Seasonal Forecasts of European Winter Storms – Skill, Dynamics and Constraints –

By

Lisa Degenhardt

A thesis submitted to
the University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY



School of Geography, Earth and Environmental Science
College of Life and Environmental Sciences
University of Birmingham
July 2023

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive
e-theses repository



This unpublished thesis/dissertation is under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:



Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Unless otherwise stated, any material in this thesis/dissertation that is cited to a third-party source is not included in the terms of this licence. Please refer to the original source(s) for licencing conditions of any quotes, images or other material cited to a third party.

© Copyright by LISA DEGENHARDT, 2023
All Rights Reserved

ABSTRACT

Severe winter windstorms are one of the most damaging extreme events for the extra-tropics, especially in Europe. Companies like the SwissRe have estimated losses of around USD 190 million for European windstorms in alone 2022 (SwissRe, 2023). These rare events are a seasonal phenomenon. Hence, they occur especially on a seasonal scale, and their amount and strength vary each season. Besides the scientific purpose, politics, insurances and the general public are interested in understanding forecasts of these extreme events. They want to know how reliable those seasonally averaged forecasts are and how they can be improved. Hence, forecasts with a few months' lead times are crucial for such events on a seasonal scale. The investigation covers multiple aspects of seasonally predicted winter windstorms over Europe. Firstly, the general forecast skill of seasonally averaged storm characteristics (frequency and intensity). Secondly, the link of these forecasts not only to known large-scale patterns like the North Atlantic Oscillation (NAO), Scandinavian Pattern (SCA) and East-Atlantic Pattern (EA) but also to more physics-based dynamical factors like the mean sea-level pressure (MSLP) gradient or measure for the instability of the atmosphere (Eady Growth Rate or equivalent potential temperature). The third part studies the constraints of the seasonal forecast model regarding windstorms and focuses on the signal-to-noise paradox.

This study investigates the seasonal forecast model of the UK Met Office, Global Seasonal forecasting system version 5, GloSea5, and compares it with different statistical measures to an observational data set, ERA5 from the European Centre for Medium-Range Weather Forecasts (ECWMF). Windstorms are tracked over the core winter season, December-January-February, with an impact-based algorithm that focuses on the 2% strongest wind events over Europe and the North Atlantic.

The results of this study can be separated by the mentioned study aims. Firstly, a positive forecast skill, as correlation, for storm frequency and for the first time for storm intensity

measures have been found over coherent regions over the British Isles and southern Scandinavia. Secondly, large-scale patterns like the NAO, SCA and EA have been linked to windstorms and explain up to 80% variability of windstorm frequency and 60% of intensity when statistically combined. But small-scale physical factors are investigated as well and it can be concluded that the seasonal forecast model is showing windstorm forecast skill out of the right physical connections between windstorms and atmospheric factors. The third part is about model constraints and therefore, the signal-to-noise paradox has been assessed within seasonal windstorm prediction. The paradox can limit the predictability of the model. For windstorm forecasts, this paradox exists strongly in the same regions as the significant forecast skill. A hypothesis that the disturbances of windstorms are causing the paradox has been investigated. Indications of a stronger signal-to-noise paradox within stormy seasons have been found in the GloSea5 model.

This all concludes that GloSea5 as a seasonal forecast model is capable of predicting the seasonal average of rare and extreme events like winter windstorms. This prediction is robust as it results from the right physical connections but is still reduced because of the existence of the signal-to-noise paradox.

To the Faculty of the University of Birmingham:

The members of the Committee appointed to examine the dissertation of LISA DEGENHARDT find it satisfactory and recommend that it be accepted.

Joshua Larsen, Dr., Chair

Ruth Geen, Dr.

Helen Dacre, Prof.

ACKNOWLEDGMENTS

First, I would like to thank UKRI, NERC and CENTA, which decided in 2019 to support this project and me financially. I am very thankful for this opportunity and couldn't have done it without this financial support. Followed by this, I would like to thank my supervisory team, Adam A. Scaife, Martin Widmann and Gregor C. Leckebusch, for all the fruitful and productive meetings throughout the 3.5 years. A very special thanks to Gregor, my main supervisor, for not only supporting my academic journey but also understanding and supporting me in situations like Covid-19 or simply being an international student. I also want to say a big thank you to my colleagues, new, constant or old, from the working group, Kelvin, Michael and Xiaoqi. They were a great support in handling data or giving constructive scientific feedback in seminars or coffee mornings.

Outside of academia, I couldn't have done this dissertation without the mental support from my family, my partner, my German friends and the amazing group of friends I met within GEES and CENTA. My whole family, my mum & dad, my brother and my grandparents, have always supported and believed in me and what I am capable of. Each phone call or visit back home to Germany gave me new energy and motivation for my studies. Each lunch, coffee break, games night, or climbing session here in Birmingham gave me a feeling of being home. Every minute of my time in Birmingham, Flo was only a phone call away and knew how to get me back on track, motivate me, calm me down and gather my thoughts. Gini was a great friend, travelling multiple times to Birmingham to bring home and girl's night a bit closer. And not to forget my Mojito-Monday Core Group. Luckily we did not have Mojitos each Monday, but going through the journey of a PhD with all of you, Adrià, Nine, Gianni, Klaske, Manon, Alex, Aaron, Andrea, Jess, Grace[2] and more, was just as joyful as having a Mojito every Monday.

And last but not least, I would like to thank the Viva board with both examiners, Ruth Geen and Helen Dacre, and Josh Larsen as chair for agreeing to take this job and taking the time to read my thesis. I am looking forward to having an interesting Viva about my studies of the last 3.5 years.

Table of Contents

Page

Abstract i

Acknowledgments vii

Table of Contents ix

List of Figures xiii

List of Tables xvii

List of Acronyms xix

1 Motivation - Storm Season 2021/2022 1

1.1 Aim of Study 4

2 Literature Overview 5

2.1 History of Seasonal Prediction 6

2.2 Seasonal Prediction Model Suites 8

2.2.1 UK Met Office - United Kingdom 11

2.2.2 European Center for Medium Weather Forecast - Europe 12

2.2.3 Météo France - France 13

2.2.4 German Weather Service - Germany 14

2.2.5 Euro-Mediterranean Center for Climate Change - Italy 14

2.2.6 National Centers for Environmental Prediction - United States of America 15

2.2.7 Environmental and Climate Change Canada - Canada 16

2.2.8 Japan Meteorological Agency - Japan 17

2.3 Seasonal Prediction Skill 19

2.3.1 Predictability of Meteorological Variables 19

2.3.2 Predictability of large-scale pattern . . . 20
2.3.3 Predictability of extreme events . . . 23
2.4 Drivers/Factors for extreme events . . . 24
2.5 Signal-to-noise paradox . . . 26

3 Data . . . 29
3.1 Reanalysis Data Set - Reanalysis Data Set ERA5 (ERA5) . . . 30
3.2 Seasonal Forecast Systems - UKMO Met Office Global Seasonal forecasting system version 5 (GloSea5) . . . 30
3.3 Data availability, use and capacity . . . 32

4 Applied and Tested Methods . . . 33
4.1 Tracking . . . 34
4.1.1 Windstorm Tracking . . . 34
4.1.2 Storm Frequency . . . 35
4.1.3 Storm Severity Measures . . . 40
4.2 EOF Analysis . . . 41
4.3 Statistics . . . 43
4.3.1 Skill Scores . . . 43
4.3.2 Correlation Coefficient . . . 47
4.3.3 Used Statistical Measures . . . 49

5 Seasonal Forecast Skill of Windstorm Characteristics & their Relation to Large-Scale Patterns . . . 51
5.1 Large-scale circulation patterns and their influence on European winter windstorm predictions . . . 52
5.1.1 Introduction . . . 53
5.1.2 Data . . . 57
5.1.3 Method . . . 58
5.1.4 Results . . . 63
5.1.5 Discussion & Summary . . . 73
5.2 Additional Outcome: Seasonal Storm Forecast Skill Dependent on Large-Scale Patterns . . . 77
5.2.1 Storm Frequency . . . 77
5.2.2 Storm Intensity . . . 79

5.2.3 Discussion 81

6 Dynamical influencing factors on seasonal winter windstorm predictions over Europe 83

6.1 Understanding Winter Windstorm Predictability over Europe 84

6.1.1 Introduction 85

6.1.2 Data 89

6.1.3 Method 91

6.1.4 Results 95

6.1.5 Discussion and Conclusion 105

6.2 Additional Outcome: Different Views on Assessing the Connection between Dynamical Factors and Windstorms 110

6.2.1 Eady Growth Rate (400hPa) 111

6.2.2 Mean Sea-Level Pressure Gradient 113

6.2.3 Conclusion 114

7 Signal-to-Noise paradox in seasonal winter windstorm predictions 115

7.1 Signal-to-noise paradox dependent on winter windstorm counts over Europe . 116

7.1.1 Introduction 117

7.1.2 Data 118

7.1.3 Method 119

7.1.4 Influence of windstorms on NAO paradox 123

7.1.5 Potential reasons 128

7.1.6 Discussion and Conclusion 129

7.2 Additional Outcome: The Signal-to-noise Paradox in Seasonal Predictions of Storm-related Dynamical Factors 133

8 Synthesis 137

8.1 Overall Summary 137

8.2 Discussion 140

8.3 Conclusion & Outlook 144

References 147

Appendices

A For Chapter 4 . . . . . A1

B For Chapter 5 . . . . . A9

C For Chapter 6 . . . . . A17

D For Chapter 7 . . . . . A27

List of Figures

1.1 Weather Maps of the three extreme storms "*Dudley*", "*Eunice*" and "*Franklin*" from 17^{th} of February 2022 to 19^{th} of February 2022. Maps are ground analysis of the German Berliner Wetterkarte (first row), from the UK Met Office (second row) and infrared 10.8 satellite images (third row). 3

2.1 Summary of historical pioneers in long-range forecast, taken from Namias (1968). 7

2.2 Meteorological time scales and predictability (combined from Simon et al. (2007) and German Weather Service (DWD)). 9

4.1 Kendall Correlation as skill measure to investigate seasonal forecast skill of severe winter windstorm frequencies for different data combinations and settings. 95% significance tested with a Kendall-correlation test and marked with dots. 37

4.2 Kendall Correlation as skill measure to investigate seasonal forecast skill of severe winter windstorm frequencies for different storm frequency counting radii, ERA-interim and Met Office Global Seasonal forecasting system version 5 (GloSea5) GC2 ensemble mean for December-January-February (DJF), 1993-2013, 6hourly resolution of 10m wind speed used for tracking. 95% significance tested with a Kendall-correlation test and marked with dots. 39

5.1 Workflow of Methodology. 58

5.2 Forecast skill for storm parameter (Frequency, ESSIa, ESSIs). Ranked Kendall-τ_b-correlation between ERA5 and GloSea5 GC2 ensemble mean (63 members) from winter season 1993/94 to 2015/16, regions with 1.9 or less storms in observation (ERA5) are shaded, dotted grid points are significant on 95% level. 63

5.3 Forecast ROC skill scores. Area under the ROC curve per grid cell between ERA5 and GloSea5 GC2 ensemble (63 member) for storm parameter (Frequency, ESSIa, ESSIs) from winter season 1993/94 to 2015/16, regions with 1.9 or less storms in observation (ERA5) are shaded, dotted grid points are significant on 95% level. 65

5.4 Forecast skill of the city scale. Mean ROC curve for 3 grid cells (London, Hamburg and Oslo) and their respective westerly and northwesterly grid cell, between ERA5 and GloSea5 GC2 ensemble (63 member) for storm parameter (Frequency, ESSIa, ESSIs) from winter season 1993/94 to 2015/16, associated ROC skill score as area under the curve in the bottom corner for the single "city"-grid cell, significance of ROCSS on 95% level marked with $*$; highest tercile (purple), middle tercile (green), lower tercile (orange). 66

5.5 Linear models of storm parameters (Frequency, ESSIa, ESSIs). Regression slopes from multiple linear regressions to the first three EOF patterns (NAO, SCA and EA), in ERA5 winter season 1993/94 to 2015/19. 68

5.6 Variance explained by linear models of storm parameters (Frequency, ESSIa, ESSIs). Determination parameter R^2 (a-c) and individual ratio of explained variance from ANOVA of first three EOF patterns (NAO – d-f, SCA – g-i, EA – j-l), in ERA5 winter season 1993/94 to 2015/16. 69

5.7 Forecast skill for storm parameter (Frequency, ESSIa, ESSIs). Ranked Kendall-τ_b-correlation between ERA5 and GloSea5 GC2 ensemble mean (63 member) (1st row), between ERA5 and regression model storm parameters (2nd row) and difference of significant correlation values as (1st row) minus (2nd and 3rd row) (4th and 5th row). Regions with 1.9 or less storms in observation (ERA5) are shaded. Dotted grid points are significant on 95% level (1st 3 rows). For the ease of comparison only those regions where one of the approaches shows positive significant skill are shown as differences and marked by a dot (last 2 rows). 71

5.8 Ratio of predictable components for storm parameter (Frequency, ESSIa, ESSIs). $RPC = r_{mo}/r_{mm}$, with r as ranked Kendall-τ_b-correlation between ERA5 and GloSea5 GC2 ensemble (63 members) from winter season 1993/94 to 2015/16, regions with 1.9 or less storms in observation (ERA5) are suppressed (grey). A ratio above 1 means that the signal-to-noise paradox exists as the model mean is better representing the reality than its individual members. . 73

5.9 Standardised Principle Component Analysis (PCA) indices of the three most dominant large-scale patterns over Europe, NAO, SCA and EA, with terciles (grey dashed line) as thresholds for the sub-samples. 77

5.10 Seasonal storm frequency forecast skill, as Kendall correlation, in different phases (positive, neutral & negative) of different large-scale patterns (NAO, SCA, EA). Regions with 1.9 or less storms in observation (ERA5) are shaded. Dotted grid points are significant on 95% level. 78

5.11 Seasonal storm ESSIa forecast skill, as Kendall correlation, in different phases (positive, neutral & negative) of different large-scale patterns (NAO, SCA, EA). Regions with 1.9 or less storms in observation (ERA5) are shaded. Dotted grid points are significant on 95% level. 80

5.12 As Fig. 5.11 for seasonal storm ESSIs forecast skill. 80

6.1 Schematic of dynamical factor connections to cyclones and windstorms (blue). Separated into different theories, quasi-geostrophic omega theory (green), potential vorticity theory (purple) and additional factors (grey). All connections and notes on arrows are based on literature. 92

6.2 Schematic map of the spatial location of factors compared to an idealised storm system. 93

6.3 Composite Anomalies, standardised on climatology, of the respective factor for strong storm seasons minus weak storm seasons in ERA5 (left column) and GloSea5 mean over all ensemble members (right column): a)&b) MSLP Gradient, c)&d) SST, e)&f) Θ_e, g)&h) EGR; dots shown for differences significant at the 90% level (p=0.9). Yellow boxes are selected regions for investigation step 3, Fig. 6.6, process-based-view (right column). 97

6.4 Correlation Maps between seasonal storm counts over the UK and dynamical factors. Only factors with a 95% significant correlation in ERA5 are shown. ERA5 connections (1^{st} row), GloSea5 member connection mean (2^{nd} row), GloSea5 individual member correlation distribution (below). The distributions are scales from -1 to 1 with 0 in the centre. The coloured line is the ERA5 correlation value within the GloSea5 member distribution, and the number represents the percentile of ERA5 in that distribution. 98

6.5 Kendall Correlation maps for selected dynamical factors, significance on 95% level marked by a dot. Yellow boxes are selected regions for investigation step 3, Fig. 6.6, factor-skill-view (left column). 100

6.6 Difference of windstorm forecast skill (as Kendall correlation) with seasonal separation by the Factor-skill-view (left column) and the Process-based-view (right column). The separation is based on spatial averages over the shown boxes from Fig. 6.5 for the left column and Fig. 6.3 for the right column, individually per factor. Dotted grid cells showed significant storm forecast skill in the bad predicted season, Triangle in the well predicted sub-sample. . 103

6.7 Different perspectives on the analysis of the connection between windstorm frequency and EGR in 400hPa, in ERA5 (top) and GloSea5 mean over all member-individual correlations (bottom). The significance correlation-test is only calculated for ERA5 (top row), as GloSea5 (bottom row) is an average over all individual members. 112

6.8 Different perspectives on analysing the connection between windstorm frequency and Mean Sea-Level Pressure Gradient, in ERA5 (top) and GloSea5 mean over all member-individual correlations (bottom). The significance correlation-test is only calculated for ERA5 (top row), as GloSea5 (bottom row) is an average over all individual members. 113

7.1 Regions for different separation per storm counts. Region 1 - -15-10°E & 48-60°N (light blue), region 2 - -15-20°E & 48-64°N (dark blue), region 3 - -40-20°E & 48-64°N (green), region 4 - 50-10°W & 40-60°N(red). 121

7.2 Ratio of Predictable component (RPC - green legend) and the individual components (r_{mo} & r_{mm} - blue-to-orange legend, σ_{total} & $\sigma_{ensmean}$ - purple legend) for NAO in the whole data set, in stormy, quiet, NAO+ and NAO– seasons. For selected C3S data sets. Separation for region 1 (-15°-10°E & 48°-60°N). Significance tested with correlation test for correlation columns, rest by 10.000 bootstrapping, * is 90%, ** is 95% significant. (Dashed) Arrows mark when two values are more than 2σ (1.7σ) apart. Grey arrows represent the test after removing 5% outliers in bootstrap distribution. 124

7.3 As Fig. 7.2 for Multi-Model-Mean approaches in region 1 (-15°-10°E & 48°-60°N). 125

7.4 As Fig. 7.2 & Fig. 7.3 for region 2 (-15°-20°E & 48°-64°N). 126

7.5 Mean sea level pressure predictable component of the observations as standard deviation of ERA5 multiplied with the forecast skill (left) and the predictable component of the model ensemble mean as standard deviation of the ensemble mean (right) for stormy (top) and quiet (bottom) seasons in GloSea5 for region 1 (-15°-10°E & 48°-60°N). 128

7.6 Ratio of Predictable component (PRC) for the four main influencing dynamical factors, only values higher than one are coloured. 135

B.1 EOF pattern of ERA5 and GloSea5 for EOF 1 to 3, correspond to NAO, SCA and EA for North Atlantic, DJF, 1993-2015. . . . . A12

B.2 Regression slopes as in Fig. 5.5 but based on GloSea5. . . . . A13

B.3 Regression model variance as in Fig. 5.6 but with GloSea5 regression. A14

B.4 Kendall Correlation Difference of direct and indirect approach for ERA5-based regression with the leave-one-out method. A15

C.1 ERA5 UK storm counts as dots and GloSea5 ensemble mean counts as bars. Bad predicted seasons (green), well predicted seasons (purple), weak ERA5 seasons (triangles) and strong ERA5 seasons (squares). A19

C.2 As Fig. 6.3 for remaining primary and secondary factors. A20

C.3 As Fig. 6.4 for remaining primary and secondary factors. A21

C.4 As Fig. 6.5 for remaining primary and secondary factors. A22

C.5 As Fig. 6.6 for remaining primary and secondary factors. A23

D.1 Ratio of Predictable component (PRC) for the remaining primary and secondary dynamical factors. A28

List of Tables

2.1 Summary of all available seasonal forecast suites within Copernicus Climate Change Service (C3S). 18

3.1 Summary of used variables in Reanalysis Data Set ERA5 (ERA5) and Met Office Global Seasonal forecasting system version 5 (GloSea5). 32

4.1 Important settings for WTrack (windstorm algorithm) established from Leckebusch et al. (2008b) 35

4.2 Settings for calculating storm frequency. 36

4.3 Summary of skill measures depending on forecast method and data format (Wilks, 2011; Mason & Stephenson, 2008). 43

4.4 2x2 Contingency Table Wilks (2011). 46

6.1 Dynamical Factors concerning storminess, cyclones or windstorms over Europe. 90

7.1 List of storm counts in different regions over North Atlantic and Europe and the allocated sub-sample, incl. NAO-separation. 120

7.2 Eddy feedback parameter (like Hardiman et al., 2022) for all 23 seasons and as sub-samples from stormy and quiet seasons in ERA5 and GloSea5 ensemble mean, incl. model member error as $2\sigma_{member}/\sqrt{n_{member}}$, for all tested regions. 129

A.1 Skill Score list of skill score measure in different forecast skill publications. . A2

C.1 Tested skilful regions of factor forecast skill. A24

C.2 Tested relevant regions of ERA5 & GloSea5 composite anomalies. A25

List of Acronyms

Θ_e equivalent potential temperature

ACC Anomaly correlation coefficient

AO Atlantic Oscillation

BlueBEAR Birmingham Environment for Academic Research

BWK Berliner Wetterkarte

C3S Copernicus Climate Change Service

CDO Climate Data Operators

CMCC Euro-Mediterranean Center for Climate Change

DePreSys Met Office decadal prediction system

DJF December-January-February

DWD German Weather Service

EA East Atlantic Pattern

EAWR East Atlantic Western Russia Pattern

ECCC Environmental and Climate Change Canada

ECMWF European Centre for Medium-Range Weather Forecasts

EFP eddy feedback parameter

EGR Eady Growth Rate

EGU European Geoscience Union General Assembly

ENSO El Nino Southern Oscillation

EOF Empirical Orthogonal Function

ERA5 Reanalysis Data Set ERA5

ESSI Event SSI

ESSIa seasonal averaged ESSI

ESSIs standardised seasonal storm-count ESSI

ETC Extra-tropical cyclones

GCFS German Climate Forecast System

GloSea4 Met Office Global Seasonal forecasting system version 4

GloSea5 Met Office Global Seasonal forecasting system version 5

GloSea6 Met Office Global Seasonal forecasting system version 6

HadGEM3 Hadley Centre Global Environment Model version 3

HPC High Performance Computing

IFS Integrated Forecasting System

IOD Indian Ocean Dipole

JMA Japan Meteorological Agency

MF Mètèo France

MJO Madden-Julian Oscillation

MSE Mean-Squared Error

MSLP Mean Sea–Level Pressure

NAO North Atlantic Oscillation

NCEP National Centers for Environmental Prediction

NOAA National Oceanic and Atmospheric Administration

PCA Principle Component Analysis

PNA Pacific–North American Pattern

QBO Quasi-Biennial-Oscillation

RMSE Root-Mean-Squared Error

ROC Relative Operating Characteristics

ROCSS Relative Operating Characteristics Skill Score

RPC Ratio of Predictable Components

RPS Ranked Probability Score

RPSS Ranked Probability Skill Score

RWB Rossby Wave Breaking

SCA Scandinavian Pattern

SEAS4 ECMWF's seasonal forecasting system 4

SEAS5 ECMWF's seasonal forecasting system 5

SPV stratospheric polar vortex

SSI Severity Storm Index

SST Sea Surface Temperature

SSW stratospheric sudden warming

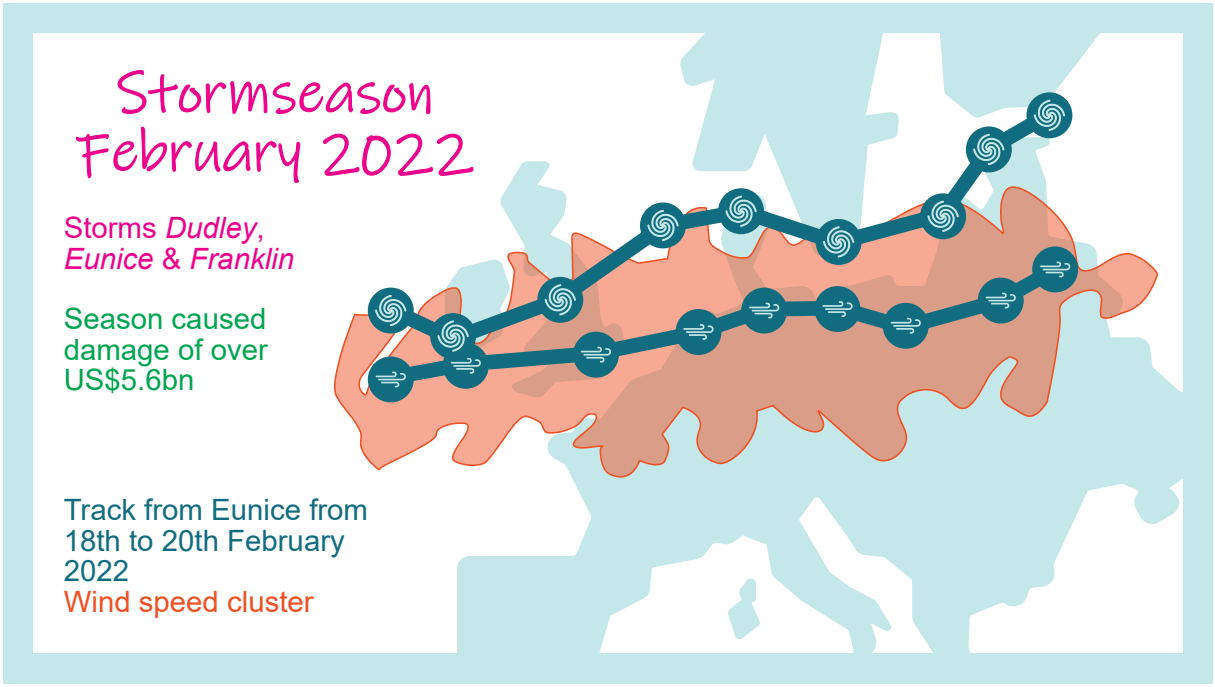
TC tropical cyclone

UK United Kingdom

UKMO UK Met Office

Chapter One

Motivation - Storm Season 2021/2022



Severe windstorms are common in the mid-latitudes, especially in Europe. They occur mostly in the winter season, December-January-February (DJF), as consequences of strong low-pressure systems, Extra-tropical cyclones (ETC). These extreme events are one of the most damaging for Europe, along with, e.g., floods (MunichRe, 2019). A recent strong season was the winter of 2021/2022. At the end of the season, in February 2022, a cyclone family reached Europe, and three extreme cyclones and hence extreme windstorm events affected Europe, especially the United Kingdom (UK), Ireland, France and Germany.

Different Meteorological institutes named these three storms, but all named them. For many institutes, like the UK Met Office (UKMO), the first criterion to name a cyclone is that the event affects Europe with a potentially high impact. The Berliner Wetterkarte (BWK) names all (low and high) pressure systems that make landfall in Europe, already since 2002, independent from the strength. They named these low pressure systems "*Ylenia*", "*Zeynep*" and "*Antonia*". The UKMO is naming extreme storms since 2015 and named them "*Dudley*", "*Eunice*" and "*Franklin*". In the following chapter, English naming will be used but includes information from German sources like maps from the BWK in Fig. 1.1. The following chapter will summarise these extreme cyclones and give an overview of their synoptical development. Meteorological extreme observations regarding wind will be summarised. Data and figures for this are collected from different sources, like Wetter3.de (2003), Berliner Wetterkarte e.V. (1998), Kendon (2022) or Ogimet (2005) and all times are in UTC.

The most extreme and damaging storm was Storm "*Eunice*". The UKMO issued multiple red wind warnings for England. A red wind warning has been issued only 7 times since 2011 for England, according to the National Severe Weather Warnings Archive (UK Met Office, 2023a). Especially cities like London reached unusually high wind gusts of over $100 km/h$ by Storm "*Eunice*", e.g., London-Heathrow with $113 km/h$ at 12 pm, 18/02/2022. Media reported about flying trampolines and that the roof of the O_2 arena got damaged (BBC, 2022). The country-wide extreme wind speed on the 18^{th} of February was measured on the Isle of Portland with $144.5 km/h$, within the highest category in the Beaufort scale (12), hurricane force. Reinsurances, like the MunichRe, have afterwards published that these three mentioned windstorms have caused damage of over US$5.6bn (MunichRe, 2023a, 2023b). This was a very rare and extreme event for the British Isles and even caused multiple case studies already to understand the dynamics and prediction of such systems (see, e.g., Volonté et al., 2023a, 2023b).

