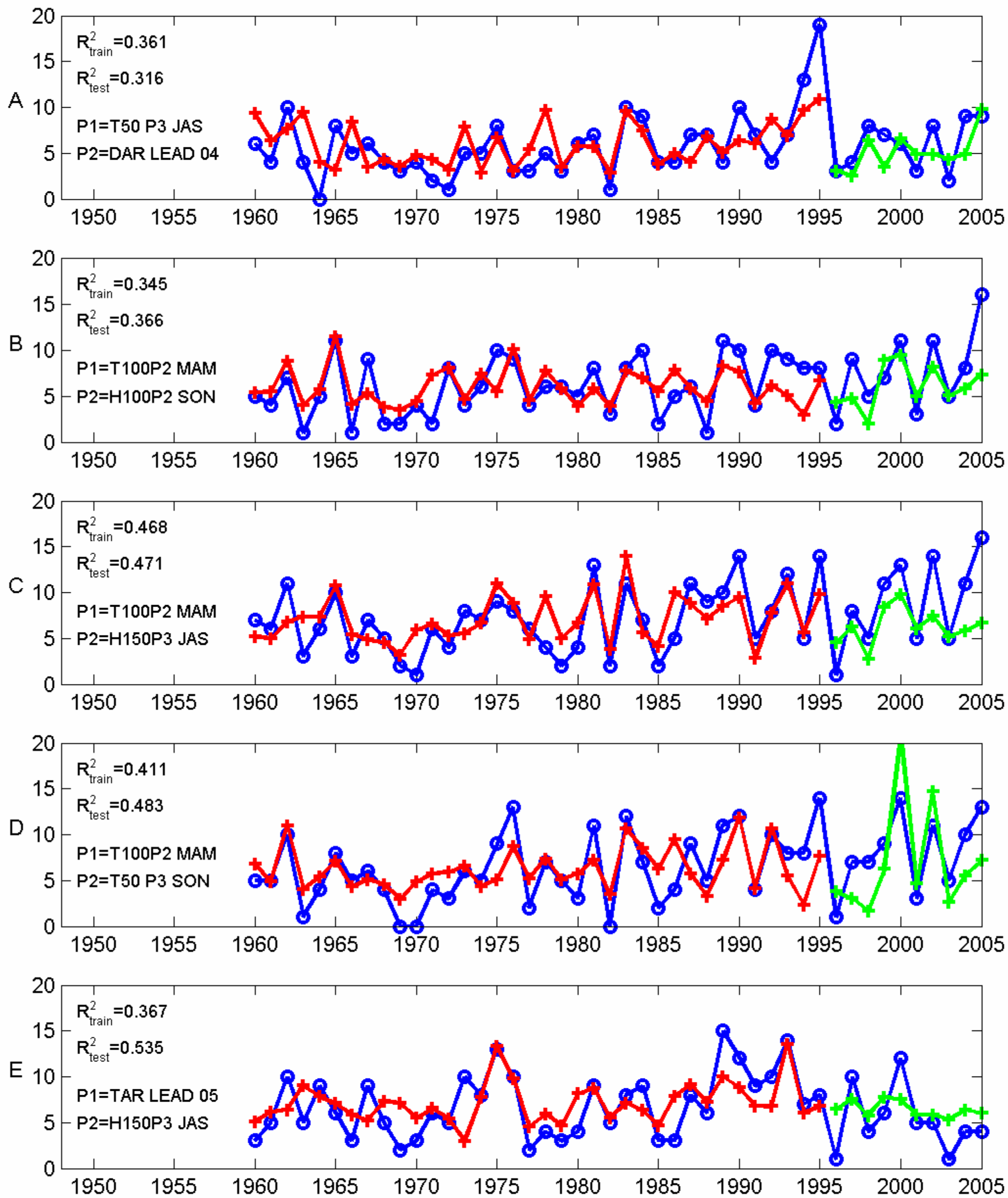

Figure 8.53. Observed and predicted DJF 90th percentile wind exceedance counts for selected locations across northern Europe, where potentially useful predictive skill persists into the validation period. Blue lines show the observed values, red shows the values predicted during the training period, and green shows the validation period predictions. R^2 values during the training and validation period are shown for each gridbox, as are the predictors for each model. Locations are as follows: A – Netherlands; B – Denmark; C – SE Sweden; D – Baltic Sea; E - Estonia

Location	Pyrenees	SE France	Czech Rep (W)	Czech Rep (E)	Serbia
Predictor1	T150 P2 ASO	TAR LEAD 05	EAP LEAD 04	SST5 LEAD 03	EPN LEAD 06
Predictor2	SST1 LEAD 02	TNH LEAD 11	SST5 LEAD 03	T50 P3 SON	
β_0	1.9012	1.8716	1.7594	1.8137	1.9695
β_1	-0.2761	-0.217	0.1906	0.1906	0.17
β_2	0.1768	-0.2332	0.195	0.2068	
R^2 training	0.2985	0.3722	0.2508	0.2836	0.2081
R^2 validation	0.3231	0.4808	0.6456	0.5662	0.3585
MAE training	2.0098	1.787	1.9326	1.9664	1.6455
MAE validation	2.8905	1.5408	2.1027	2.3368	2.7047

Table 8.9. Selected 90th percentile wind model specifications and skill for the DJF season over southern Europe. Five gridbox locations are included where potentially useful skill persists into the validation period. Gridbox locations are given in the first row, and can be compared with the top right panel of Figure 8.36. Locations chosen are a sample of those that retain skill in the south of the domain. Predictor names, together with the predictor season or lead time, and the associated coefficients complete the model specifications, and R^2 and MAE values from the training and validation periods are included for comparison.

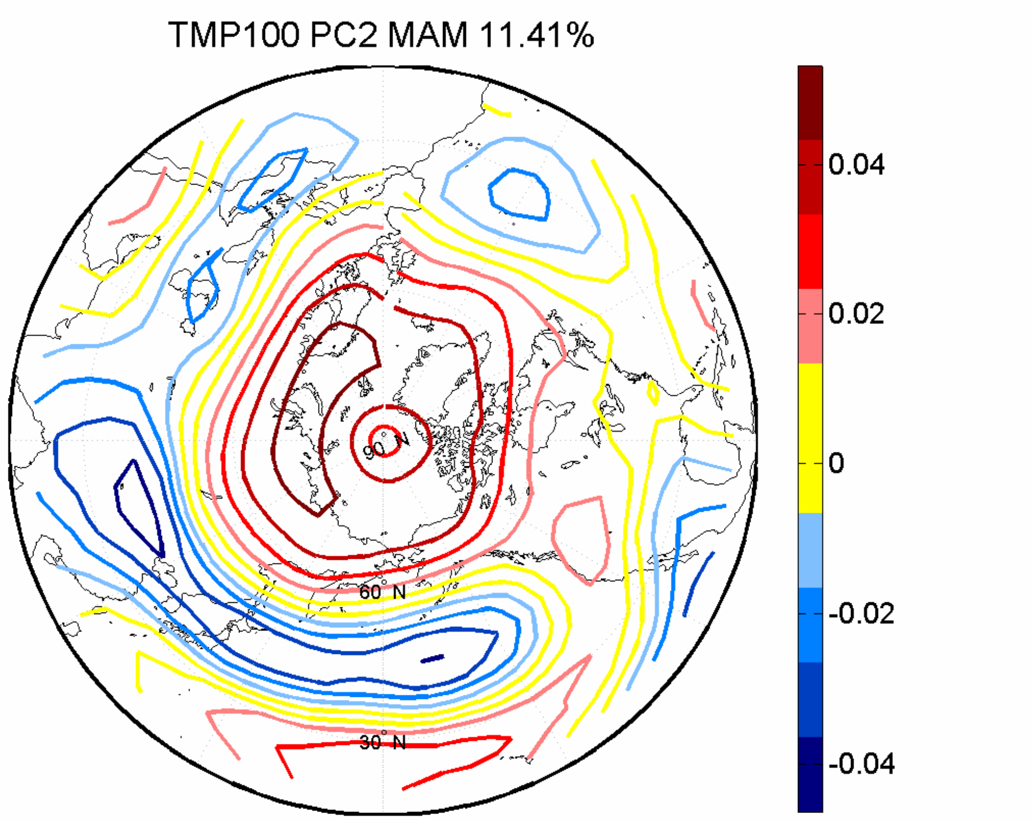

Figure 8.54. The second principal component of MAM 100hPa temperature anomalies. This predictor shows potentially useful skill in predicting 90th percentile wind exceedances over regions of northern Europe during the following DJF season. The pattern is almost annular, and leads the Arctic Oscillation and to a lesser extent the NAO by several months.

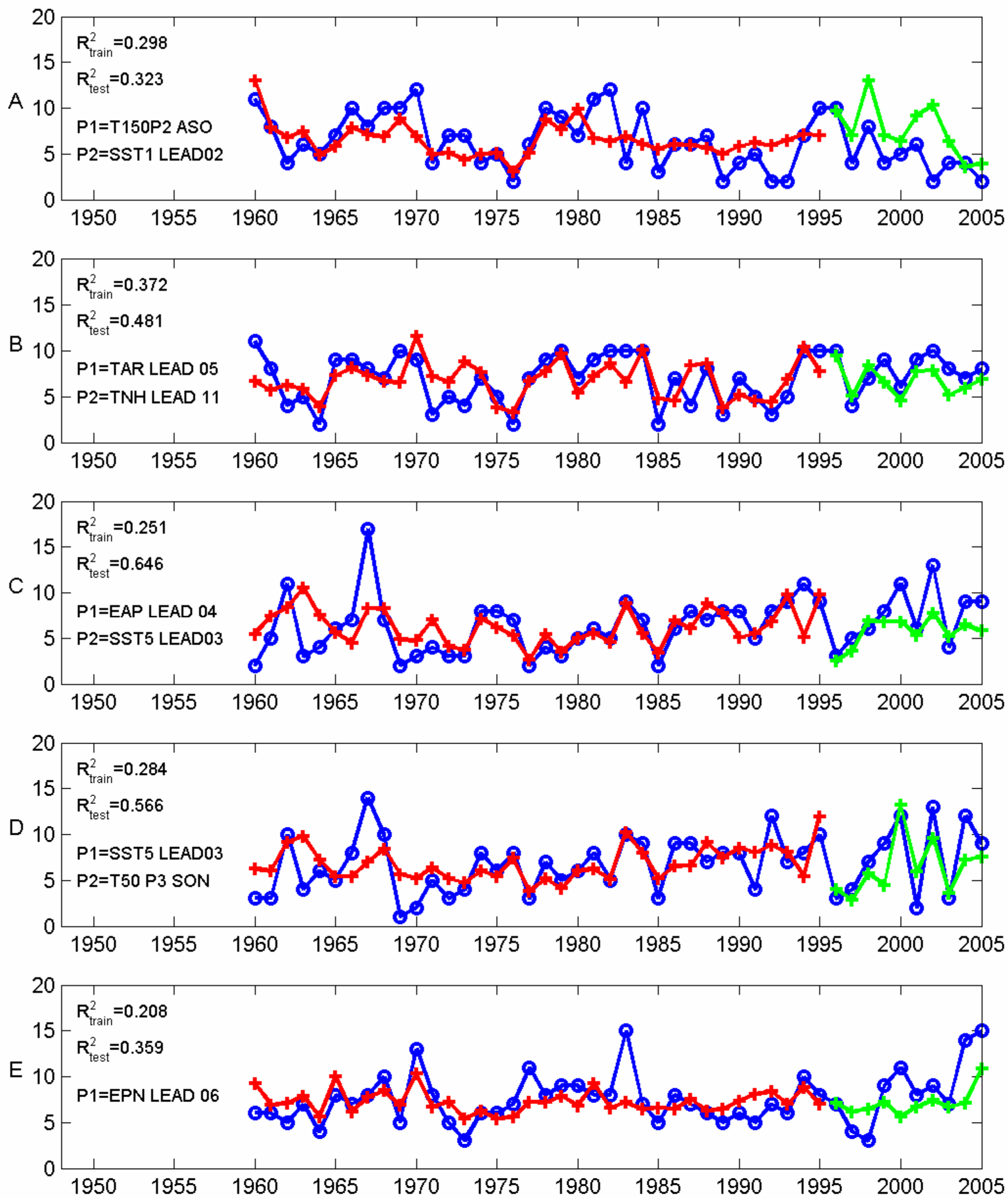

Figure 8.55. Observed and predicted DJF 90th percentile wind exceedance counts for selected locations across southern Europe, where potentially useful predictive skill persists into the validation period. Blue lines show the observed values, red shows the values predicted during the training period, and green shows the validation period predictions. R^2 values during the training and validation period are shown for each gridbox, as are the predictors for each model. Locations are as follows: A – Pyrenees; B – SE France; C – Czech Republic (W); D – Czech Republic (E); E - Serbia

8.3.8 95th Percentile Wind Model Validation: Skillful Models

The 95th percentile wind models exhibit a similar seasonal cycle and spatial configuration of persistent skill to the 90th percentile wind models. The season with apparently the most skill is JFM, and the spatial pattern of this skill is very similar to the DJF pattern observed for the 90th percentile models, although the regions of highest skill are shifted to the east, particularly in the southern region. This is illustrated in Figure 8.37. Table 8.10 describes a sample of five models within this region of validation period skill. As for the 90th percentile models, the T50 P3 SON predictor appears to be the dominant one, and other indices of stratospheric temperature also feature. For the NE Germany (bordering the Czech Republic) model, T50 P1 SON is one of the predictors, along with the October AO index. This model results in a relatively high R^2 value of 0.68, and it can be seen from the second panel of Figure 8.56 that the predicted series corresponds well with the observed throughout the training and validation periods. The stratospheric predictor for this model accounts for 47% of the variability in the SON 50hPa temperature field, and is represented spatially by a strong positive loading centred over the North Pole, and weakening towards the midlatitudes. This pattern is highly correlated with the October PNA, and to a lesser extent with the September and October NAO and AO indices.