The whole of February 2022 was dominated by multiple low-pressure systems intensifying over the North Atlantic. In the first half of February 2022, the systems decreased in strength

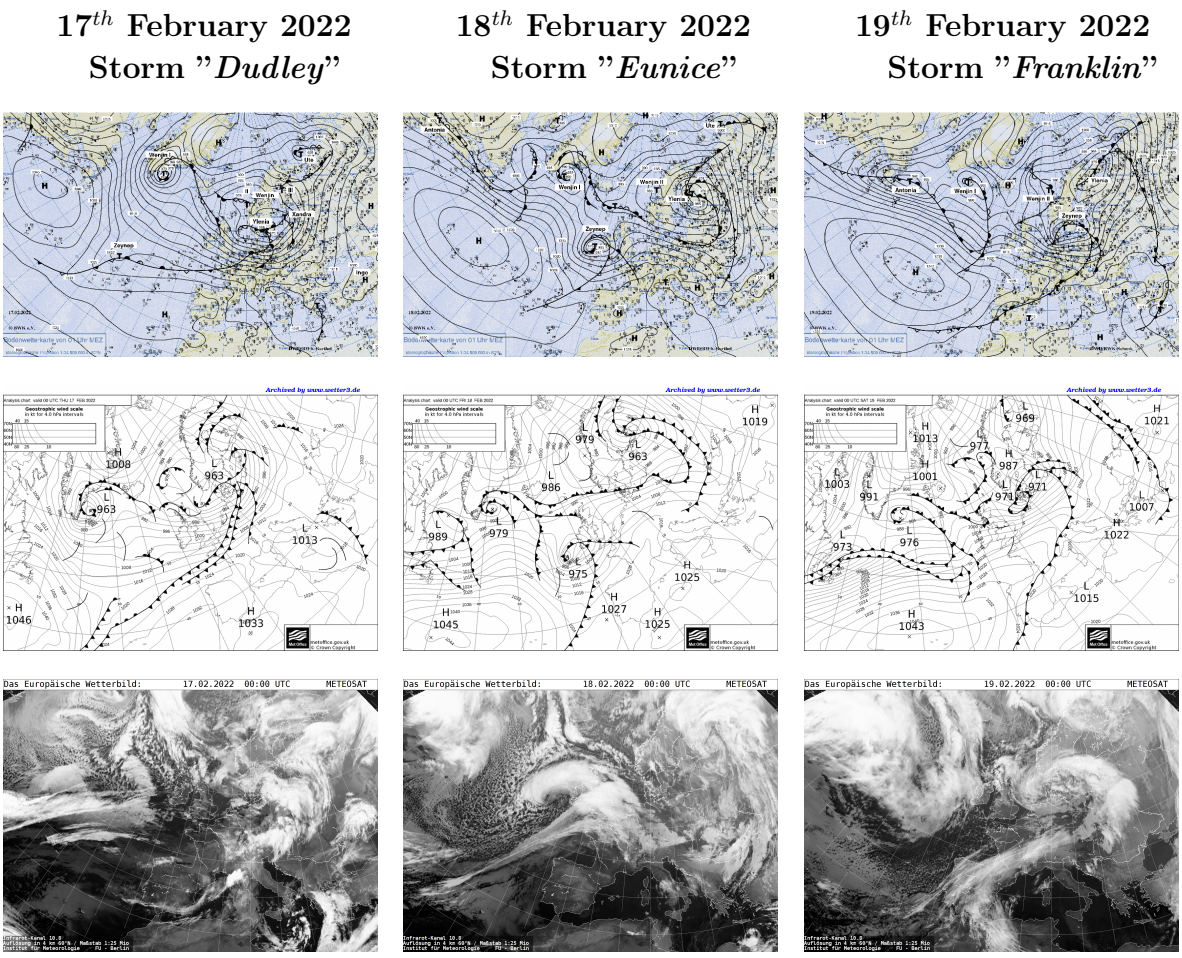


Figure 1.1 Weather Maps of the three extreme storms "*Dudley*", "*Eunice*" and "*Franklin*" from 17^{th} of February 2022 to 19^{th} of February 2022. Maps are ground analysis of the German Berliner Wetterkarte (first row), from the UK Met Office (second row) and infrared 10.8 satellite images (third row).

over the European continent. The first cyclonic storm hitting Europe with a centred low pressure of 970hPa was Storm "*Dudley*" on the 16^{th} of February. This extreme cyclone reached wind speeds with a peak of $129km/h$ at 4 pm in Capel Curig, Wales, and on average over $85km/h$, Beaufort scale 9 - strong/severe gale, for the whole UK. The weather chart analysis of the 17^{th} of February (Fig. 1.1) shows already another low-pressure system generating over the North Atlantic at the frontal system of Storm "*Dudley*". This Storm, now called "*Eunice*", reaches the British Isles by the 18^{th} of February, roughly around 6 am, with a central pressure of 975hPa. The frontal system of Storm "*Eunice*" is fully developed and

brings wind speeds of over $100km/h$ in gust winds over the whole of South-West England. The centre of cyclone "*Eunice*" passed directly over the south of Ireland and caused maximum winds of $136km/h$, 9 am, 18/02/2022, at Roches Point and on average $98km/h$ on the 18^{th} of February for whole Ireland. Storm "*Eunice*" persists as a strong low-pressure system over the next 2 to 3 days while moving east and covering Europe. By the end of the 20^{th} of February, Storm "*Eunice*" is located over Russia and reduced its intensity. Storm "*Eunice*" caused much wind-related damage in other European countries, like Germany, with an average wind speed of $98km/h$ on the 18^{th} of February. The third strong cyclone, Storm "*Franklin*", reached the UK with a low pressure of 960hPa north of Scotland by the end of the 20^{th} of February.

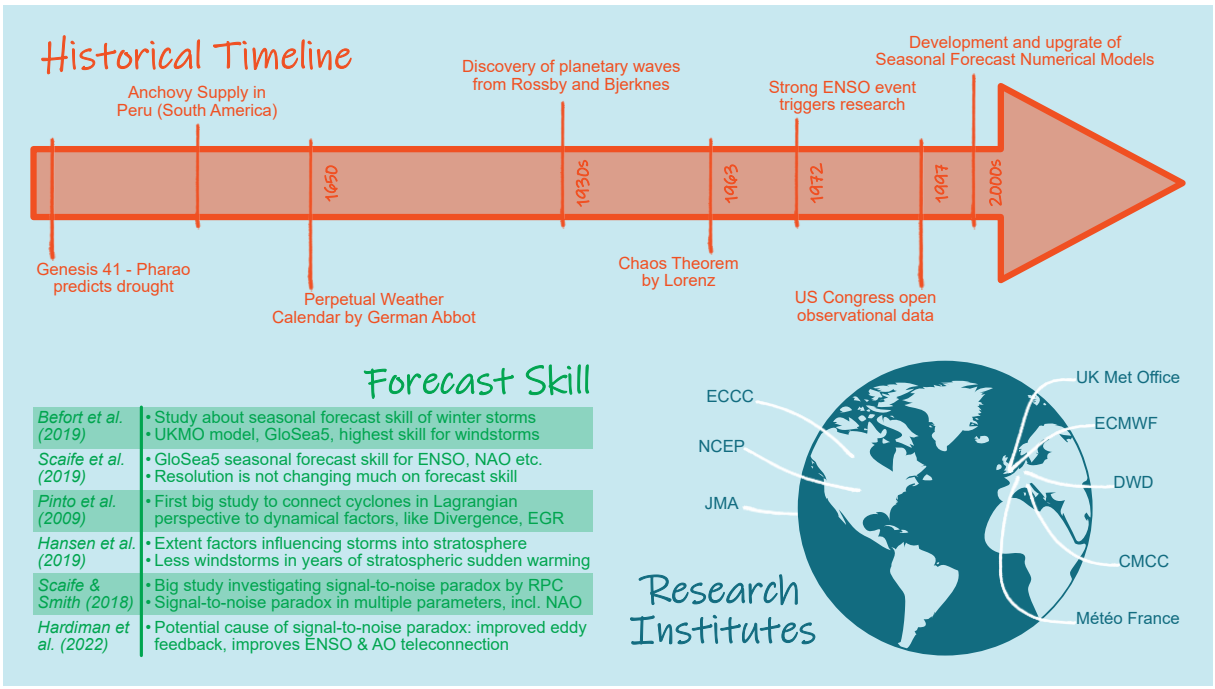
1.1 Aim of Study

Events like Storm "*Eunice*" are rare but cause billions of Euros in damage. E.g. up to 1.4 billion just for Germany caused by the three storms in February 2022 (AXA Agentur Koller & Baur, 2022). These events occur on a seasonal time scale, and their damage is mostly measured on a seasonal average. This study aims to investigate, assess and better understand available predictions of winter windstorms on a seasonal scale. New investigations are:

- Investigate seasonal forecast skill of winter windstorm characteristics (e.g., storm frequency and intensity)
- Identify the connection of large-scale patterns to windstorm characteristics
- Assessing physical drivers of windstorms and connecting them to the prediction performance
- Understanding the reasons for numerical ensemble forecast model limits (signal-to-noise paradox) concerning windstorms

Chapter Two

Literature Overview



The following chapter aims to summarise the literature on the general topic of seasonal predictions. It will start with a historical introduction to the long history of seasonal predictions and use outside numerical ensemble model suites (chapter 2.1). A range of modern numerical seasonal prediction suites will be described in chapter 2.2, summarising the current operating seasonal forecast models. Chapter 2.3 will introduce the current state-of-the-art of seasonal forecast prediction skills for different variables: general meteorological variables, large-scale patterns like North Atlantic Oscillation (NAO) or El Nino Southern Oscillation (ENSO) and

existing literature about seasonal extreme event predictions. The second result part of the thesis (chapter 6) will analyse dynamical factors connected to storms. Current literature about potential factors will be summarised in chapter 2.4. The literature overview ends with a summary of the current knowledge about the signal-to-noise paradox (chapter 2.5), which will be further investigated in the results chapter 7.

2.1 History of Seasonal Prediction

The interest in numerical dynamical forecast models and their utilisation as operational systems on a seasonal scale is recent for some research centres, and the interest has grown in the last decades. But the interest in seasonal predictions and observing weather phenomena on the seasonal time scale started quite early, centuries ago (Troccoli et al., 2008). The scientific approach of numerically modelling the seasonal scale only started end of the 20^{th} century (Stockdale, 2000).

It is well known how fishermen in South America, especially Peru, have realised seasonal variability centuries ago. They have not studied the weather on purpose, but their fish (anchovy) supply varied around their coasts. This happened especially around Christmas, which is why they gave this phenomenon the name "El Niño" (Troccoli et al., 2008). There is little literature on the far-back history of seasonal forecasts before numerical models, but books like Troccoli et al. (2008) and Board et al. (2016) summarise some key facts in their introductions. For example, the oral tradition from the Peruvians and the anchovy fishing or about tribes that used flora and fauna to know about the current states of the atmosphere and the development of the upcoming seasons (Troccoli et al., 2008). Even the bible includes long-range predictions already. In Genesis 41, it says that Pharaoh had a dream about a 7-year-long drought. The people prepared for it with food supplies etc. and survived the drought thanks to the Pharaoh's dream (Stockdale, 2000). Hence, it is not surprising that a German abbot observed the weather on a longer time scale in the 1650s. Dr Knauer was one

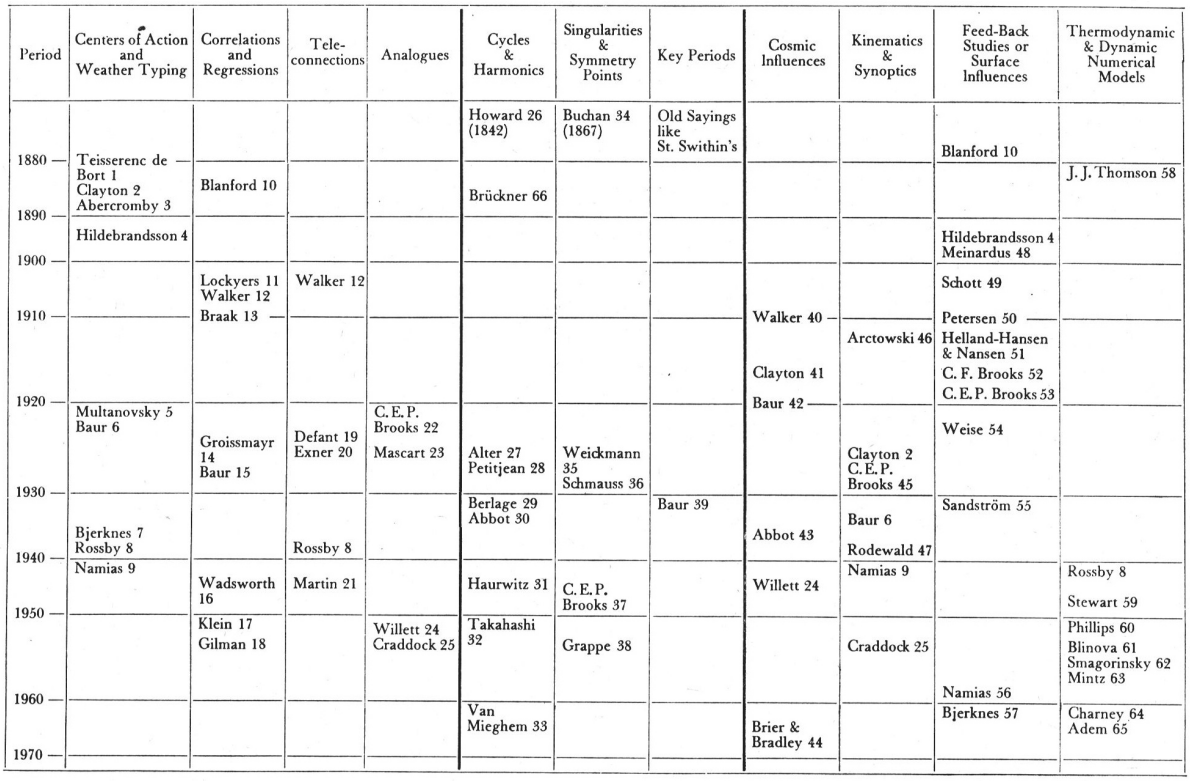


Figure 2.1 Summary of historical pioneers in long-range forecast, taken from Namias (1968).

of the first to observe weather and created a perpetual weather calendar (Stockdale, 2000). The first scientific observations and attempts to understand the longer time scales for weather started in the late 19^{th} century with Teisserenc de Bort, who started investigating the surface pressure and categorised the first weather types (Namias, 1968). This had further development in 1920 with the Grosswetterlagen from Mulanovsky and Baur (see Fig. 2.1, Namias, 1968). Fig. 2.1 summarised by Namias (1968) shows the key research in long-range forecasts ongoing from the late 1800s to the mid-1900s. This table mainly focuses on Northern Hemispheric research, like the discovery of planetary waves from Rossby and Bjerknes in the late 1930s (Namias, 1968). Rossby's research in planetary waves and vorticity conservation was important for understanding large-scale phenomena. But researchers also realised the importance of longer times scales and larger circulations for the Southern Hemisphere. So did Walker, for example, investigate the Southern Oscillation as opposite pressure anomalies

in Sydney and Buenos Aires in 1924 (Stockdale, 2000). In the mid-1950s, large-scale patterns continued to be studied in the Australian Weather Bureau (Board et al., 2016).

Many attempts at understanding and maybe even reproducing the discovered long-range phenomenon, like Rossby Waves, ENSO etc., failed and demotivated scientists (Namias, 1968), but with the discovery of the chaos theorem of Lorenz (1963), these failures were explained (Stockdale, 2000). A strong ENSO event in 1972/73 increased the interest in understanding this phenomenon, and many scientists investigated again (Troccoli et al., 2008). This also triggered the first development of dynamical forecast models in the 1980s, which mainly aimed to predict ENSO correctly (Board et al., 2016). 1997 the US Congress released their observational data with open access, which was a huge break point for the model performance (Troccoli et al., 2008). Since then, research centres have worked and developed their long-range or seasonal forecast suites. A summary of a few recent examples follows in the next chapter 2.2.

2.2 Seasonal Prediction Model Suites

The interest in seasonal predictions has increased in the last few decades. An overlap of two major model conceptions is one of the main reasons seasonal forecasts are still developing, and this time scale is still difficult to predict. While weather forecasts are already quite well predicted but have to resolve the initial condition problem, climate projections are still less predictable because of the boundary-value problem (see Fig. 2.2). Seasonal forecasts have both issues to overcome (Simon et al., 2007), illustrated in Fig. 2.2.

This study mainly focuses on the seasonal forecast model of the UKMO. Still, more research institutes are working on seasonal forecast suites to use in their operational work and communicate well-predicted meteorological parameters to the general public. This chapter introduces more available seasonal forecast models, especially available on the Copernicus Climate Change Service (C3S) archive (Copernicus Climate Change Service, 2023b). As

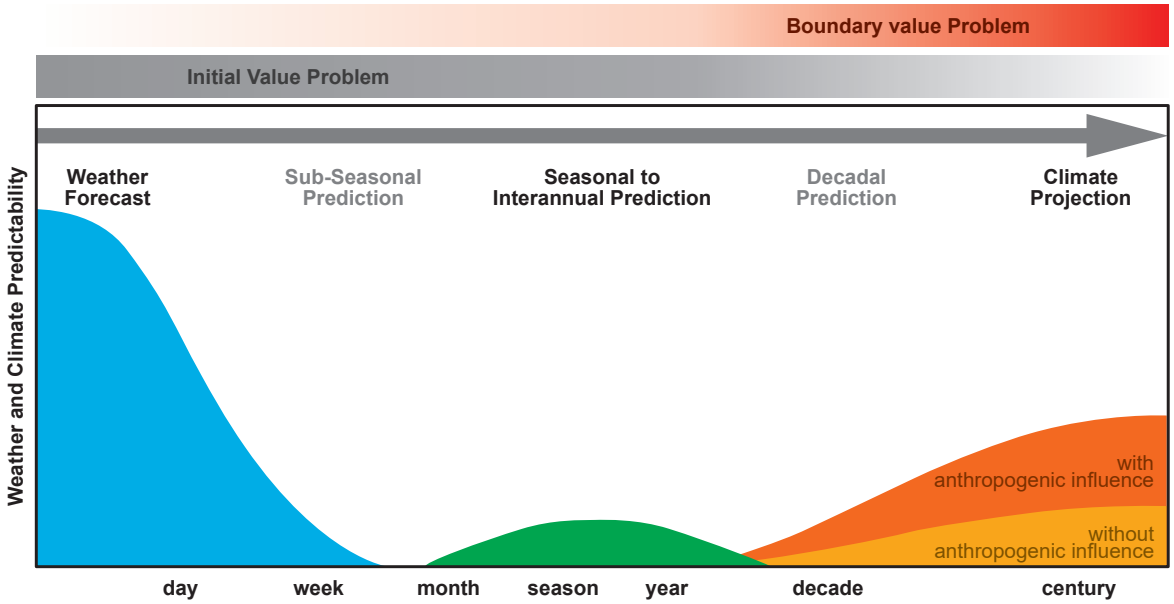


Figure 2.2 Meteorological time scales and predictability (combined from Simon et al. (2007) and German Weather Service (DWD)).

shortly described in chapter 2.1, seasonal forecasting has a long history, but the use of numerical models and operational use is recent.

Numerical models in general are used to predict the atmospheric state of a later time step by using the knowledge of the dynamics of the atmosphere (Holton, 2004). High Performance Computing (HPC) are used to run dynamical complex equations that represent physical processes in the atmosphere. Essential for a numerical prediction is a correct representation of these dynamics, the current state of the atmosphere and a good temporal development of the different processes Holton (2004). Institutes, like the UKMO or the European Centre for Medium-Range Weather Forecasts (ECMWF) develop and maintain their own global numerical forecasting systems (Hadley Centre Global Environment Model version 3 (HadGEM3) current version of UKMO and Integrated Forecasting System (IFS) for ECMWF). These systems are used for all different forecast time scales as "background" model, followed by more specific data assimilation, model coupling or downscaling for different forecasts purposes (UK Met Office, 2023b; ECMWF, 2023b, 2023a):

- **Resolution:** The models' final resolution is a computational power and storage issue.

Downscaling to higher resolutions can be used and is needed for shorter forecast time scales, like weather predictions, as these are only produced for specific smaller regions. Longer time scales, more than a month, need the whole interaction of the global model. Hence, the final resolution will be lower as more capacity is needed.

- **Complexity/Coupling:** Longer time scales need more complexity in terms of the full interaction of the atmosphere with other global systems, like ocean or sea ice.
- **Data Assimilation:** Data assimilation is the concept of adjusting and comparing a forecast output with newly measured observations. These observations are mainly satellite data or weather station measurements. The model will be adjusted to these new observations. This process happens on different scales for different forecast times. A weather prediction is initialised more often to be closer to reality, while longer time scales are less often initialised because of computational power.

Some of these mentioned points are also reasons for the different model performances of different institutes. As described the model resolution is mainly higher in shorter forecast times, but depending on the institutional computational power the aim is to have the highest resolution possible in each time scale. The representation of the dynamics and the coupling of different systems can vary from each research institute and lead to different forecast performances. The data assimilation is an important part of dynamical numerical prediction and with that the initial conditions. The atmosphere is a chaotic system and slightly different starting conditions can result in different outcomes (butterfly effect, Lorenz, 1963). This is mainly regulated by using an ensemble forecast, meaning the model runs multiple times with slightly different starting conditions to cover a wider range of possible realities. The number of used model members can also influence the final forecast.

The following chapter presents different research centres that work with their own seasonal forecast model and develop it further. Tab 2.1 summarises the individual technical key facts of the described forecast suites at the end but the following literature overview focuses on

the models' output performance rather than the technical details.

2.2.1 UK Met Office - United Kingdom

The UKMO has a long history with seasonal forecasts. At the current state, the UKMO is operating version 6 of their seasonal forecast model started in February 2021 (Copernicus Climate Change Service, 2023b), the Global Seasonal Forecasting System version 6 (Copernicus Climate Change Service, 2023a; Williams et al., 2018). The precursors were Met Office Global Seasonal forecasting system version 5 (GloSea5), operational from January 2012 to January 2021, and Met Office Global Seasonal forecasting system version 4 (GloSea4), operational from September 2009 to December 2011 (MacLachlan et al., 2015). GloSea5 is the mainly used seasonal forecast model in this study. More details on how the model is used in this thesis can be found in chapter 3, Data.

As Met Office Global Seasonal forecasting system version 6 (GloSea6) data is just recently published, not many studies regarding its forecast skill are available yet. More studies about GloSea5 or comparisons to GloSea4 are published (e.g., MacLachlan et al., 2015). The UKMO publishes frequent studies about specific winter or summer seasons and how well their prediction was for these seasons (e.g., Dunstone et al., 2018a; Scaife et al., 2017a). So is winter 2015/16 interesting because it was a very wet winter over the UK (Scaife et al., 2017a). They show that the model, GloSea5, represented the pressure anomalies in December very well, as well as the strong NAO phase, which all led to extreme precipitations over the UK. The opposite of this was the following winter season, 2016/17, which was a specific dry winter for the UK (Dunstone et al., 2018a). This season was less well captured by GloSea5, as the model shows a skill gap of Mean Sea–Level Pressure (MSLP) near central Europe, and the important high pressure over the North Sea only emerged end of November.

The UKMO is a well-established research centre and is also included in the C3S of the ECMWF. This data set is used widely for multi-model studies. The publications of Lledó

et al. (2020), Portal et al. (2022), and Lee et al. (2020) are examples of using this data set to investigate different variables and will appear more often in the following parts of this chapter. A general study about European large-scale patterns in different seasonal forecast suites is done by Lledó et al. (2020). They investigate NAO, East Atlantic Pattern (EA), Scandinavian Pattern (SCA) and East Atlantic Western Russia Pattern (EAWR) in all four seasons and different lead times to the initialisation of the models. The UKMO model, here GloSea5, performs very well with significant correlation skill and Ranked Probability Skill Score (RPSS) for NAO and SCA. It performs best in winter, spring and summer, as all tested models. None of the models shows a skilful forecast for EAWR, and GloSea5 has difficulties with predicting winter EA as well. Another study investigated the stratosphere and analysed how different seasonal forecast suites predict the stratospheric polar vortex (SPV) (Portal et al., 2022). GloSea5 is one of the best-performing models and shows a significant skill for the polar vortex. They also investigated further into reasoning and showed that UKMO overestimates the connection between the stratosphere and Quasi-Biennial-Oscillation (QBO). Another seasonal case study is from Lee et al. (2020). They study the winter season 2019/20 because of a strong SPV in regards to a strong positive Atlantic Oscillation (AO). GloSea5 is the most consistent with the AO prediction and has a high skill for the polar vortex prediction, especially for January to March 2020 (Lee et al., 2020).

2.2.2 European Center for Medium Weather Forecast - Europe

The ECMWF model is within its fifth version, ECMWF's seasonal forecasting system 5 (SEAS5) and has been operational since November 2017 (Johnson et al., 2019). This version has a higher resolution and includes a sea-ice model compared to its precursor, ECMWF's seasonal forecasting system 4 (SEAS4) (Johnson et al., 2019). First predictability investigations in the first SEAS5 publication from Johnson et al. (2019) show an improvement in forecast skill of ENSO, the temperature over the tropical Pacific, the north Pacific subtropic

high and winter troughs over the UK. On the other hand, some variables did not improve or even decrease performance in comparison between SEAS5 and SEAS4, like the skill of the NAO index. Other studies also investigate the difference between SEAS5 and SEAS4 and summarise SEAS5 "outperform" SEAS4 because the temperature forecast skill is almost perfect, especially for stakeholder important regions, both temperature and precipitation are well predicted in all four seasons, but precipitation has poor reliability (Manzanas et al., 2022). SEAS5 has high forecast skills for all large-scale patterns important for Europe, NAO, AO, EA, SCA (Lledó et al., 2020; Portal et al., 2022) and for the SPV, as most of the other models as well (Lee et al., 2020). Only the connection of the SPV to the QBO is underestimated for the ECMWF model (Portal et al., 2022).

The seasonal forecast model series, SEAS, is the official and operational model of the ECMWF, but Weisheimer et al. (2020) also created a longer seasonal hindcast simulation to investigate the importance of the time period of the ensemble on predictability. CSF-20C is based on the ECMWF reanalysis CERA-20C and includes over 100 years, 1901-2010. This data set shows that even a longer hindcast agrees with the skill of current seasonal forecast models with the current hindcast length of 20-30 years (Weisheimer et al., 2020).

2.2.3 Météo France - France

Mètèo France (MF) is working on seasonal forecasts since the late 1990s and is operating their System 8 since 2021 (Batté et al., 2021). The model runs with the atmospheric ARPEGE-IFS model developed by MF in collaboration with the ECMWF. This atmospheric model was specially designed for climate projections and seasonal predictions (Déqué et al., 1994). The MF model is not much investigated (by the knowledge of the author) but included in the C3S data set and hence, included in some multi-model studies. Lledó et al. (2020) found no skill in the MF model for either EA or EAWR and Lee et al. (2020) for AO. However, on a lead time of 0, the winter NAO forecast is skilful (Lledó et al., 2020). Portal et al. (2022)

and Lee et al. (2020) agree that the MF model does not capture the polar vortex, either in general or in the winter season of 2019/20.

2.2.4 German Weather Service - Germany

The German Weather Service (DWD) uses their German Climate Forecast System (GCFS) for seasonal predictions. The GCFS is developed with the Max-Planck Institute (MPI) and the University of Hamburg (Fröhlich et al., 2021). The MPI-part of the DWD model is investigated in more detail by Domeisen et al. (2015). They summarise that the model has a moderate prediction skill for ENSO up to 6 months ahead, but stratospheric sudden warming (SSW) are not predictable after 2 weeks. As these two factors, ENSO and SSW, are very important predictors for predictability, years with a strong El Niño and SSW events are better predicted over Europe (Domeisen et al., 2015). The DWD model is among the forecast suites that predict the three most dominant large-scale patterns over Europe well, NAO, EA and SCA (Lledó et al., 2020). However, the magnitude of AO was underestimated in winter 2019/20, but the extreme state of the polar vortex was captured in this year (Lee et al., 2020). The model shows, in general, a significant seasonal forecast skill for the polar vortex (Portal et al., 2022). The winter forecast skill (Anomaly correlation coefficient (ACC)) of the NAO index has improved from 0.2 in GCFS1 to 0.4 in GCFS2 (Fröhlich et al., 2021).

2.2.5 Euro-Mediterranean Center for Climate Change - Italy

The Italian research institute Euro-Mediterranean Center for Climate Change (CMCC) is already with its third version of a season forecast model, which was updated with slight changes to CMCC SPS3.5 (Seasonal Prediction System, Version 3.5) in October 2020 (Gualdi et al., 2020). A detailed investigation regarding the CMCC model is mainly about the precursor SPS3 by Sanna et al. (2016). They summarise the updates to previous model versions, like an improved resolution in ocean and atmosphere, better representation of

ocean, sea ice, and stratosphere and an updated dynamical core of the model by an updated Earth System Model. They also summarise the first improvements in seasonal forecast skill, like a persistent high skill for ENSO after initialisation or a NAO correlation skill of 0.5. Variables like 2m temperature have not improved significantly compared to previous model versions (Sanna et al., 2016).

The already mentioned studies that include the C3S data also include the CMCC model. Lledó et al. (2020) investigates the four main large-scale patterns over Europe for the core winter, DJF, with model initialisation in November and found that CMCC has difficulties predicting the EAWR or EA pattern. The study about the SPV summarises that CMCC is one of three models that have significant forecast skill (Portal et al., 2022). Lee et al. (2020) investigated as well a large-scale pattern over Europe, AO, in connection to the SPV in winter 2019/20. The CMCC model underestimated the AO magnitude but captured the extreme state of the polar vortex for this particular winter season.

2.2.6 National Centers for Environmental Prediction - United States of America

National Centers for Environmental Prediction (NCEP), known for its reanalysis products, also developed its seasonal forecast model, which was implemented in 2004 and called CFS version 1 (Saha et al., 2010). The model is based on the same global model as the know NCEP reanalysis and follows its updates (Saha et al., 2006). The updated version 2, operational since March 2011, runs 4 times every 5 days for 9 months and even daily for a sub-seasonal resolution (see Tab. 2.1, Copernicus Climate Change Service, 2023b; Saha et al., 2014). The CFSv1 model was the first dynamical model in the United States that reached the same forecast skill for ENSO as previously used statistical seasonal prediction models (Saha et al., 2010). The study of Saha et al. (2010) investigates some improvements due to the update of the global model. This shows the Nino3.4 Sea Surface Temperature (SST) is better predicted

in the new model than in previous statistical models, but the temperature and precipitation of the Northern Hemisphere are only equal to these statistical models. The temperature results in a slightly higher skill in summer, probably because of a good representation of soil moisture flux and the precipitation skill increases in winter because of a closer relationship to ENSO (Saha et al., 2010). They also suggest that the updated seasonal model could be of great use for a seasonal hurricane prediction which is of great interest, especially for the US. NCEP has the model version 2 still in operation and no further updates, to the author's knowledge.

Besides the model's initial publication, other studies used the NCEP seasonal forecast model and investigated the forecast skill. Like Kim et al. (2012), they compared the CFSv2 with the UKMO model SEAS4 (precursor of SEAS5). They summarised that both models have difficulties capturing the winter temperature over the US and Europe. The latter might be connected to the low forecast skill in NAO which is only 0.16 (correlation) for SEAS4 and 0.25 (correlation) for CFSv2. Lee et al. (2020) included the NCEP model for the first part of their multi-model study and found the right signal and ensemble spread of the NCEP model for the winter 2019/20 AO forecast.

2.2.7 Environmental and Climate Change Canada - Canada

The Canadian organisation Environmental and Climate Change Canada (ECCC) has been using seasonal forecasts since 1995, but just in 2011, they developed their model system called "Canadian Seasonal to Inter-annual Prediction System" (CanSIPS, Merryfield et al., 2013). CanSIPS is developed by the Canadian Centre for Climate Modelling and Analysis (CCCma) and the Numerical Research Division (RPN-Dorval). By changing one of the two coupled ocean models to the newer GEM-NEMO, they doubled the model ensemble members to 20 runs (Copernicus Climate Change Service, 2023b). As mentioned in chapter 2.1, the focus of many seasonal forecast suites is the prediction of ENSO. Hence, also for the

ECCC model, an improvement in ENSO forecast is seen from one model version to another (Merryfield et al., 2013). Other studies, like Kharin et al. (2009), compared different model versions of either CCCma or RPN and found that all of them have a high seasonal forecast skill for temperature, especially with zero lead time, in winter and early summer, and the improvements are mainly seen over the ocean.