Compared to the 90th percentile wind models, the NDJ and DJF seasons also show coherent regions of persistent skill, implying that for both sets of wind predictands, significant predictability may exist in the late autumn and early winter seasons. A number of the models capture the peaks in exceedance counts rather well, as illustrated in Figure 8.56. Further investigation is required into the physical basis of this potential predictability.

Location	NE Germany (B)	NE Germany (C)	N Poland	Latvia	NE Poland
Predictor1	T50 P3 SON	AOS Lead 04	T50 P3 SON	T50 P3 SON	T50 P3 SON
Predictor2	POL Lead 04	T50 P1 SON	EAP Lead 03	PV30 P3 JFM	T30 P3 FMA
β_0	1.0412	1.1487	0.9316	0.9589	0.9483
β_1	0.3998	0.237	0.5059	0.4961	0.3731
β_2	0.2348	0.2728	-0.2681	0.2171	-0.2036
R^2 training	0.489	0.3594	0.3178	0.4041	0.3543
R^2 validation	0.5421	0.6824	0.7053	0.5524	0.4513
MAE training	1.3457	1.369	1.761	1.5621	1.6078
MAE validation	1.7165	2.2462	1.4528	1.9517	2.1843

Table 8.10. Selected 95th percentile wind model specifications and skill for JFM season. Five gridbox locations are included where potentially useful skill persists into the validation period. Gridbox locations are given in the first row, and can be compared with the top right panel of Figure 8.37. Predictor names, together with the predictor season or lead time, and the associated coefficients complete the model specifications, and R^2 and MAE values from the training and validation periods are included for comparison. The two gridboxes for NE Germany are situated on the Baltic Coast (B), and neighbouring the Czech Republic (C), respectively.

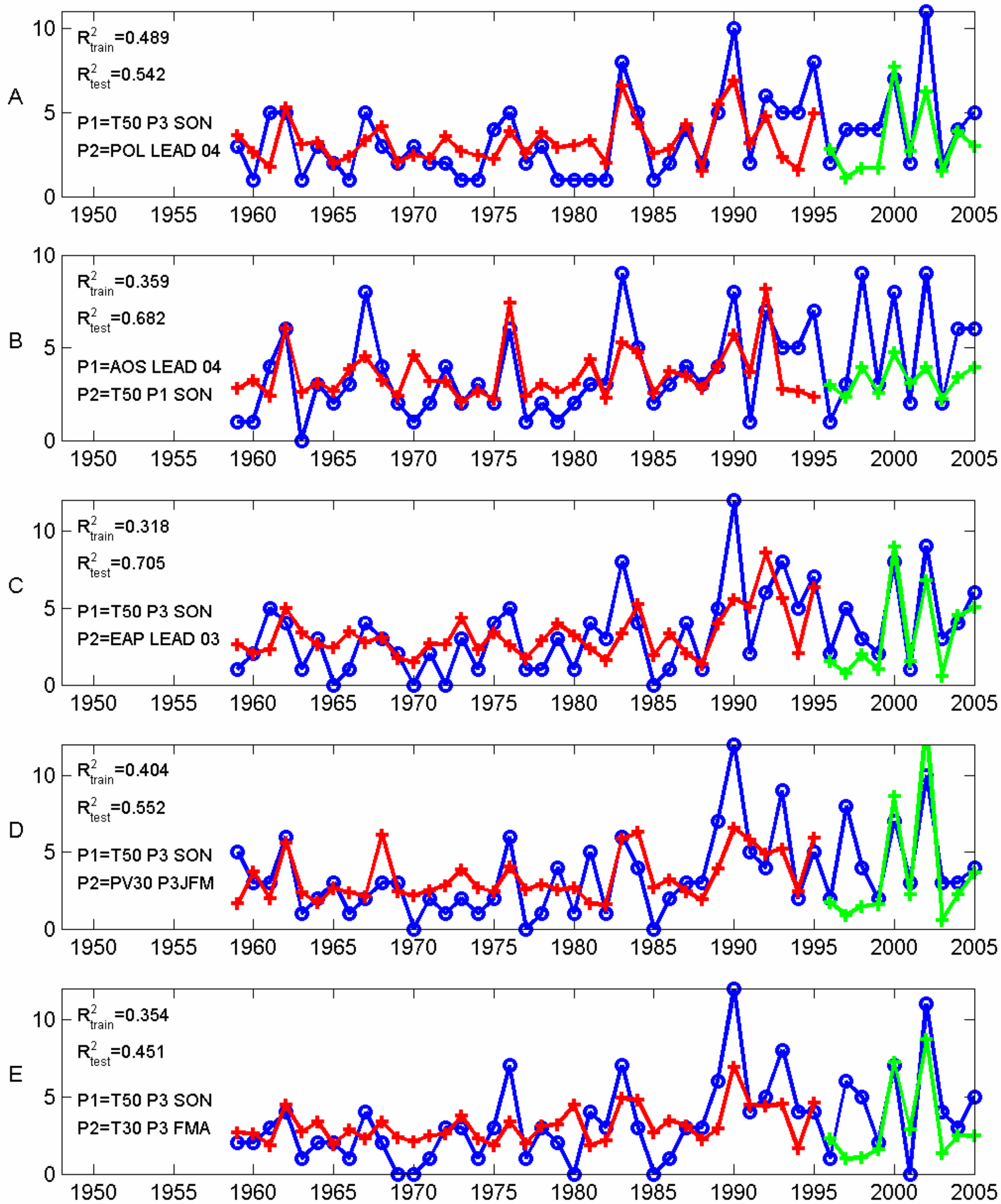

Figure 8.56 Observed and predicted JFM 95th percentile wind exceedance counts for selected locations in the southern Baltic Sea region, where potentially useful predictive skill persists into the validation period. Blue lines show the observed values, red shows the values predicted during the training period, and green shows the validation period predictions. R^2 values during the training and validation period are shown for each gridbox, as are the predictors for each model. Locations are as follows: A – NE Germany (Baltic Coast); B –NE Germany (Czech border); C – N Poland; D – Latvia; E – NE Poland.

8.4 Summary

This chapter presents the results of the model validation experiment applied to the models fitted in Chapter 7. The models have been selected based on predictors which exhibit field significance and combine to result in the lowest MAE based on a cross-validation test.

The results of further testing on an independent validation period, presented here indicate that in the great majority of cases, the levels of skill apparent during the training period and as quantified by R^2 and the MAE, do not hold for the validation dataset.

There are a number of possible causes for this degradation in skill levels. Firstly, from a purely statistical point of view, the establishment of reliable predictive models in the context of an exploratory study is inherently hazardous, and the models are prone to overfitting. That is, the models are conditioned on variability in the training data which may not relate to any physical processes. Secondly, the relationships observed during the training period may be nonlinear, for example as postulated in Pozo-Vazquez *et al*. (2005) with respect to the ENSO influence on European winter precipitation, and Wu and Hsieh (2004) on nonlinearities in the ENSO relationship with winter SLP over the North Atlantic and Europe. In this case, the linear fit obtained in the training period is not representative of the true influence of the predictor variable. The main constraints to further investigation of this aspect are the length of the observational record and shortcomings in the ability of numerical models to capture all the subtleties of interactions within and between the ocean, troposphere, stratosphere and land surface. On the basis of the work carried out in this study, among others, it is clear that the statistical investigation of nonlinear climate interactions requires a dataset substantially longer than the 37 year training period available here. Thirdly, the relationships observed during the training period may be physically valid, and linear within the training period, but nonstationary in the longer term. For example, Zanchettin *et al.* (2008) find that the Pacific Decadal Oscillation (PDO), which is a Pacific-wide mode varying on decadal timescales, affects the impact of ENSO on European winter precipitation. Similarly to the problem of nonlinearity, a much longer dataset is required to fully assess this problem. Broennimann (2007) uses reconstructed indices of ENSO and European climate variables over a 500 year period to investigate the nonstationarity of the ENSO influence on Europe in more detail – although at a coarser resolution than would allow an assessment of the effect on extremes of either precipitation or wind. An important influence in modulating the nature of the ENSO signal in Europe is found in the north Pacific, which may be related to variability of the PDO. It is worth noting that the period 1996-2005 was characterised by a number of remarkable events, including the exceptionally strong El Niño event of 1997-98, the 2003 summer heatwave across Europe, and also a number of major flooding events – for example autumn 2000 in the UK, and summer 2002 in central Europe. It is possible that these events have contributed to the breakdown in skill during the validation period. For stratospheric predictors, the issue of nonstationarity is perhaps less important, since there is less evidence that the stratosphere varies on decadal timescales (although van Loon and Meehl, 2008, note that there is a statistically discernible solar influence on the Southern Oscillation and in the stratosphere at decadal timescales), but is still more difficult to address, owing to the lack of data or proxy data over a sufficient time period.

There are a relatively small number of cases where potentially useful skill does persist, particularly for the wind predictands during winter, and to a lesser extent the precipitation predictands during spring. Both of these results correspond with other research into European seasonal predictability – although not relating directly to the frequency of extremes. For example, Lloyd-Hughes and Saunders (2002) with respect to spring precipitation, and Thompson *et al*. (2002) with respect to winter circulation anomalies both identify similar opportunities for potential predictability.

For precipitation, the validation period degradation in skill is the greatest, and no large coherent regions of persistent skill exist in any season, for either of the variables. The most notable case where there is something approaching a coherent response is during AMJ, for both the 90th and 95th percentile models. This skill is located over Western Europe – in particular the Bay of Biscay and the UK regions. The key predictors driving this skill appear to be related to ENSO – in particular the atmospheric descriptors thereof, at lead times of up to ten months, and also to some extent local SST anomalies, and stratospheric predictors. However, the validation period skill levels are still relatively low, although this is not unexpected as a general feature of precipitation predictability. Furthermore, in the cases where skill persists, as identified here, there appears to be an upwards trend in some of the observed precipitation predictand variables during the validation period, resulting in large increases in the MAE values of the models.

The wind models exhibit similar degradation in validation period skill throughout the year, except for the late autumn and winter seasons, where some coherent regions of persistent skill are identified. In particular, during NDJ a large region over the Benelux countries shows some skill for 90th percentile precipitation and during DJF for the 90th percentile models, two distinct regions retain some skill. The northern of these regions reaches from the North Sea to the Baltic States, and the southern region from the Pyrenees to the Balkans. This pattern is repeated during JFM for the 95th percentile wind models, albeit shifted eastward. In all cases, stratospheric predictors (and in particular temperature) appear to be the most important predictors. Other potentially useful predictors include those derived from ENSO, teleconnection patterns such as the AO, and snow cover anomalies. On this basis, it appears that some potentially useful predictive skill may be offered for winter extremes of wind in Europe. For both the wind and precipitation models where skill is apparent during the validation period the nature of the possible physical processes causing predictability at these timescales is unclear. A comprehensive assessment of these relationships would ideally include the reproduction in a dynamical model of these effects as observed, and a detailed analysis of the physics which enable predictable relationships at such long lead times.

On a practical note, if the skill identified here as useful were to be developed into an operational scheme, a number of refinements would be necessary. Primarily, the use of predictors with different lead times might well be constrained in an operation setting, and depending on the lead of the forecast, some of the predictors may not be available. Additionally, if the lead time of a predictor as given here is longer than that of the forecast, it may be possible to enhance the forecast with the addition of new information. Both of these adjustments would to some extent compromise the model selection process as outlined here, in which case there would need to be a sound physical understanding of the predictability on which to base further predictor selection and refinement.

9 Conclusions and Suggestions for Further Work

9.1 Summary of data for model development

This thesis presents an empirical analysis of the potential seasonal predictability of extremes of precipitation and wind in Europe. The predictand data are derived from the ERA-40 reanalysis (Uppala *et al*. 2005), which are found to compare favourably with the available station data and gridded monthly means. The ERA-40 analysis is extended to include years up to 2005 by including the ECMWF Operational Analysis data up to this date. ERA-40 is selected on the basis that temporal and spatial coverage is continuous, which is conducive to the type of exploratory analysis undertaken here. The predictands are constructed from percentile exceedance counts above the 90th and 95th threshold for both wind and precipitation, where the thresholds are determined by first filtering the gridbox timeseries so that for the precipitation data only rain days (with more than 0.1mm of precipitation) are included, and for the wind gust data only the larger of the values is included where consecutive days exceed the threshold. The thresholds are determined by fitting a gamma distribution to the data. Twelve overlapping three-month seasons are considered in order to best represent any sensitivity deriving from seasonality.

A wide range of potential predictors are considered, on the basis that a three stage process of predictor and model refinement clarifies potential predictive skill. The initial stage considers a ‘full set’ of predictors, including a wide range of atmospheric teleconnection indices; large scale indices of SST variability for all the major ocean basins; indices of local European coastal SST anomalies; indices of stratospheric geopotential height, temperature and potential vorticity; Northern Hemisphere snow cover and Solar output. Altogether the full set comprises approximately 850 predictors when each index is considered at a range of lead times from two to eleven months ahead of the middle month of each predictand season. Since the research is concerned with interannual variability, and the observation period for model development is not necessarily long enough to capture decadal variability, each predictor timeseries is low-pass filtered to remove trends and decadal variability. The predictors are selected on the basis that they either represent an important component of atmospheric, oceanic or land-surface variability, or there exists previous work which has identified potentially useful predictive skill for some aspect of the European climate for that predictor.