2.2.8 Japan Meteorological Agency - Japan

The Japan Meteorological Agency (JMA) has an older seasonal model in operation. Their system 2 was developed already in 2015 (Takaya et al., 2018). Version 2 greatly improved compared to version 1, especially in ENSO forecast skill, because of improved resolution and better physics for atmosphere, ocean and sea ice. More improvements between version 1 and 2 are described in Takaya et al. (2018), which are, e.g., a higher ACC and lower Root-Mean-Squared Error (RMSE) for NINO3 SST forecasts. The overall seasonal SST, 2m temperature and precipitation are better predicted. The winter NAO forecast skill (ACC) of this seasonal forecast model is roughly 0.2 but increases to 0.65 for the spring season (Saito et al., 2017). Since 2015 their seasonal model had no update, and not many studies about the seasonal performance of the JMA seasonal model are available.

Table 2.1 Summary of all available seasonal forecast suites within C3S.

	GloSea5/ GloSea6	SEAS5	Météo France	DWD	CMCC	NCEP	ECCC	JMA
Horizontal resolution of the atmospheric model	N216	L319	TL359	T127	0.5°	T128	T63	TL159
Number of atmospheric model levels	85	91	137	95	46	64	35	60
Top of the model	85 km	0.01 hPa	0.01 hPa	0.1 hPa	0.3 hPa	0.02 hPa	1 hPa	0.1 hPa
Forecast length	215 days	215 day	7 months	6 months	6 months	215 days	214 days	215 days
Run frequency forecast	Each day	1 per month	1 per month	1 per month	1 per month	Each day	1 per month	Each day
Number of perturbed ensemble members	2 per day	51	51	50	50	4 per day	20	51
Coupled (ocean/sea ice)	Yes/ Yes	Yes/ Yes	Yes/ Yes	Yes/ No	Yes/ Yes	Yes/ Yes	Yes/ No	Yes/ Yes
Hindcast size	7 per run	25	25	30	40	4 per run	20	10
Run frequency hindcast	4 per month	1 per month	1 per month	1 per month	1 per month	6-7 per month	1 per month	2 per month
Hindcast production	On-the-fly	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
Hindcast period	1993-2016	1981-2016	1993-2018	1990-2019	1993-2016	1981-2010	1981-2010	1979-2015

2.3 Seasonal Prediction Skill

The previous chapters introduced the history of seasonal forecast models (chapter 2.1) and the current state-of-the-art available seasonal models (chapter 2.2). The following chapter will summarise the existing studies about seasonal prediction skills in three groups, general meteorological parameters like temperature or precipitation (chapter 2.3.1), large-scale patterns like NAO or ENSO (chapter 2.3.2) and extreme events like storms (chapter 2.3.3).

2.3.1 Predictability of Meteorological Variables

Temperature

Dynamical models (as described in chapter 2.2) are assessed for their performance in predicting parameters like temperature on a seasonal scale. Clark et al. (2017) investigated seasonal forecasts of temperature and wind speed in GloSea5 and how reliable they are for the energy sector. They summarise that the direct forecast of these parameters shows a reliable skill, but also the NAO can be used to predict the required energy parameters. Multi-model studies find skill for temperatures (Hemri et al., 2020; Mishra et al., 2019). The latter only find limited skill but with stronger predictions in summer. A robustness test of the discovered forecast skill made with the "Signal-to-noise" Paradox (see chapter 2.5) shows that the temperature skill is reliable (Eade et al., 2014).

Not only numerical dynamical ensemble models are used to investigate temperature on seasonal scales. Previous studies have also used long observation series to analyse temperatures on a monthly or longer time (e.g., Toreti et al., 2010; Degenhardt & Ólafsson, 2019). Both studies are regionally restricted, Toreti et al. (2010) for Italy and Degenhardt and Ólafsson (2019) for Iceland.

Precipitation

Precipitation is another meteorological factor that is of great interest. Not only because it is generally difficult to predict, but in the context of this study, tropical precipitation can be a potential cause for predictability (Scaife et al., 2017b). This tropical precipitation forecast is strongly skilful on a seasonal scale, especially for the Pacific region (Scaife et al., 2017b). But also a wet bias has been found, especially for the Indian Ocean and Western Pacific (Scaife et al., 2019b). Extratropical precipitation is more difficult to forecast than tropical (Kim et al., 2012). Mishra et al. (2019) find seasonal skill for different models over different European regions, like the Météo France model is skilful over the UK, Scandinavia and Italy in winter, SEAS4 and GloSea5 over eastern Europe. The summer precipitation has been investigated, e.g., in decadal forecast suites and summarised as skilful (Dunstone et al., 2018b).

Other Parameter

MSLP is also very important, especially for the extra-tropical weather. Hence, models that predict the seasonally averaged MSLP well, can be useful. Studies like Scaife et al. (2017a) investigated whole winter seasons, incl. the general pressure pattern, and found a good representation for December. The prediction of this parameter is as well as temperature and precipitation reliable (Eade et al., 2014). The pressure pattern is related to wind speeds, which are skilful and reliable in a seasonal forecast model, like GloSea5, but tend to be underconfident (Clark et al., 2017). Other winter seasons, 2009/10, are used for case studies in the stratosphere and found skilful seasonal forecasts of SSW (Fereday et al., 2012).

2.3.2 Predictability of large-scale pattern

Large-scale patterns or teleconnection patterns are the main drivers of the atmosphere. Hence, a skilful prediction of these is the key to a skilful prediction in general. As described

in chapter 2.1, ENSO was one of the first phenomenons that triggered the investigation in the seasonal time scale. But other large-scale patterns, like NAO, are important and required to be skilful in a seasonal forecast suite to ensure skilful forecasts of the extra-tropics, Europe (Smith et al., 2020).

ENSO

One main focus in each research centre is the right prediction of ENSO or an improvement of the previously existing forecast of ENSO in new model versions. This is documented for the UKMO model GloSea5 (Scaife et al., 2014) and, for example, the DWD model (Domeisen et al., 2015). Reaching a confident forecast skill for a specific factor can also be used to investigate changes in this skill in other settings. Scaife et al. (2019a) used the known ENSO forecast skill and changed the models' resolution to investigate its influence. They show that both resolutions (higher and lower) show a similar ENSO forecast skill, and both are still overconfident with the amplitude of ENSO. Jin et al. (2008) shows that a multi-model approach of ten different coupled general circulation models already has a NINO3.4 SST skill of 0.86 (anomaly correlation coefficient) but still highly depends on the season, ENSO phase and strength. With further model improvements, like higher resolutions or fully coupled ocean-atmosphere systems, the ENSO seasons of 2002 to 2011 were highly predictable (Barnston et al., 2012). Even the prediction of the different individual ENSO types has been investigated (Ren et al., 2019), but have not been very well predicted after a month lead time.

NAO

The AO and NAO are the most dominant large-scale patterns for the North Hemispheric extratropics. Sometimes the forecast skill is investigated together as they are strongly linked (Ambaum et al., 2001). Both NAO and AO are mainly well predicted in seasonal forecast ensemble models (Athanasiadis et al., 2017; Nie et al., 2019b; Tang et al., 2007). This well-

predicted parameter can be used to investigate more questions, like what the forecast skill depends on. Athanasiadis et al. (2017) summarise with the AO/NAO skill, that forecast skill depends on hindcast period and ensemble size. Nie et al. (2019b) connected the AO/NAO prediction to the stratosphere and found an influence from upper stratospheric zonal winds on the AO/NAO prediction. But also, the reliability of the forecast skill is still under investigation (Tang et al., 2007). Some studies investigate the NAO individually. A previous study with GloSea4 and ERA-interim find a forecast skill correlation of 0.62 for winter NAO (Scaife et al., 2014), which is consistent in GloSea5 (Walz et al., 2018b). An ECMWF model had only an ACC of 0.31 for over 110 winters (Weisheimer et al., 2017). Baker et al. (2018) and Eade et al. (2014) found that the NAO forecast is underconfident in its prediction, which means the unpredictable noise is too high in the model. These skill values also depend on the different available definitions of the NAO index (see Baker et al., 2018; Lledó et al., 2022). The seasonal forecast model, SEAS4, has a NAO forecast skill of 0.32 (as correlation) if calculated grid-cell-based with Lisbon and Reykjavik and 0.49 (correlation) if the Empirical Orthogonal Function (EOF) definition is used (Walz et al., 2018b).

Other large-scale patterns

Some studies investigated more than the first EOF representing NAO, hence, EA, SCA and EAWR (Lledó et al., 2022; Lledó et al., 2020). They used a multi-model approach and summarised that mostly winter, spring and summer are well-predicted seasons, but mainly for NAO, EA and SCA. Another study also used the EOF approach to calculate the large-scale patterns but focused on the EOF patterns, which are different for the reanalysis and the used seasonal model, SEAS4 (Walz et al., 2018b).

2.3.3 Predictability of extreme events

A very early study by Renggli et al. (2011) investigated the seasonal forecasts of extreme events, like windstorms. They used the DEMETER and ENSEMBLE project data sets to investigate the seasonal forecast skill of storm track densities. Already different skill scores were used, like RPSS or Relative Operating Characteristics (ROC) (see chapter 4.3.1). The storm frequency per season was well represented. Hence, a significant skill score was only found for this storm parameter and not intensities (Renggli et al., 2011). However, a slight overestimation and south-westward shift of the model tracks were found. Recent studies used the mentioned seasonal forecast suites from chapter 2.2. While Befort et al. (2019) investigated the forecast skill of cyclone and windstorm frequencies in different models directly, other publications focused more on the influence of different factors on seasonal windstorm predictions (Walz et al., 2018a; Walz et al., 2018b; Hansen et al., 2019). The direct forecast skill of cyclones and windstorms shows that models like GloSea5 overestimate the numbers of all tracked cyclones and shows the highest skill for windstorm frequency over Europe. Not only tracked events but also the dynamical components for a rare extreme event, cyclone clustering, are well forecasted on a seasonal scale (Priestley et al., 2018). Factors like the 10 dominant EOFs of geopotential height over Europe or sea ice have been used to find the dominant drivers for windstorms over Europe (Walz et al., 2018a). They created a so-called "map of drivers" which shows that SCA is the dominant driver for windstorms over a majority of Europe with an explanatory variance of over 40% (see Walz et al., 2018a, their Fig. 3a & b). But also, the stratosphere influences the surface windstorms over Europe. Hansen et al. (2019) found a perfect forecast of the stratosphere can improve the seasonal forecast of severe storms from Greenland over the British Isles to Northern Europe.

Other studies have investigated different meteorological extreme events, like heatwaves, wildfires or flooding, on slightly different time scales, like sub-seasonal (Domeisen et al., 2022; Marshall et al., 2022) or decadal (Moulds et al., 2023) scales. Domeisen et al. (2022) sum-

marise the predictability of individual case studies of multiple extreme events, like heatwaves, cold spells, extreme precipitation and tropical and extratropical cyclones. They find a good forecast of these extreme events on their time scale roughly 2 to 3 weeks in advance. Marshall et al. (2022) investigates more regionally based on Australian wildfires and their connection to large-scale patterns, like the Indian Ocean Dipole (IOD), South Annual Mode, ENSO, Madden-Julian Oscillation (MJO). Moulds et al. (2023) used a large model ensemble of CMIP5 and CMIP6 models to investigate the flooding predictions for UK. They find more skill in flood predictions over the UK using a member-sub-sample with well-represented NAO. These studies have not been summarised further, as sub-seasonal and decadal data sets differ from seasonal forecast data. Still, the author wanted to mention at least other extreme event studies close to seasonal predictions.

2.4 Drivers/Factors for extreme events

The second focus of this thesis is the investigation of dynamical factors that influence windstorms and windstorm prediction skills in seasonal forecasts. Previous studies have already analysed large-scale patterns or factors influencing cyclones or windstorms (see Befort et al., 2019; Walz et al., 2018b; Walz et al., 2018a). This is a general literature overview of influencing factors for cyclones or windstorms. A more detailed description of the used factors can be found in the introduction of chapter 6.1.

Pinto et al. (2008) investigated (extreme) cyclones in connection to the NAO and summarized that extreme cyclones have higher intensities (deeper cores) and longer lifetimes (longer track lengths) in a strongly positive NAO phase. This agrees with findings for windstorms, as a strong positive NAO occurs on 20% of storm days (Donat et al., 2010). The general weather type can be of influence for cyclones and windstorms (Donat et al., 2010), as 80% of storms appear in a westerly flow over Europe. But Donat et al. (2010) investigated that the westerly flow has a higher impact on the location of the storm, and hence the frequency and north-

westerly flow influence the intensity of storms more. But not only the NAO is of importance for European weather. The NAO is the dominant driver over the northern North Atlantic and northern UK, and SCA is the main reason for the windstorm variability over Central Europe, including southern England and the Northern Atlantic (Walz et al., 2018a). Besides the large-scale patterns and weather types, dynamical factors are important for the development and process of storms. Pinto et al. (2008) investigated a summary of factors, like the jet stream in 250hPa (jet_{250}), Eady Growth Rate (EGR) in 400hPa (σ_{400}), Divergence in 250hPa (div_{250}) and equivalent potential temperature (Θ_e), which influence the growth of extreme cyclones. These factors are investigated on a regional scale (jet_{250} & σ_{400}) or with a Lagrangian view (div_{250} & Θ_e). Another factor that might influence storms, specifically cyclone clustering, is the Rossby Wave Breaking (RWB) (Pinto et al., 2014; Priestley et al., 2017a; Karremann & Pinto, 2020). Studies showed that an extended eddy-driven jet towards Western Europe for at least one-week favours clustering of extreme cyclones (Pinto et al., 2014). Priestley et al. (2017a) describes in more detail that this extended jet stays in the position because of the RWB north and south of it, especially around 55°N. If the RWB is stronger in the south of the jet, the extreme cyclone clustering happens further north (around 65°N), and a stronger RWB north of the jet causes the clustering at around 45°N (Priestley et al., 2017a). This has also been shown in a case study for the winter season 2013/14, which was a very intense season regarding cyclone clustering (Priestley et al., 2017b). The influence of cyclones and windstorms is not restricted to the troposphere. Hansen et al. (2019) found that seasons with SSW are seasons with less severe windstorms and vice versa. They connected this link between windstorms and SSW to seasonal forecast predictions and find an improvement in windstorm prediction over Greenland and the British Isles, when the stratosphere is well predicted. Some of these mentioned factors above are also shown in individual case studies. For example, the explosive intensification of Storm "*Kyrill*" in January 2007 was caused by a very strong and zonal jet (Fink et al., 2009). Additional factors like dry air intrusions over the occlusion, additional cold fronts and a strong pre-existing MSLP gradient over Europe

kept Storm "*Kyrill*" strong until Eastern Europe and increased the total damage (Fink et al., 2009).

2.5 Signal-to-noise paradox

A very interesting phenomenon in long-range model forecast data was discovered by Eade et al. (2014) and further investigated by Scaife and Smith (2018). They found the so-called "signal-to-noise paradox". This paradox shows with statistical measures (see chapter 4.3.1) that the models' ensemble mean predicts the real world better than their individual model members. This means the model would be limited and not reach the potential forecast skill. Scaife and Smith (2018) analysed the NAO and AO in the seasonal forecast model GloSea5. They show that, especially over the North Atlantic, the correlation between the model mean, and observations is higher than between the model mean and one individual member of the model itself. That means the Ratio of Predictable Components (RPC), as a measure for the signal-to-noise paradox, is much higher than 1 for the NAO. Their summary shows that not every ensemble member is an equivalent representation of the real world. The Scaife and Smith (2018) paper is based on the investigation from Eade et al. (2014). They established the RPC equation (Eq. 4.12) and found that the RPC is not significantly different to 1 (perfect) for 70% of grid points in GloSea5 and 57% in Met Office decadal prediction system (DePreSys) (year 2-5) for surface temperature, MSLP and precipitation. Other papers found some discrepancy in the skill/RPC as well, e.g., Dunstone et al. (2018b) for European summer precipitation, Jin (2022) in the ENSO prediction and Scaife et al. (2019b) for tropical precipitation. The latter even find the signal-to-noise paradox stronger in the tropics, especially the Indian Ocean, than in the extratropics. Sévellec and Drijfhout (2019) investigated a regional difference in the signal-to-noise paradox and found that over 70% of the grid cells with a high RPC for surface temperature are located over the ocean. Decadal data sets have been tested, and the signal-to-noise paradox has been found (e.g., Zhang

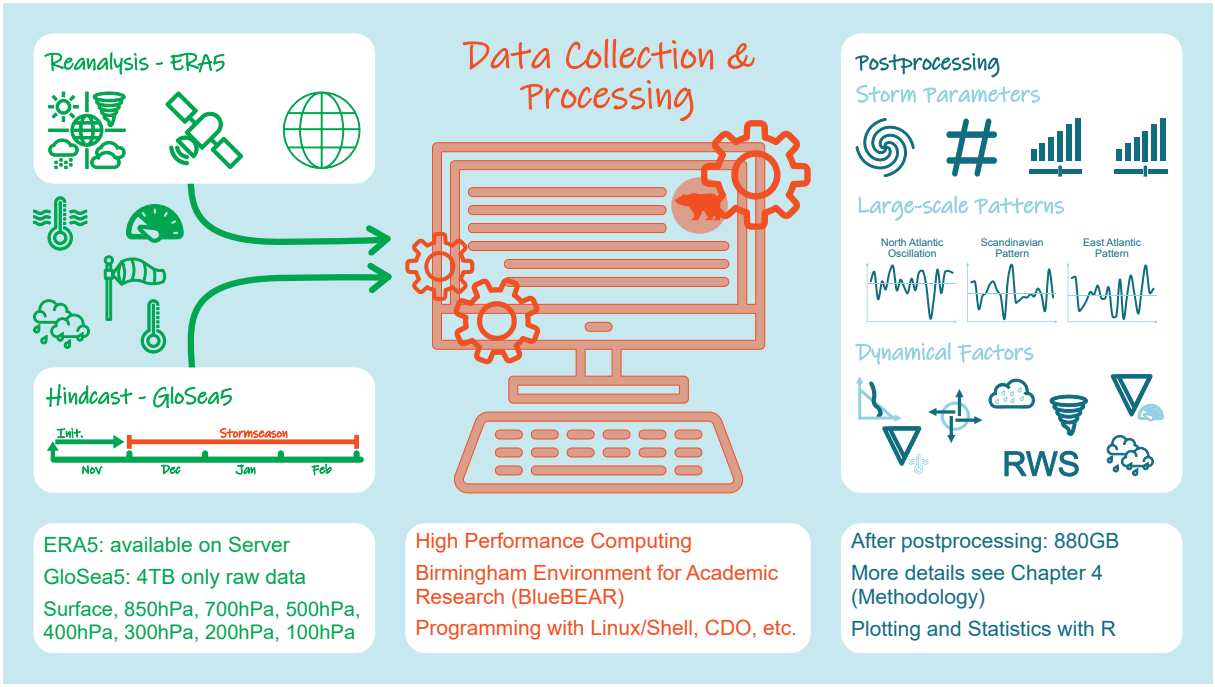
et al., 2021b).

Another study about the signal-to-noise paradox is published by Weisheimer et al. (2019). They compared the RPC method from Scaife and Smith (2018) with other statistical measures. They used a perfect model approach and agreed that a perfect model would have a RPC equals 1, but also a RMSE that equals the spread of the model mean ($\sigma_{ensmean}$). Hence, Weisheimer et al. (2019) analysed the signal-to-noise paradox with the relation between RMSE and $\sigma_{ensmean}$.

One way of studying the signal-to-noise paradox is to investigate its existence in different data sets, different parameters or different time scales. Another part is to investigate the cause or solve the problem. Multiple hypotheses exist about what could cause or even increase the signal-to-noise paradox. For example, Parker et al. (2019) mention that the link between jet location and NAO index might not be completely right in the model. Scaife et al. (2019b) suggests that improving the tropical-extratropical interaction might improve (increase) the seasonal forecast skill and (decrease) the RPC. Others suggest that the signal-to-noise paradox is influenced by the ENSO growth rate (Jin, 2022), by too weak eddy feedback (Hardiman et al., 2022), by mesoscale ocean features (Zhang et al., 2021b) or by the initial ensemble spread relative to observational spread (Mayer et al., 2021). The latter has investigated the signal-to-noise paradox in a theoretical Lorenz model (Lorenz, 1963) and also discovered that the signal-to-noise paradox is not appearing in the model when the ensemble spread equals the observational spread. This agrees with other investigations that conclude the signal-to-noise paradox is more a general model issue than, e.g., an initialisation problem (Zhang & Kirtman, 2019; Zhang et al., 2021b). However, multiple studies have investigated the signal-to-noise paradox, but it is not fully understood or solved yet.

Chapter Three

Data



The following chapter summarises the two core data sets used for this study and compares them to previous versions. The reference data is the latest reanalysis of the ECMWF, Reanalysis Data Set ERA5 (ERA5), for all predictability studies and used as a possible representation of reality (Chapter 3.1). The investigation aims to understand the seasonal forecast of severe winter windstorms. This study is focused on the seasonal forecast model of the UKMO, GloSea5 (Chapter 3.2).

3.1 Reanalysis Data Set - ERA5

ERA5 is the latest observational reanalysis of the ECMWF. It is an update of the reanalysis ERA-interim (Dee et al., 2011). Hersbach et al. (2019) shows the differences between these two data sets. The main improvements which could lead to an even more accurate representation of the real-world state are:

- a higher spatial resolution (0.7 degree - ERAint vs 0.25 degree - ERA5)
- new and more observations

These three improvements might lead to a better resolution of real-world storms in the reference data set. This means the reanalysis represents reality better. Hence, a skilful prediction would be more meaningful because it is compared to a close-to-reality data set. A more detailed explanation of ERA5 is given in Hersbach et al. (2019). ERA5 is available from 1940 to the present on the C3S server. The spatial resolution is a T639 grid, which resolves in 0.25°x0.25°. First publications compare ERA5 to ERAint or other observational data sets to understand the new representation of ERA5, e.g. Hersbach et al. (2019), Olauson (2018), and Hoffmann et al. (2019). Hersbach et al. (2019) summarises that the higher resolution can represent smaller patterns of, e.g., precipitation or vertical temperature profile better than other data sets. The study from Olauson (2018) compares ERA5 to another commonly used reanalysis, MERRA-2, for predicting wind energy parameters. They summarise that ERA5 performs in multiple measures better than the other data set, with more skill and fewer errors for wind power generations.

3.2 Seasonal Forecast Systems - UKMO GloSea5

The seasonal forecast model, which is used in this study, is GloSea5 from the UKMO based on the HadGEM3. GloSea5 replaced its precursor GloSea4 in January 2013 (MacLachlan et al., 2015). Same as described between ERAint and ERA5, GloSea5 was an enhancement of

the before used GloSea4. The higher spatial resolution of the model and a higher resolution of the orography inside the model leads to an increase of the forecast skill in GloSea5 for parameters like SST or surface temperature (MacLachlan et al., 2015). The GloSea5 data set is explained in MacLachlan et al. (2015). They also show a distribution map of tropical storms over the Atlantic and see a well-represented pattern in GloSea5, but the amount of storms is still too low in the model mean. Many publications investigated the predictability skill in GloSea5 for different parameters, e.g. Scaife et al. (2014), Befort et al. (2019), and Scaife et al. (2019a). While MacLachlan et al. (2015) summaries general parameters like the SST or MSLP, studies like Scaife et al. (2014) show a high forecast skill in the large-scale pattern NAO in GloSea5. Storms (extreme cyclones and windstorms) have been investigated by Befort et al. (2019). They find high forecast skill of windstorm frequency south of the British Isles in GloSea5. But in general, all publications which compare different data sets with GloSea5 show a higher predictability skill in GloSea5 for different parameters. As one example, Scaife et al. (2014) found more skill in ENSO and NAO forecast for GloSea5 than in GloSea4. UKMO released a new version of their seasonal forecast model, GloSea6. This is now the operational seasonal forecast model of the UKMO. It is based on the latest version of the HadGEM3. This study continued with investigating GloSea5, even after GloSea6 was released, to be consistent throughout the study.

GloSea5 hindcasts are produced on-the-fly, meaning close to the real-time forecast (Copernicus Climate Change Service, 2023b). Four initialisation dates exist per month (1^{st}, 9^{th}, 17^{th} & 25^{th}). As this study is investigating the core winter, DJF, the three initialisations around the 1^{st} of November are used (25^{th} October, 1^{st} & 9^{th} November). These initialisation are used to consider the model noise right after initialisation. The first few weeks of a seasonal forecast model are more influenced by the initial conditions, afterwards by the boundary forcing (Kharin et al., 2001; Kharin & Zwiers, 2001). Both Kharin et al. (2001) and Kharin and Zwiers (2001) show that the later time period, after 15 days, shows higher forecast skill. The skill is also highly sensitive to the number of ensemble members (Kharin et al., 2001).

They find higher skill with up to 24 members. Hence, the data set used in this study will be used with 63 members and with a lead time of 2 to 4 months. GloSea5 is available from 1993 to 2016 on the C3S server with a spatial resolution of 1°x1° but originally is produced on an N216 grid (0.83°x0.55°).

3.3 Data availability, use and capacity

All data are stored, and the calculations are run on the HPC of the University of Birmingham, Birmingham Environment for Academic Research (BlueBEAR). GloSea5 data is available on the C3S server (Copernicus Climate Change Service, 2023c) hosted by the ECMWF. The used GloSea5 data is roughly 4 TB of raw data.

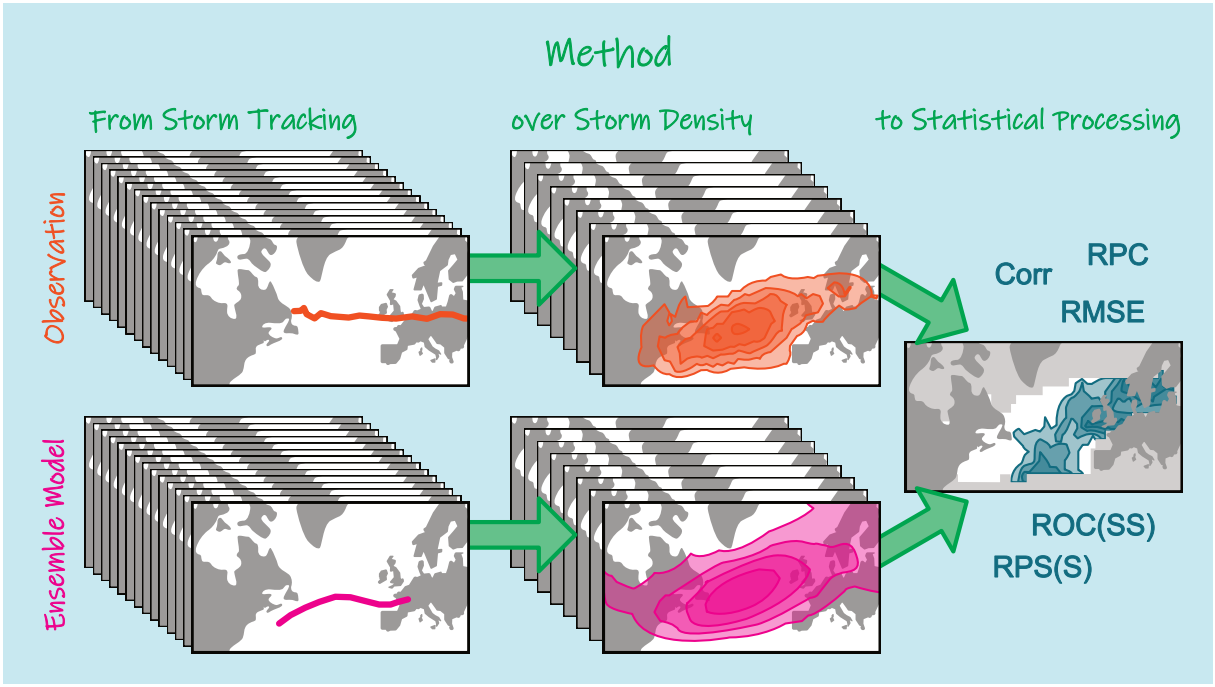
After the first processing step, tracking windstorms in ERA5 and GloSea5, storm data of around 880 GB were available on BlueBEAR. The used and downloaded different variables on respective pressure levels for the whole study can be found in Tab 3.1 and will be named in the respective chapters.

Table 3.1 Summary of used variables in ERA5 and GloSea5.

	ERA5 & GloSea5
single level	10m Wind Speed, MSLP, SST, Precipitation
multiple levels	Geopotential Height, Temperature, U- & V-Wind Component
	850hPa, 700hPa, 500hPa, 400hPa, 300hPa, 200hPa, 100hPa

Chapter Four

Applied and Tested Methods



The following chapter will introduce the general methods used in all individual studies. More result-chapter-related detailed methods can be found in the respective chapters. The first part will describe the algorithms used to track windstorms (chapter 4.1). The following chapter will explain the used Empirical Orthogonal Function analysis used for large-scale patterns (chapter 4.2). Chapter 4.3 will summarize the used statistics, different skill measures/scores, and different types of correlations.

4.1 Tracking

The work focuses on severe synoptic-scale European windstorms. Many different definitions of cyclones as storms are used in the scientific community (Neu et al., 2013), but this study investigates the 2% strongest wind events near the surface. This allows the tracking of impact-based extreme events, right at the target region (the surface) which is more relevant for end-users like insurance companies. These events will be identified for each time step in the different data sets by creating tracks and calculating a severity measure per storm.