9.2 Summary of initial predictor selection phase

The full predictor set is reduced substantially by assessing each predictor individually for a field-significant response at 95% confidence over the training period of 1958 to 1995 in each predictand dataset. Only those predictors which are associated with a field significant response are retained to be considered in the model selection phase. A wide range of the full set of potential predictors is retained. Overall for the 90th and 95th percentile precipitation predictors 9.4% and 8.5% of the full predictor set is retained respectively. For the 90th and 95th percentile wind predictands 5.8% and 5.9% are retained respectively. Given that a confidence level of 95% is taken as the threshold for significance it might be inferred that a considerable proportion of the retained predictors are selected due to chance, particularly when taking into account cross-correlation between the predictors. However, it is noted that there are frequent instances of correspondence between the field-significant relationships observed here, and relationships identified in the literature. For example, Northern Hemisphere snow cover anomalies are seen to be associated with variability in the wind predictands, although contrary to other findings (e.g. Cohen *et al.*, 2001), the North American component seems to be more important than Eurasia and the associated Siberian High. ENSO-related indices are also found to be important, and corroborate work by (for example) Pozo-Vazquez *et al. (*2005) on the observed relationship between ENSO and winter precipitation in Europe. Here there are found to be statistical links between SST and atmospheric indices of ENSO variability at the full range of lead times from two to eleven months and for most of the predictand seasons considered. However, typically the Niño regional SST indices and the SOI-related indices are found to be of more interest than the PCA-based indices, and in particular due to the model selection and validation phases, SOI-related indices are found to be more important as potentially useful predictors than the Niño regional SST indices. In many cases the predictor-predictand relationships have field-significance over remarkably large ranges of lead times, and large portions of the annual cycle. This is particularly true of the ENSO-related predictors, and in the case of the wind predictands, for some of the stratospheric predictors.

There also appears to be an annual cycle component to the predictor retention, with the precipitation analysis resulting in seasonal peaks of predictor density in the spring and autumn, and the wind analysis resulting in seasonal peaks in from late summer to early winter. It is unclear as to whether this results from spatial properties of the predictands during these peak seasons, or cross-correlation between the predictor indices, or whether a 'real' feature of seasonal predictability is illustrated.

It is important to note that for this analysis the results should be interpreted strictly within the context of the model training period, and as a purely empirical illustration of potential predictability. Nevertheless, a number of potentially interesting relationships are identified.

9.3 Summary of model skill during the training period

The second stage of model development uses the reduced predictor set and an all-subsets model selection algorithm is implemented, using the cross-validated MAE to select the best model over the training period. This results in a final model comprising one or two predictors for each gridbox over the predictand domain. It is found that potentially useful levels of skill – as measured using the R^2 and MAE statistics – exist for all predictands, although there is substantial variability by season and spatially. R^2 values (describing the proportion of variation explained by the model) range from zero to in excess of 0.5 – the latter corresponding to 50% of variation explained. This is a surprisingly high value for European seasonal predictability – particularly for wind and precipitation – and it is speculated that some of the models may be over-fitted to the data.

Some predictand gridboxes do not have a viable model – largely due to the lack of field-significant predictors at that location. The frequency of this occurrence varies by predictand and by season, corresponding to the predictor density pattern discussed in 9.2 above.

Again, notwithstanding the possibility of overfitting, and the existence of some models which do not have useful levels of skill, several points relating specifically to the predictors retained at this stage are worthy of further mention. Generally, since on average a larger number of predictors per gridbox are retained than are allowed in the final model, the spatial structure of the retained predictors loses some coherence compared to the field-significant fit of each predictor. This is more notable for the precipitation models, partly because overall a larger number of predictors are retained for the precipitation predictands, and presumably partly also because the spatial correlation distance of precipitation extremes is considerably less than that for wind extremes.

9.3.1 Summary of predictor selection for the precipitation models

Owing to the similarities between the 90^{th} and 95^{th} percentile precipitation predictands, the results are summarised together, for clarity. The most frequently selected predictors – in other words those resulting in the most cross-validation skill within the training period – are as follows. Indices of stratospheric potential vorticity are the most frequently selected as contributing to the greatest cross-validation skill. Selection of these indices tends to follow a seasonal cycle, peaking in the summer and early autumn. Local SST indices are also important throughout the annual cycle, but in particular peak in the DJF and JFM seasons. This is counter to the prevailing view that warm summer SST anomalies might contribute to local precipitation anomalies (for example as discussed in Zheng and Frederiksen, 2006 relating to seasonal prediction of New Zealand summer rainfall). Additionally, the exceptionally wet summer of 2007 in the UK is thought to be in part due to anomalously warm SSTs surrounding the UK. However, the local SST anomalies typically show a positive fit with the precipitation extremes, consistent with the principle that warm SSTs would be expected to contribute greater quantities of moisture to the atmosphere. It may be that the nature of the predictand – being somewhat coarsely gridded relative to extremes of convective precipitation – is more conducive to the identification of relationships with large-scale precipitation, which is dominant in the winter. Furthermore, Bengtsson (2008) finds that the ECHAM5 GCM simulates up to a 50% increase in extreme precipitation associated with North Atlantic extra-tropical cyclones over the 21st Century when the model is forced with the IPCC scenario A1B (Nakicenovic *et al*., 2000). No large changes in wind speeds are found, and although the study does not relate the increased precipitation directly to warmer SSTs, it is plausible that along with a warmer atmosphere, the SSTs contribute to enhanced precipitation. It may therefore be that local SST anomalies are a useful guide to extremes of rainfall on seasonal timescales.

Atlantic and Pacific indices of large-scale SST variability are the third and fourth most frequently selected sets of predictors respectively. There is no clearly defined seasonal cycle in the response to either of these sets, and no particular season or mode in the SSTs appears to be dominant. The leading modes in each basin are selected rarely – typically the third, fourth or fifth modes are found to be more important here. In the Pacific these modes tend to be related to variability in the North Pacific, and correlate with the PNA and EPNP patterns, suggesting a possible influence on the midlatitude planetary-scale flow upstream of the North Atlantic/European sector. Indices describing the atmospheric component of ENSO (the SOI and its local components at Darwin and Tahiti) feature prominently. The Darwin indices are most important in the FMA and MAM season, while the Tahiti indices are most important in the autumn and winter. The SOI response is relatively important from the autumn through to the spring. There are almost no instances where ENSO-related predictors are selected in the summer months, and the Niño regional SST indices are selected very rarely relative to the atmospheric indices.

9.3.2 Summary of predictor selection for the wind models

As with the precipitation models, there is considerable similarity between the most frequently selected predictors for the 90th and 95th percentile wind predictands. They are therefore considered together here. Indices of Pacific large-scale SST variability comprise the most frequently selected family of predictors for the wind extremes. The first and second PCs describing ENSO variability are almost never selected, and the fifth PC is the most frequently selected by some margin, particularly in the OND and NDJ seasons. However there is no apparent likely explanation for the link between these indices and the wind extremes. The AO is the second most frequently selected predictor, appearing throughout the year – including somewhat unexpectedly in the summer months. It is worth noting here that extremes of wind in the summer are of relatively less importance as a climate hazard.

Indices of large-scale Atlantic SST variability are frequently selected in the autumn and early winter; in particular during OND, the second and fifth PCs from the preceding OND season account for a large number of models. To some extent this supports the idea that temperature anomalies within the mixed layer of the ocean may become isolated from the shallow summer mixed layer and re-emerge the following autumn and winter as the mixed layer deepens as discussed in Timlin *et al*. (2002). The precise mechanism for any subsequent coupling between the ocean and atmosphere remains unclear.

The snow cover predictors are more important to the model cross-validation skill for wind than for precipitation. However, it is interesting to note that the Eurasian snow cover index is of considerably lesser importance than those for North America or the Northern Hemisphere as a whole. This is contrary to findings by – for example – Saito and Cohen (2003) who link Eurasian snow cover with interannual and decadal variability of the NAO/AO.

ENSO indices derived from SST anomalies do not feature prominently in the final model selection phase as discussed above. However, in common with the precipitation models, atmospheric indices of ENSO – specifically the Tahiti MSLP index – are frequently selected during the winter months from OND to JFM. The SOI is selected to a lesser extent during the same seasons, and the Darwin index is hardly selected at all. This raises interesting questions about the way in which ENSO variability is summarised with respect to its possible influence on the European climate. Traditionally, studies have looked at indices describing SST anomalies, or the SOI as a whole (Fraedrich, 1994), or in some cases have used indices expressing the coupled ocean-atmosphere variability (as reviewed in Broennimann, 2007). Jia *et al*. (2008) identify distinct responses to tropical Pacific SSTs in the Northern Hemisphere, where western Pacific SST anomalies tend to be associated with AO responses, and eastern Pacific anomalies tend to be associated with the PNA. It is not known whether further studies support the finding in this research that atmospheric anomalies in the eastern tropical Pacific appear to relate more strongly to extreme wind events in Europe than do other components of the ENSO coupled system.

In general, the model selection process for the wind predictands results in more coherent predictor selection in space. This is likely due to the larger spatial correlation scale of extreme wind events – which is also a likely cause of the smaller reduced predictor subset obtained for the wind predictors due to the field-significance testing. Absolute skill levels are not notably different between wind and precipitation, although the skill of the wind models is smoother in space.

9.4 Summary of model validation skill

As a final assessment of the potential usefulness of the models from a statistical perspective, further testing is carried out on an independent validation period from 1996 to 2005. This results in widespread degradation in the levels of skill, implying that the models developed on the training period data are either over-fitted, or depict nonlinear or nonstationary relationships. The extent to which each of these factors applies is not clear. It is likely that in a substantial number of cases the models are overfitted, due to the relatively large number of predictors available in the model training phases. However, there is also evidence to support the existence of both nonlinear and nonstationary relationships – particularly between ENSO and the European climate (for example Pozo-Vazquez *et al*. (2005) and Broennimann (2007). It is also well understood that decadal variability in the North Atlantic is an important component of the climate variability, and the period of observations in this study is not sufficient to account for this – hence only interannual variability is considered. It may be that post 1995 (comprising the validation period) the apparent shift in the NAO trend compared to the last three decades comprises a more important influence than the observed interannual variability from the 1960s to the 1990s. This is documented in Scaife *et al*. (2008).

Despite the widespread degradation in skill, there are a small number of cases where skill persists into the validation period. There are fewer instances of this for the precipitation models than for the wind models, but some of the models for the AMJ season for both 90th and 95th percentile precipitation models seem to retain some skill. This skill is based on local SST anomalies, the SOI indices and some stratospheric predictors. The physical basis of the observed predictability is not determined and would likely require substantial further investigation – using both empirical and numerical methods.

Overall the wind models appear to perform better in the validation period. In particular, the winter seasons of NDJ, DJF and JFM show large regions of coherent skill situated over the Benelux countries in NDJ, and over two distinct regions in both DJF and JFM, where the northerly region reaches from the North Sea to the Baltic States, and the southern region from the Pyrenees to the Balkans. Stratospheric temperature predictors from the preceding summer appear to be the main cause of this potential predictability, with some of the model fit also deriving from AO, ENSO and snow cover predictors. On the basis of this research the potential predictability of extreme winter wind events over these regions is the most promising finding of this study, and warrants further investigation to determine the likely mechanisms which facilitate this apparent predictability.

9.5 Conclusions

A comprehensive exploratory investigation into the linear empirical predictability of wind and precipitation extremes in Europe is carried out. The emphasis is on identifying potential predictive skill at seasonal timescales rather than the precise quantification of predictive skill. It is found that the potential predictability is very limited based on the criteria of model validation on independent data. This implies that either nonlinear processes are important, nonstationary processes affect the model training and validation periods differently, or the models are overfitted and no useful skill can be obtained from the predictors used here. Further investigation would be required to clarify this. However, the length of the observational record available to carry out such a study is not conducive to nonlinear methods or those considering decadal signals in the climate system.

Some instances of potentially useful predictability are found on the basis that skill persists during the validation period. In particular, the frequency of extreme wind events during the winter can be related significantly to stratospheric temperature predictors from the previous summer. The mechanisms causing this potential predictability are as yet undetermined. Precipitation appears to be inherently less predictable within the context of this research. There are few scattered instances – most notably in the spring – where potentially useful skill appears to persist, although no clear physical explanation for this is apparent.

Independently of the model validation period, many instances of apparently significant relationships are found between the predictands and predictors at lead times of up to eleven months. These relationships include responses to ENSO, large-scale SST modes in all the major ocean basins, atmospheric teleconnection patterns, modes of variability in the stratosphere, and Northern Hemisphere snow cover. In many cases there is considerable empirical and theoretical support for these observed relationships in the literature conditioned on empirical model training, independent validation and numerical modelling. Although the findings presented here do not discount any of this work, it is clear from this work and a review of the literature that evidence for any skill in the seasonal prediction of European climate is marginal at best, and highly sensitive to the data and methods employed, both for empirical and model studies. Furthermore, it is recognised that for any exploratory study of this nature it is imperative that any apparently significant results are presented in the full context of all those results which do not yield apparently significant skill.

9.6 Further Work

The seasonal predictability of European wind and precipitation extremes is a subject which requires extensive further research. The findings presented here constitute a wide ranging study, and a number of issues relating to the predictability of European extremes have been addressed. However, the study is not exhaustive. Based on the findings, it is possible to identify what may well be key avenues for further research in this area.