4.1.1 Windstorm Tracking

The algorithm used in this work was developed by Leckebusch et al. (2008b). Other algorithms like the cyclone tracking from Murray and Simmonds (1991) are based on the pressure in the physical description of extra-tropical cyclones. The windstorm algorithm from Leckebusch et al. (2008b) focuses on the wind speed near the surface. Thus, the Severity Storm Index (SSI) from Leckebusch et al. (2008b) is a more impact-based scale. The SSI is calculated as:

$$SSI_{T,K} = \sum_{t}^{T} \sum_{k}^{K} \left[\left(max \left(0, \frac{v_{k,t}}{v_{Perc,k}} - 1 \right) \right)^3 \cdot A_k \right] \tag{4.1}$$

A local cubic exceedance of the near-surface wind speed (v) over a local percentile (v_{Perc}) is used to determine the severity of the storm. These exceedances are weighted with the area (A_k) and added up over every time step (T) and grid point (K).

The 10m wind is the recommended variable to use, as it represents the wind speed which causes damage on the surface. The 98% percentile for each grid point individually is calculated as the local threshold. This is a valid reference to loss-causing wind speeds (Klawa & Ulbrich, 2003). The first step of the algorithm from Leckebusch et al. (2008b) is locating areas with high wind speeds (top 2%). If an area exceeds a certain minimum threshold (see Tab. 4.1), the "centre of mass" will be set as windstorm track location. This detection of high wind fields is calculated per time step. A cohesive windstorm track is built with the

Table 4.1 Important settings for WTrack (windstorm algorithm) established from Leckebusch et al. (2008b)

Shortcut	Description	Value
minarea	Minimum coherent clustered area, the wind needs to exceed the 98% Percentile	130000 km^2
mindur	The minimum time steps a windstorm has to be tracked	8 (=2 days)
max_translatdist	The maximum distance a windstorm can travel between two timesteps	600 km
maxlength	The maximum length in time steps of a windstorm	120

nearest neighbour approach and a minimum overlapping region as criterion.

The output of this algorithm includes two results, which together describe the whole characteristic of the windstorm. One result is the track table which includes all information about the storm tracks, identification number, time step, location (lon, lat) of the centre of mass and the storm's severity at this time step. The second output is the footprint data which is a NetCDF file which includes storm information on a grid-cell level. This means each grid cell has the storm track identification number to show all the grid cells that this windstorm included or the grid-cell-individual storm's severity on this location at each time step of the storm. Both track table and footprint are used in this study to create storm parameters for the different studies.

4.1.2 Storm Frequency

The track table includes all individual tracks of a windstorm season and is used to create seasonal climatologies called storm frequency. This is further developed and explained in more detail by Kruschke et al. (2014) and Kruschke (2015). The storm frequency shows the number of storms per season (in this study, the main winter season - DJF) per grid cell. The calculation is re-gridding the study area (parameter res in Tab. 4.2) to reduce the noise

of neighbouring grid cells. Storms are counted per grid cell with an extended, pre-defined radius around each specific grid cell (parameter maxdist in Tab. 4.2) to cover the total impact area of the storm. The storm frequency results from the number of storms crossing this radius per grid cell. The settings for this study are shown in Tab. 4.2.

The following validation study assesses these settings and the basic data sets used in this study. The figures are presented and published at the European Geoscience Union General Assembly (EGU) conference 2020 (see Degenhardt et al., 2020). The validation study was used to investigate two core questions for the method.

Table 4.2 Settings for calculating storm frequency.

Shortcut	Description	Value
res	resolution, the new grid size	2.5°
maxdist	The maximum distance for counting events (radius)	700km

1. Which influence have different data set parameters, like different observational data sets, different tracking levels or different temporal resolutions on the seasonal forecast skill of windstorm frequency?
2. What influence does the storm frequency radius have on the seasonal forecast skill of severe winter windstorms?

The validation study investigates storm frequency prediction skill dependent on different data sets and storm frequency settings (Fig. 4.1). The following changes have been assessed:

- Change of the seasonal forecast model version (GloSea5 GA3 → GC2)
- Change of the tracking level (12h, 925hPa → 6h, 10m)
- Change of the storm count radius (700km → 1000km)
- Change in the observational reanalysis data (ERAint → ERA5)

The first change from GloSea5 version GA3 to GC2 indicates directly that the data set and even a slight model change can be very influential on the forecast skill. The GA3 forecast skill is mainly over the Biscay and Scandinavia, while the GC2 has the strongest skill over the

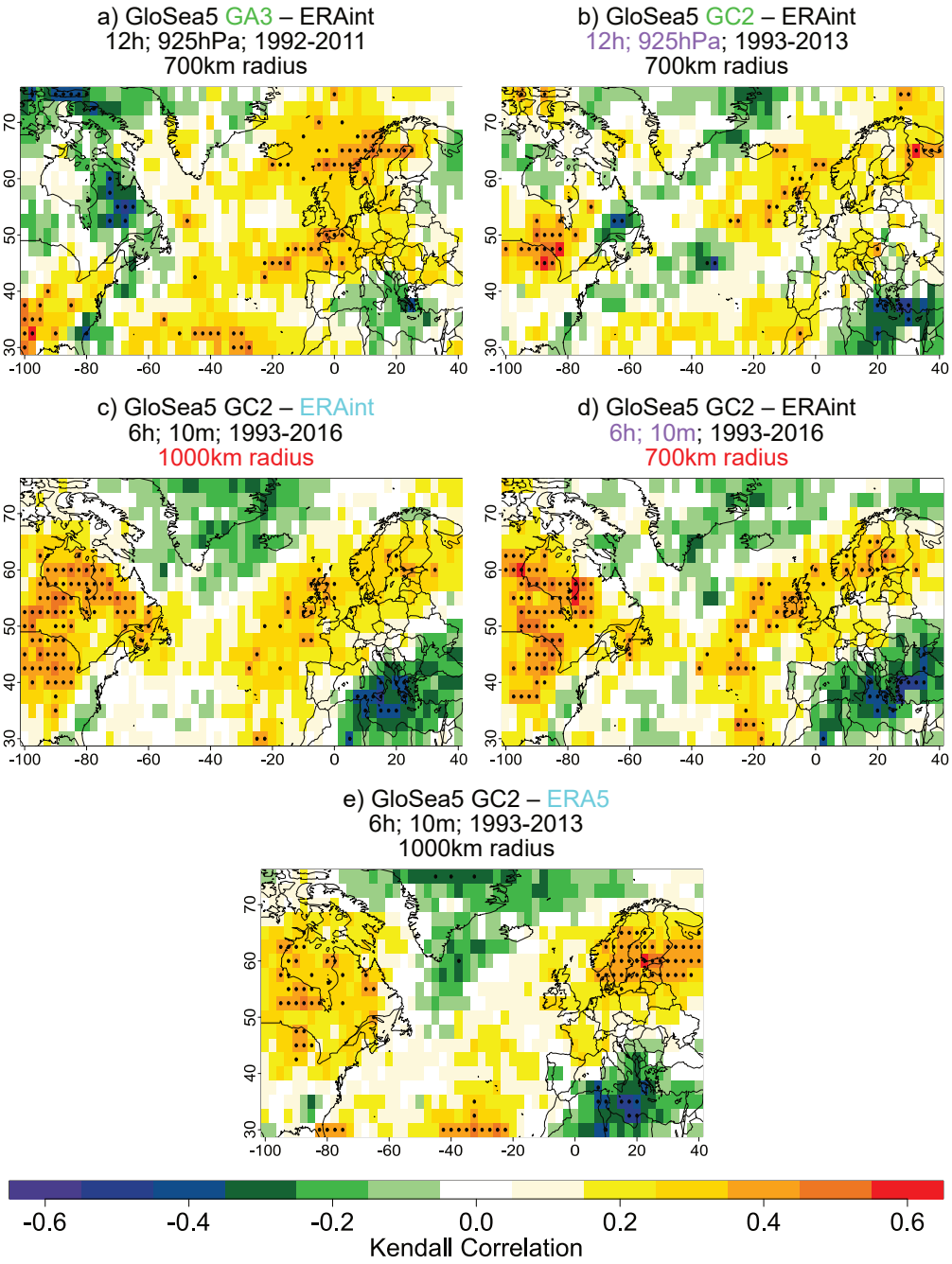


Figure 4.1 Kendall Correlation as skill measure to investigate seasonal forecast skill of severe winter windstorm frequencies for different data combinations and settings. 95% significance tested with a Kendall-correlation test and marked with dots.

British Isles (Fig. 4.1). A change in the tracking level and temporal resolution of the data (12h, 925hPa → 6h, 10m, see Fig. 4.1 panel b & d) is causing an increase in the significant forecast skill over the British Isles. But the general pattern is still similar, which shows that

the tracking algorithm from Leckebusch et al. (2008b) is also suitable to track 12-hourly data on 925hPa (e.g., used in Befort et al., 2019). Initially, the algorithm is established for 6-hourly data and 10m wind speed, which results in stronger forecast skill (see Fig. 4.2d). The change in radii is investigated in smaller steps in Fig. 4.2. Fig. 4.1 shows the change in forecast skill from 700km to 1000km radius, which is only minor compared to the other changes. The last change is within the observational data set used as a reference for the forecast skill. The forecast skill compared to ERAint shows a coherent region of positive skill downstream of the windstorm track. When ERA5 is used as a reference, this is slightly decreasing, but the skill over Scandinavia and Russia is increasing.

The second question is investigated using the data sets GloSea5 GC2 and ERA-interim from 1993 to 2013, in 6-hourly resolution and with 10m wind speed as a tracking variable. Storm counts with a lower radius include fewer storms per grid cell because the target area is smaller. This is also causing more noise from one grid cell to another. A very high radius would count a lot of storms on one particular grid cell, maybe even when the storm is too far away to impact that grid cell location. This frequency pattern would be more spatially smoothed than with smaller radii.

Fig. 4.2 shows the different resulting Kendall correlations for the different radii, used to count storms. With a focus on the British Isles and central Europe, which is the target region for this study, it is seen that the forecast skill varies over the different radii. While a small radius, like 300 km, shows very scattered signals and few significant values, a high radius, like 1000 km, shows a big coherent region of positive skill with significant grid cells. Multiple potential explanations exist for this. A smaller radius is counting fewer storms per grid cell, hence, has a lower interannual variability than a bigger radius at an individual grid cell. It is more difficult for the statistical measure to predict smaller anomalies. A larger radii, count more storms within one grid cell, but the bigger radius also smooths the spatial variability, which could lead to less spatial variability per grid cell. Another problem occurs if using a too-big radius, as this could include storms which are more distant and which do

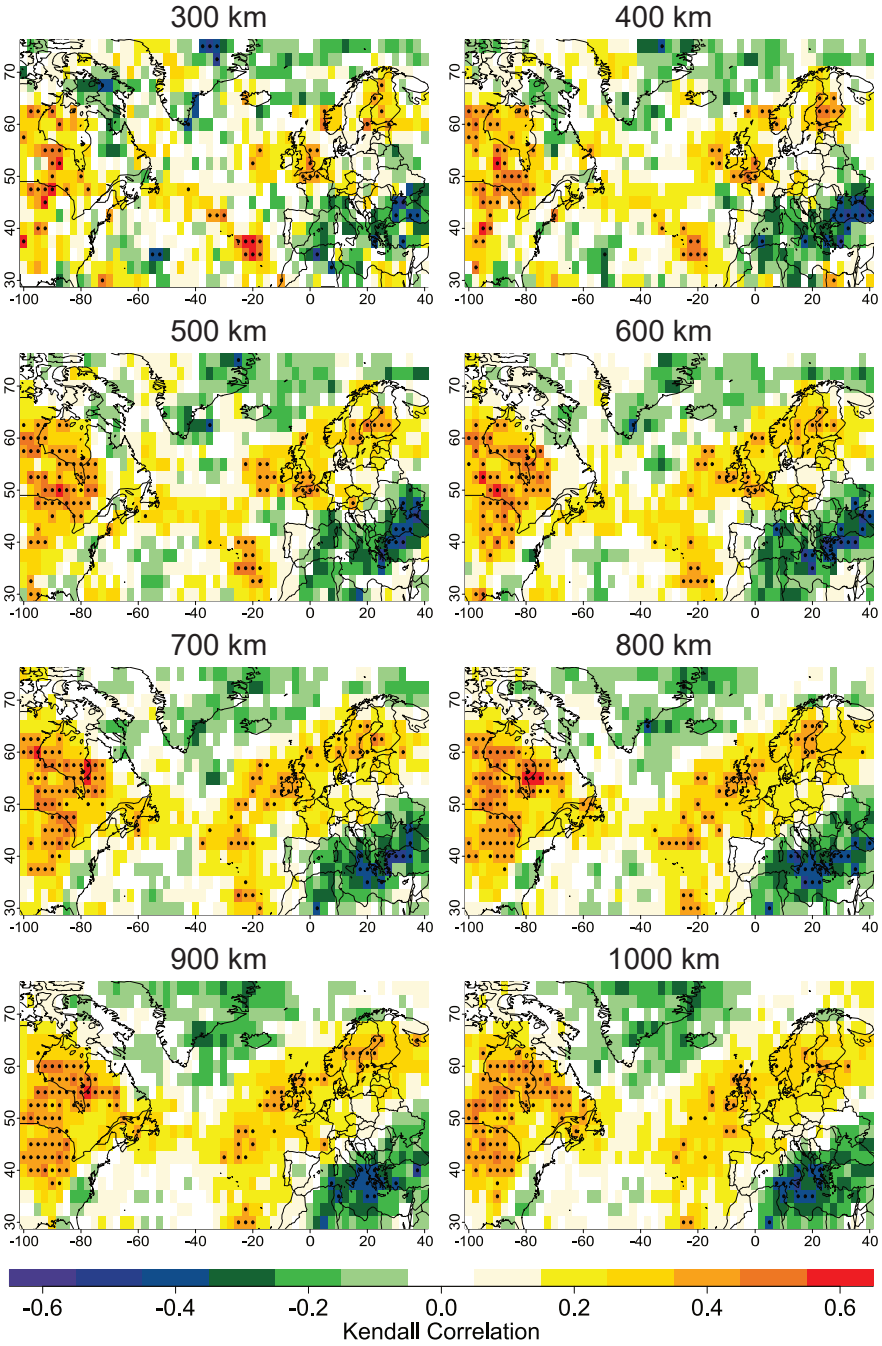


Figure 4.2 Kendall Correlation as skill measure to investigate seasonal forecast skill of severe winter windstorm frequencies for different storm frequency counting radii, ERA-interim and GloSea5 GC2 ensemble mean for DJF, 1993-2013, 6hourly resolution of 10m wind speed used for tracking. 95% significance tested with a Kendall-correlation test and marked with dots.

not have a big impact on the grid cell anymore. This validation of radii concludes that a too-small or too-large radius for counting windstorms can result in different forecast skills. To avoid over- or under-prediction and to take the physics of a storm into account, this study continues with a radius of 700 km. This is in line with previous studies (e.g. Kruschke, 2015; Befort et al., 2019).

None of the validated results in this chapter are the final used combination for the results chapters. The following study is continued with GloSea5 GC2, as this was the latest available version and the operational version of the model when the study started (see Fig. 5.2, chapter 5.1.4). This validation aimed, before using the final combination, to verify and assess different settings and understand the influence of different settings on the resulting seasonal forecast skill. The observational data set is ERA5, as this is the latest available re-analysis and with a higher resolution potentially the more accurate reality. The windstorm tracking is calculated with a 6-hourly 10m wind speed, as the initial algorithm intended, and the storm count is done with a 700km radius to represent the spatial extent of severe windstorms most adequately (Kruschke, 2015).

4.1.3 Storm Severity Measures

The previously mentioned footprint of each storm is used to develop two storm intensity measures that can be investigated for seasonal forecast skill. Therefore, the total storm severity (calculated by Eq. 4.1) of the individual storm (Event SSI (ESSI)) is assigned to each grid cell of the storm's footprint. All these grid cells are individually accumulated over each winter season and smoothed as an average with the 8 surrounding grid cells. This process of smoothing the ESSI has the same intention as the re-gridding of the storm counts, to reduce the single grid-cell noise. The simple accumulation over all time steps per season is the first severity measure in the following study, called seasonal averaged ESSI (ESSIa). The second severity measure is the ESSIa standardised by the number of storms that crossed

this grid cell, standardised seasonal storm-count ESSI (ESSIs). The ESSIa is a measure for the total severity of the storm season. The ESSIs is an average storm severity of an average single storm of that season. Both measures are used in the following studies to investigate the seasonal forecast skill of severe winter windstorm characteristics.

4.2 EOF Analysis

Atmospheric large-scale patterns are analysed concerning their connection to windstorms and how they influence storm prediction skill. It is known that the dominant large-scale pattern for Europe is the NAO. The NAO is described by Hurrell et al. (2001) as a climate variability from the east cost North America to Siberia and from the Arctic to the subtropical Atlantic, which is visible in surface parameters but also the troposphere and stratosphere. There are different ways to calculate the NAO. Leckebusch et al. (2008a) summarize some commonly used ways for the NAO index. The station-based method (pressure anomaly between Iceland and Azores, Hurrell, 1995) is more useful for investigations including long time series but does not take into account that the centre of action of the NAO can vary (Leckebusch et al., 2008a). Another method discussed in Leckebusch et al. (2008a) is the EOF method. This method leads to two different results, the index of the pattern, due to the Principle Component Analysis (PCA), and a spatial view, resulting from the EOF. This method is used in this study as the EOF leads to multiple large-scale patterns. Hence, this makes large-scale patterns more comparable as they are based on the same calculation. The resulting large-scale patterns are more robust, as even with a potential model bias in the patterns (EOF), the large-scale pattern index (PCA) can be extracted. Scientific institutes, like National Oceanic and Atmospheric Administration (NOAA), also use a way of EOF and PCA (National Weather Service Climate Prediction Center, 2005).

A more theoretical explanation of the EOF analysis can be found in Wilks (2011), who explains the EOF or PCA as reducing the number of variables in a data set with linear

combinations of these variables, to get uncorrelated new variables representing the same amount of temporal variability. The first PCA contains the most variability of the data set, the second PCA the second most and so on. The EOF patterns, as eigenvectors of the covariance matrix, aim to summarise the maximum amount of total temporal variance (Ambaum et al., 2001).

A more meteorology-connected explanation of EOF analysis in regards to the NAO is given by Hurrell et al. (2003), Ambaum et al. (2001), and Ulbrich and Christoph (1999). They explain more specifically that the analysis of the MSLP field shows in the first EOF, with the most variability, the pattern of the AO/NAO (Ambaum et al., 2001; Ulbrich & Christoph, 1999). This leads to two quantities, the NAO index, as in the station-based approach, and the model's individual spatial pattern of the NAO.

The scripts to calculate the EOF analysis are based on Climate Data Operators (CDO) (Schulzweida et al., 2019) and were previously used similarly, e.g., in Walz et al. (2018b). They investigated large-scale drivers influencing windstorms. Therefore, they calculated 9 EOFs of monthly MSLP for the Northern Hemisphere. The nomenclature was used as given from NOAA, and this will be used in this study as well. The EOF method has been chosen for calculating the large-scale patterns over Europe and the North Atlantic region (100°W – 40°E, 30° – 75°N, see Fig. B.1 in Appendix B from chapter 5.1). It is the most suitable method to investigate multiple large-scale patterns at once and the most suitable for shorter time series, as only 23 hindcast years are used. Due to the EOF analysis, it is also possible to make both data sets, reanalysis and ensemble model, comparable. Therefore, both EOF patterns can be compared, and if they are similar, hence, no model bias exists, the reanalysis pattern can be used to project model data onto the same pattern.

4.3 Statistics

The following descriptions of used statistics are based on Wilks (2011). The chapter is separated into two parts. At first, different skill scores are summarized. Followed by an explanation of general correlation and different possible methods. The described measures are chosen by summarising the methods of published papers which also investigate forecast skills, listed in Tab. A.1 in Appendix A.

4.3.1 Skill Scores

The main purpose of this study is to analyse the predictability of extreme windstorms in long-time forecasts. The best way to assess the variance between a model prediction and the observed data set (reference) is forecast skills (Wilks, 2011). These are mostly presented as skill scores of the individual skill factor (A), in comparison to a reference (A_{Ref}) and the perfect result for this individual measure (A_{Perf}):

Table 4.3 Summary of skill measures depending on forecast method and data format (Wilks, 2011; Mason & Stephenson, 2008).

	Probabilistic	Non-Probabilistic
Discrete	ROC(SS), RPS(S)	2x2 Contingency tab
Continuous	CRPS	MSE, RMSE

$$SS = \frac{A - A_{Ref}}{A_{Perf} - A_{Ref}}. \tag{4.2}$$

Skill scores show in percentage how much the forecast is improved over the reference forecast, scaled from 0 to 1 (Wilks, 2011). The following parts will describe different forecast skills in more detail. The skill score used depends on the given categories of forecast and data. Tab. 4.3 summarises and categorises the following skill scores according to Wilks (2011).

Mean-Squared Error (MSE)

The MSE is a measure mainly for non-probabilistic and continuous forecasts (WMO, 2023, or see Tab. 4.3). The advantage over the MSE of the Mean Absolute Error (MAE) is the MSE is more sensitive for larger errors and outliers using the squaring of the forecast error ($(y_m - o_m)^2$) which results from the forecasted time series (y) and the observations (o) (Wilks, 2011).

$$MSE = \frac{1}{M}\sum_{m=1}^{M}(y_m - o_m)^2 \tag{4.3}$$

The skill score for the MSE can be described as Mean Squared Skill Score (MSSS).

$$MSSS = 1 - \frac{MSE_{model}}{MSE_{clim}} \tag{4.4}$$

The MSSS compares the MSE of the forecast (MSE_{model}) and the observation (MSE_{clim}). MSE_{model} in Equation 4.4 is the same as Equation 4.3 with y as forecast, o as observations and M the number of time steps in the data set.

Root-Mean-Squared Error

The RMSE is an advancement of the previously described MSE (Eq. 4.3 with the same parameters). The square root of the MSE leads to the same dimension of data and RMSE as output. This makes it easier to interpret the results.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(y_m - o_m)^2} \tag{4.5}$$

Spread/standard deviation

In some publications, the spread and the standard deviation are utilised to investigate the forecast performance. Sometimes they are treated as the same, but their definitions differ slightly. The standard deviation s describes how much the time series (for n-times individual

x_is) variate from the total mean ($\bar{x}$):

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{4.6}$$

The spread can be described with different equations, e.g. the Interquartile Range (IQR). It describes only the difference between the 0.25 and the 0.75 quartiles ($q_{0.25}$ & $q_{0.75}$). This means the IQR ignores the top and bottom 25% of the data, which can be good in case of outliers (Wilks, 2011; Aguinis et al., 2013).

$$IQR = q_{0.75} - q_{0.25} \tag{4.7}$$

This means the IQR as spread does not include all data as the standard deviation would. Another alternative for the spread assessment is, for example, the median absolute deviation.

Ranked Probability Score

The Ranked Probability Score (RPS) is a skill score for probabilistic forecast (Murphy, 1969), to verify forecasts in predefined categories (Weigel et al., 2007) and is listed under the group "probabilistic forecast and discrete predictands" (see Tab. 4.3, Wilks, 2011).

$$RPS = \sum_{k=1}^{K}\left(F_k - O_k\right)^2 \tag{4.8}$$

Eq. 4.8 is comparing if a specific category k is the same in the model forecast F and the observational data set O. The respective skill score of the RPS, the RPSS, is built with Equation 4.2 and used in studies like Renggli et al. (2011) or Hansen et al. (2019) (see Tab. A.1).

ROC Diagram

The ROC diagram shows the relation between the False Alarm Rate F (horizontal axis) and the Hit Rate H (vertical axis) and Wilks (2011) mentioned it under the verification of

non-probabilistic and discrete data sets (see Tab. 4.3).

$$F = \frac{b}{b+d} \qquad H = \frac{a}{a+c} \tag{4.9}$$

The equations for F and H result from the 2x2 Contingency table (see Tab. 4.4), where a is the amount of right predicted events, b is how often a not occurred event was predicted, c is how often the prediction failed the occurrence and d is the amount of right predicted not occurred events. For the skill score (Eq. 4.2), the diagram needs to be represented as a single value. This is typically done by the area (A) under the curve. The perfect ROC diagram would be 100% Hit Rate and no False Alarms. This would mean the whole diagram area ($A_{Perf} = 1$). A random forecast would be represented as a bisecting line in the curve. Hence, the reference value A_{Ref} equals 0.5 (Wilks, 2011).

Table 4.4 2x2 Contingency Table Wilks (2011).

		Observed		
		Yes	No	
Forecast	Yes	a	b	a+b
	No	c	d	c+d
		a+c	b+d	n

The ROC curves are one measure for forecast verification (Kharin & Zwiers, 2003), but Relative Operating Characteristics Skill Score (ROCSS) are also commonly used in the verification of operational seasonal forecasts. They summarise in one value how likely a probabilistic forecast captured a special event (Weisheimer et al., 2017). The ROCSS is defined with the area A under the ROC curve as e.g., used by Renggli et al. (2011):

$$ROCSS = 2A - 1 \tag{4.10}$$

Ratio of Predictable Components

The RPC is used as the score for the signal-to-noise paradox in Scaife and Smith (2018) and Eade et al. (2014). Eade et al. (2014) presented the following equation for the RPC

resulting from the predictable component (PC) of the observational data set (*obs*) and the model (*mod*):

$$RPC = \frac{PC_{obs}}{PC_{mod}} \leq \frac{r}{\sqrt{\sigma^2_{sig}/\sigma^2_{tot}}} \tag{4.11}$$

with r as the (Pearson's) correlation between ensemble mean and observations, σ^2_{sig} as the variance of the model ensemble mean in time and σ^2_{tot} the variance of all members and time steps. Scaife et al. (2014) simplifies the equation to:

$$RPC^2 = \frac{r^2_{\overline{m}o}}{\overline{r}^2_{\overline{m}m}} \tag{4.12}$$

with $r_{\overline{m}o}$ as the simple correlation between the observations o and the model mean $\overline{m}$ and $\overline{r}_{\overline{m}m}$ as the averaged correlation between the model mean and each individual model member m. The RPC measures the relation between the predictable component of the model and the observation (Scaife et al., 2014).

- RPC = 1 - perfect
- RPC > 1 - underconfidence - model underestimates the predictability of the real world, ensemble member contains too much noise
- RPC < 1 - overconfidence - variance/correlation between model-observation smaller than expected due to predictable part in model

4.3.2 Correlation Coefficient

The correlation coefficient (in the following only correlation) is a very often used measurement for a relation between two variables. The "standard" correlation is mostly the **Pearson's correlation** which uses the covariance and the standard deviations of each variable (Wilks, 2011) as in Equation 4.13, with r being the correlation coefficient, COV the covariance and s the standard deviation of the two variables x and y.

$$r_{x,y} = \frac{COV(x,y)}{s_x s_y} \tag{4.13}$$

The Pearson's correlation is one method to calculate the relation between two time series. Another method would be the **Spearman rank correlation**. As the name mentioned, this is a correlation calculated over a ranked time series. The used equation is the same as for Pearson's correlation (Eq. 4.13), but the time series are ranked before calculation. Two examples are shown here:

$$A : c(2,3,4,5,6,7,8,9,10,11) \longrightarrow c(1,2,3,4,5,6,7,8,9,10)$$

$$B : c(4,2,5,6,8,1,3,9,7,15) \longrightarrow c(4,2,5,6,8,1,3,9,7,10)$$

Both Pearson's and Spearman's correlation assume a Gaussian distribution of the used data sets.

A third and more common method for analysing the predictability is the **Kendall's τ_b rank correlation** (in the following only Kendall's correlation). Wilks (2011) suggests Kendall's correlation as a robust alternative to the commonly used Pearson's correlation. This method is different to Pearson's or Spearman's correlations because it is calculated by comparing the time series pair-by-pair with the following equation:

$$\tau_b = \frac{N_C - N_D}{\sqrt{(n_0 - n_{tA})(n_0 - n_{tB})}} \tag{4.14}$$

The number of concordant (N_C), discordant (N_D) and tied (t) pairs are counted. Therefore, each pair of A and B vectors is compared with the next higher pairs. If both values decrease or increase in the same direction, this counts as concordant - N_C. A development in different directions is a discordant pair - N_D. The difference between N_C and N_D is divided by the total amount of possible pair combinations $n_0 = \binom{n}{2} = n(n-1)/2$ and the amount of ties $n_t = \sum t(t-1)/2$ (Kendall, 1945). For example, the vectors A and B (shown above) have 33 concordant, 12 discordant and 0 tie pairs. The Kendall correlation does not require a Gaussian distribution of the data set. A direct comparison of the three different correlation methods can be seen using the example vectors A and B. Pearson's correlation between A and B would result in a correlation coefficient of 0.620, Spearman of 0.564 and Kendall of 0.467. Besides the mentioned assumptions for the different methods, these results also

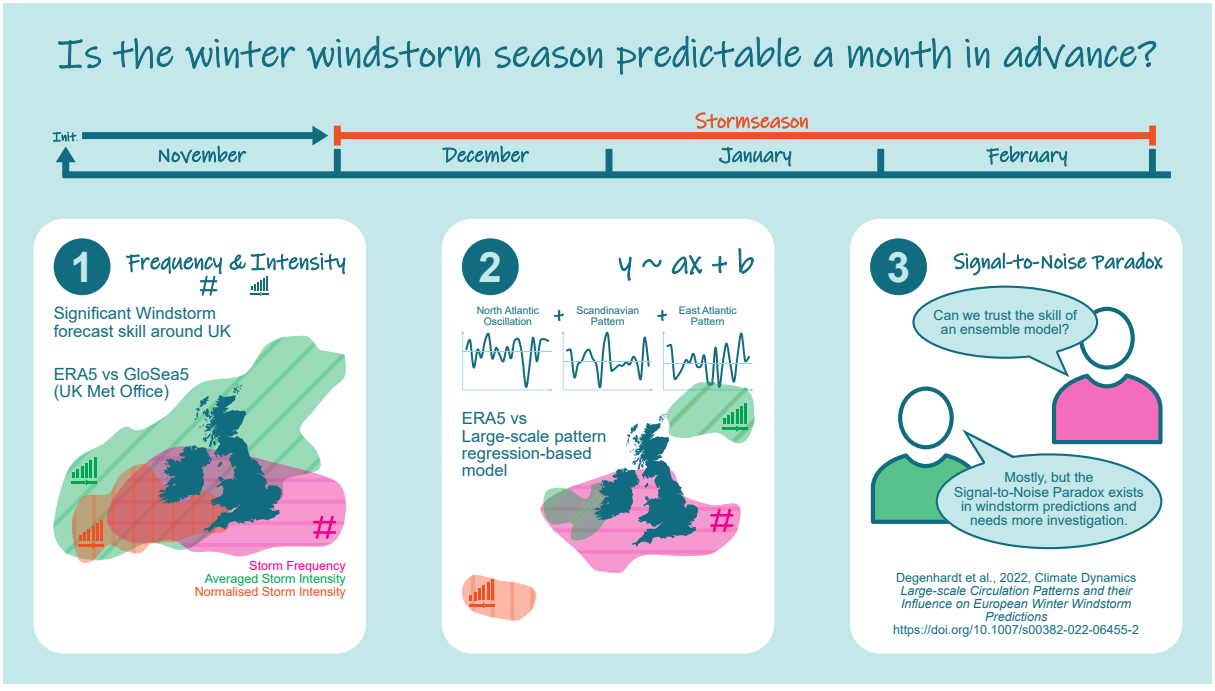
show that Pearson's correlation leads to the highest correlation coefficient but hence, could overestimate the real correlation as it is not given that A or B is normally distributed (Armstrong, 2019; Pearson & Lee, 1903). Kendall Correlation is mainly used in the following studies as a forecast skill measure. This has two reasons, firstly it cannot be assumed that any tested variables, e.g. the storm parameters (frequency or intensity), are normally distributed, and secondly, this skill measure was used in Befort et al. (2019) which is the closest existing publication to the following study. Hence, using the same skill measure leads to a better comparison with that study.

4.3.3 Used Statistical Measures

The following result chapters include studies that focus on some of the mentioned skill measures. The basic skill measure in the whole study is the **Kendall's τ_b rank correlation**. Because the data do not need to be normally distributed, and a correlation is easy to follow. But also other measures are published, like ROC/ROCSS or RPC. Other mentioned factors were assessed but did not show significant results, like ACC or RPSS. Hence, these measures are described here and used in other publications, as seen in Tab. A.1 in the Appendix (Chapter A), but the results are not presented in this thesis.

Chapter Five

Seasonal Forecast Skill of Windstorm Characteristics & their Relation to Large-Scale Patterns



Seasonal forecasts are of great importance in the scientific community. Forecasts of basic parameters like temperature, precipitation or the NAO index on a seasonal scale and as area average are highly skilful within the current operational models. But what about extreme

events which happen very rarely? This chapter is dedicated to a vital skill study. The first results chapter is published as a peer-reviewed paper in Climate Dynamics.

Degenhardt, L., Leckebusch, G. C., & Scaife, A. A. (2022). Large-scale circulation patterns and their influence on European winter windstorm predictions. Climate Dynamics, 1-15. https://doi.org/10.1007/s00382-022-06455-2.

The results of the following publication are fully calculated by Lisa Degenhardt, with the professional supervision of Gregor C. Leckebuch and Adam A. Scaife. The first full draft of the manuscript is written by Lisa Degenhardt, with further support from both co-authors before submission and suggestions by two anonymous reviewers in the review process.

5.1 Large-scale circulation patterns and their influence on European winter windstorm predictions

Abstract

Severe winter windstorms are amongst the most damaging weather events for Europe and show significant interannual variability. While surface variables (temperature, precipitation) have been successfully predicted for some time now, predictability of severe windstorms caused by extra-tropical cyclones remains less well explored. This study investigates windstorm prediction skill of the UK Met Office Global Seasonal Forecast System Version 5 (GloSea5) for the Northeast-Atlantic and European region. Based on an objective Lagrangian tracking of severe, damage relevant windstorms, three storm parameters are analysed: windstorm frequency and two intensity measures. Firstly, skill based on direct tracking of simulated windstorms is diagnosed. Significant positive skill for storm frequency and intensity is found over an extended area at the downstream end of the storm track, i.e., from the UK to southern Scandinavia. The skill for frequency agrees well with previous stud-

ies for older model versions, while the results of event-based intensity are novel. Receiver Operating Characteristic Curves for three smaller regions reveal significant skill for high and low storm activity seasons. Second, skill of windstorm characteristics based on their multi-linear regressions to three dominant large-scale circulation patterns (i.e., the NAO, the SCA, and the EA) are analysed. Although these large-scale patterns explain up to 80% of the interannual variance of windstorm frequency and up to 60% for intensity, the forecast skill for the respectively linear-regressed windstorms do not show systematically higher skill than the direct tracking approach. The signal-to-noise ratio of windstorm characteristics (frequency, intensity) is also quantified, confirming that the signal-to-noise paradox extends to windstorm predictions.

5.1.1 Introduction

Winter Windstorms and Seasonal Forecasting

Severe winter windstorms are amongst the most damaging natural hazards in Europe and have a large impact on society. A potential increase in these extreme events over densely populated regions of Europe can result in tragic loss. Consequently, this leads to a crucial interest from society (e.g., government, contingency planners), industry (e.g., the insurance sector) and scientists in order to create and to benefit from better prediction capabilities. Previous studies about mid-latitude cyclones and storminess are numerous and reviews of past and potential future developments can be found e.g., in Ulbrich et al. (2009), Feser et al. (2015) or more recently in Catto et al. (2019). Multiple studies were conducted to quantify objectively extra-tropical cyclones (e.g., Murray & Simmonds, 1991; Roberts et al., 2014; Trigo, 2006). A more detailed overview of cyclone tracking methods and their performances can be found in Neu et al. (2013). Nevertheless, to allow for an analysis of the extreme, most damaging events at the tail of the intensity distribution, those objective trackers may not be focussed enough on potential impact to society. Consequently, Leckebusch et

al. (2008b) introduced an objective percentile-based, near-surface wind tracking to identify those windstorms which have high impact potential. This method is widely used and has successfully been applied to different data sets: synoptic forecasts (e.g., Ng & Leckebusch, 2021), seasonal-to-decadal data (e.g., Befort et al., 2019; Kruschke et al., 2016; Schuster et al., 2019), as well as for climate time scales (e.g., Nissen et al., 2013) and in relation to potential damages (Walz & Leckebusch, 2019).

Studies of seasonal predictability have already shown skill for mean seasonal conditions in Europe e.g., for temperature or precipitation (cf. e.g., Fereday et al., 2012; Folland et al., 2012; Kim et al., 2012; Palmer et al., 2004) and have gained increasing interest (Domeisen, 2020; Simon et al., 2007). Recent studies focussing on seasonal predictability of European average temperature and precipitation has shown them to be usable, but with limited forecast skill. Yet, forecast skill of these parameters by using circulation patterns as predictors are higher (Athanasiadis et al., 2017; Baker et al., 2018; Scaife et al., 2014). Seasonal forecast skill can be further improved using larger ensemble sizes (e.g., Mishra et al., 2019). In winter, the extratropical skill for temperature and precipitation appears to originate in the tropics (Scaife et al., 2017b) where precipitation shows strong seasonal forecast skill especially in the East Pacific (Scaife et al., 2019b). For other seasons, skill is generally more limited but there has been some skill found for summer rainfall and circulation (Beverley et al., 2019; Dunstone et al., 2018b).

One of the first studies to investigate systematically the seasonal forecast skill of severe winter windstorms was published by Renggli et al. (2011). Their study revealed overall only marginally usable skill, but a better skill for high and low frequency storm seasons than for neutral seasons. Recent studies investigated newer versions of operational seasonal forecast suites, showing now significant positive forecast skill for storm frequencies over North-West Europe and the North-East Atlantic region (e.g., Befort et al., 2019; Hansen et al., 2019). The former investigated the seasonal forecasts of the UK Met Office (Global Seasonal Forecast System Version 5 - GloSea5) and two forecasts from the ECMWF (System

3 and 4). They found positive skill for windstorm frequency especially over Western Europe with a high influence of the NAO. Hansen et al. (2019) revealed a strong connectivity of storm count predictions to SSW. Other storminess measures (not event based/tracked) show strong and significant skill on a seasonal scale in GloSea5 (Scaife et al., 2014). They also investigated the NAO as the main driver of European storminess and showed that signal-to-noise ratios were anomalously low and that the model predicted the real world better than its own members. Subsequent studies verified this for other timescales and this 'signal-to-noise paradox' appears to be widespread in long range predictions when there is high enough skill (Eade et al., 2014; Dunstone et al., 2016; Scaife & Smith, 2018; Weisheimer et al., 2019). This signal-to-noise paradox will be investigated for the first time for storm forecasts here.

Forecast skill of large-scale patterns important for extra-tropical cyclones and European storminess

The NAO is the predominant variability pattern for European weather and climate, including extra-tropical cyclones and severe winter storms (Ambaum et al., 2001; Hurrell et al., 2003; Hurrell et al., 2001; Leckebusch et al., 2008a). Studies of the seasonal forecast skill of the NAO now show agreement that there is predictability of the NAO index at lead times of to 2 to 4 months, calculated with different methods and in various data sets (Athanasiadis et al., 2017; Baker et al., 2018; Domeisen et al., 2018; Dunstone et al., 2016; Hansen et al., 2019; Scaife et al., 2014; Weisheimer et al., 2019; Weisheimer et al., 2017).

Besides the NAO, other dominant large-scale patterns are of interest for European weather (Zubiate et al., 2017), for example, the Scandinavian Pattern (SCA: Barnston & Livezey, 1987; Bueh & Nakamura, 2007) and the East-Atlantic Pattern (EA: Nesterov, 2009; Wallace & Gutzler, 1981).

The influence of NAO phases on the development of cyclones and extreme windstorms were investigated by Pinto et al. (2008) and Donat et al. (2010). They conclude a strong positive NAO phase leads to a higher number of extreme cyclones. The connection between

windstorms and large-scale patterns was further investigated e.g., by Walz et al. (2018b). They developed a so-called "map of drivers" revealing the spatial distribution of the most dominant large-scale pattern influencing the inter-annual variability of windstorm frequency through a multiple linear regression model. Related to this, a later publication from these authors showed also a link between the NAO and potential windstorm losses (Walz & Leckebusch, 2019). Befort et al. (2019) and Scaife et al. (2014) used the NAO as the main driver over Europe for an indirect approach to statistically forecast windstorm numbers or general storminess, respectively, based solely on the model forecasted NAO. This indirect regression-based approach results in significantly positive forecast skill of windstorm frequency over the British Isles but with a lack of skill at around 45°N. This is in line with the "map of drivers" (Walz et al., 2018b), where for this region the EA index (rather than the NAO) is identified as the leading influence factor. The main dominant large-scale patterns influencing storm frequency over the North Atlantic/European region were identified to be the NAO, SCA and EA. For regions with high damage potential, Walz et al. (2018b) could successfully model the interannual variability of severe windstorms based on these three steering factors.

The present study will expand on the investigations from Befort et al. (2019) and Walz et al. (2018b) in two major aspects. For the first time, the forecast skill of the seasonal windstorm intensity over Europe will be analysed. Secondly, the impact of important large-scale patterns (NAO, SCA, and EA) for skilful seasonal windstorm predictions is analysed more systematically and compared to the skill of forecasts based on explicitly modelled storms. The paper is structured as follows: Data description is given in chapter 5.1.2. The methodology is described in detail in chapter 5.1.3. In chapter 5.1.4, results for the direct and indirect approach and the signal-to-noise paradox are presented. The paper finishes with a discussion and conclusion in chapter 5.1.5.

5.1.2 Data

The UK Met Office "Global Seasonal Forecast System Version 5" (GloSea5, MacLachlan et al., 2015) is used to quantify the seasonal prediction skill of objectively identified and tracked windstorms. Previous studies already showed positive skill in GloSea5 for predicting various parameters (Befort et al., 2019; Scaife et al., 2014; Scaife et al., 2019a). The GloSea5 hindcast data are available from the C3S for 1993 to 2016 and in this study we used 6 hourly resolution for 10m wind speed and MSLP. The spatial resolution of GloSea5 is 0.83° in longitude and 0.55° in latitude. The hindcast has four different initialisation dates per month (1st, 9th, 17th and 25th of each month), runs over 7 months with 7 members for each initialisation date. The members for the same date differ only by use of a stochastic physics scheme (MacLachlan et al., 2015).

This study investigates the main winter storm season from DJF. Consequently, the initialisations around the 1st of November (i.e., the 25th October, 1st November and 9th November) were used. 3 system versions are available for GloSea5 which refer to small model updates. This results in an ensemble of 63 members (3 initialisations x 3 system versions x 7 members) for GloSea5 in this study. As observational reference, the ECMWF reanalysis ERA5 data set (Hersbach et al., 2019) is used for the same years (1993 – 2016) as available for the hindcast data set. The 10m wind speed is used for windstorm identification and tracking and the MSLP for calculating large-scale patterns (for details cf. chapter 5.1.3) both in 6 hourly steps with the 0.25° spatial resolution of ERA5.

The storm frequency data for both data sets is analysed on a 2.5°x2.5° grid resolution as identified tracks are counted from the storm tracking outputs. This extrapolation is part of the frequency algorithm described in the Method section and retained for a better comparison with previous windstorm frequency studies. It leads to a better reflection of the true available track information, as a finer defined grid would contain data gaps. For storm intensities, the model resolution is unaffected. Hence, for a grid-based comparison, ERA5 data were

re-gridded onto the GloSea5 resolution by a bilinear interpolation.

5.1.3 Method

An overview of the investigation approach is shown in Fig. 5.1. For analysing predictive skill of severe windstorm events, all individual windstorm events in all ensemble members are identified, tracked, and then analysed with respect to seasonal frequency and the two intensity measures. In order to quantify the individual role of large-scale variability patterns, an indirect (regression-based) approach uses the large-scale patterns to build a statistical link (regression) to storm counts and intensity, which is then used to statistically model storm parameters. These regressions are built in two different ways: a) ERA5-based with the observed large-scale patterns and storm parameters and b) GloSea5-based with forecast large-scale patterns and storm parameters. After creating these links, they are used with GloSea5 large-scale patterns to predict windstorms within the indirect approach. Consequently, in both indirect approach settings, the predictive skill comes solely from the predictive skill of the large-scale patterns of the forecast ensemble. In a) a real-world link to local storm parameters is utilised and in b) a model-world link between large-scale patterns and storm parameters is used. The windstorm prediction

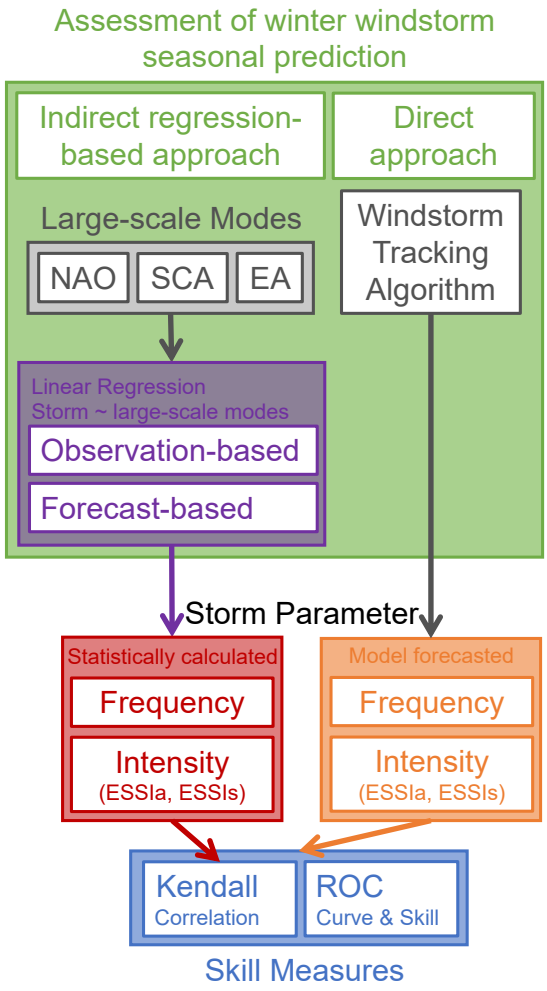


Figure 5.1 Workflow of Methodology.

skill from these statistical regressions is validated with Kendall correlation as for the direct approach. Kendall correlations and ROC are used to measure skill (please see the appendix for details).

Severe Windstorm Tracking

Severe windstorm events are objectively identified and tracked following the approach by Leckebusch et al. (2008b). Windstorms are defined as clustered (minimum 130.000 km^2) exceedance of the local 98^{th} percentile of the 10m wind speed per grid cell. This local 98^{th} percentile was identified as a suitable threshold for potential damages from extra-tropical cyclones (Klawa & Ulbrich, 2003). A storm track output was created if this exceedance lasted for at least 8 time steps (∼42h, Angus & Leckebusch, 2020; Priestley et al., 2020). Three core parameters are analysed for windstorm activity with these tracking outputs: the frequency, the season integrated intensity and the season-averaged event intensity.

Windstorm frequency was established by Leckebusch et al. (2008b) as track densities (following Leckebusch & Ulbrich, 2004) where the latter is further defined and described by Kruschke (2015). The storm frequency calculation uses a specific radius to be representative for the affected area for counting events. Similar to Befort et al. (2019), in this study a radius of 700 km is selected in line with the principal area of influence of severe extra-tropical cyclones. A validation study with different radii can be found in Degenhardt et al. (2020). The storm counts per grid cell are accumulated per season.

Windstorm intensity per season is quantified based on respective storm footprints (all grid cells the windstorm is affecting) and the related objective intensity measure, the Storm Severity Index (SSI, introduced by Leckebusch et al., 2008b) in two ways. Firstly, the so-called Event SSI (ESSI) is used as the total sum of all grid cell SSI values of an individual storm track over all time steps. This ESSI is consequently attributed to all grid cells of the storm footprint. To avoid scattered extreme values and to gain a homogeneous severity area, the individual grid cells are smoothed with their eight surrounding grid cells. Grid-cell-based

intensities per individual storm are accumulated over the whole winter season to result in a seasonal intensity measure. The seasonal accumulated ESSI (ESSIa) is one intensity measure in this study and represents the total severity of the season. Secondly, as a further intensity measure, the ESSIa has been standardised by the number of storms per grid cell. Hence, this measure is not a season integral, but storm-count normalized ESSI (ESSIs) per season and represents thus an average severity of a storm in that season.

Calculation of Large-Scale Patterns

Besides the validation of the principle forecast skill of winter windstorms derived from the model, our study utilizes the three dominant large-scale patterns over Europe as predictors to statistically calculate storm parameters. Our definition and calculation of large-scale patterns follows the NOAA definitions (Barnston & Livezey, 1987) also used e.g., in Walz et al. (2018b). Large-scale patterns are calculated as monthly means with an EOF analysis over a standardized monthly mean anomaly of MSLP for the North Atlantic/European region (100°W – 40°E, 30° – 75°N). The first EOF is identified as the NAO, the second as the SCA and the third is the EA. These patterns for ERA5 and GloSea5 respectively, can be found in the Appendix (Fig. B.1). To enable a better comparison between the reanalysis data set and the forecast model, the PCA calculation for both data sets is based on the monthly EOF patterns from ERA5. This method of projection can be done because the EOF patterns for the three selected large-scale patterns are spatially well represented in the GloSea5 hindcast data set (cf. Fig. B.1): the climatological EOF patterns are similar in ERA5 and GloSea5. The resulting indices are used as seasonal means (DJF) to compare with the seasonal windstorm parameters, seasonal frequency and intensity.

Skill analysis for Direct and Indirect Approaches

In the first, direct approach, predictive skill of windstorm parameters (frequency, ESSIa and ESSIs) are diagnosed directly based on storms identified and tracked in individual members

of the seasonal hindcast ensembles (GloSea5). This approach was also utilised in Befort et al. (2019) for windstorm frequency forecasts, consequently our results should be directly comparable with this previous study. In addition, windstorm intensity is investigated in a similar way.

In the second, indirect approach, storm parameters are calculated via a statistical regression, based on a combination of large-scale patterns as predictors. This method was previously used (e.g. in Befort et al., 2019; Scaife et al., 2014). To do this, a multiple linear regression model is built for each of the three windstorm parameters as predictands (frequency, ESSIa or ESSIs) and with the three selected large-scale patterns (NAO, SCA and EA) as predictors (see equations 5.1 & 5.2). For GloSea5 the ensemble mean is used to build the regression model.

Step 1 to build the regressions for the indirect approach is to use the large-scales patterns, either from ERA5 or GloSea5 and link them to the storm parameters in the respective data set. As seen in Equations 5.1 & 5.2 this leads by the definition of a multi-linear regression to the regression coefficients β.

$$StormParameter_{ERA5} \sim \beta_{NAO:ERA5} NAO_{ERA5} + \beta_{SCA:ERA5} SCA_{ERA5} + \beta_{EA:ERA5} EA_{ERA5} \quad (5.1)$$

$$StormParameter_{GloSea5} \sim \beta_{NAO:GloSea5} NAO_{GloSea5} + \beta_{SCA:GloSea5} SCA_{GloSea5} + \beta_{EA:GloSea5} EA_{GloSea5} \quad (5.2)$$

These regressions will be validated by an ANOVA (analysis of variance) testing (Von Storch & Zwiers, 2001) to investigate the influence of the three large-scale patterns on the storm parameter prediction individually and in combination.

In step 2 of this approach, these coefficients are used and connected with GloSea5 large-scale pattern indices to statistically predict the storm parameters by using the large-scale patterns

(see equation 5.3 & 5.4).

$$\beta_{NAO:ERA5} NAO_{GloSea5} + \beta_{SCA:ERA5} SCA_{GloSea5} + \beta_{EA:ERA5} EA_{GloSea5} = StormParameter_{stat.forcasted} \quad (5.3)$$

$$\beta_{NAO:GloSea5} NAO_{GloSea5} + \beta_{SCA:GloSea5} SCA_{GloSea5} + \beta_{EA:GloSea5} EA_{GloSea5} = StormParameter_{stat.forcasted} \quad (5.4)$$

ERA5-based regression as well as the GloSea5-based regression to windstorms are applied to finally deduce windstorm parameters; more details can be found in the Appendix.

Studies have also used a generalised linear Poisson model for the investigation of severe windstorm event counts (e.g., Walz et al., 2018a). Here we are not just investigating storm counts, but also continuous storm intensity values, thus a simple multi-linear model was chosen. Nevertheless, the performance of a simple multi-linear model in comparison to a Poisson model was tested but revealed no significant difference in performance. More detailed descriptions of the 2 different regression settings and skill measures can be found in the Appendix.

Signal-to-noise paradox

The Signal-to-noise paradox is investigated as described by Eade et al. (2014) and Scaife and Smith (2018) by using the RPC. The RPC is the ratio of the direct forecast skill of the model ensemble mean (r_{mo}) and the average correlation of ensemble members with the ensemble mean (r_{mm}). If the $RPC(= r_{mo}/r_{mm})$ is higher than 1 then the skill of the model in predicting observations is higher than the skill of the model in predicting its own ensemble members.

5.1.4 Results

Direct forecasts of frequency and intensity

The windstorm frequency forecast skill following the direct approach is assessed by the ranked Kendall-τ_b-correlation between ERA5 and GloSea5 ensemble mean (Fig. 5.2a). Highly significant positive correlations over extended areas over Europe at the end of the Atlantic storm track are revealed: the British Isles, northern France, the Scandinavian region and over the Azores. Significant skill at the 95% level is found with a significance test for Kendall correlation. These results corroborate similar patterns in Befort et al. (2019) and Scaife et al. (2014) but within the latest GloSea5 model version and on event-based storm counts.

For the integrated seasonal intensity (ESSIa, Fig 2b), similarly positive significant skill over western central Europe especially over the UK, the North Sea and large parts of Scandinavia is found. A secondary positive skill area is also found over the Azores. In comparison to the skill pattern for storm frequency, the significant signal is shifted slightly towards the northwest. The skill pattern for the standardised storm intensity (ESSIs, Fig. 5.2c) shows a very coherent pat-

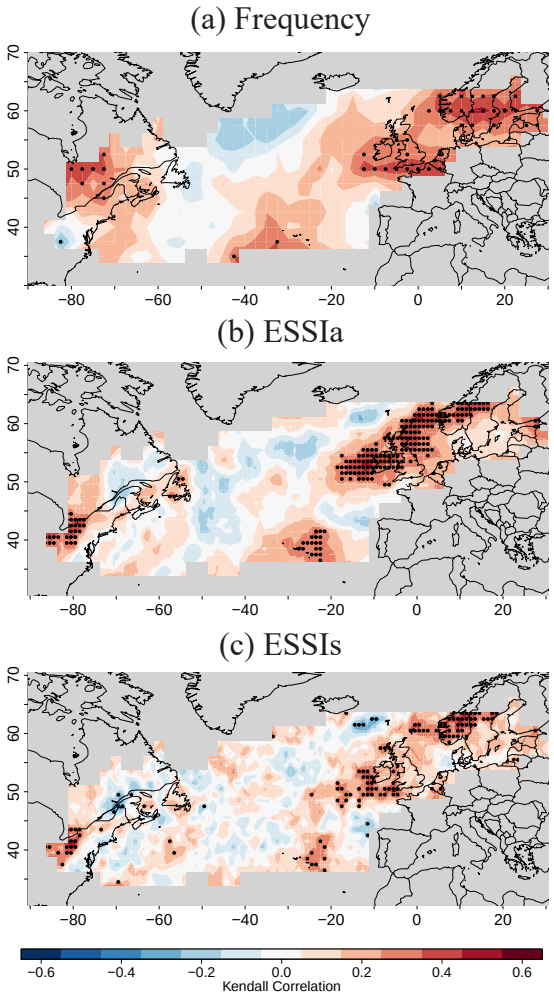


Figure 5.2 Forecast skill for storm parameter (Frequency, ESSIa, ESSIs). Ranked Kendall-τ_b-correlation between ERA5 and GloSea5 GC2 ensemble mean (63 members) from winter season 1993/94 to 2015/16, regions with 1.9 or less storms in observation (ERA5) are shaded, dotted grid points are significant on 95% level.

tern, although with slightly reduced areas of significance. Interestingly, a core region at the end of the North Atlantic storm track, mainly the region upstream of Ireland and over the south of Norway shows significant positive forecast skill. Overall, for the first time, coherent and significant skill for a dedicated, objective storm intensity measure is identified in seasonal forecasts for highly relevant regions over Europe and the Northeast Atlantic.

In addition to correlations, the ROC curve statistic is used as forecast verification. The ROCSSs are presented as spatial distributions (Fig. 5.3). Skill for storm parameters is assessed using terciles. ROCSS values are increased over similar regions as already revealed by the correlation-based analysis: For frequency and both intensity measures, regions over western central Europe and Scandinavia (e.g., the British Isles, the North Sea, northern France, Norway, Sweden and the Baltic region) show high and significant ROCSS with values up to 0.9, depending on the specific region, tercile and storm parameter. Overall, the lower and higher frequency/intensity terciles seem to be better forecast than the middle tercile as is often seen in seasonal forecasts (Mason et al., 2021).

The spatial patterns of positive and significant Kendall correlations (Fig. 5.2) are similar to the pattern of ROCSS. This means the high correlation skill in windstorm frequencies over the UK, Azores and east Scandinavia are shown as high ROCSS in the lower frequency seasons and the correlation skill over Norway in the higher frequency cases. Same for ESSIa where the correlation skill over the UK (Norway) is seen spatially coherent with the significant areas of the lower (higher) intensity seasons. For ESSIs, the significant ROCSS areas are seen in line with the correlation skill within the lower (higher) intensity season over Ireland (Norway). This suggests that the ROC and correlation scores contain largely similar information about the forecast skill (cf. Yang et al., 2018) and that the skilful forecast originate mainly from the extreme seasons.

Seasonal forecasts are highly likely to be skilful on greater spatial scales. With the ROC curves (Fig. 5.4), an investigation on a city scale was made to ensure skilful predictions for densely populated regions. Three cities were selected as small-scale representative and

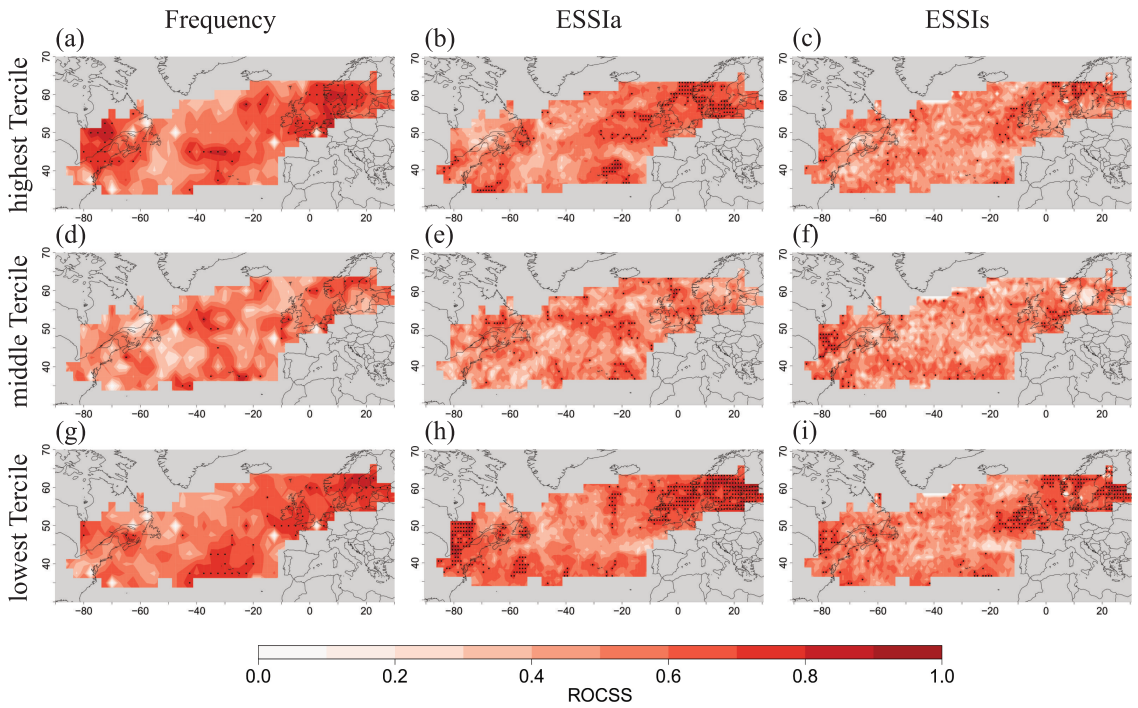


Figure 5.3 Forecast ROC skill scores. Area under the ROC curve per grid cell between ERA5 and GloSea5 GC2 ensemble (63 member) for storm parameter (Frequency, ESSIa, ESSIs) from winter season 1993/94 to 2015/16, regions with 1.9 or less storms in observation (ERA5) are shaded, dotted grid points are significant on 95% level.

averaged with the adjacent westerly and north-westerly grid cell. The British Isles are represented by a London (0°E, 52.5°N) region. An area for Hamburg (10°E, 52.5°N), Northern Germany, was picked as representative for Central Europe. And for the Scandinavian region, Oslo (10°E, 60°N) is selected.