9.6.1 Data and methods

By definition the observation data are the key component of any empirical study into the seasonal predictability of climate. As discussed in Chapter 4, a range of datasets are available to derive both the predictor and predictand indices. Here the predictand indices are derived from the ERA-40 reanalysis data (including four years of ECMWF

Operational Analysis data), which each present a spatially and temporally complete and homogenous record. However, in the study of extremes, the sensitivity to spatial scale is important, particularly for precipitation, and it may be that ERA-40 as used here does not provide the optimal representation of this process. Although extensive and largely favourable comparisons were carried out between ERA-40 and station records for both wind and precipitation where available, it is not clear whether station timeseries or a gridded dataset such as the ENSEMBLES high resolution daily dataset for Europe (Haylock *et al*. in preparation) might identify potential predictability more effectively. Another potential issue is the possibility of significant inhomogeneities between the ERA-40 and ECMWF Operational Analysis. The use of station timeseries results in susceptibility to local effects which may not be representative of the large-sale circulation, and the gridded dataset option does not include observations over the ocean, which compromises the approach taken here of assessment for field-significance. Nevertheless, it would be worthwhile further to assess the sensitivity of the observed relationships to different predictand data. In summary, the specification of spatial scale in this study is subject to a number of limitations. At one end of the scale, in order to represent the extremes as accurately as possible it is necessary to use a very high resolution in order to capture all the relevant local effects. The disadvantage here is that when considering synoptic-scale forcing of the predictand this approach is highly sensitive to translational shifts in the extremes – in other words a particular predictor may be useful in describing variability at a larger scale than that at which the extreme event typically takes place. At the other end of the scale – that at which the mean synoptic effects of potential predictors might be felt, the ability to identify extremes is severely compromised since they typically take place at a smaller scale within this domain. From the work carried out in this thesis, it is implied that the gridbox size adopted here is firmly in the middle of this scale – that is, there is still good correspondence with the interannual variability of extremes at the point scale, but there is also significant noise in the model selection and parameter estimation, to the extent that it might be desirable to explore potential predictability at a still coarser spatial scale. This would be a key goal of future work. Such an undertaking would contribute towards simpler, more spatially coherent model selection. Given that in many cases the best model as selected in this study is only marginally better than other models, this is a highly desirable outcome, and should lead to a more consistent set of predictors, which in turn might be expected to aid in the validation process.

The form of the predictand indices presented here as seasonal counts of exceedance over a percentile threshold might also impose a limitation on the study. That is, in order to gain a sufficient sample size for each predictand timeseries, at least a three month season is required, as used here. However, it is apparent from the observed relationships that many of them are sensitive to the seasonal cycle, and some form of monthly index – for example of precipitation intensity – might be more suitable. Furthermore, Pezzulli *et al.* (2005) find that significant departures occur in the seasonality of responses to Niño 3.4 SSTs, such that the traditional approach as adopted in this study – of fixed three month seasons – may well hide important insights into seasonal predictability. The exact formulation of a more flexible approach in this context is unclear, but future efforts might benefit from seeking to develop appropriate methods.

The predictor data generally represent large-scale phenomena, and many of the indices are likely to be less sensitive to the exact formulation of the indices – although Hu and Huang (2006) find that the response of the NAO to leading SST anomalies is highly sensitive to the precise definition of the index to the point where only one index responds with statistical significance. There are a number of ways of addressing this. One would be to present multiple formulations of each predictor index, although this is likely to further compound the risk of overfitting the models. Other approaches might include multivariate techniques such as canonical correlation analysis (CCA) which is applied to the analysis of large-scale circulation controls on extremes in Haylock and Goodess (2004). Here the predictor data fields are related to the predictand in such a way as to maximise the correlation between the canonical 'modes' in each dataset. It might be expected, for example to see a stronger fit between some of the predictands and MSLP in the eastern tropical Pacific based on the findings documented here. An additional potential benefit of multivariate techniques is that the spatial response in the predictands might be more coherent, and not subject to large qualitative changes in the predictor selection between neighbouring gridboxes as is the case here (van den Dool, 2007). In other words, the CCA modes correspond approximately to the spatial scale at which variability is observed in the predictands, as discussed above. However the predictand data used in this study results in some obstacles to standard multivariate techniques – being approximately Poisson distributed. Furthermore, it is also worth noting that in a purely exploratory study such as this, while signals or potentially important predictors may be masked by noise in the model selection process, in cases where predictors have been selected across coherent regions of the domain despite the lack of constraint in the selection algorithm, it follows that this may highlight a predictor of particular importance.

Another area where significant further work could be carried out is in exploring different training and validation periods. In this study a relatively short (10 year) validation period is used, since the objective is more a binary appraisal of whether skill persists in the validation period, rather than a precise quantification of the validation period skill. However, such a short validation period also gives rise to the possibility that non-stationary features of interannual to decadal climate variability might compromise the assessment of validation period skill. For example the occurrence of a major El Niño event in 1997, as well as the 2003 European heatwave and a number of severe flood events might significantly bias the results. Possible solutions to this problem are either to use a longer validation period, or to use multiple independent validation periods. For practical reasons this would have to take place in conjunction with a simplification of the model selection process – perhaps limiting the number of predictors by selecting ‘common’ predictors from subsets which are highly correlated, or by using a coarser spatial resolution in the predictand datasets. The skill metrics used in this study are the mean absolute error (MAE) and coefficient of determination (R^2). These are used as they both represent simple appraisals of skill – namely the mean error of seasonal percentile exceedance counts, and the proportion of variance explained by the predictors. It might be informative to use additional measures of skill, such as percentage improvement over climatology (e.g. Saunders and Qian, 2003).

One potentially interesting validation of the approach presented in this study would be to apply it to a different region or predictand variable, where predictability is better quantified. An example might be South American precipitation (e.g. Folland *et al*., 2001) or other variability closely associated with ENSO. This might allow additional insight into the potential for over-fitting for example.

If the techniques applied in this thesis, and incorporating some of the suggestions for further work outlined above were applied to an operational forecast scheme, a number of further issues might need to be addressed. Currently there is no constraint on the lead times of the selected predictors – the only criteria is that the best model is selected based on the training data. In an operational scheme, there is frequently a delay in predictor data being made available, which might compromise some of the models – particularly those with predictors at short lead-times. Conversely, if for example a forecast with a three month lead time were to be made, it might be thought inappropriate to use a predictor with a six month lead time, when more recent data are available – unless there is a sound physical reason for assuming that a six month lead gives the best skill. Furthermore, in general the greater the extent to which simplification in the models (i.e. a reduction in the overall number of predictors across the domain), the greater the ease of use will be of an operational scheme.

9.6.2 The Physical Basis for Predictability

For the models where skill is retained through the validation period, a source of seasonal predictability must be said to exist that is potentially useful, given that a viable theoretical explanation can be identified and verified. The noteworthy case where further investigation might be useful is that of the winter wind extremes over the North Sea and Baltic regions, and over southern Europe. It appears from the specifications of the models that stratospheric predictors from the preceding summer are important to the model skill, although to date the means by which this signal is transferred to the winter troposphere are not clear. Perhaps the most useful approach to this problem would be to investigate the forced response to stratospheric temperature anomalies such as those observed in this study of the subsequent winter wind extremes in a climate model. To date research has typically focussed on interactions between the stratosphere and troposphere at shorter lead times, although it has been shown for example by Broennimann *et al*. (2004, 2007) and Kodera *et al*. (1996) that a link exists between ENSO and the Northern Hemisphere stratosphere, and Broennimann (2007) shows how this might affect the European climate in a schematic which relates ENSO-driven anomalies in the planetary wave activity to changes in the zonal and meridional flow in the stratosphere, which in turn affects the strength of the polar vortex. Baldwin and Dunkerton (2001) show how anomalies in the polar vortex can propagate downwards on a timescale of weeks to affect the Northern Hemisphere midlatitude circulation. Further work on assessing the potential seasonal predictability of the European climate through better understanding of the propagation of the ENSO signal to Europe is highly desirable. The ability to model successfully all the major components of this process is not yet fully realised, but may well be the most useful tool to move beyond a merely empirical appraisal of forecast skill.

References

Ambaum, M. H. P., B. J. Hoskins, and D. B. Stephenson, 2001: Arctic oscillation or North Atlantic oscillation? *Journal of Climate*, **14**(16)**,** 3495-3507.

Ambaum, M. H. P. and B. J. Hoskins, 2002: The NAO troposphere-stratosphere connection. *Journal of Climate*, **15**(14)**,** 1969-1978.

Anderson, J., H. van den Dool, A. Barnston, W. Chen, W. Stern, and J. Ploshay, 1999: Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction. *Bulletin of the American Meteorological Society*, **80**(7)**,** 1349-1361.

Bader, J. and M. Latif, 2005: North Atlantic oscillation response to anomalous Indian Ocean SST in a coupled GCM. *Journal of Climate*, **18**(24)**,** 5382-5389.

Baldwin, M. P. and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294**(5542)**,** 581-584.

Baldwin, M. P., L. J. Gray, T. J. Dunkerton, K. Hamilton, P. H. Haynes, W. J. Randel, J. R. Holton, M. J. Alexander, I. Hirota, T. Horinouchi, D. B. A. Jones, J. S. Kinnersley, C. Marquardt, K. Sato, and M. Takahashi, 2001: The quasi-biennial oscillation. *Reviews of Geophysics*, **39**(2)**,** 179-229.

Baldwin, M. P., 2003: Comment on "Tropospheric response to stratospheric perturbations in a relatively simple general circulation model" by Lorenzo M. Polvani and Paul J. Kushner. *Geophysical Research Letters*, **30**(15)**,** art. no.-1812.

Baldwin, M. P., D. B. Stephenson, D. W. J. Thompson, T. J. Dunkerton, A. J. Charlton, and A. O'Neill, 2003: Stratospheric memory and skill of extended-range weather forecasts. *Science*, **301**(5633)**,** 636-640.

Barnett, T. P., 1981: Statistical Prediction of North-American Air Temperatures from Pacific Predictors. *Monthly Weather Review*, **109**(5)**,** 1021-1041.

Barnett, T. P. and R. Preisendorfer, 1987: Origins and Levels of Monthly and Seasonal Forecast Skill for United-States Surface Air Temperatures Determined by Canonical Correlation-Analysis. *Monthly Weather Review*, **115**(9)**,** 1825-1850.

Barnett, T. P., L. Dumenil, U. Schlese, and E. Roeckner, 1988: The Effect of Eurasian Snow Cover on Global Climate. *Science*, **239**(4839)**,** 504-507.

Barnett, T. P., L. Dumenil, U. Schlese, E. Roeckner, and M. Latif, 1989: The Effect of Eurasian Snow Cover on Regional and Global Climate Variations. *Journal of the Atmospheric Sciences*, **46**(5)**,** 661-685.

Barnston, A. G. and R. E. Livezey, 1987: Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns. *Monthly Weather Review*, **115**(6)**,** 1083-1126.

Barnston, A. G., 1994: Linear Statistical Short-Term Climate Predictive Skill in the Northern-Hemisphere. *Journal of Climate*, **7**(10)**,** 1513-1564.

Battisti, D. S., U. S. Bhatt, and M. A. Alexander, 1995: A modeling study of the interannual variability in the wintertime North Atlantic Ocean. *Journal of Climate*, **8**(12)**,** 3067-3083.

Benestad, R. E. and A. Melsom, 2002: Is there a link between the unusually wet autumns in southeastern Norway and sea-surface temperature anomalies? *Climate Research*, **23**(1)**,** 67-79.

Bengtsson, L., 2008: Extreme European Windstorms and Expected Changes in a Warmer Climate. Presentation to the Lighthill Risk Network, London, May 2nd 2008.

Bjerknes, J. A. B., 1959: North Atlantic Sea Surface Temperatures and Atmospheric Circulation. *Journal of Geophysical Research*, **64**(6)**,** 689-689.

Bjerknes, J., 1966: A Possible Response of Atmospheric Hadley Circulation to Equatorial Anomalies of Ocean Temperature. *Tellus*, **18**(4)**,** 820-&.

——, 1969: Atmospheric Teleconnections from Equatorial Pacific. *Monthly Weather Review*, **97**(3)**,** 163-&.

Black, R. X. and B. A. McDaniel, 2004: Diagnostic case studies of the northern annular mode. *Journal of Climate*, **17**(20)**,** 3990-4004.

Blade, I., 1997: The influence of midlatitude ocean-atmosphere coupling on the low-frequency variability of a GCM .1. No tropical SST forcing. *Journal of Climate*, **10**(8)**,** 2087-2106.

Blender, R., U. Luksch, K. Fraedrich, and C. C. Raible, 2003: Predictability study of the observed and simulated European climate using linear regression. *Quarterly Journal of the Royal Meteorological Society*, **129**(592)**,** 2299-2313.

Bodri, L. and V. Cermak, 2000: Prediction of extreme precipitation using a neural network: application to summer flood occurrence in Moravia. *Advances in Engineering Software*, **31**(5)**,** 311-321.

Bodri, L., 2001: Neural network prediction of monthly precipitation: Application to summer flood occurrence in two regions of Central Europe. *Studia Geophysica Et Geodaetica*, **45**(2)**,** 155-167.

Bojariu, R. and L. Gimeno, 2003: Predictability and numerical modelling of the North Atlantic Oscillation. *Earth-Science Reviews*, **63**(1-2)**,** 145-168.

——, 2003: The role of snow cover fluctuations in multiannual NAO persistence. *Geophysical Research Letters*, **30**(4)**,** art. no.-1156.

Brabson, B. B. and J. P. Palutikof, 2000: Tests of the generalized pareto distribution for predicting extreme wind speeds. *Journal of Applied Meteorology*, **39**(9)**,** 1627-1640.

Brankovic, C., T. N. Palmer, and L. Ferranti, 1994: Predictability of Seasonal Atmospheric Variations. *Journal of Climate*, **7**(2)**,** 217-237.

Breiteig, T. Extra-tropical synoptic cyclones and downward propagating anomalies in the Northern Annular Mode. *Geophysical Research Letters* **35**, art. no. L07809.