At all 3 locations and for all 3 storm parameters the middle tercile (green) never shows significant ROC scores and they are mainly close to the zero-skill line (Fig. 5.4). This means that neutral seasons are not skilfully predicted. For the highest and lowest tercile, the ROCSS reveal an overall capability for skilful prediction (Fig. 5.4). In detail, the London area forecast has significant positive prediction skill for the higher tercile of storm frequency, but not for lower or middle tercile seasons. However, for ESSIa and ESSIs: the ROC curves are close to the no-skill line, especially for medium and high intensity seasons, although the low intensity seasons are close to significant. The Hamburg grid cells show significant

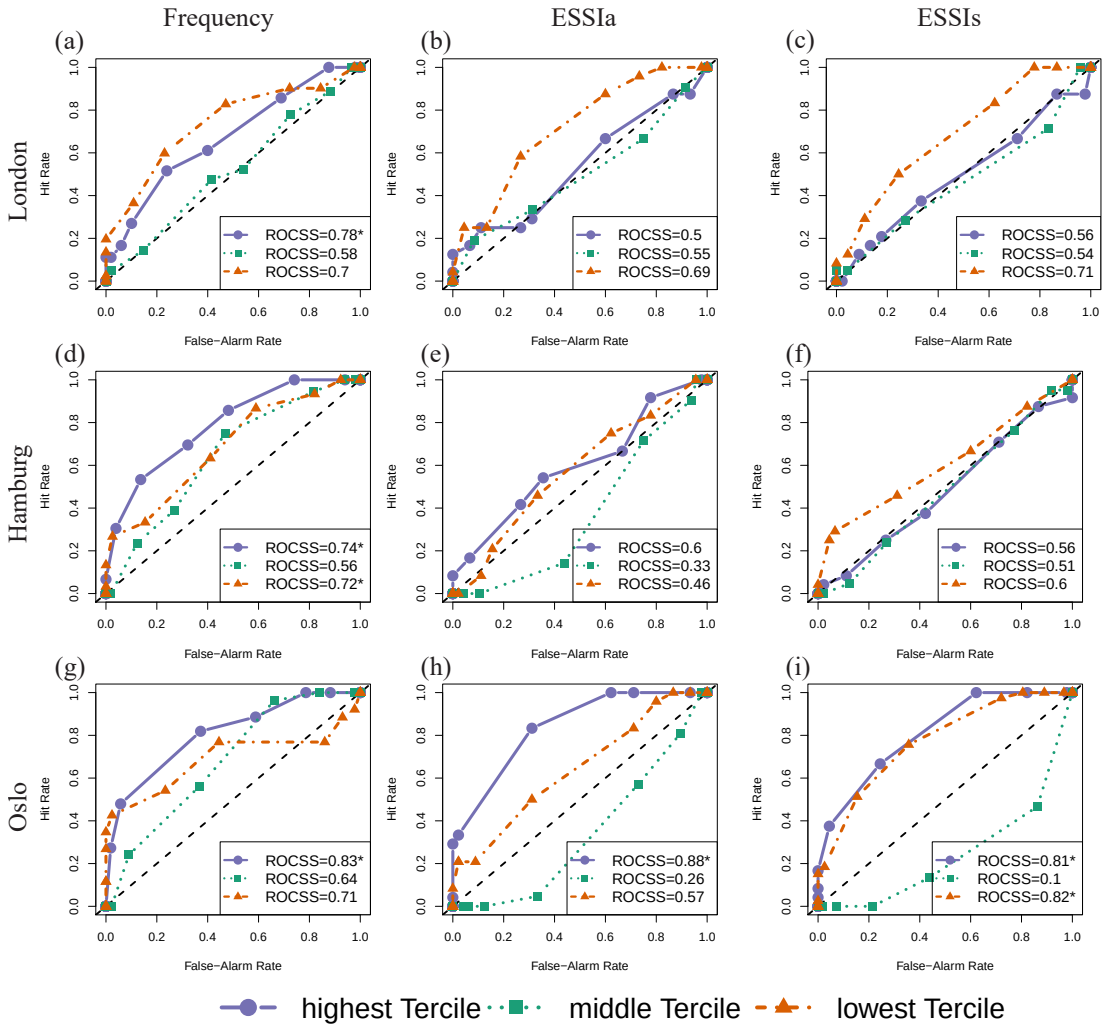


Figure 5.4 Forecast skill of the city scale. Mean ROC curve for 3 grid cells (London, Hamburg and Oslo) and their respective westerly and northwesterly grid cell, between ERA5 and GloSea5 GC2 ensemble (63 member) for storm parameter (Frequency, ESSIa, ESSIs) from winter season 1993/94 to 2015/16, associated ROC skill score as area under the curve in the bottom corner for the single "city"-grid cell, significance of ROCSS on 95% level marked with $*$; highest tercile (purple), middle tercile (green), lower tercile (orange).

forecast skill for high and low frequency seasons. The total accumulated intensity measure (ESSIa) and storm-count standardised intensity (ESSIs) show no significant ROC scores for either category. For the Oslo region, the highest tercile (purple in Fig. 5.4) of all three storm parameters is significantly well predicted, as well as the lower activity tercile for the standardised intensity (ESSIs). The storm frequency has the highest and significant forecast

skill in all three city regions in the higher tercile seasons.

Indirect forecasts of frequency and intensity

Here the role of the relevant large-scale circulation patterns in producing skilful predictions of windstorm parameters is analysed over the Northeast-Atlantic and Europe by the two different algorithms described above (chapter 5.1.3).

The indirect approach with ERA5-based regression is generated with ERA5 storm parameters and ERA5 large-scale pattern indices. The derived regression slopes per grid cell are presented in Fig. 5.5 for the 3 different windstorm parameters, respectively. These slopes represent the connection between windstorms and large-scale patterns. For windstorm frequency, the regression slopes show, as expected, a strong relation between the windstorm frequency and the large-scale pattern, NAO, SCA and EA. Both windstorm intensity measures, ESSIa and ESSIs, show similar patterns in the relation between storms and large-scale pattern, but with less strong slopes, and therefore with less strong linear relations. For all three windstorm parameters, the large-scale patterns cover different regions which potentially could increase the area of skilful windstorm forecast using this combination of predictors, although skill in patterns other than the NAO on seasonal timescales is only suggested in a few studies (e.g., Lledó et al., 2020; Baker et al., 2018).

The ability of this combined linear ERA5-based regression to explain observed storms is shown in Fig. 5.6. The coefficient of determination, R^2, shows the percentage of the explained variance . A high R^2 value of the model for storm frequency is seen especially over the storm track region from North Atlantic to the British Isles, mid-Europe, North Sea and Baltic Sea. The explained variance for the entire regression is up to almost 80%, meaning the calculated regression with NAO, SCA and EA covers up to 80% of the variance found in seasonal windstorm frequencies. For both intensity measures, less variance is explained. However, over the UK and the North Sea, the combined regression model can explain up to 50-60% of the observed storm intensity. The intensity measures have a higher proportion of

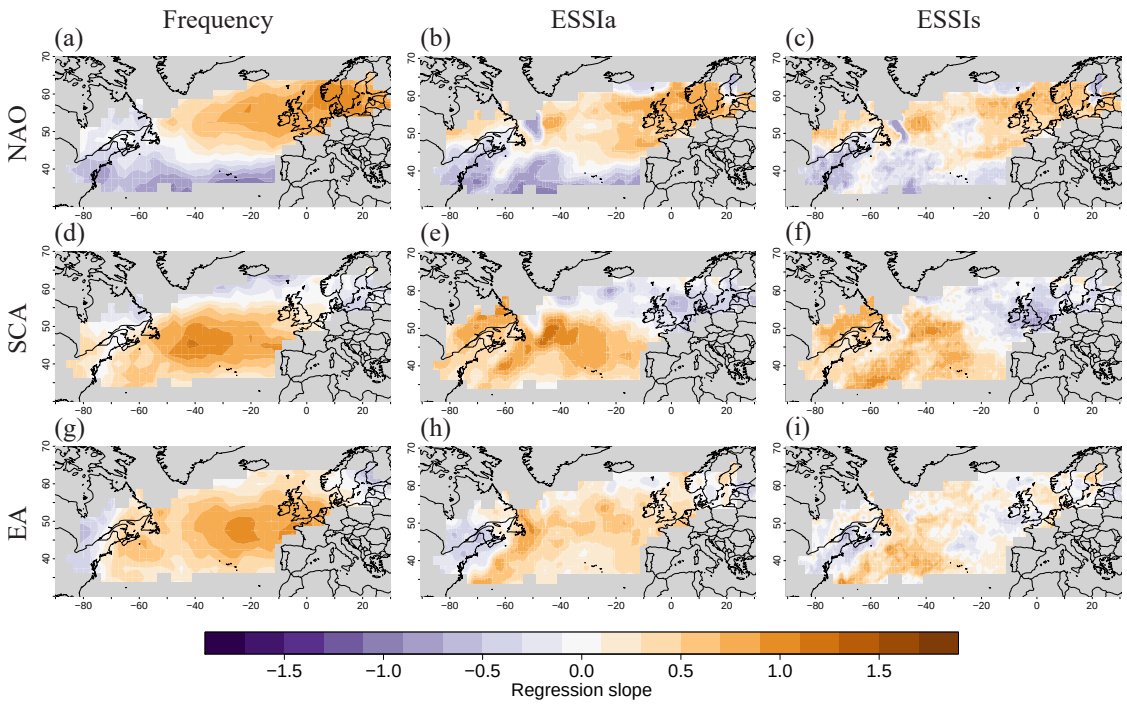


Figure 5.5 Linear models of storm parameters (Frequency, ESSIa, ESSIs). Regression slopes from multiple linear regressions to the first three EOF patterns (NAO, SCA and EA), in ERA5 winter season 1993/94 to 2015/19.

variance which is not explained by the large-scale pattern indices, perhaps suggesting it is less predictable.

ANOVA is used to quantify the influence of the three large-scale patterns (Fig. 5.6d-l). This means the individual figures for each large-scale pattern show the percentage of explained variance that contributes to the total explained variance R^2. For all three storm parameters, the NAO is the dominant variability mode to explain the variance in the storm parameters over most of the European region north of 45°N. It is even closer to 1 over Scandinavia which means the NAO is mainly responsible for the statistical performance of the regression. But the ANOVA shows an absence of connection between NAO and windstorm parameters between 40° and 50°N. Exactly in this region, SCA and EA show high values of explained variances for all three windstorm parameters and are thus potentially good candidates to improve skill if they are predictable.

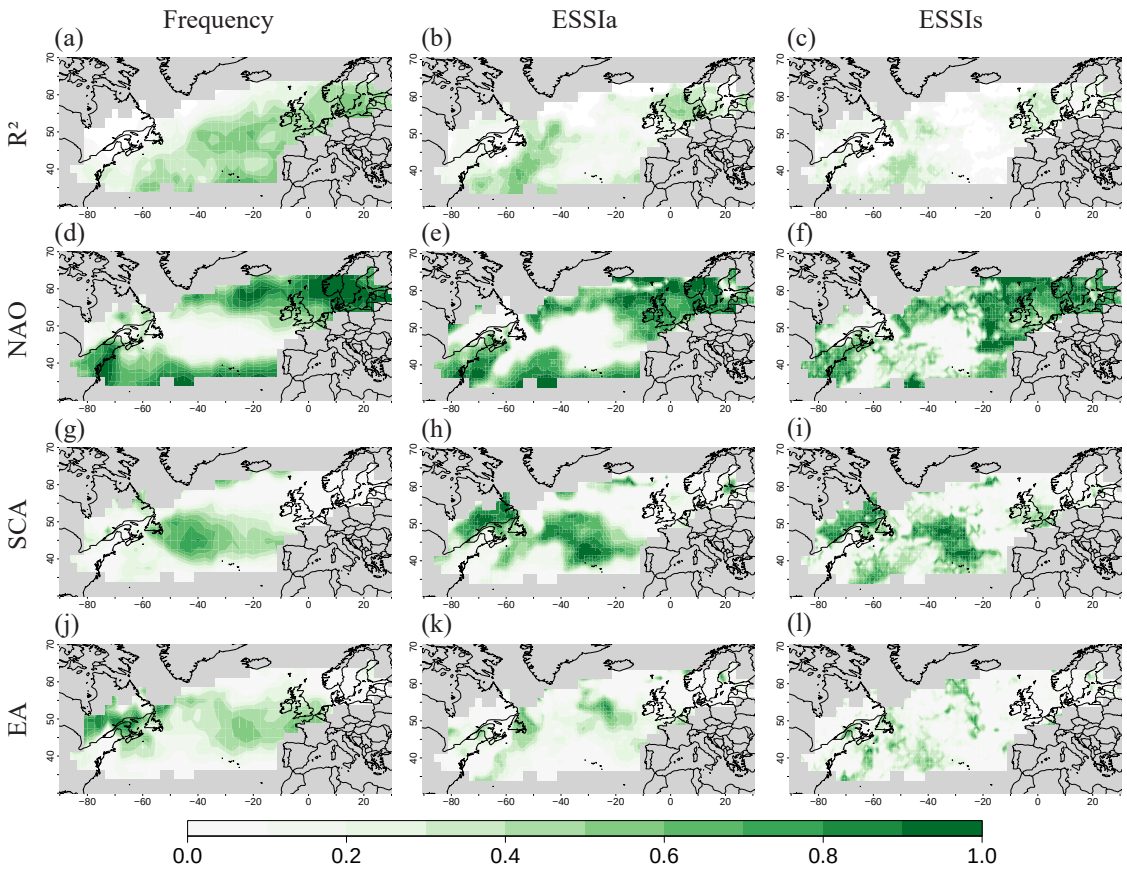


Figure 5.6 Variance explained by linear models of storm parameters (Frequency, ESSIa, ESSIs). Determination parameter R^2 (a-c) and individual ratio of explained variance from ANOVA of first three EOF patterns (NAO – d-f, SCA – g-i, EA – j-l), in ERA5 winter season 1993/94 to 2015/16.

Besides the ERA5-based regression, the link between storm parameters and large-scale patterns is investigated within the model (ensemble mean), thus based on the link between large-scale circulation and windstorm parameters purely from GloSea5 data. The resulting figures, like GloSea5-based regression slopes and verification (R^2 and ANOVA), can be found in the Appendix. The regression slopes, respectively for each large-scale pattern in the regression (Appendix Fig. B.2), show, similar to Fig. 5.5, a very strong connection between the windstorm parameter and the individual large-scale patterns. Only the spatial pattern from the SCA slopes is slightly different compared to the equivalent from the ERA5 regression as the maximum slope is shifted towards the northwest (model bias). The validation

of the GloSea5 regression is calculated the same way as the ERA5 regression validation and can be found in Fig. B.3 in the Appendix. The R^2 (Fig. B.3, a-c) shows a high percentage of explained variance of the total regression over similar regions to ERA5 (Fig. 5.6). The R^2 for storm frequency shows a slight extension of high values over larger parts of the North Atlantic. This means the GloSea5 large-scale patterns are explaining more of GloSea5 windstorm counts in this region than the ERA5 patterns explain of the ERA5 storm counts in the reanalysis data set. The individual contribution to the explained variance (ANOVA; Fig. B.3, d-l) for NAO shows similar patterns as the ERA5-based regression. The GloSea5-based regression seems to show a stronger influence of EA over the region that is independent of the NAO. The results of the ERA5-based regression lead to a similar ratio of explained variance from SCA and EA. As mentioned for the GloSea5-based regression slopes, the dominant SCA pattern is shifted towards the northwest and the same is seen in the ANOVA. The regions where SCA explains a significant part of the variance of storm parameters are shifted west-northwest and hence, play a smaller part in complementing the NAO as a source of variability.

Skill comparison between direct and indirect forecast approach

For the ease of comparison, Fig. 5.7 compares the Kendall correlations between windstorm parameters of ERA5 and those directly derived from GloSea5 with the skill resulting from the indirect method. These windstorm parameters are forecast using the ERA5-based regression, thus reflecting the real-world link between large-scale patterns and storms, but the prediction is coming from using the forecast GloSea5 large-scale patterns. Following this ERA5-based regression link, all three storm parameters reveal a very similar spatial pattern to the direct approach, which corroborates the principal usefulness of such a regression-based forecast approach. Nevertheless, overall, the area of significant positive skill is less pronounced. For the storm frequency over northern France, the English Channel, and southern Scandinavia, a slightly higher forecast skill is provided by the direct approach and for most areas over the

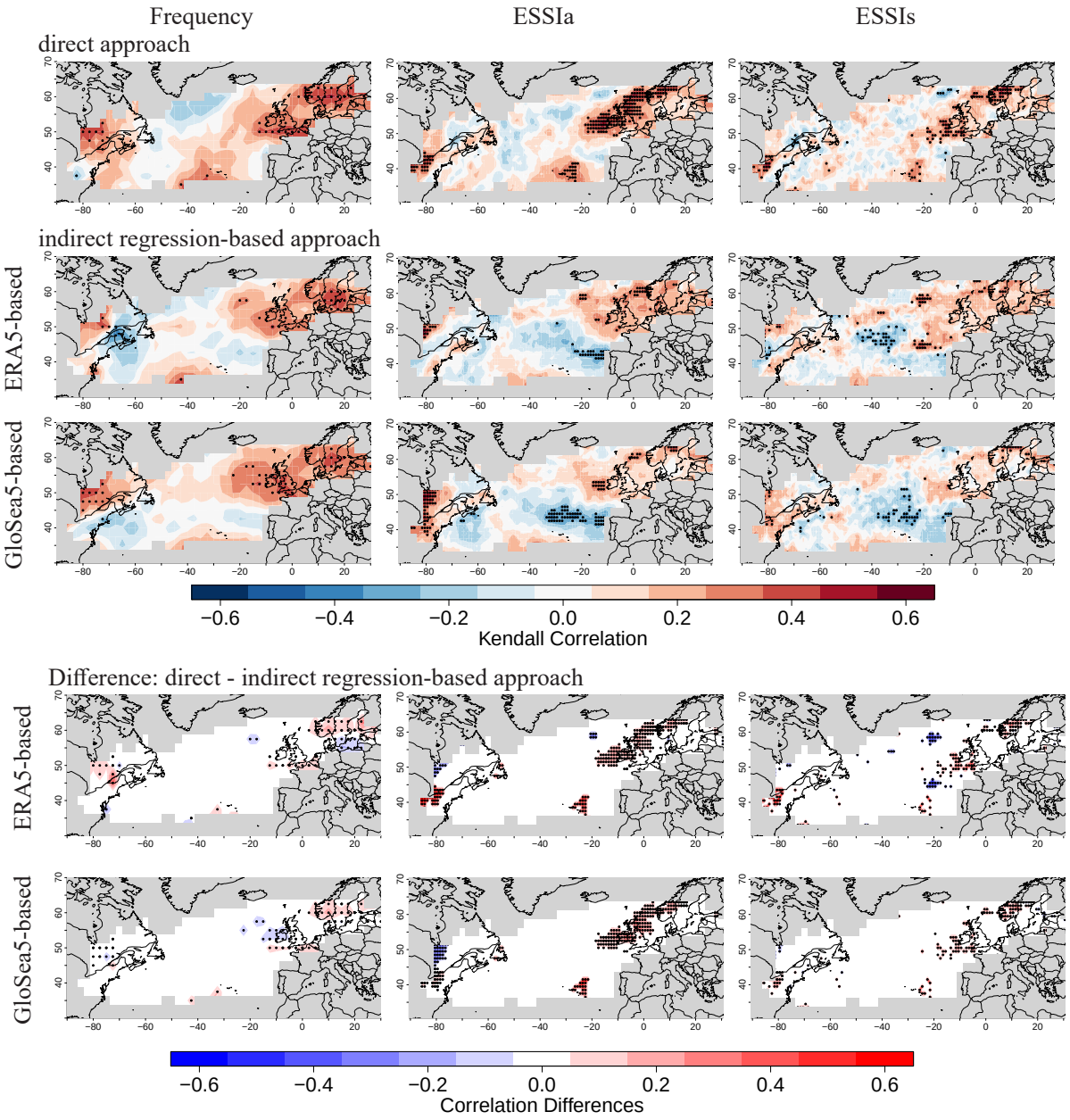


Figure 5.7 Forecast skill for storm parameter (Frequency, ESSIa, ESSIs). Ranked Kendall-τ_b-correlation between ERA5 and GloSea5 GC2 ensemble mean (63 member) (1st row), between ERA5 and regression model storm parameters (2nd row) and difference of significant correlation values as (1st row) minus (2nd and 3rd row) (4th and 5th row). Regions with 1.9 or less storms in observation (ERA5) are shaded. Dotted grid points are significant on 95% level (1st 3 rows). For the ease of comparison only those regions where one of the approaches shows positive significant skill are shown as differences and marked by a dot (last 2 rows).

British Isles there are no significant differences. In addition, the accumulated storm intensity measure, ESSIa, shows higher skill for the direct approach over the whole region stretching from the Northeast-Atlantic across the British Isles to the Norwegian Sea and Scandinavia. For the ESSIs the differences are very scattered, but upstream of Ireland the direct approach shows some stronger skill than the indirect. The corresponding skill using entirely GloSea5-based regression by applying GloSea5 large-scale pattern indices to the model link also shows overall very similar skill patterns but fewer significant grid cells than the direct approach. Only a few grid cells downstream of Ireland show more skill in the GloSea5-based regression approach than in the direct approach.

The GloSea5-based regression skill is lower in most regions than the direct approach and the ERA5-based regression. Both ERA5- and GloSea5-based regression relations cannot predict windstorm frequency or intensity as well as the explicit storms in the forecast model itself in storm relevant regions (Fig. 5.7, 4th and 5th row). These regressions are calculated using the total time series with all data points, to eliminate any potential internal effects. The leave-one-out method with one season as test data was tested for the ERA5-based regressions and the differences show the same patterns for all three storm parameters (see Appendix, Fig. B.4). The windstorm relevant areas with significant correlations in the direct approach show strong positive differences from the other methods. For storm frequencies this difference corresponds to a skill increase of up to 75% and for the intensity measure even more than 100% compared to the indirect approach. As in Befort et al. (2019) the focus is on significant changes in skill and the differences show that the direct approach has higher skill over north-western Europe than in the indirect approach.

Skill Verification over signal-to-noise paradox (only direct approach)

The possibility of a signal-to-noise paradox in the seasonal winter windstorm forecast, for tracked windstorms and their parameters (frequency and intensity measures), is investigated

using the direct approach. The RPC for the storm parameters shows very similar patterns (Fig. 5.8) as the direct forecast skill (Fig. 5.2) itself. The RPC values are high and exceed 1 in the same regions where the direct approach shows significant forecast skill. The RPC is up to 3 for the storm frequency and up to 5 for the intensity measure over the British Isles. This implies that the forecast windstorm signals, like the large-scale flow, are too weak in amplitude given the high correlation skill. Hence, from the presented findings, it can be concluded that seasonal forecasts underpredict the amplitude of year-to-year fluctuations in windstorms. This finding will be investigated in more depth in an independent study.

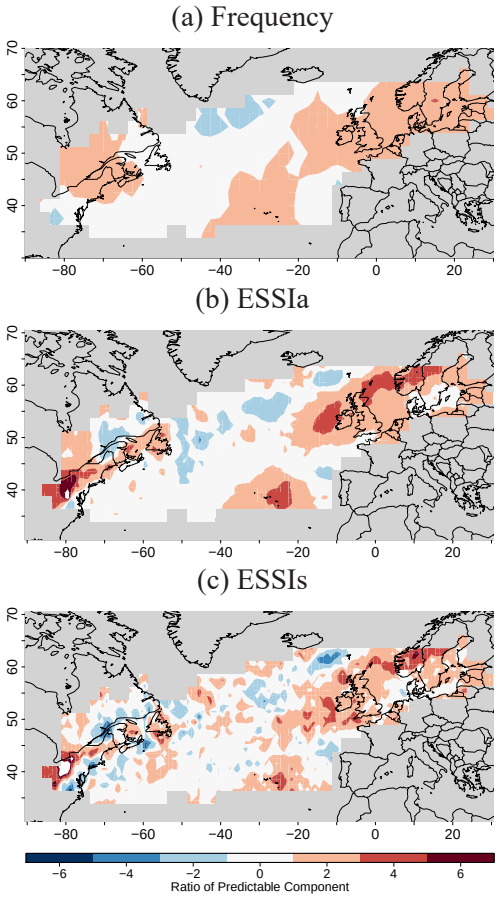


Figure 5.8 Ratio of predictable components for storm parameter (Frequency, ESSIa, ESSIs). $RPC = r_{mo}/r_{mm}$, with r as ranked Kendall-τ_b-correlation between ERA5 and GloSea5 GC2 ensemble (63 members) from winter season 1993/94 to 2015/16, regions with 1.9 or less storms in observation (ERA5) are suppressed (grey). A ratio above 1 means that the signal-to-noise paradox exists as the model mean is better representing the reality than its individual members.

5.1.5 Discussion & Summary

This study investigated three storm parameters that define extra-tropical winter windstorms: the storm frequency (Befort et al., 2019; Kruschke, 2015), the accumulated storm severity per season (ESSIa) and the average storm severity per season (ESSIs).

Two approaches to assess forecast skill are applied for relevant storm characteristics. The "direct approach" diagnoses the forecast skill by identifying and tracking windstorms di-

rectly within the hindcasts. The indirect approach uses multi-linear regression to large-scale circulation patterns, where the regression relation can be taken from observations or the model.

The storm frequency and other storminess definitions show similar significance patterns over Europe to those found by Befort et al. (2019) and Scaife et al. (2014). The highest positive skill for frequency is found over southern Great Britain, northern France and southern Scandinavia. In addition to existing studies which exclusively investigate windstorm frequency, the intensity forecast is evaluated: the combined storm intensity measure (ESSIa) shows similar high skill but is shifted towards the northwest, hence positive skill is identified more over the North Atlantic and the exit region of the North Atlantic storm track. The event-averaged storm severity measure (ESSIs) shows positive skill only in a region downstream of Ireland.

These regions are all located at the end of the North Atlantic storm track. The region over the central North Atlantic which contains the climatological maximum windstorm activity does not show strong forecast skill. This may be due to the fact that large scale modes, which explain much of the skill, have a smaller influence there. This study shows that NAO, SCA and EA are responsible for forecast storm variations at the end of the storm track. SCA and EA show higher explained variabilities over the central North Atlantic than NAO, but it seems this is not enough to strengthen the forecast skill. The addition of SCA and EA helps in regions where these large-scale patterns show their centres of action (Fig. B.1), but not in the central North Atlantic, this region would need further investigations with additional factors.

A separate skill analysis for seasons with positive/neutral/negative large-scale patterns has been done by Renggli et al. (2011) and shows stronger skill for extreme NAO seasons.

The study from Clark et al. (2017) additionally investigated the reliability of the GloSea5 forecast for 1.5m temperature and 10m wind speed. They summarized for air temperature a reliable forecast over the same region as found in this study for forecast skill of windstorms.

Interestingly, for wind speed most of this region (storm track area over North Atlantic and the UK) was found to be "underconfident" in its prediction. As seen in the RPC results (Fig. 5.8), the signal-to-noise paradox is present in the forecasts for all three windstorm parameters, suggesting that forecast ensemble mean signals are too weak which would explain the underconfidence. Further investigations of the signal-to-noise paradox in storm predictions are required and are planned in a separate study.

The ROCSS maps and selected ROC curves allow skill estimates for different storm parameter categories. The lowest and highest terciles of the storm parameter are more predictable than the middle tercile. This is in agreement with Renggli et al. (2011), who also found better prediction skill for the lower and higher terciles of storm frequency. Their study investigated just storm frequency, but here we show similar results for better prediction of lower and higher terciles of storm intensity.

The second approach used in this study is the "indirect approach". A combined linear regression with the three dominant large-scale patterns (EOF-based definition) for the North Atlantic and Europe (Walz et al., 2018b) - NAO, SCA and EA. The resulting EOF patterns of ERA5 and GloSea5 MSLP do not reveal any significant spatial difference. The applied PCA is done by using the ERA5 EOF pattern for both data sets without loss of generality to ensure a robust comparison. Using the ERA5-based regression, 80% of the interannual storm frequency variance can be explained by the three large-scale patterns. For storm intensity measures the linear regression model is less well defined but is still significant. The results for the ERA5-based regression show very clearly the main influence of the NAO around 60°N over Europe. A gap of explained variance from NAO is seen at around 45°N for all storm parameters as expected given this is close to the neutral line of the NAO influence. SCA is the leading pattern for the western part of the gap and EA for the more eastern part in the ERA5-based regressions. For the GloSea5-based regression, the same area shows EA as the dominant contributor for the variance. SCA is less important for the storm track region over the North Atlantic in the model climate. The magnitude of explained variance is very high

for the NAO which is in line with this mode being the main large-scale mode influencing the North Atlantic region (Hurrell et al., 2001) and the result by Scaife et al. (2014), showing the NAO is the strongest predictable factor for seasonal forecasts for Europe. The link between storm frequency and the NAO has also been shown by Befort et al. (2019), with the same kind of lack of storm predictions around 45°N. This study showed SCA and EA can be used to increase storm forecast skill by including those two additional large-scale patterns compared to Befort et al. (2019).

The different regression approaches show on one hand that there is a link between the large-scale patterns: NAO, SCA and EA, and storm count and storm intensity. This link also exists in the model climate with the SCA pattern having a smaller influence in the model than the real climate. The comparison between the direct method and the indirect method shows that the direct forecast model skill is higher than for the regression approach. A comparison of prediction skill for intensity shows a similar higher skill in the direct method. Compared to storm frequency, the increase in skill for intensity is more pronounced. This study improved the forecast skill based on large-scale pattern regressions for storm frequency in comparison to Befort et al. (2019) and show similar high difference between regression-based and direct-model skill for storm intensity measures.

In summary, this study showed that in addition to positive seasonal windstorm count prediction skill, there is also positive event-based windstorm intensity skill in the GloSea5 seasonal forecasts. A statistical regression approach combining the three dominant large-scale patterns over Europe improves the indirect windstorm forecast compared with NAO-based regression alone, but it still does not show higher forecast skill than the direct approach.

5.2 Additional Outcome: Seasonal Storm Forecast Skill Dependent on Large-Scale Patterns

Besides the published study, more questions and ideas about seasonal windstorm forecast skills have been assessed. One question has been: "How strong is the seasonal forecast skill of windstorms related to different phases of the most dominant large-scale patterns?". Therefore the three most dominant large-scale patterns (NAO, SCA and EA from chapter 5.1) were selected. The standardised PCA (see chapter 4.2) of the individual pattern has been used as criteria to sub-sample the entire 23 seasons of data. The sub-sampling has been done over terciles to create three similar-sized data sub-samples, which represent the positive (+, eight seasons), neutral (n, seven seasons) and negative (–, eight seasons) phase of the individual large-scale pattern. Fig. 5.9 shows the respective large-scale pattern indices and marks the individual tercile thresholds. Each sub-sample has been used to calculate the seasonal forecast skill with the same method as in chapter 5.1, as Kendall-τ_b-correlation.

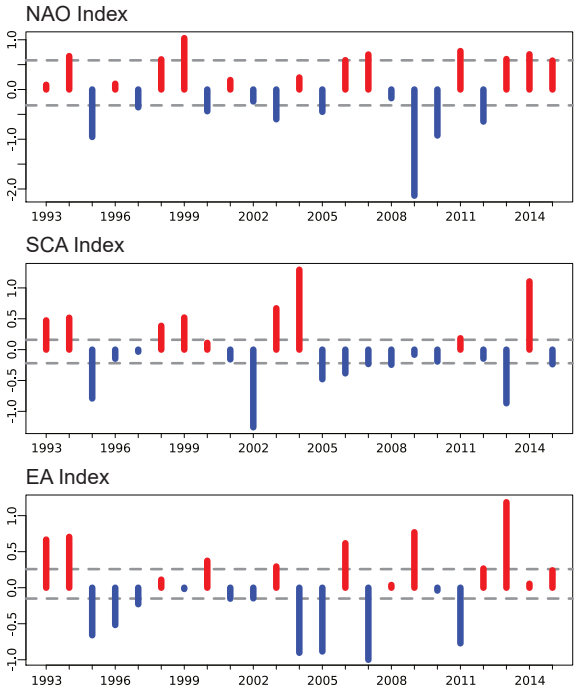


Figure 5.9 Standardised PCA indices of the three most dominant large-scale patterns over Europe, NAO, SCA and EA, with terciles (grey dashed line) as thresholds for the sub-samples.

5.2.1 Storm Frequency

Fig. 5.10 includes the seasonal forecast skill of storm frequencies depending on the phase of three large-scale patterns. The total forecast skill of the whole data set can be seen in

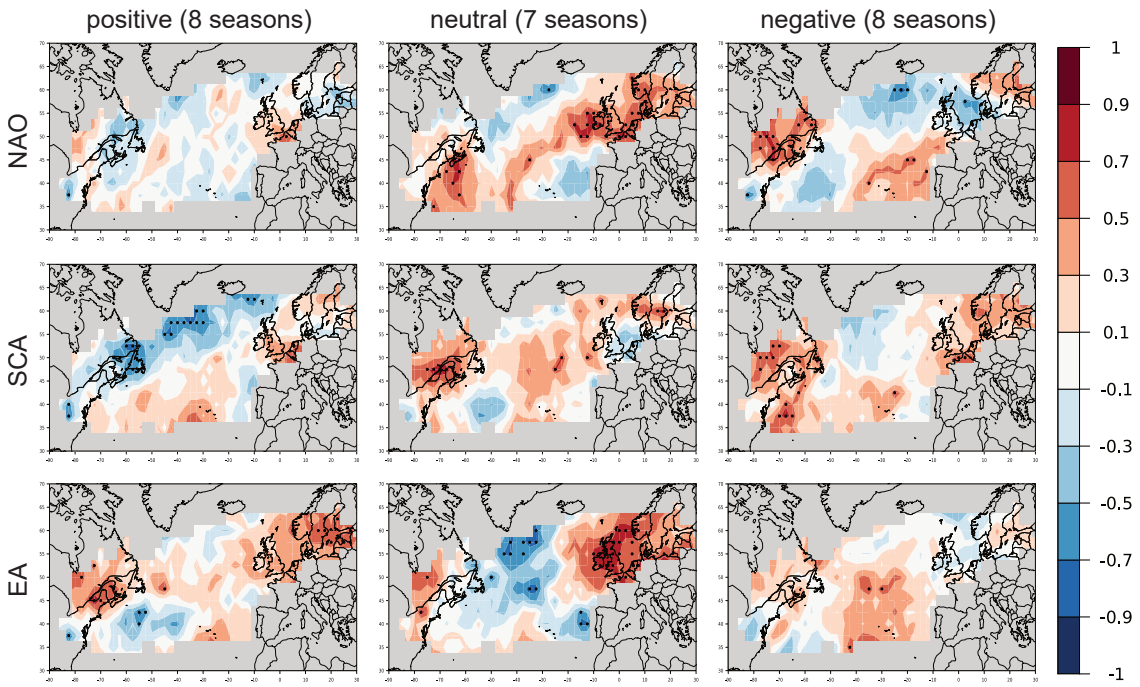


Figure 5.10 Seasonal storm frequency forecast skill, as Kendall correlation, in different phases (positive, neutral & negative) of different large-scale patterns (NAO, SCA, EA). Regions with 1.9 or less storms in observation (ERA5) are shaded. Dotted grid points are significant on 95% level.

chapter 5.1, Fig. 5.2.

The storm frequency is well predicted over the British Isles and Scandinavia in the entire data set (Fig. 5.2). The separation by NAO phases reveals that only NAOn has significant forecast skill over the British Isles and Scandinavia. The extreme phases of NAO, + and -, show less or negative forecast skill over these regions. The SCA pattern is the second dominant pattern over Europe. The windstorm forecast skill dependent on the different phases shows that SCA– has storm forecast skill, but not significant. The results for EA are similar to the ones of NAO, and mainly seasons in a neutral phase result in significant forecast skill for windstorms. This leads to the idea that seasons of NAOn, EAn and EA+ phases are mainly responsible for the found forecast skill of winter windstorm frequencies over the British Isles and Scandinavia. But previous studies find that windstorms are highly connected to NAO+ phases (Donat et al., 2010) and that the extreme seasons (strong/low storm seasons) are

more predictable (Renggli et al., 2011). This investigation shows where the forecast skill in GloSea5 originates, and therefore the neutral phases seem more important. The forecast skill of the negative phases of all three large-scale patterns over the southern North Atlantic is higher than the total skill (Fig. 5.2). As fewer seasons are in each sub-sample (positive has 8 seasons, negative has 8 seasons, and neutral has 7 seasons) than in the whole data set, the threshold for a significant correlation is higher.

5.2.2 Storm Intensity

The two introduced storm intensity measures (see chapter 4.1.3) have been investigated on how the total windstorm forecast skill (see Fig. 5.2) changes in different large-scale pattern phases. Fig. 5.11 shows the different results for ESSIa and Fig. 5.12 for ESSIs.

The forecast skill of the intensity measure ESSIa for the whole data set (Fig. 5.2) has a positive and significant forecast skill over the British Isles. The same pattern is represented in the storm forecast skill of NAOn seasons, while NAO+ and NAO– seasons result in a lower or even negative storm intensity forecast skill. The SCAn phase shows the positive and significant forecast skill of windstorms only southwest of the British Isles and SCA– seasons only northeast, covering the same area as in the overall ESSIa forecast skill. Storm forecast skills dependent on the EA phase show individual skilful regions in all three phases. EA+ seasons are mostly skilful forecasted over Scandinavia for windstorm intensities, EAn over the north of the British Isles and EA– seasons are skilful for storm intensities over the south of the North Atlantic. As ESSIs is the standardised ESSIa, the windstorm forecast skill for ESSIs (Fig. 5.12) is very similar to ESSIa (Fig. 5.11), but more scattered and smaller regions significant. The forecast skill of all intensity measures is increasing in the neutral large-scale pattern phase. But also SCA– and EA+ show a stronger correlation of windstorm intensity over the British Isles.

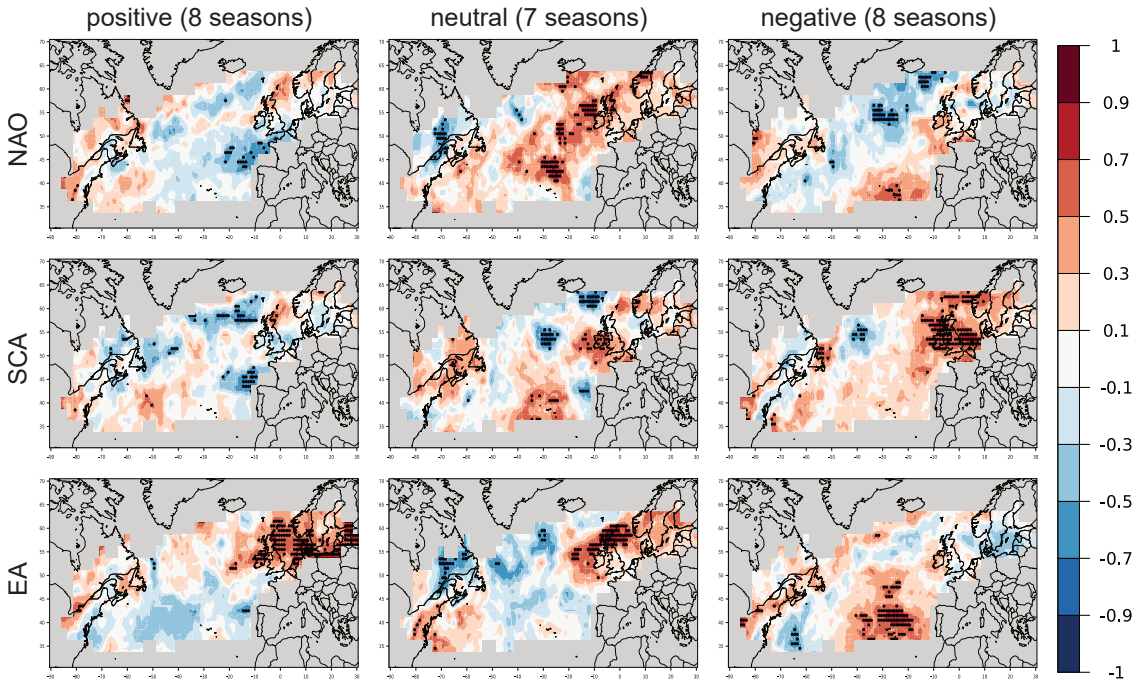


Figure 5.11 Seasonal storm ESSIa forecast skill, as Kendall correlation, in different phases (positive, neutral & negative) of different large-scale patterns (NAO, SCA, EA). Regions with 1.9 or less storms in observation (ERA5) are shaded. Dotted grid points are significant on 95% level.

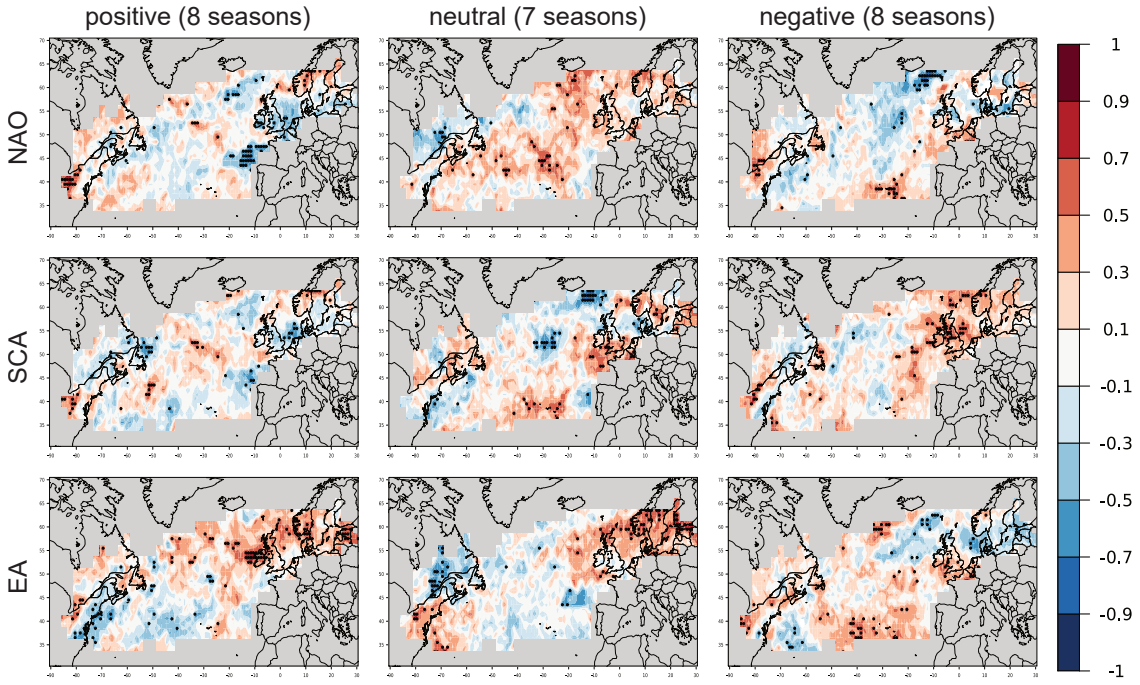


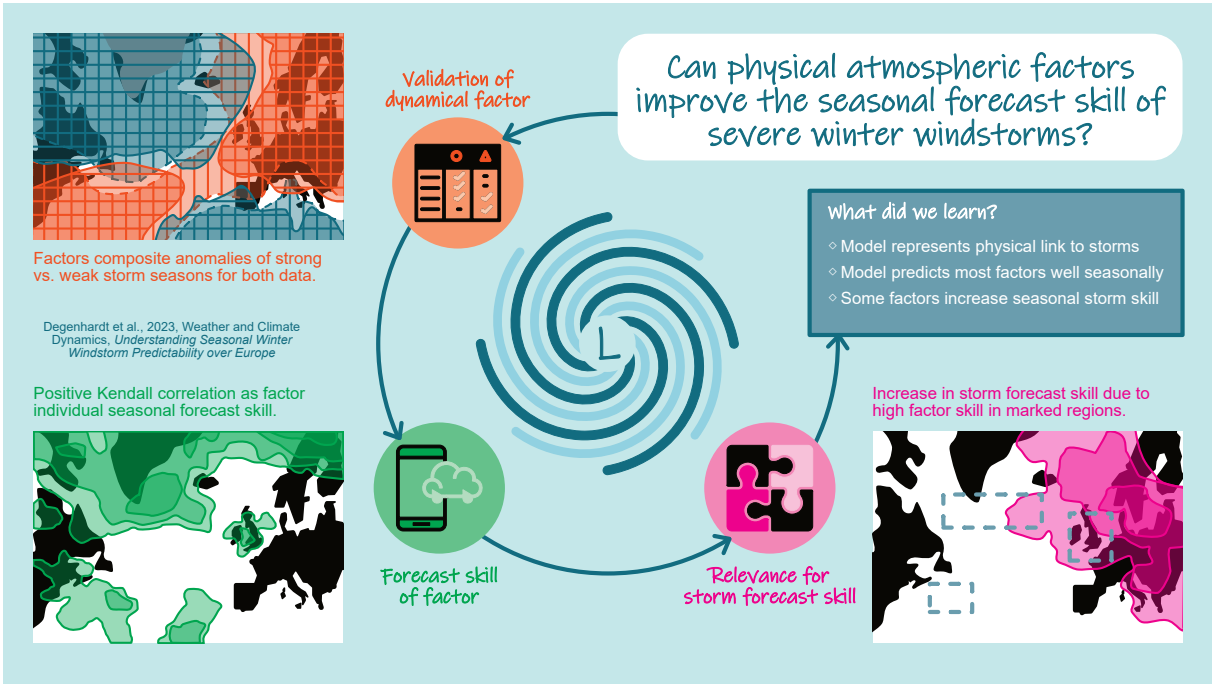
Figure 5.12 As Fig. 5.11 for seasonal storm ESSIs forecast skill.

5.2.3 Discussion

No studies have investigated the forecast skill of windstorms dependent on the phase of large-scale patterns. Studies show that more storms occur during NAO+ phases and only a few in NAO– and even none in very strong negative NAO phases (Donat et al., 2010). Storm skill studies like Renggli et al. (2011) or Befort et al. (2019) did not separate into positive or negative NAO phases and mainly looked into storm frequency. One suggestion related to phases of large-scale patterns is that a more substantial relation between NAO and Pacific–North American Pattern (PNA) could lead to more predictability in the extra-tropics (Renggli et al., 2011). Other studies showed the influence of the NAO phase on other forecast skills, like the MJO (Lin et al., 2021). They find a higher subseasonal-to-seasonal forecast skill for initialisation with NAO–. Lin et al. (2021) also investigated the influence of NAO phases on the u-wind component forecast in 200hPa and 850hPa, but in the tropical regions and found here as well that NAO– shows higher predictions for these parameters after 30 days. This additional study summarises that the GloSea5 seasonal forecast skill of severe winter windstorm characteristics mainly originates from seasons in NAO or EA neutral phases or for intensities in SCA positive phase.

Chapter Six

Dynamical influencing factors on seasonal winter windstorm predictions over Europe



The second result chapter investigates small-scale dynamical factors which are on the synoptic scale important but still influence synoptic-scaled winter windstorms or cyclones as their "precursor" system. This study investigates how predictable these dynamical and physical

factors are on a seasonal time scale and in which connection they are to winter windstorms and their predictability.

This study is submitted (22/05/2023) as a preprint in the initial version (online since 01/06/2023) in Weather and Climate Dynamics with the reference:

Degenhardt, L., Leckebusch, G. C., and Scaife, A. A. (2023). Understanding Winter Windstorm Predictability over Europe, Weather Clim. Dynam. Discuss. [preprint], https://doi.org/10.5194/wcd-2023-12, in review.

The idea and concept of the study are from all authors of the paper. Lisa Degenhardt did all calculations and main manuscript writing with the editorial support of both supervisors, Gregor C. Leckebusch and Adam A. Scaife. The following manuscript version is a first revision following the first reviewer's comments (submitted 18/07/2023, https://doi.org/10.5194/wcd-2023-12-RC1) in the online discussion.

6.1 Understanding Winter Windstorm Predictability over Europe

Abstract

Winter windstorms are one of the most damaging meteorological events in the extra-tropics. Their impact on society makes it essential to understand and improve the seasonal forecast of these extreme events. Skilful predictions on a seasonal time scale have been shown in previous studies by investigating hindcasts from various forecast centres. This study aims to connect forecast skills to relevant dynamical factors. Therefore, 10 factors have been selected which are known to influence either windstorms directly or their synoptic systems, cyclones. These factors are tested with ERA5 and GloSea5 seasonal hindcasts for their relation to windstorm forecast performance.

Following GloSea5 factors' are (1) validated on the physical connections to windstorms, (2) investigated on the seasonal forecast skill of the factors themselves, and (3) assessed on the relevance and influence of their forecast quality to windstorm forecast quality. Factors like mean-sea-level pressure gradient, sea surface temperature, equivalent potential temperature and Eady Growth Rate show coherent results within these three steps. Their physical connection is well represented in the model. These factors are skilfully predicted in storm-relevant regions. And this skill leads to increased forecast skill of winter windstorms over Europe. Nevertheless, not all factors show this clear signal of forecast skill improvement for winter windstorms, and this might indicate the potential for further model improvements or understanding to improve seasonal winter windstorm predictions.

6.1.1 Introduction

Severe winter windstorms are one of the most damaging and loss-bringing events in the extra-tropics, especially for the European region (MunichRE, 2010). Hence, it is of great scientific interest for stakeholders and the general public to understand these rare extreme events. Studies use various algorithms to identify and track cyclones (Neu et al., 2013). This study aims at understanding an even more extreme event, the surface-near windstorm, which is produced by the strongest of extra-tropical cyclones. Windstorms in this study are thus more related to the direct impacts of a cyclonic system. Leckebusch et al. (2008b) developed an objective tracking algorithm for these strongest wind events. They used a threshold that intentionally relates to observed losses (Klawa & Ulbrich, 2003) and detects about the top 2% strongest, coherent extreme events in the extra-tropics. This objective windstorm tracking has been used for multiple different studies in the past, spanning from different regions and hazards (Ng & Leckebusch, 2021; Nissen et al., 2013), individual event analysis (Donat et al., 2011b) over climate (Donat et al., 2011a; Schuster et al., 2019) and seasonal studies (Befort et al., 2019; Renggli et al., 2011; Walz et al., 2018a; Degenhardt

et al., 2022).

Seasonal hindcasts have been investigated in multiple studies for different storm-relevant aspects, like the forecast skill of the North Atlantic Oscillation (NAO; Parker et al., 2019; Athanasiadis et al., 2017; Scaife et al., 2019a; Scaife et al., 2014), stratospheric conditions (Nie et al., 2019b) or connections between tropical cyclones and extra-tropical storms (Angus & Leckebusch, 2020). In addition, different regions and events were investigated concerning their seasonal forecast skill (Dunstone et al., 2018b; Scaife et al., 2017a). For extreme European winter windstorms, one of the first studies was published in Renggli et al. (2011) based on DEMETER (Palmer et al., 2004) and ENSEMBLES (Weisheimer et al., 2009) pilot seasonal hindcasts. More recent studies investigated later operational systems, like the ECMWF systems (SEAS 3 and 4) and the UK Met Office's GloSea5 (Befort et al., 2019). They found forecast skill in windstorm frequencies and their relation to the large-scale pattern of the NAO. Following this, Degenhardt et al. (2022) found a strong positive and significant forecast for windstorm frequency and (for the first time) intensity. A connection to the three dominant large-scale patterns over Europe showed the NAO, Scandinavian Pattern and East-Atlantic Pattern together explain between 60% and 80% of interannual variability of windstorms over Europe in this seasonal hindcasts, corroborating results from Walz et al. (2018a) based on century-long reanalysis data. This leads to the motivation for this study, which is to understand which dynamical factors drive the seasonal winter windstorm prediction skill, whether as primary or secondary related factors.

Multiple studies investigated dynamical factors influencing cyclone and storm generation and intensification in the past. The Eady Growth Rate (EGR) parameter (Eady, 1949) is used as a standard measure for baroclinic instability of the atmospheric flow and is known as a source and intensifying factor for extra-tropical cyclones (Hoskins & Valdes, 1990).

$$\sigma_{max} = 0.31 \cdot \left(\frac{f}{N}\right) \cdot \left|\frac{\partial(u,v)}{\partial z}\right| \tag{6.1}$$

Eq. 6.1 represents the used equation for EGR, with f as the Coriolis parameter, N the

Brunt-Väisälä frequency ($N = \sqrt{\frac{g}{\Theta}\frac{d\Theta}{dz}}$) and $\left|\frac{\partial(u,v)}{\partial z}\right|$ the vertical wind shear. This equation is applied in the 2-layer approach of Eady (1949). This means, two pressure levels are used as boundaries and the instability in between these two layers is investigated by Eq. 6.1. The equation includes that the atmosphere closer to the equator is more stable, as f becomes smaller. N includes the potential temperature (Θ) difference with height, which is stronger in an instable than in a stable atmosphere. Later, e.g. Pinto et al. (2008) investigated important dynamical factors and their connection to strong cyclones over Europe for future climate change scenarios, based on previously identified contributors like EGR in the upper troposphere (Hoskins & Hodges, 2002), upper-troposphere divergence (Ulbrich et al., 2001), the jet stream speed (Kurz, 1990; Hoskins et al., 1983; Shaw et al., 2016) and the equivalent-potential temperature (Θ_e), as another stability measure (Chang et al., 1984). These variables were also used in other studies (Pinto et al., 2008; Hansen et al., 2019; Walz et al., 2018b; Priestley et al., 2023).

For EGR, this study uses the same diagnostic level of 400hPa as in Pinto et al. (2008) for the upper troposphere and 700hPa (resulting from 2 available model levels) to diagnose lower troposphere baroclinicity. The location and strength of the jet stream are important for whether the end of the North Atlantic storm track reaches Europe (Parker et al., 2019). Θ_e is a parameter that describes the temperature of a fully dried air parcel dry-adiabatically lowered onto a reference level, usually 1000hPa (Bolton, 1980). It is not only a measurement of the moisture content in the atmosphere and its static stability but links to the concept of the isentropic Potential Vorticity (PV; e.g. Hoskins, 2015; Hoskins et al., 1985). Thus, Raymond (1992) could demonstrate that latent heat release leads to a redistribution of PV, with positive PV tendencies below the maximum heating level and negative tendencies above. It is known that the downwards propagation of upper tropospheric positive PV anomaly favours the strengthening of cyclones (Hoskins et al., 1985; Büeler & Pfahl, 2017). Hence, it is connected to cyclonic systems and can indicate their strength and location over

the North Atlantic. Hoskins et al. (1985) compared different isentropic levels for the PV, including 350K, which is used in this study as it is a good average representative for the synoptic scales in the troposphere. They have also connected this concept with the Rossby Wave transition, by showing PV on the 350K-level for a clear identification of Rossby waves and by consecutive time steps identifying negative/positive PV advections into higher/lower latitudes, which agrees with the concept of Rossby wave breaking (Postel & Hitchman, 1999). Upper-troposphere Divergence is also part of the equation for the Rossby Wave Source (RWS), a measure of developing Rossby waves which are transporting cyclones and potentially transporting predictability from the tropics to the extra-tropics (Beverley et al., 2019; Dunstone et al., 2018b; Scaife et al., 2017b).

Other influencing factors for the generation and intensification of cyclones are the general environmental conditions which are thus indirectly connected to windstorms like the sea surface temperature (SST) distribution, SST Gradient and mean sea level pressure (MSLP) gradient (Shaw et al., 2016). These contributing environmental factors will be called secondary factors in the following, while factors like EGR or PV, which have a direct influence on cyclones and windstorms, are called primary factors. Recently, the SST and the jet stream have been identified as drivers for storm track biases in CMIP6 data (Priestley et al., 2023). Beyond those generally well-established factors, other studies identify the important role of tropical precipitation as an indicator for European climate predictability (e.g. Scaife et al., 2017b): tropical convective precipitation triggers enhanced vertical lifting, which again leads to the establishment of Rossby Waves trains to impact on Europe. Further on, Wild et al. (2015) discovered a dependency of the windstorm frequency over Europe on the temperature gradient between North American surface temperature anomalies and those of the SST over the western North Atlantic.

This study investigates primary and secondary dynamical factors connected to windstorms in seasonal forecasts from the UK Met Office, GloSea5 (MacLachlan et al., 2015), and the respective seasonal windstorm forecast skill. This could lead to a better knowledge of the

origin of the seasonal forecast skill and hence confidence in real-time forecasts.

This study uses a 3-step approach to understand the role of different primary and secondary dynamical factors for the winter windstorm predictability over Europe.

Step 1: Validation of dynamical factors: Is the observed physical link between factor and storm well represented in the model?

Step 2: Skill of Factors: Is the dynamical factor skilfully predicted on a seasonal scale?

Step 3: Relevance of Factors for Storm forecast skill: Is the forecast skill of windstorms related to the factor's forecast skill or factor-related "centres of activity"?

The study will first introduce the data sets used in section 6.1.2, followed by a description of applied methods in section 6.1.3. In section 6.1.4, the results are presented and structured within the above-mentioned 3-step approach. The study finishes with a discussion and conclusion presented in section 6.1.5.

6.1.2 Data

This study investigates the seasonal forecast of the UK Met Office's Global Seasonal Forecasting System version 5 (GloSea5; MacLachlan et al., 2015), in comparison to ECMWF re-analysis, ERA5 (Hersbach et al., 2019). Both data sets are used for a consistent time from 1993 to 2016. GloSea5 is a multi-member ensemble model with 4 initialisations per month (on the 1^{st}, 9^{th}, 17^{th} & 25^{th} of each month) and 7 members per initialisation. Currently, 3 different model versions are available, which differ in small system updates. This study investigates the northern hemisphere winter (December to February, DJF) and therefore uses initialisation around the 1st of November (25^{th} Oct., 1^{st} and 9^{th} Nov.). This leads to 63 ensemble members for GloSea5 (3 system updates x 7 members x 3 initialisations). The seasonal model output has a spatial resolution of 0.83° longitude x 0.56° latitude. ERA5 is a commonly used re-analysis and provides observation-near data, which are used as a reference in this study. The reference data set has a resolution of 0.25x0.25°. Further details of ERA5

Table 6.1 Dynamical Factors concerning storminess, cyclones or windstorms over Europe.

Factor		Version	Level	Parameter (ERA5/GloSea5)	Analysis Regions
Temperature Dipole index	Secondary		Surface	sea surface temperature (6h/6h)	North America (105°-80° W, 38°-55° N) North Atlantic (85°-50° W, 15°-35° N)
Sea-Surface Temperature		Original			Boxes of 10°x10° over North Atlantic
		Gradient			
		meridional Gradient			
Mean Sea-Level Pressure		Gradient		mean sea level pressure (6h/6h)	
		Meridional Gradient			
Total precipitation		mean		total precipitation (1h/daily)	
		Only December mean			
Jet		Location	200hPa	u- & v-wind component (6h/12h)	60°-0° W, 30°-75° N
		Speed			
Potential Vorticity	Primary	original	350K	u- & v-wind component, temperature T (6h/12h)	Boxes of 10°x10° over North Atlantic
		Bandpass 2-8d			
		Advection	400hPa		
Equivalent potential Temperature Θ_e			850hPa		
Eady Growth Rate		original	400hPa	u, T, Geopotential (6h/12h)	
		3d variability			
		Bandpass 2-4d			
		original	700hPa		
Divergence			200hPa	u- & v-wind component (6h/12h)	
Rossby Wave Source					

can be found in Hersbach et al. (2019). All factors are calculated as described in the method section, and the used variables and levels are summarised in Tab. 6.1. The windstorm tracking is based on 10m wind speeds for the calculation (details cf. below). In the case of a grid-cell by grid-cell comparison of both data sets, a re-gridding from ERA5 to the spatial resolution of GloSea5 has been done by a bilinear interpolation using Climate Data Operators (Schulzweida et al., 2019).