Broennimann, S., J. Luterbacher, J. Staehelin, T. M. Svendby, G. Hansen, and T. Svenoe, 2004: Extreme climate of the global troposphere and stratosphere in 1940-42 related to El Nino. *Nature*, **431**(7011)**,** 971-974.

Broennimann, S., 2007: Impact of El Nino Southern Oscillation on European climate. *Reviews of Geophysics*, **45**(2).

Broennimann, S., E. Xoplaki, C. Casty, A. Pauling, and J. Luterbacher, 2007: ENSO influence on Europe during the last centuries. *Climate Dynamics*, **28**(2-3)**,** 181-197.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *Journal of Geophysical Research-Atmospheres*, **111**(D12).

Burnham, K.P. and D.R. Anderson, 2002: Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach. Springer

Campolo, M., A. Soldati, and P. Andreussi, 2003: Artificial neural network approach to flood forecasting in the River Arno. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, **48**(3)**,** 381-398.

Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental Forecasts of El-Nino. *Nature*, **321**(6073)**,** 827-832.

Carson, D. J., 1998: Seasonal forecasting. *Quarterly Journal of the Royal Meteorological Society*, **124**(545)**,** 1-26.

Cassou, C., L. Terray, J. W. Hurrell, and C. Deser, 2004: North Atlantic winter climate regimes: Spatial asymmetry, stationarity with time, and oceanic forcing. *Journal of Climate*, **17**(5)**,** 1055-1068.

Cavazos, T., 1999: Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. *Journal of Climate*, **12**(5)**,** 1506-1523.

Cayan, D. R., 1992: Latent and Sensible Heat-Flux Anomalies over the Northern Oceans - Driving the Sea-Surface Temperature. *Journal of Physical Oceanography*, **22**(8)**,** 859-881.

——, 1992: Latent and Sensible Heat-Flux Anomalies over the Northern Oceans - the Connection to Monthly Atmospheric Circulation. *Journal of Climate*, **5**(4)**,** 354-369.

Cayan, D. R., K. T. Redmond, and L. G. Riddle, 1999: ENSO and hydrologic extremes in the western United States. *Journal of Climate*, **12**(9)**,** 2881-2893.

Chaboureau, J. P. and C. Claud, 2003: Observed variability of North Atlantic oceanic precipitating systems during winter. *Journal of Geophysical Research-Atmospheres*, **108**(D19)**,** art. no.-4435.

Chandler, R. E., 2005: On the use of generalized linear models for interpreting climate variability. *Environmetrics*, **16**(7)**,** 699-715.

Charlton, A. J., A. O'Neill, D. B. Stephenson, W. A. Lahoz, and M. P. Baldwin, 2003: Can knowledge of the state of the stratosphere be used to improve statistical forecasts of the troposphere? *Quarterly Journal of the Royal Meteorological Society*, **129**(595)**,** 3205-3224.

Cohen, J. and D. Rind, 1991: The Effect of Snow Cover on the Climate. *Journal of Climate*, **4**(7)**,** 689-706.

Cohen, J. and D. Entekhabi, 1999: Eurasian snow cover variability and Northern Hemisphere climate predictability. *Geophysical Research Letters*, **26**(3)**,** 345-348.

Cohen, J., K. Saito, and D. Entekhabi, 2001: The role of the Siberian high in Northern Hemisphere climate variability. *Geophysical Research Letters*, **28**(2)**,** 299-302.

Cohen, J. and D. Entekhabi, 2001: The influence of snow cover on Northern Hemisphere climate variability. *Atmosphere-Ocean*, **39**(1)**,** 35-53.

Cohen, J., D. Salstein, and K. Saito, 2002: A dynamical framework to understand and predict the major Northern Hemisphere mode. *Geophysical Research Letters*, **29**(10)**,** art. no.-1412.

Cohen, J. L. and K. Saito, 2003: Eurasian snow cover, more skillful in predicting US winter climate than the NAO/AO? *Geophysical Research Letters,* **30**(23)**,** art. no.-2190.

Colman, A., 1997: Prediction of summer central England temperature from preceding North Atlantic winter sea surface temperature. *International Journal of Climatology,* **17**(12)**,** 1285-1300.

Colman, A. and M. Davey, 1999: Prediction of summer temperature, rainfall and pressure in Europe from preceding winter North Atlantic ocean temperature. *International Journal of Climatology*, **19**(5)**,** 513-536.

Colman, A. W. and M. K. Davey, 2003: Statistical prediction of global sea-surface temperature anomalies. *International Journal of Climatology,* **23**(14)**,** 1677-1697.

Copsey, D., R.T. Sutton and J. Knight, 2006: Recent trends in sea level pressure in the Indian Ocean region. *Geophysical Research Letters*, **33**(19), L19712.

Czaja, A. and J. Marshall, 2001: Observations of atmosphere-ocean coupling in the North Atlantic. *Quarterly Journal of the Royal Meteorological Society,* **127**(576)**,** 1893-1916.

Czaja, A. and C. Frankignoul, 2002: Observed impact of Atlantic SST anomalies on the North Atlantic oscillation. *Journal of Climate*, **15**(6)**,** 606-623.

Davies, J. R., D. P. Rowell, and C. K. Folland, 1997: North Atlantic and European seasonal predictability using an ensemble of multidecadal atmospheric GCM simulations. *International Journal of Climatology*, **17**(12)**,** 1263-1284.

Dawson, C. W. and R. L. Wilby, 1999: A comparison of artificial neural networks used for river flow forecasting. *Hydrology and Earth System Sciences*, **3**(4)**,** 529-540.

de la Torre, L., L. Gimeno, J. A. Anel, and R. Nieto, 2006: Solar influence on Northern Annular Mode spatial structure and QBO modulation. *Particle Acceleration, Space Plasma Physics, Solar Radiation and the Earth's Atmosphere and Climate*, **37**(8)**,** 1635-1639.

DeGaetano, A. T., M. E. Hirsch, and S. J. Colucci, 2002: Statistical prediction of seasonal East Coast winter storm frequency. *Journal of Climate,* **15**(10)**,** 1101-1117.

DelSole, T., 2004: Predictability and information theory. Part I: Measures of predictability. *Journal of the Atmospheric Sciences*, **61**(20)**,** 2425-2440.

Delworth, T. L., 1996: North Atlantic interannual variability in a coupled ocean-atmosphere model. *Journal of Climate*, **9**(10)**,** 2356-2375.

Delworth, T. L. and M. E. Mann, 2000: Observed and simulated multidecadal variability in the Northern Hemisphere. *Climate Dynamics*, **16**, 661-676.

Delworth, T. L. and V. M. Mehta, 1998: Simulated interannual to decadal variability in the tropical and sub-tropical North Atlantic. *Geophysical Research Letters*, **25**(15)**,** 2825-2828.

Deser, C. and M. L. Blackmon, 1993: Surface Climate Variations over the North-Atlantic Ocean During Winter - 1900-1989. *Journal of Climate*, **6**(9)**,** 1743-1753.

Deser, C., G. Magnusdottir, R. Saravanan, and A. Phillips, 2004: The effects of North Atlantic SST and sea ice anomalies on the winter circulation in CCM3. Part II: Direct and indirect components of the response. *Journal of Climate*, **17**(5)**,** 877-889.

Doblas-Reyes, F. J., V. Pavan, and D. B. Stephenson, 2003: The skill of multi-model seasonal forecasts of the wintertime North Atlantic Oscillation. *Climate Dynamics*, **21**(5-6)**,** 501-514.

Drevillon, M., C. Cassou, and L. Terray, 2003: Model study of the North Atlantic region atmospheric response to autumn tropical Atlantic sea-surface-temperature anomalies. *Quarterly Journal of the Royal Meteorological Society*, **129**(593)**,** 2591-2611.

Drobot, S., 2003: Long-range statistical forecasting of ice severity in the Beaufort-Chukchi Sea. *Weather and Forecasting*, **18**(6)**,** 1161-1176.

Enloe, J., J. J. O'Brien, and S. R. Smith, 2004: ENSO impacts on peak wind gusts in the United States. *Journal of Climate*, **17**(8)**,** 1728-1737.

Feddersen, H., 2000: Impact of global sea surface temperature on summer and winter temperatures in Europe in a set of seasonal ensemble simulations. *Quarterly Journal of the Royal Meteorological Society*, **126**(567)**,** 2089-2109.

——, 2003: Predictability of seasonal precipitation in the Nordic region. *Tellus Series a-Dynamic Meteorology and Oceanography*, **55**(5)**,** 385-400.

——, 2003: Impact of tropical SST variations on the linear predictability of the atmospheric circulation in the Atlantic/European region. *Annals of Geophysics*, **46**(1)**,** 109-124.

Feldstein, S. B. and C. Franzke, 2006: Are the North Atlantic Oscillation and the Northern Annular Mode distinguishable? *Journal of the Atmospheric Sciences*, **63**(11)**,** 2915-2930.

Folland, C., J. Owen, M. N. Ward, and A. Colman, 1991: Prediction of Seasonal Rainfall in the Sahel Region Using Empirical and Dynamic Methods. *Journal of Forecasting*, **10**(1-2)**,** 21-56.

Folland, C. K., A. W. Colman, D. P. Rowell, and M. K. Davey, 2001: Predictability of northeast Brazil rainfall and real-time forecast skill, 1987-98. *Journal of Climate*, **14**(9)**,** 1937-1958.

Fowler, A. M. and K. J. Hennessy, 1995: Potential Impacts of Global Warming on the Frequency and Magnitude of Heavy Precipitation. *Natural Hazards*, **11**(3)**,** 283-303.

Fraedrich, K. and K. Muller, 1992: Climate Anomalies in Europe Associated with Enso Extremes. *International Journal of Climatology*, **12**(1)**,** 25-31.

Fraedrich, K., K. Muller, and R. Kuglin, 1992: Northern-Hemisphere Circulation Regimes During the Extremes of the Elnino Southern Oscillation. *Tellus Series a-Dynamic Meteorology and Oceanography*, **44A**(1)**,** 33-40.

Fraedrich, K., 1994: An Enso Impact on Europe - a Review. *Tellus Series a-Dynamic Meteorology and Oceanography*, **46**(4)**,** 541-552.

Fraedrich, K., C. Ziehmann, and F. Sielmann, 1995: Estimates of Spatial Degrees of Freedom. *Journal of Climate*, **8**(2)**,** 361-369.

Frankignoul, C., 1978: Generation of Sea-Surface Temperature Anomalies. *Transactions-American Geophysical Union*, **59**(4)**,** 293-293.

Frankignoul, C. and R. W. Reynolds, 1983: Testing a Dynamical Model for Mid-Latitude Sea-Surface Temperature Anomalies. *Journal of Physical Oceanography*, **13**(7)**,** 1131-1145.

Frankignoul, C., P. Friederichs, and E. Kestenare, 2003: Influence of Atlantic SST anomalies on the atmospheric circulation in the Atlantic-European sector. *Annals of Geophysics*, **46**(1)**,** 71-85.

Frankignoul, C. and E. Kestenare, 2005: Air-sea interactions in the tropical Atlantic: A view based on lagged rotated maximum covariance analysis. *Journal of Climate*, **18**(18)**,** 3874-3890.

Frei, A. and D. A. Robinson, 1999: Northern hemisphere snow extent: Regional variability 1972-1994. *International Journal of Climatology*, **19**(14)**,** 1535-1560.

Frich, P., L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. Tank, and T. Peterson, 2002: Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, **19**(3)**,** 193-212.

Friederichs, P. and C. Frankignoul, 2003: Potential seasonal predictability of the observed Euro-Atlantic atmospheric variability using SST forced ECHAM4-T42 simulations. *Quarterly Journal of the Royal Meteorological Society*, **129**(594)**,** 2879-2896.

Friederichs, P. and A. Hense, 2003: Statistical inference in canonical correlation analyses exemplified by the influence of North Atlantic SST on European climate. *Journal of Climate*, **16**(3)**,** 522-534.

Gallego, M. C., J. A. Garcia, and J. M. Vaquero, 2005: The NAO signal in daily rainfall series over the Iberian Peninsula. *Climate Research*, **29**(2)**,** 103-109.

Garcia-Herrera, R., N. Calvo, R. R. Garcia, and M. A. Giorgetta, 2006: Propagation of ENSO temperature signals into the middle atmosphere: A comparison of two general circulation models and ERA-40 reanalysis data. *Journal of Geophysical Research-Atmospheres*, **111**(D6).

Gershunov, A., 1998: ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Implications for long-range predictability. *Journal of Climate*, **11**(12)**,** 3192-3203.

Gershunov, A. and T. P. Barnett, 1998: ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous United States: Observations and model results (vol 11, pg 1575, 1998). *Journal of Climate*, **11**(11)**,** 3062-3065.

Gershunov, A. and D. R. Cayan, 2003: Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. *Journal of Climate*, **16**(16)**,** 2752-2765.

Gibson, J. K., A. Hernandez, P. Kallberg, A. Nomura, E. Serrano, and S. Uppala, 1996: Current status of the ECMWF Re-Analysis (ERA) project. *Seventh Symposium on Global Change Studies*112-115.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology*, **21**(9)**,** 1111-1152.

Gong, G., D. Entekhabi, and J. Cohen, 2002: A large-ensemble model study of the wintertime AO-NAO and the role of interannual snow perturbations. *Journal of Climate*, **15**(23)**,** 3488-3499.

——, 2003: Relative impacts of Siberian and north American snow anomalies on the winter Arctic Oscillation. *Geophysical Research Letters*, **30**(16)**,** art. no.-1848.

——, 2003: Modeled Northern Hemisphere winter climate response to realistic Siberian snow anomalies. *Journal of Climate*, **16**(23)**,** 3917-3931.

Gong, G., D. Entekhabi, J. Cohen, and D. Robinson, 2004: Sensitivity of atmospheric response to modeled snow anomaly characteristics. *Journal of Geophysical Research-Atmospheres*, **109**(D6)**,** art. no.-D06107.

Gong, G. V., D. Entekhabi, and J. Cohen, 2004: Orographic constraints on a modeled Siberian snow-tropospheric- stratospheric teleconnection pathway. *Journal of Climate*, **17**(6)**,** 1176-1189.

Goswami, B. N., M. S. Madhusoodanan, C. P. Neema, and D. Sengupta, 2006: A physical mechanism for North Atlantic SST influence on the Indian summer monsoon. *Geophysical Research Letters*, **33**(2).

Graham, N. E., J. Michaelsen, and T. P. Barnett, 1987: An Investigation of the El-Nino-Southern Oscillation Cycle with Statistical-Models .2. Model Results. *Journal of Geophysical Research-Oceans*, **92**(C13)**,** 14271-14289.

Gueremy, J. F., M. Deque, A. Braun, and J. P. Piedelievre, 2005: Actual and potential skill of seasonal predictions using the CNRM contribution to DEMETER: coupled versus uncoupled model. *Tellus Series a-Dynamic Meteorology and Oceanography*, **57**(3)**,** 308-319.

Haigh, J. D., M. Blackburn, and R. Day, 2005: The response of tropospheric circulation to perturbations in lower-stratospheric temperature. *Journal of Climate*, **18**(17)**,** 3672-3685.

Halliwell, G. R. and D. A. Mayer, 1996: Frequency response properties of forced climatic SST anomaly variability in the North Atlantic. *Journal of Climate*, **9**(12)**,** 3575-3587.

Halliwell, G. R., 1997: Decadal and multidecadal North Atlantic SST anomalies driven by standing and propagating basin-scale atmospheric anomalies. *Journal of Climate*, **10**(10)**,** 2405-2411.

Hamilton, K., 1993: An Examination of Observed Southern Oscillation Effects in the Northern-Hemisphere Stratosphere. *Journal of the Atmospheric Sciences*, **50**(20)**,** 3468-3473.

Hampson, J. and P. Haynes, 2006: Influence of the equatorial QBO on the extratropical stratosphere. *Journal of the Atmospheric Sciences*, **63**(3)**,** 936-951.

Harris, I., 2005: Generalised pareto methods for wind extremes. Useful tool or mathematical mirage? *Journal of Wind Engineering and Industrial Aerodynamics*, **93**(5)**,** 341-360.

Hartley, D. E., J. T. Villarin, R. X. Black, and C. A. Davis, 1998: A new perspective on the dynamical link between the stratosphere and troposphere. *Nature*, **391**(6666)**,** 471-474.

Hasselmann, K. and T. P. Barnett, 1981: Techniques of Linear Prediction for Systems with Periodic Statistics. *Journal of the Atmospheric Sciences*, **38**(10)**,** 2275-2283.

Hastenrath, S. and L. Greischar, 1993: Changing Predictability of Indian Monsoon Rainfall Anomalies. *Proceedings of the Indian Academy of Sciences-Earth and Planetary Sciences*, **102**(1)**,** 35-47.

Haylock, M. R. and C. M. Goodess, 2004: Interannual variability of european extreme winter rainfall and links with mean large-scale circulation. *International Journal of Climatology*, **24**(6)**,** 759-776.

Haylock, M.R., N. Hofstra, A.M.G. Klein Tank, E.J. Klok, P.D. Jones, M. New, 2007: A European daily high-resolution gridded dataset of surface temperature and precipitation (in prep).

Heino, R., R. Brazdil, E. Forland, H. Tuomenvirta, H. Alexandersson, M. Beniston, C. Pfister, M. Rebetez, G. Rosenhagen, S. Rosner, and J. Wibig, 1999: Progress in the study of climatic extremes in northern and central Europe. *Climatic Change*, **42**(1)**,** 151-181.

Hellstrom, C. and B. A. Malmgren, 2004: Spatial analysis of extreme precipitation in Sweden 1961-2000. *Ambio*, **33**(4-5)**,** 187-192.

Holt, T., J. P. Palutikof, T. J. Osborn, T. D. Davies , 2001: Investigation of the risk from climate change and variability over Northern Europe. Climatic Research Unit, University of East Anglia, UK. 315pp

Hirsch, M. E., A. T. DeGaetano, and S. J. Colucci, 2001: An East Coast winter storm climatology. *Journal of Climate*, **14**(5)**,** 882-899.

Hsieh, W. W. and B. Y. Tang, 1998: Applying neural network models to prediction and data analysis in meteorology and oceanography. *Bulletin of the American Meteorological Society*, **79**(9)**,** 1855-1870.

Hsieh, W. W., 2000: Nonlinear canonical correlation analysis by neural networks. *Neural Networks*, **13**(10)**,** 1095-1105.

——, 2001: Nonlinear canonical correlation analysis of the tropical Pacific climate variability using a neural network approach. *Journal of Climate*, **14**(12)**,** 2528-2539.

——, 2001: Nonlinear principal component analysis by neural networks. *Tellus Series a-Dynamic Meteorology and Oceanography*, **53**(5)**,** 599-615.

Hsieh, W. W. and A. M. Wu, 2002: Nonlinear multichannel singular spectrum analysis of the tropical Pacific climate variability using a neural network approach. *Journal of Geophysical Research-Oceans*, **107**(C7)**,** art. no.-3076.

Hu, Y. Y. and K. K. Tung, 2002: Interannual and decadal variations of planetary wave activity, stratospheric cooling, and Northern Hemisphere Annular mode. *Journal of Climate*, **15**(13)**,** 1659-1673.

Hu, Z. Z. and B. H. Huang, 2006: On the significance of the relationship between the North Atlantic Oscillation in early winter and Atlantic sea surface temperature anomalies. *Journal of Geophysical Research-Atmospheres*, **111**(D12).

Huang, B. H., P. S. Schopf, and J. Shukla, 2004: Intrinsic ocean-atmosphere variability of the tropical Atlantic Ocean. *Journal of Climate*, **17**(11)**,** 2058-2077.

Hurrell, J. W., 1995: Decadal Trends in the North-Atlantic Oscillation - Regional Temperatures and Precipitation. *Science*, **269**(5224)**,** 676-679.

Hurrell, J. W., M. Visbeck, A. Busalacchi, R. A. Clarke, T. L. Delworth, R. R. Dickson, W. E. Johns, K. P. Koltermann, Y. Kushnir, D. Marshall, C. Mauritzen, M. S. McCartney, A. Piola, C. Reason, G. Reverdin, F. Schott, R. Sutton, I. Wainer, and D. Wright, 2006: Atlantic climate variability and predictability: A CLIVAR perspective. *Journal of Climate*, **19**(20)**,** 5100-5121.

Itoh, H. and K. Harada, 2004: Coupling between tropospheric and stratospheric leading modes. *Journal of Climate*, **17**(2)**,** 320-336.

Iwasaki, T., 1991: Year-to-Year Variation of Snow Cover Area in the Northern-Hemisphere. *Journal of the Meteorological Society of Japan*, **69**(2)**,** 209-217.

Jagger, T. H., X. F. Niu, and J. B. Elsner, 2002: A space-time model for seasonal hurricane prediction. *International Journal of Climatology*, **22**(4)**,** 451-465.

Jia, X., H. Lin, and J. Derome, 2008: The influence of tropical Pacific forcing on the Arctic Oscillation. *Climate Dynamics*, **In Press**.

Johansson, A., A. Barnston, S. Suranjana, and H. van den Dool, 1998: On the Level and Origin of Seasonal Forecast Skill in Northern Europe. *Journal of the Atmospheric Sciences*, **55**, 103-127

Johansson, A. 2007: Prediction Skill of the NAO and PNA from Daily to Seasonal Time Scales. *Journal of Climate*, **20**, 1957-1975

Jones, P. D., T. Jonsson, and D. Wheeler, 1997: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *International Journal of Climatology*, **17**(13)**,** 1433-1450.

Jones, P. D., E. B. Horton, C. K. Folland, M. Hulme, D. E. Parker, and T. A. Basnett, 1999: The use of indices to identify changes in climatic extremes. *Climatic Change*, **42**(1)**,** 131-149.

Jones, P.D., Goodess, C., Davies, T. 2001: Foreword (ACCORD Special Issue). *Climate Research*, **17**(3)

Kaczmarek, Z., 2003: The impact of climate variability on flood risk in Poland. *Risk Analysis*, **23**(3)**,** 559-566.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph, 1996: The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77**(3)**,** 437-471.

Kanamitsu, M., A. Kumar, H. M. H. Juang, J. K. Schemm, W. Q. Wang, F. L. Yang, S. Y. Hong, P. T. Peng, W. Chen, S. Moorthi, and M. Ji, 2002: NCEP dynamical seasonal forecast system 2000. *Bulletin of the American Meteorological Society*, **83**(7)**,** 1019-+.

Karl, T. R., N. Nicholls, and A. Ghazi, 1999: CLIVAR/GCOS/WMO Workshop on Indices and Indicators for Climate Extremes - Workshop summary. *Climatic Change*, **42**(1)**,** 3-7.

Kharin, V. V. and F. W. Zwiers, 2002: Climate predictions with multimodel ensembles. *Journal of Climate*, **15**(7)**,** 793-799.

Kim, G. and A. P. Barros, 2001: Quantitative flood forecasting using multisensor data and neural networks. *Journal of Hydrology*, **246**(1-4)**,** 45-62.

Kingston, D. G., D. M. Lawler, and G. R. McGregor, 2006: Linkages between atmospheric circulation, climate and streamflow in the northern North Atlantic: research prospects. *Progress in Physical Geography*, **30**(2)**,** 143-174.

Klein Tank, A., J. B. Wijngaard, G. P. Konnen, R. Bohm, G. Demaree, A. Gocheva, M. Mileta, S. Pashiardis, L. Hejkrlik, C. Kern-Hansen, R. Heino, P. Bessemoulin, G. Muller-Westermeier, M. Tzanakou, S. Szalai, T. Palsdottir, D. Fitzgerald, S. Rubin, M. Capaldo, M. Maugeri, A. Leitass, A. Bukantis, R. Aberfeld, A. F. V. Van Engelen, E. Forland, M. Mietus, F. Coelho, C. Mares, V. Razuvaev, E. Nieplova, T. Cegnar, J. A. Lopez, B. Dahlstrom, A. Moberg, W. Kirchhofer, A. Ceylan, O. Pachaliuk, L. V. Alexander, and P. Petrovic, 2002: Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology*, **22**(12)**,** 1441-1453.

Klein Tank, A. and G. P. Konnen, 2003: Trends in indices of daily temperature and precipitation extremes in Europe, 1946-99. *Journal of Climate*, **16**(22)**,** 3665-3680.

Kodera, K., 1994: Influence of Volcanic-Eruptions on the Troposphere through Stratospheric Dynamical Processes in the Northern-Hemisphere Winter. *Journal of Geophysical Research-Atmospheres*, **99**(D1)**,** 1273-1282.

Kodera, K., 1995: On the Origin and Nature of the Interannual Variability of the Winter Stratospheric Circulation in the Northern-Hemisphere. *Journal of Geophysical Research-Atmospheres*, **100**(D7)**,** 14077-14087.

Kodera, K., M. Chiba, H. Koide, A. Kitoh, and Y. Nikaidou, 1996: Interannual variability of the winter stratosphere and troposphere in the Northern Hemisphere. *Journal of the Meteorological Society of Japan*, **74**(3)**,** 365-382.

Kumar, A. and F. L. Yang, 2003: Comparative influence of snow and SST variability on extratropical climate in northern winter. *Journal of Climate*, **16**(13)**,** 2248-2261.

Kuroda, Y., 2002: Relationship between the Polar-Night Jet Oscillation and the Annular Mode. *Geophysical Research Letters*, **29**(8)**,** art. no.-1240.

Kushnir, Y., W. A. Robinson, I. Blade, N. M. J. Hall, S. Peng, and R. Sutton, 2002: Atmospheric GCM response to extratropical SST anomalies: Synthesis and evaluation. *Journal of Climate*, **15**(16)**,** 2233-2256.

Kushnir, Y., W. A. Robinson, P. Chang, and A. W. Robertson, 2006: The physical basis for predicting Atlantic sector seasonal-to-interannual climate variability. *Journal of Climate*, **19**(23)**,** 5949-5970.

Kvamsto, N. G., P. Skeie, and D. B. Stephenson, 2004: Impact of labrador sea-ice extent on the North Atlantic oscillation. *International Journal of Climatology*, **24**(5)**,** 603-612.

Labitzke, K., 1982: On the Interannual Variability of the Middle Stratosphere During the Northern Winters. *Journal of the Meteorological Society of Japan*, **60**(1)**,** 124-139.

Lamb, H. H., 1950: Types and Spells of Weather around the Year in the British-Isles - Annual Trends, Seasonal Structure of the Year, Singularities. *Quarterly Journal of the Royal Meteorological Society*, **76**(330)**,** 393-438.

Lau, N. C., 1997: Interactions between global SST anomalies and the midlatitude atmospheric circulation. *Bulletin of the American Meteorological Society*, **78**(1)**,** 21-33.

Lebeaupin, C., V. Ducrocq, and H. Giordani, 2006: Sensitivity of torrential rain events to the sea surface temperature based on high-resolution numerical forecasts. *Journal of Geophysical Research-Atmospheres*, **111**(D12).

Lee, Y., J.A. Nelder, Y. Pawitan, 2006: Generalized linear models with random effects: a unified approach via h-likelihood, Chapman and Hall, London

Li, S. L., M. P. Hoerling, and S. L. Peng, 2006: Coupled ocean-atmosphere response to Indian Ocean warmth. *Geophysical Research Letters*, **33**(7).

Lin, H. and J. Derome, 2003: The atmospheric response to North Atlantic SST anomalies in seasonal prediction experiments. *Tellus Series a-Dynamic Meteorology and Oceanography*, **55**(3)**,** 193-207.

Livezey, R. E. and W. Y. Chen, 1983: Statistical Field Significance and Its Determination by Monte-Carlo Techniques. *Monthly Weather Review*, **111**(1)**,** 46-59.

Lloyd-Hughes, B. and M. A. Saunders, 2002: Seasonal prediction of European spring precipitation from El Nino-Southern Oscillation and local sea-surface temperatures. *International Journal of Climatology*, **22**(1)**,** 1-14.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner, 2004: European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, **303**(5663)**,** 1499-1503.

Magnusdottir, G., C. Deser, and R. Saravanan, 2004: The effects of North Atlantic SST and sea ice anomalies on the winter circulation in CCM3. Part I: Main features and storm track characteristics of the response. *Journal of Climate*, **17**(5)**,** 857-876.

Mailier, P. J., D. B. Stephenson, C. A. T. Ferro, and K. I. Hodges, 2006: Serial Clustering of Extratropical Cyclones. *Monthly Weather Review*, **134,** 2224-2240.

Mariotti, A., J. Ballabrera-Poy, and N. Zeng, 2005: Tropical influence on Euro-Asian autumn rainfall variability. *Climate Dynamics,* **24**(5)**,** 511-521.

Marshall, J., Y. Kushner, 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. *International Journal of Climatology*, **21**(15)**,** 1863-1898.

Martin, E., 2004: Validation of Alpine snow in ERA-40. *ERA-40 Project Report Series*, **14**. ECMWF, Reading, UK.

Martineu, C., J. Y. Caneill, and R. Sadourny, 1999: Potential predictability of European winters from the analysis of seasonal simulations with an AGCM. *Journal of Climate,* **12**(10)**,** 3033-3061.

Mathieu, P. P., R. T. Sutton, B. Dong, and M. Collins, 2004: Predictability of winter climate over the North Atlantic European region during ENSO events. *Journal of Climate*, **17**(10)**,** 1953-1974.

McCullagh, P., J.A. Nelder, 1989: Generalized Linear Models, 2nd edR Chapman and Hall, London.

McDaniel, B. A. and R. X. Black, 2005: Intraseasonal dynamical evolution of the Northern annular mode. *Journal of Climate*, **18**(18)**,** 3820-3839.

McGregor, G. R. and I. D. Phillips, 2004: Specification and prediction of monthly and seasonal rainfall over the south-west peninsula of England. *Quarterly Journal of the Royal Meteorological Society*, **130**(596)**,** 193-210.

Mehta, V. M., M. J. Suarez, J. V. Manganello, and T. L. Delworth, 2000: Oceanic influence on the North Atlantic Oscillation and associated Northern Hemisphere climate variations: 1959-1993. *Geophysical Research Letters*, **27**(1)**,** 121-124.

Miller, A. J., 1992: Large-Scale Ocean Atmosphere Interactions in a Simplified Coupled Model of the Midlatitude Wintertime Circulation. *Journal of the Atmospheric Sciences*, **49**(4)**,** 273-286.

Mitchell, T. D. and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, **25**(6)**,** 693-712.

Moberg, A. and P. D. Jones, 2005: Trends in indices for extremes in daily temperature and precipitation in central and western Europe, 1901-99. *International Journal of Climatology*, **25**(9)**,** 1149-1171.

Mudelsee, M., M. Borngen, G. Tetzlaff, and U. Grunewald, 2003: No upward trends in the occurrence of extreme floods in central Europe. *Nature*, **425**(6954)**,** 166-169.

Munoz-Diaz, D. and F. S. Rodrigo, 2006: Seasonal rainfall variations in Spain (1912-2000) and their links to atmospheric circulation. *Atmospheric Research*, **81**(1)**,** 94-110.

Murnane, R. J., 2004: Climate research and reinsurance. *Bulletin of the American Meteorological Society*, **85**(5)**,** 697-707.

Nakicenovic, N. *et al.*, 2000: Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, U.K., 599 pp.

Namias, J. and D. R. Cayan, 1981: Large-Scale Air-Sea Interactions and Short-Period Climatic Fluctuations. *Science*, **214**(4523)**,** 869-876.

Namias, J., 1985: Some Empirical-Evidence for the Influence of Snow Cover on Temperature and Precipitation. *Monthly Weather Review*, **113**(9)**,** 1542-1553.

——, 1988: Abrupt Change in Climate Regime from Summer to Fall 1985 and Stability in the Fall. *Meteorology and Atmospheric Physics*, **38**(1-2)**,** 34-41.

Nicholls, N. and W. Murray, 1999: Workshop on Indices and Indicators for Climate Extremes, Asheville, NC, USA, 3-6 June 1997 - Breakout Group B: Precipitation. *Climatic Change*, **42**(1)**,** 23-29.

Okumura, Y. and S. P. Xie, 2006: Some overlooked features of tropical Atlantic climate leading to a new Nino-like phenomenon. *Journal of Climate*, **19**(22)**,** 5859-5874.

Paciorek, C. J., J. S. Risbey, V. Ventura, and R. D. Rosen, 2002: Multiple indices of Northern Hemisphere cyclone activity, winters 1949-99. *Journal of Climate*, **15**(13)**,** 1573-1590.

Paeth, H., M. Latif, and A. Hense, 2003: Global SST influence on twentieth century NAO variability. *Climate Dynamics*, **21**(1)**,** 63-75.

Palmer, T. N. and D. L. T. Anderson, 1994: The Prospects for Seasonal Forecasting - a Review Paper. *Quarterly Journal of the Royal Meteorological Society*, **120**(518)**,** 755-793.

Palmer, T. N. and J. Ralsanen, 2002: Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature*, **415**(6871)**,** 512-514.

Palmer, T. N., A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Delecluse, M. Deque, E. Diez, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J. F. Gueremy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J. M. Terres, and M. C. Thomson, 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction (demeter). *Bulletin of the American Meteorological Society*, **85**(6)**,** 853-+.

Palutikof, J. P., B. B. Brabson, D. H. Lister, and S. T. Adcock, 1999: A review of methods to calculate extreme wind speeds. *Meteorological Applications*, **6**(2)**,** 119-132.

Palutikof, J. P., J. A. Halliday, G. M. Watson, T. Holt, R. J. Barthelmie, J. P. Coelingh, L. Folkerts, and J. W. Cleijne, 2002: Predicting the wind energy resource in the offshore seas of Europe. *Third Symposium on Environmental Applications: Facilitating the Use of Environmental Information*70-73.

Parker, D., C. Folland, A. Scaife, J. Knight, A. Colman, P. Baines, and B. W. Dong, 2007: Decadal to multidecadal variability and the climate change background. *Journal of Geophysical Research-Atmospheres*, **112**.

Pavan, V. and F. J. Doblas-Reyes, 2000: Multi-model seasonal hindcasts over the Euro-Atlantic: skill scores and dynamic features. *Climate Dynamics*, **16**(8)**,** 611-625.

Pavan, V., F. Molteni, and C. Brankovic, 2000: Wintertime variability in the Euro-Atlantic region in observations and in ECMWF seasonal ensemble experiments. *Quarterly Journal of the Royal Meteorological Society*, **126**(567)**,** 2143-2173.

Pavan, V., S. Tibaldi, and C. Brankovic, 2000: Seasonal prediction of blocking frequency: Results from winter ensemble experiments. *Quarterly Journal of the Royal Meteorological Society*, **126**(567)**,** 2125-2142.

Peng, S. L. and L. A. Mysak, 1993: A Teleconnection Study of Interannual Sea-Surface Temperature-Fluctuations in the Northern North-Atlantic and Precipitation and Runoff over Western Siberia. *Journal of Climate*, **6**(5)**,** 876-885.

Peng, S. L., W. A. Robinson, and S. L. Li, 2002: North Atlantic SST forcing of the NAO and relationships with intrinsic hemispheric variability. *Geophysical Research Letters*, **29**(8)**,** art. no.-1276.

——, 2003: Mechanisms for the NAO responses to the North Atlantic SST tripole. *Journal of Climate*, **16**(12)**,** 1987-2004.

Peng, S. L., W. A. Robinson, S. L. Li, and M. P. Hoerling, 2005: Tropical Atlantic SST forcing of coupled north Atlantic seasonal responses. *Journal of Climate*, **18**(3)**,** 480-496.

Penland, C. and T. Magorian, 1993: Prediction of Nino 3 sea surface temperatures using linear inverse modeling. *Journal of Climate*, **6**, 1067-1076.

Pezzulli, S., D. B. Stephenson, and A. Hannachi, 2005: The variability of seasonality. *Journal of Climate*, **18**(1)**,** 71-88.

Piedelievre, J. P., 2000: Numerical seasonal predictions using global climate models. Researches on the long range forecasting at Meteo France. *Stochastic Environmental Research and Risk Assessment*, **14**(4-5)**,** 319-338.

Plaut, G. and E. Simonnet, 2001: Large-scale circulation classification, weather regimes, and local climate over France, the Alps and Western Europe. *Climate Research*, **17**(3)**,** 303-324.

Pozo-Vazquez, D., S. R. Gamiz-Fortis, J. Tovar-Pescador, M. J. Esteban-Parra, and Y. Castro-Diez, 2005: North Atlantic winter SLP anomalies based on the autumn ENSO state. *Journal of Climate*, **18**(1)**,** 97-103.

——, 2005: El Nino Southern oscillation events and associated European winter precipitation anomalies. *International Journal of Climatology*, **25**(1)**,** 17-31.

Qian, B. D. and M. A. Saunders, 2003: Seasonal predictability of wintertime storminess over the North Atlantic. *Geophysical Research Letters*, **30**(13)**,** art. no.-1698.

——, 2003: Summer UK temperature and its links to preceding Eurasian snow cover, North Atlantic SSTs, and the NAO. *Journal of Climate*, **16**(24)**,** 4108-4120.

Quadrelli, R., M. Lazzeri, C. Cacciamani, and S. Tibaldi, 2001: Observed winter Alpine precipitation variability and links with large-scale circulation patterns. *Climate Research*, **17**(3)**,** 275-284.

Quadrelli, R. and J. M. Wallace, 2004: A simplified linear framework for interpreting patterns of Northern Hemisphere wintertime climate variability. *Journal of Climate*, **17**(19)**,** 3728-3744.

Ramaswamy, V., M. L. Chanin, J. Angell, J. Barnett, D. Gaffen, M. Gelman, P. Keckhut, Y. Koshelkov, K. Labitzke, J. J. R. Lin, A. O'Neill, J. Nash, W. Randel, R. Rood, K. Shine, M. Shiotani, and R. Swinbank, 2001: Stratospheric temperature trends: Observations and model simulations. *Reviews of Geophysics*, **39**(1)**,** 71-122.

Randall, D. A., R. D. Cess, J. P. Blanchet, S. Chalita, R. Colman, D. A. Dazlich, A. D. Delgenio, E. Keup, A. Lacis, H. Letreut, X. Z. Liang, B. J. McAvaney, J. F. Mahfouf, V. P. Meleshko, J. J. Morcrette, P. M. Norris, G. L. Potter, L. Rikus, E. Roeckner, J. F. Royer, U. Schlese, D. A. Sheinin, A. P. Sokolov, K. E. Taylor, R. T. Wetherald, I. Yagai, and M. H. Zhang, 1994: Analysis of Snow Feedbacks in 14 General-Circulation Models. *Journal of Geophysical Research-Atmospheres*, **99**(D10)**,** 20757-20771.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research-Atmospheres*, **108**(D14).

Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **126**(563)**,** 649-667.

Robertson, A. W., C. R. Mechoso, and Y. J. Kim, 2000: The influence of Atlantic sea surface temperature anomalies on the north Atlantic oscillation. *Journal of Climate*, **13**(1)**,** 122-138.

Robinson, D. A., K. F. Dewey, and R. R. Heim, 1993: Global Snow Cover Monitoring - an Update. *Bulletin of the American Meteorological Society*, **74**(9)**,** 1689-1696.

Robinson, W. A., S. L. Li, and S. L. Peng, 2003: Dynamical nonlinearity in the atmospheric response to Atlantic sea surface temperature anomalies. *Geophysical Research Letters*, **30**(20)**,** art. no.-2038.

Robock, A., 2001: Stratospheric forcing needed for dynamical seasonal prediction. *Bulletin of the American Meteorological Society*, **82**(10)**,** 2189-2192.

Rodwell, M. J., D. P. Rowell, and C. K. Folland, 1999: Oceanic forcing of the wintertime North Atlantic Oscillation and European climate. *Nature*, **398**(6725)**,** 320-323.

Rodwell, M. J. and C. K. Folland, 2002: Atlantic air-sea interaction and seasonal predictability. *Quarterly Journal of the Royal Meteorological Society*, **128**(583)**,** 1413-1443.

——, 2003: Atlantic air-sea interaction and model validation. *Annals of Geophysics*, **46**(1)**,** 47-56.

Rodwell, M. J. and F. J. Doblas-Reyes, 2006: Medium-range, monthly and seasonal prediction for Europe and the use of forecast information. *Journal of Climate*, **19**(23)**,** 6025-6046.

Ropelewski, C. F. and M. S. Halpert, 1987: Global and Regional Scale Precipitation Patterns Associated with the El-Nino Southern Oscillation. *Monthly Weather Review*, **115**(8)**,** 1606-1626.

Ropelewski, C. F. and P. D. Jones, 1987: An Extension of the Tahiti-Darwin Southern Oscillation Index. *Monthly Weather Review*, **115**(9)**,** 2161-2165.

Ross, B. and J. E. Walsh, 1986: Synoptic-Scale Influences of Snow Cover and Sea Ice. *Monthly Weather Review*, **114**(10)**,** 1795-1810.

Saito, K., J. Cohen, and D. Entekhabi, 2001: Evolution of atmospheric response to early-season Eurasian snow cover anomalies. *Monthly Weather Review*, **129**(11)**,** 2746-2760.

Saito, K. and J. Cohen, 2003: The potential role of snow cover in forcing interannual variability of the major Northern Hemisphere mode (vol 30, pg 1302, 2003). *Geophysical Research Letters*, **30**(11)**,** art. no.-1594.

——, 2003: The potential role of snow cover in forcing interannual variability of the major Northern Hemisphere mode. *Geophysical Research Letters*, **30**(6)**,** art. no.-1302.

Saito, K., T. Yasunari, and J. Cohen, 2004: Changes in the sub-decadal covariability between northern hemisphere snow cover and the general circulation of the atmosphere. *International Journal of Climatology*, **24**(1)**,** 33-44.

Saravanan, R., 1998: Atmospheric low-frequency variability and its relationship to midlatitude SST variability: Studies using the NCAR Climate System Model. *Journal of Climate*, **11**(6)**,** 1386-1404.

Saunders, M. A., and B. D. Qian, 2002: Seasonal predictability of the winter NAO from north Atlantic sea surface temperatures. *Geophysical Research Letters*, **29**(22), art. no. 2049

Saunders, M. A., B. D. Qian, and B. Lloyd-Hughes, 2003: Summer snow extent heralding of the winter North Atlantic Oscillation. *Geophysical Research Letters,* **30**(7)**,** art. no.-1378.

Scaife, A. A., J. R. Knight, G. K. Vallis, and C. K. Folland, 2005: A stratospheric influence on the winter NAO and North Atlantic surface climate. *Geophysical Research Letters*, **32**(18).

Scaife, A. A., C. K. Folland, L. V. Alexander, A. Moberg, and J. R. Knight, 2008: European climate extremes and the North Atlantic Oscillation. *Journal of Climate*, **21**72-83.

Schlosser, C. A. and D. M. Mocko, 2003: Impact of snow conditions in spring dynamical seasonal predictions. *Journal of Geophysical Research-Atmospheres*, **108**(D16)**,** art. no.-8616.

Schmith, T., 2001: Global warming signature in observed winter precipitation in Northwestern Europe? *Climate Research*, **17**(3)**,** 263-274.

Scott, R. K. and D. G. Dritschel, 2005: Downward wave propagation on the polar vortex. *Journal of the Atmospheric Sciences*, **62**(9)**,** 3382-3395.

Seager, R., Y. Kushnir, M. Visbeck, N. Naik, J. Miller, G. Krahmann, and H. Cullen, 2000: Causes of Atlantic Ocean climate variability between 1958 and 1998. *Journal of Climate*, **13**(16)**,** 2845-2862.

Shamseldin, A. Y. and K. M. O'Connor, 2001: A non-linear neural network technique for updating of river flow forecasts. *Hydrology and Earth System Sciences*, **5**(4)**,** 577-597.

Shongwe, M.E., C.A.T. Ferro, C.A.S. Coelho and G.J. van Oldenborgh, 2007: Predictability of cold spring seasons in Europe. *Monthly Weather Review*, **135**, 4185-4201

Shukla, J., J. Anderson, D. Baumhefner, C. Brankovic, Y. Chang, E. Kalnay, L. Marx, T. Palmer, D. Paolino, J. Ploshay, S. Schubert, D. Straus, M. Suarez, and J. Tribbia, 2000: Dynamical seasonal prediction. *Bulletin of the American Meteorological Society*, **81**(11)**,** 2593-2606.

Smith, I., 1994: Indian-Ocean Sea-Surface Temperature Patterns and Australian Winter Rainfall. *International Journal of Climatology*, **14**(3)**,** 287-305.

Sutton, R. T. and M. R. Allen, 1997: Decadal predictability of North Atlantic sea surface temperature and climate. *Nature*, **388**(6642)**,** 563-567.

Sutton, R. T., S. P. Jewson, and D. P. Rowell, 2000: The elements of climate variability in the tropical Atlantic region. *Journal of Climate*, **13**(18)**,** 3261-3284.

Sutton, R. T. and D. L. R. Hodson, 2003: Influence of the ocean on North Atlantic climate variability 1871-1999. *Journal of Climate*, **16**(20)**,** 3296-3313.

Taguchi, M. and S. Yoden, 2002: Internal interannual variability of the troposphere-stratosphere coupled system in a simple global circulation model. Part II: Millennium integrations. *Journal of the Atmospheric Sciences*, **59**(21)**,** 3037-3050.

Thompson, D. W. J. and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophysical Research Letters*, **25**(9)**,** 1297-1300.

Thompson, D. W. J., M. P. Baldwin, and J. M. Wallace, 2002: Stratospheric connection to Northern Hemisphere wintertime weather: Implications for prediction. *Journal of Climate*, **15**(12)**,** 1421-1428.

Timlin, M. S., M. A. Alexander, and C. Deser, 2002: On the reemergence of North Atlantic SST anomalies. *Journal of Climate*, **15**(18)**,** 2707-2712.

Timmermann, A., M. Latif, R. Voss, and A. Grotzner, 1998: Northern Hemispheric interdecadal variability: A coupled air- sea mode. *Journal of Climate*, **11**(8)**,** 1906-1931.

Trigo, R. M., T. J. Osborn, and J. M. Corte-Real, 2002: The North Atlantic Oscillation influence on Europe: climate impacts and associated physical mechanisms. *Climate Research*, **20**(1)**,** 9-17.

Troccoli, A., M. Harrison, D. L. T. Anderson, and S. J. Mason (eds), 2008: Seasonal Climate: Forecasting and Managing Risk. NATO Science Series IV: Earth and Environmental Sciences – Vol. 82. Springer, Dordrecht.

Uppala, S. M., P. W. Kallberg, A. J. Simmons, U. Andrae, V. D. Bechtold, M. Fiorino, J. K. Gibson, J. Haseler, A. Hernandez, G. A. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. P. Allan, E. Andersson, K. Arpe, M. A. Balmaseda, A. C. M. Beljaars, L. Van De Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Holm, B. J. Hoskins, L. Isaksen, P. Janssen, R. Jenne, A. P. McNally, J. F. Mahfouf, J. J. Morcrette, N. A. Rayner, R. W. Saunders, P. Simon, A. Sterl, K. E. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen, 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131**(612)**,** 2961-3012.

Van den Dool, H., 2007: Empirical Methods in Short Term Climate Prediction. Oxford University Press, Oxford.

Ventura, V., C. J. Paciorek, and J. S. Risbey, 2004: Controlling the proportion of falsely rejected hypotheses when conducting multiple tests with climatological data. *Journal of Climate*, **17**(22)**,** 4343-4356.

Visbeck, M., H. Cullen, G. Krahmann, and N. Naik, 1998: An ocean model's response to North Atlantic Oscillation-like wind forcing. *Geophysical Research Letters*, **25**(24)**,** 4521-4524.

Von Storch, H. and F.W. Zwiers, 1999: Statistical Analysis in Climate Research. Cambridge University Press, Cambridge.

Von Storch, H. and A. Navarra, 1999: Analysis of Climate Variability: Applications of Statistical Techniques. Springer

Walker, G. T. and E. W. Bliss, 1932: World Weather *V. Mem. Roy. Meteor. Soc*., **4**(36), 53-84

Wallace, J. M., C. Smith, and Q. R. Jiang, 1990: Spatial Patterns of Atmosphere Ocean Interaction in the Northern Winter. *Journal of Climate*, **3**(9)**,** 990-998.

Wallace, J. M., 2000: North Atlantic Oscillation/annular mode: Two paradigms - one phenomenon. *Quarterly Journal of the Royal Meteorological Society*, **126**(564)**,** 791-805.

Walland, D. J. and I. Simmonds, 1996: Modelled atmospheric response to changes in Northern Hemisphere snow cover. *Climate Dynamics*, **13**(1)**,** 25-34.

——, 1997: Association between modes of variability of January Northern Hemisphere snow cover and circulation. *Theoretical and Applied Climatology*, **58**(3-4)**,** 197-210.

Walsh, J. E., 1984: Snow Cover and Atmospheric Variability. *American Scientist*, **72**(1)**,** 50-57.

Wang, X. C. and S. S. Shen, 1999: Estimation of spatial degrees of freedom of a climate field. *Journal of Climate*, **12**(5)**,** 1280-1291.

Wang, C. C. and J. C. Rogers, 2001: A composite study of explosive cyclogenesis in different sectors of the North Atlantic. Part I: Cyclone structure and evolution. *Monthly Weather Review*, **129**(6)**,** 1481-1499.

Ward, M. N. and C. K. Folland, 1991: Prediction of Seasonal Rainfall in the North Nordeste of Brazil Using Eigenvectors of Sea-Surface Temperature. *International Journal of Climatology*, **11**(7)**,** 711-743.

Wedgbrow, C. S., R. L. Wilby, H. R. Fox, and G. O'Hare, 2002: Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales. *International Journal of Climatology*, **22**(2)**,** 219-236.

Weng, W. J. and J. D. Neelin, 1998: On the role of ocean-atmosphere interaction in midlatitude interdecadal variability. *Geophysical Research Letters*, **25**(2)**,** 167-170.

Wilby, R. L., 2001: Seasonal forecasting of river flows in the British Isles using North Atlantic pressure patterns. *Journal of the Chartered Institution of Water and Environmental Management*, **15**(1)**,** 56-63.

Wilby, R. L., C. S. Wedgbrow, and H. R. Fox, 2004: Seasonal predictability of the summer hydrometeorology of the River Thames, UK. *Journal of Hydrology*, **295**(1-4)**,** 1-16.

Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences. Academic Press, San Diego.

Willmott, C. J. and K. Matsuura, 2005: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, **30**(1)**,** 79-82.

Wu, P. L. and C. Gordon, 2002: Oceanic influence on North Atlantic climate variability. *Journal of Climate*, **15**(14)**,** 1911-1925.

Wu, A. M. and W. W. Hsieh, 2004: The nonlinear northern hemisphere winter atmospheric response to ENSO. *Geophysical Research Letters*, **31**(2)**,** art. no.-L02203.

Wyrtki, K., 1975: El Nino - Dynamic-Response of Equatorial Pacific Ocean to Atmospheric Forcing. *Journal of Physical Oceanography*, **5**(4)**,** 572-584.

——, 1979: Comments on the Variability of the Tropical Ocean. *Dynamics of Atmospheres and Oceans*, **3**(2-4)**,** 209-212.

Xie, P. P. and P. A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *Journal of Climate*, **9**(4)**,** 840-858.

Xie, S. P., 2004: Satellite observations of cool ocean-atmosphere interaction. *Bulletin of the American Meteorological Society*, **85**(2)**,** 195-+.

Yan, Z. W., S. Bate, R. E. Chandler, V. Isham, and H. Wheater, 2002: An analysis of daily maximum wind speed in northwestern Europe using generalized linear models. *Journal of Climate*, **15**(15)**,** 2073-2088.

Ye, H. C., 2000: Decadal variability of Russian winter snow accumulation and its associations with Atlantic sea surface temperature anomalies. *International Journal of Climatology*, **20**(14)**,** 1709-1728.

——, 2001: Quasi-biennial and quasi-decadal variations in snow accumulation over northern Eurasia and their connections to the Atlantic and Pacific Oceans. *Journal of Climate*, **14**(24)**,** 4573-4584.

Zanchettin, D., S. W. Franks, P. Traverso, and M. Tomasino, 2008: On ENSO impacts on European wintertime rainfalls and their modulation by the NAO and the Pacific multi-decadal variability described through the PDO index. *International Journal of Climatology*, **28**995-1006.

Zebiak, S. E. and M. A. Cane, 1987: A Model El-Nino Southern Oscillation. *Monthly Weather Review*, **115**(10)**,** 2262-2278.

Zhang, X. B., G. Hegerl, F. W. Zwiers, and J. Kenyon, 2005: Avoiding inhomogeneity in percentile-based indices of temperature extremes. *Journal of Climate*, **18**(11)**,** 1641-1651.

Zheng, X. G. and J. A. Renwick, 2003: A regression-based scheme for seasonal forecasting of New Zealand temperature. *Journal of Climate*, **16**(11)**,** 1843-1853.

Zhou, S. T., A. J. Miller, J. Wang, and J. K. Angell, 2002: Downward-propagating temperature anomalies in the preconditioned polar stratosphere. *Journal of Climate*, **15**(7)**,** 781-792.

Zolina, O., A. Kapala, C. Simmer, and S. K. Gulev, 2004: Analysis of extreme precipitation over Europe from different reanalyses: a comparative assessment. *Global and Planetary Change*, **44**(1-4)**,** 129-161.

Zorita, E., V. Kharin, and H. Vonstorch, 1992: The Atmospheric Circulation and Sea-Surface Temperature in the North-Atlantic Area in Winter - Their Interaction and Relevance for Iberian Precipitation. *Journal of Climate*, **5**(10)**,** 1097-1108.

Zorita, E. and C. Frankignoul, 1997: Modes of north Atlantic decadal variability in the ECHAM1/LSG coupled ocean-atmosphere general circulation model. *Journal of Climate*, **10**(2)**,** 183-200.

Zveryaev, II, 2004: Seasonality in precipitation variability over Europe. *Journal of Geophysical Research-Atmospheres*, **109**(D5)**,** art. no.-D05103.