6.1.3 Method

Storm Tracking

The windstorm analysis is done via an impact-based algorithm developed by Leckebusch et al. (2008b). This objective identification and tracking uses a clustered exceedance of the 98^{th} percentile of surface wind speeds. These synoptic-scale wind clusters are tracked following a nearest-neighbour approach. Only events above a minimum size and duration will be considered: a coherent wind cluster must persist for at least 48 hours and reach at least a size of 130.000 km^2 (cf. details e.g., in Leckebusch et al., 2008b). Consequently, an individual storm track and a grid cell-based footprint of each storm is created. This footprint is used to count the number of storms over a defined region. The target area in this study is the extended area of the British Isles (-15° to 10° E & 48° to 60° N). Recently, the authors showed significantly skilful seasonal windstorm predictions for this area (Degenhardt et al., 2022). The individual windstorm tracks are also used to calculate the track density (used in section 6.1.4; Kruschke, 2015).

Factors

Dynamical factors are selected by previously known connections to windstorms or cyclones. The selected factors can be separated into primary and secondary dynamical factors in regard to their connection to windstorms. Hence, primary factors, like EGR or PV, are

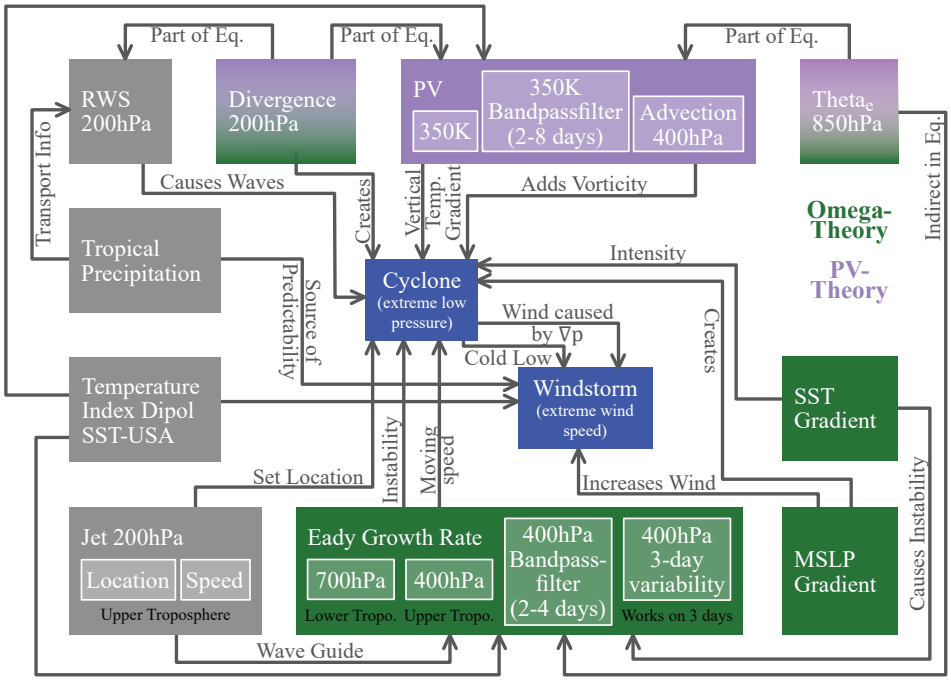


Figure 6.1 Schematic of dynamical factor connections to cyclones and windstorms (blue). Separated into different theories, quasi-geostrophic omega theory (green), potential vorticity theory (purple) and additional factors (grey). All connections and notes on arrows are based on literature.

dynamical factors which act on a smaller and shorter scale but can influence the cyclone or windstorm directly/primarily. Secondary factors are acting on a larger and longer scale. These are, for example, MSLP gradient or SST, and they have a more indirect/secondary link to windstorms as they reflect the general state of atmospheric conditions. A summary of all factors and the way they are used can be found in Tab. 6.1. Individual factors are used as seasonal (3-month) averages in the following analysis.

More details about the different ways of calculating the factors can be found in the appendix. The standard calculations have been used, e.g., the gradient of MSLP and SST, the jet characteristics (Parker et al., 2019), or the divergence in 200hPa. Other factors have been calculated following original studies, like EGR (Eady, 1949), or PV (Hoskins et al., 1985). More unique factors like Rossby Wave Source (RWS) have been calculated as described in, e.g. Beverley et al. (2019) or the Temperature Dipole used from Wild et al. (2015).

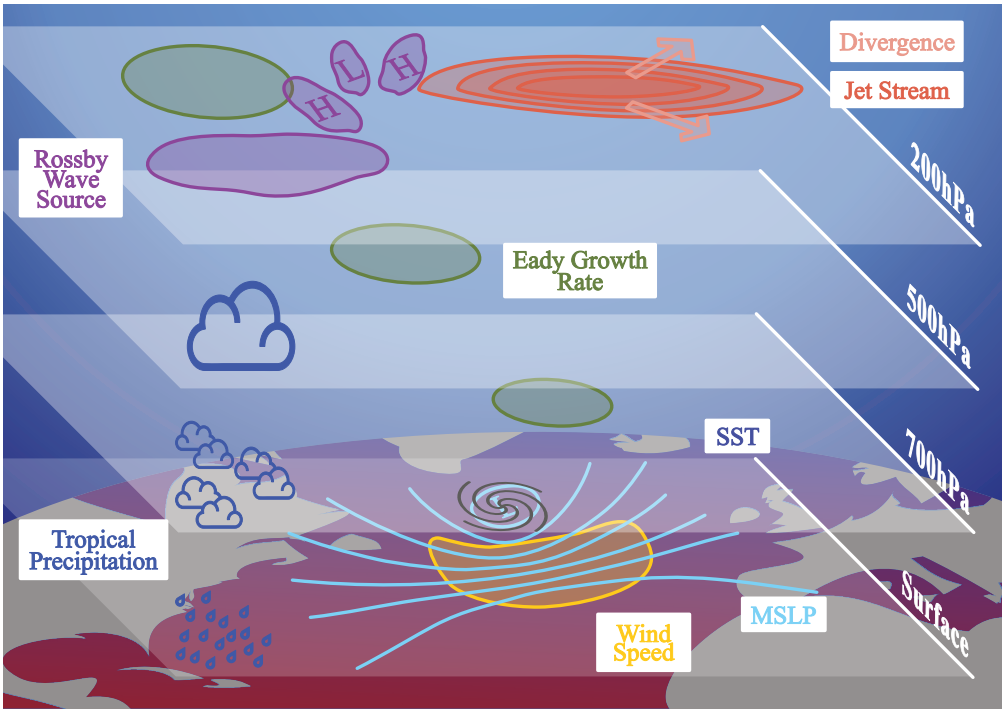


Figure 6.2 Schematic map of the spatial location of factors compared to an idealised storm system.

A schematic highlighting the different connections and interactions is presented in Fig. 6.1 and 6.2, illustrating the physical connectivity between different factors to each other and cyclones and windstorms in general. The coloured boxes indicate which physical view (Quasi-geostrophic Omega- and PV-theory) these factors are included. Each arrow represents how the factors influence or relate to cyclones/windstorms, based on literature (Fig. 6.1). Fig. 6.2 is a more exemplary schematic of an idealised storm-cyclone system, highlighting where the respective factors would be expected to be of relevance. EGR (green), as one of the most important factors to strengthen cyclones, is located northeast of the storm centre (at the lowest level) and slopes towards the northwest with decreasing pressure levels. A high upper tropospheric baroclinicity (EGR 400hPa) triggers vertical rising, which results in respective upper-level divergence (peach) at the upper boundary and hence, can enhance the jet stream (orange). The counterpart to this is the SST (ocean colour) which influences the low-level baroclinicity (EGR 700hPa), which impacts the MSLP gradient (light blue) and hence, the

wind speed (yellow). Another process related to the potential predictability of windstorms is caused by convective tropical precipitation (dark blue) via vertical lifting, triggering a Rossby wave train (purple) formation to the North Atlantic region in higher altitudes.

Composite Analysis

To understand how and when those factors influence the windstorm forecast quality, a composite analysis has been done by separating data sets into two different anomaly categories depending on storm frequency and prediction skill, respectively.

Firstly, separation is done by the number of storms, thus the seasons' overall activity. The storm counts over the extended area of the British Isles (-15° - 10° E & 48° - 60° N) in ERA5, and each GloSea5 ensemble members are used and separated into 3 categories, the 10 strongest seasons, the 10 weakest seasons and the 3 neutral seasons (10-3-10). A separation into 10-3-10-splitting has the aim of still using data sets with at least 10 years of data to achieve representative results but also to ignore the 3 neutral seasons to reduce the noise. The separation is done individually per model ensemble member to ensure that each composite compares strong vs weak storm seasons internally. This might lead to different seasons within the sub-samples. The strong-weak-composites are presented as (member-individual) standardised composite anomalies to allow for a clear comparison between the ERA5 and GloSea5 data sets.

Secondly, a categorisation concerning the forecast skill is used: well (badly) forecasted years are identified using the absolute difference of the respective seasonally averaged factor over an individually defined region in the GloSea5 ensemble mean and ERA. These categories are built for consistency according to the 10-3-10 approach again, i.e., the 10 seasons with the lowest (greatest) absolute difference are used as well- (badly-) predicted seasons. An example categorisation for individual years can be seen in the appendix (Fig. C.1) for ERA5 and GloSea5 ensemble mean windstorm counts in the UK region. Both composite methods are presented as composite anomaly differences, tested for significance via a student's t-Test.

Statistical metric of prediction skill

All approach steps include correlations, here performed using ranked τ_b-Kendall correlations (Kendall, 1945). Kendall correlation is a similar measure to the commonly used Pearson's correlation but investigates ranked time series and is less subject to normally distributed data. In more detail, correlation is used in step 1: the verification for the member individual verification (section 6.1.4), in step 2: the skill analysis (section 6.1.4) for the factor individual forecast skill and in step 3: relevance (section 6.1.4) for the storm forecast skill for different data samples. Correlations are a straightforward statistic to use for either relationship between two time series or even forecast skill (e.g., Befort et al., 2019; Athanasiadis et al., 2014; Scaife et al., 2014). Kendall correlation is used because it cannot be assumed that the data are normally distributed. The same correlation method is used as in Degenhardt et al. (2022) for a better comparison, as this study builds upon their results.

6.1.4 Results

In the results section, the focus will be on 4 factors (2 primary and secondary factors, respectively), which highlight the postulated link to forecast skill of winter storms clearly and best, MSLP Gradient, SST, Θ_e (850hPa), EGR (400hPa). More factors (see Tab. 6.1) have been tested within the 3-step approach, but not all the required links could be clearly identified. Reasons may vary from factor to factor and will be discussed in the discussion section (section 6.1.5). Additional results for five moderate performing factors can be found in the supplementary material (appendix Fig. C.2-C.5), EGR (700hPa), MSLP Meridional Gradient, Precipitation, Divergence (200hPa) & PV (350K).

Validation of dynamical factors in GloSea5 via anomaly composite analysis

Does the model represent the same physical connections between causal factors as the re-analysis? Composite anomalies of the dynamical factors separated into strong and weak

storm seasons in the observational and model data are compared. Standardised composite anomalies for ERA5 and GloSea5 (mean over each ensemble member composite) are used to validate the individual factors on their connection to windstorms in both data sets (Fig. 6.3; for 4 selected factors discussed in more detail here). The composite anomalies between strong and weak storm seasons give a useful indication of how the factors are connected to windstorms.

It is clearly identified for the MSLP-Gradient that a stronger storm season is characterised by a stronger MSLP-Gradient over the northern part of the North Atlantic, as expected. This pattern is coherent in observations and the model. The SST-pattern (Fig. 6.3c & d) shows a clear tripole (positive-negative-positive anomaly) structure over the North Atlantic in ERA5 as well as in GloSea5 (Fig. 6.3d). The GloSea5 mean signal (mean overall ensemble mean composites) is less strong but still reveals a similar pattern. The three centres of action in the SST composite of ERA5 are also reflected in the composite pattern of Θ_e. The model means of composites result in a quadrupole pattern for Θ_e but with a stronger influence of potential latent heat release over the centre of the North Atlantic than in ERA5. Also, EGR (400hPa) shows a clear and significant pattern over the North Atlantic, with higher baroclinicity in a latitudinal band around 50° N during strong storm seasons over the UK. The assessed factors are known to have a link with cyclones and windstorms, but the former also gets influenced by the latter. Nonetheless, the investigated windstorm systems (max. 2% of days per grid cell) will influence the seasonal average of the factors only marginally. This has been tested (but not presented) by removing time steps with storms affecting the UK from a factor example time series and the differences to the original factor time series are not significant.

The appendix includes the composites for more factors (Fig. C.2) like EGR (700hPa), MSLP meridional gradient or PV (350K), which are showing in principle similar results as the previous factors, with a strong and coherent increase of the factor itself for stronger storm season over the UK in ERA5 and a good representation of a similar pattern in GloSea5.

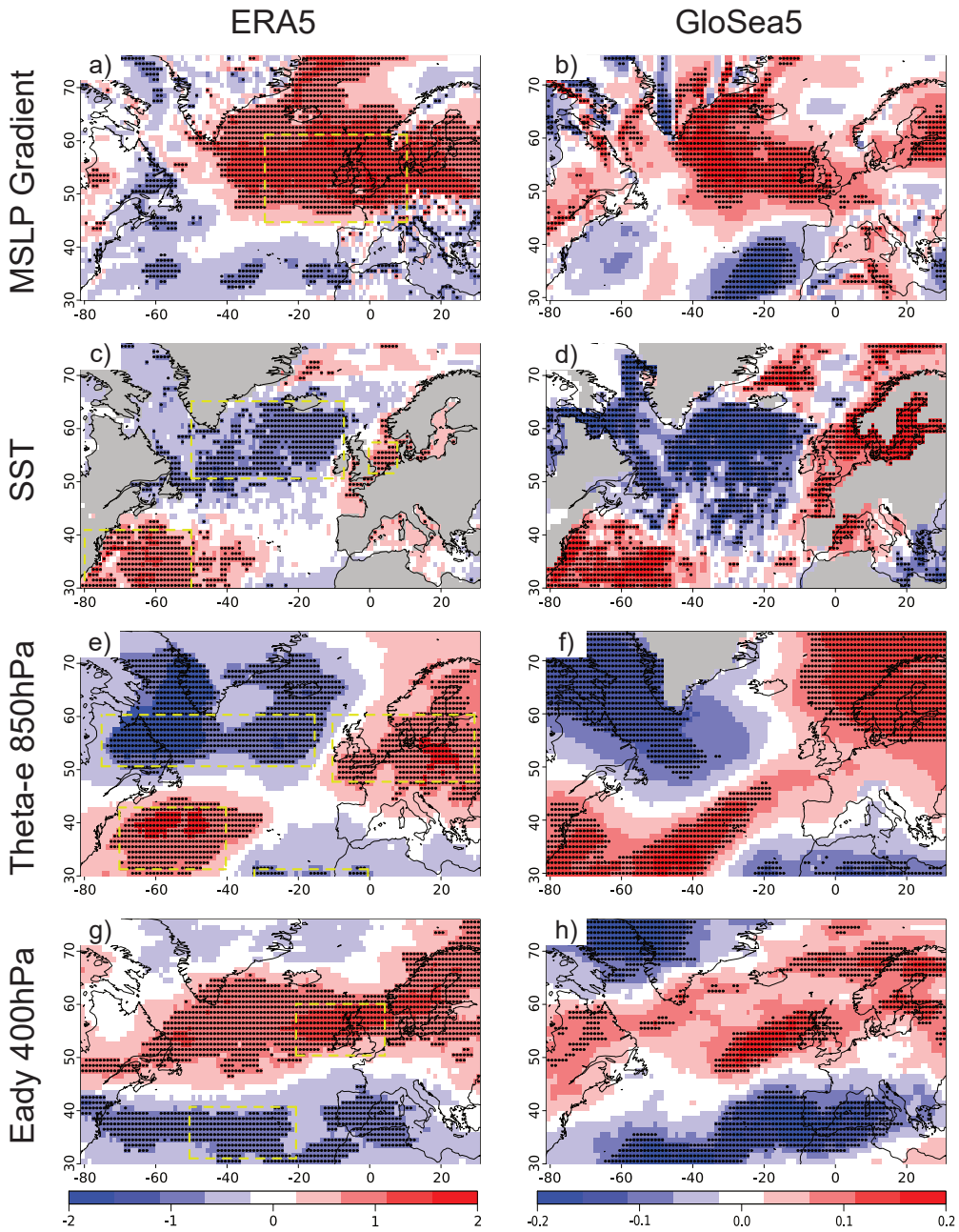


Figure 6.3 Composite Anomalies, standardised on climatology, of the respective factor for strong storm seasons minus weak storm seasons in ERA5 (left column) and GloSea5 mean over all ensemble members (right column): a)&b) MSLP Gradient, c)&d) SST, e)&f) Θ_e, g)&h) EGR; dots shown for differences significant at the 90% level (p=0.9). Yellow boxes are selected regions for investigation step 3, Fig. 6.6, process-based-view (right column).

Nevertheless, precipitation shows a north-south dipole in ERA5 downstream of the British Isles and Iberian Peninsula, which is less dominant in GloSea5, but also less relevant for

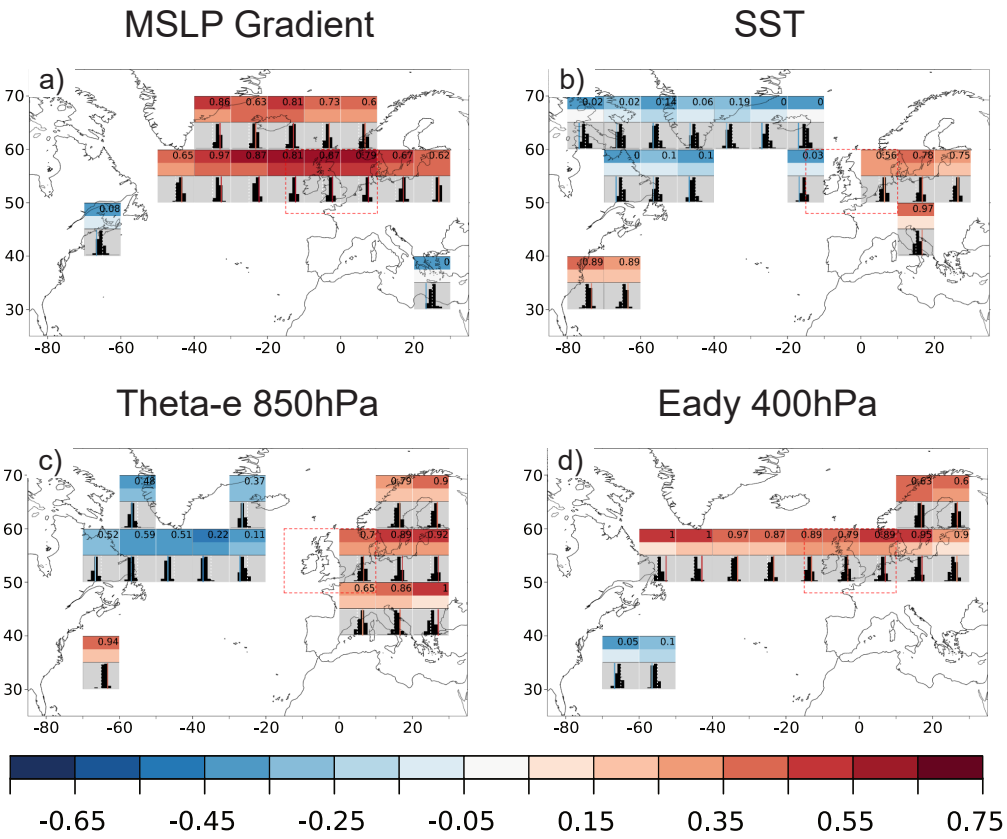


Figure 6.4 Correlation Maps between seasonal storm counts over the UK and dynamical factors. Only factors with a 95% significant correlation in ERA5 are shown. ERA5 connections (1^{st} row), GloSea5 member connection mean (2^{nd} row), GloSea5 individual member correlation distribution (below). The distributions are scales from -1 to 1 with 0 in the centre. The coloured line is the ERA5 correlation value within the GloSea5 member distribution, and the number represents the percentile of ERA5 in that distribution.

windstorm forecasts. As Scaife et al. (2017b) suggest, tropical precipitation is important for European forecast skill. The model has a strong signal and clear dipole around the equator, revealing more precipitation in the tropics in strong UK storm seasons.

Composites are categorical separations of data sets, which is useful for identifying the difference between two data sub-samples clearly, but the time-coherent link between storms and factors is also of great interest. Hence a correlation analysis between the factors' time development and windstorm frequency is used for validation. Maps are created to show the correlation link between the windstorm target region (the extended area of the British Isles)

and systematic (10°x10° boxes over the whole North Atlantic) regions of the factor over the North Atlantic. Fig. 6.4 presents the four focused factors as examples, with the remaining in the appendix (Fig. C.3).

These results show the regions where a factor is relevant to windstorms over the extended area of the British Isles (red dotted box in each panel) and how this connection is represented in the different ensemble members (histograms). The results can be separated into two parts. First, only the ERA5 connection (1^{st} row of each box) compared to the GloSea5 member mean connection (2^{nd} row of each box). All factors show in each factor box the correlation results in the same sign in ERA5 and GloSea5 member mean. Factor regions which are further outside of the storm-related area have some discrepancies, such as the MSLP gradient (Fig. 6.4a) or PV (Fig. C.3e) region over Newfoundland or the eastern Mediterranean Sea. In these regions, GloSea5 members are not in agreement with the observational relation. This can be seen in the second part of this figure's interpretation, the percentile of ERA5 within the GloSea5-member distribution. For example, the region around Newfoundland of the EGR (400hPa, Fig. 6.4d) has a percentile for the ERA5-correlation in the GloSea5-member-distribution of 1, which means the significant correlation in ERA5 is far outside the GloSea5 member correlation distribution and hence statistically different. Another example, the SST box over the North Sea (Fig. 6.4b), has an ERA5-percentile of 0.56, so the GloSea5-member distribution covers the ERA5 correlation.

This second validation concludes that GloSea5 represents the physical link between different factors and windstorms over the target region in the same way as in ERA5. This agrees with the results from Fig. 6.3. Additional information from Fig. 6.4 present that in storm-relevant regions all members agree with the physical link of ERA5, as there ERA5 correlation is not significantly different to the GloSea5 member correlation distribution.

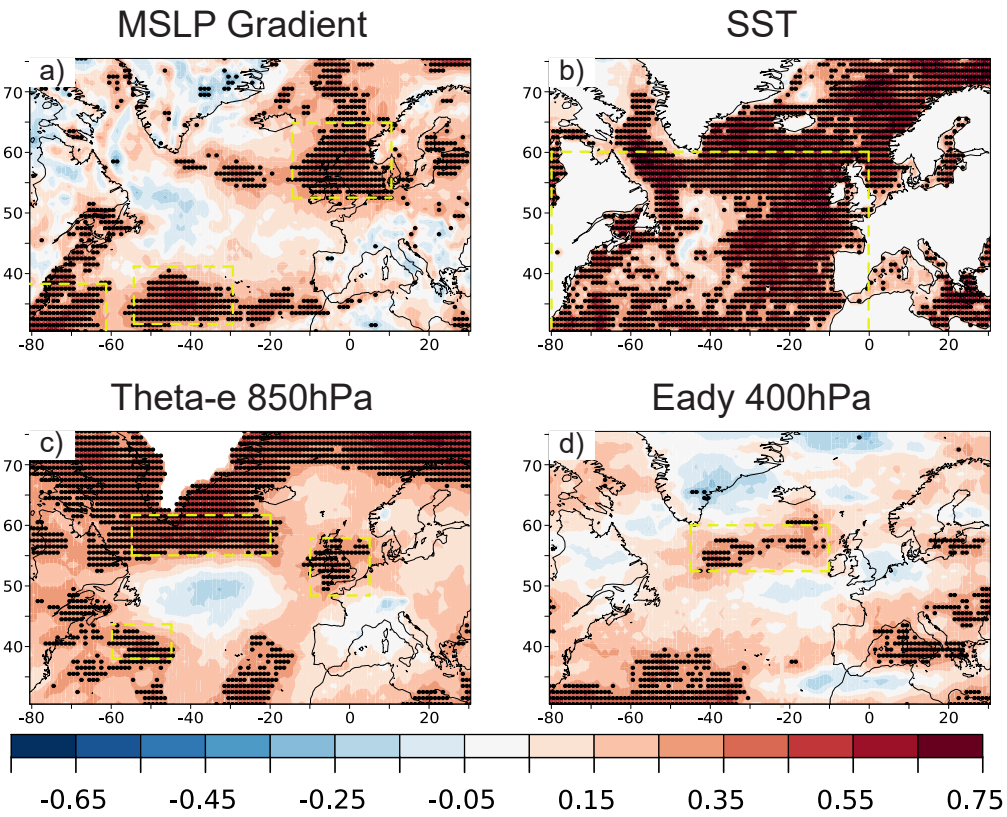


Figure 6.5 Kendall Correlation maps for selected dynamical factors, significance on 95% level marked by a dot. Yellow boxes are selected regions for investigation step 3, Fig. 6.6, factor-skill-view (left column).

Skill of Factors

Is the dynamical factor skilfully predicted? The previous results summarise that relevant factors are well represented in their connection to windstorms not only from an ensemble mean perspective but also within individual ensemble members and thus representing a consistent physical development. The next step tests if these factors themselves are well predicted. Thus, this step evaluates how far the model suite can forecast the necessary ingredients for storm development. Thus, in those regions of important connections between factors and windstorms (section 6.1.4), they should be well predicted to influence the windstorm forecast performance. The Kendall correlation is used to assess the skill of the model's ensemble mean compared to ERA5. Fig. 6.5 shows this correlation skill for the main four dynamical factors. MSLP-Gradient has a skilful and coherent region of predictability over the North

Atlantic and the British Isles. The SST is well predicted overall, with a small gap upstream of Newfoundland. The same gap but larger and stronger negative correlation is also identified for Θ_e. EGR (400hPa) correlates significantly in the region upstream of the British Isles, northeast of the Atlantic storm track. Beyond the four main factor variables discussed thus far, EGR (700hPa) reveals the same area of skill as 400hPa (cf. appendix Fig. C.4). The MSLP meridional gradient shows an extended region of skilful forecasts over the North Atlantic compared to the total gradient but not the coherent skilful region over the British Isles. Precipitation, divergence and PV 350K show very little to no skilful prediction close to the target region, the British Isles and Europe. However, precipitation is skilfully predicted in the tropics (cf. appendix Fig. C.4), which is the region Scaife et al. (2017b) suggest to be important for European predictability, as this convective precipitation would trigger Rossby Waves which propagate towards the extra-tropics.

Relevance of Factors for Storm forecast skill

Is the storm forecast skill (found by Degenhardt et al., 2022) related to the forecast skill of the factor or the regions that show a strong connection to windstorms? To answer this final question, the previous results about factors have been related to windstorm forecast skills. This step aims to find factors and individual regions influencing the seasonal forecast skill of windstorms. Therefore, the storm seasons data has been split into two sub-samples to generate two storm forecast skills depending on each sub-sample. These different characterised storms season sub-samples are separated by two approaches, one by the factor individual forecast skill (Factor-skill-view, results from section 6.1.4, Fig. 6.3) and one by the centre-of-action from the composite analysis (Process-based-view, results from section 6.1.4, Fig. 6.5).

In more detail:

a) The Factor-skill-view answers the question: "Does the existing factor's forecast skill improve the windstorm forecast?" Therefore, for the sub-samples of forecast skill, regions are

selected that show strong forecast skill for the individual factors, resulting from the approach-step-2: forecast skill (section 6.1.4, Fig. 6.5). This first view focuses on the regions with already existing and highest factor skill to assess whether the existing positive factor forecast skill in these regions is a source of the existing model's windstorm forecast and a potential improvement. If this is the case, it would mean that the correct factor prediction leads to higher storm forecast skill. Thus, the storm seasons are split between well and bad predicted factor seasons.

b) The Process-based-view focuses on the question: "Does an improved factor forecast in areas of a strong connection (centre of action) improve the windstorm forecast?" This second view uses regions that appeared most relevant in the connection between factor and windstorms (centre of action – section 6.1.4, Fig. 6.3) to create the sub-samples for the different windstorm forecast skills. This view aims to assess if a better prediction of these centres of activity would improve the seasonal windstorm forecast skill. For this, the difference in storm forecast skill (based on correlations) is calculated between sub-samples created by well- and bad-predicted factors seasons in the "centre of action"-regions. These regions have not been selected because they are skilfully predicted in GloSea5 but because they show a physical link between storms and factors.

Fig. 6.6 shows differences in the storm skill separated by both approaches, based on successful/bad predictions (factor-skill-view) and the process-based view, respectively. The region used for separation is marked individually in each panel, some boxes might be out of the mapping area, but all box details can be found in the appendix (Tab. C.1 & C.2). For both views, the selected regions (which can be multiple) are spatially averaged, and well- and bad-predicted seasons are detected by the absolute difference between the resulting ERA5- and GloSea5-time series in the used regions. The regions for the factor-skill-view are all skilful for the respective factor. For the process-based view, this is not a criterion.

For factor MSLP Gradient, three boxes were identified from the factor forecast skill analysis (cf. Fig. 6.5a) and this correlation difference in Fig. 6.6a shows the storm forecast skill

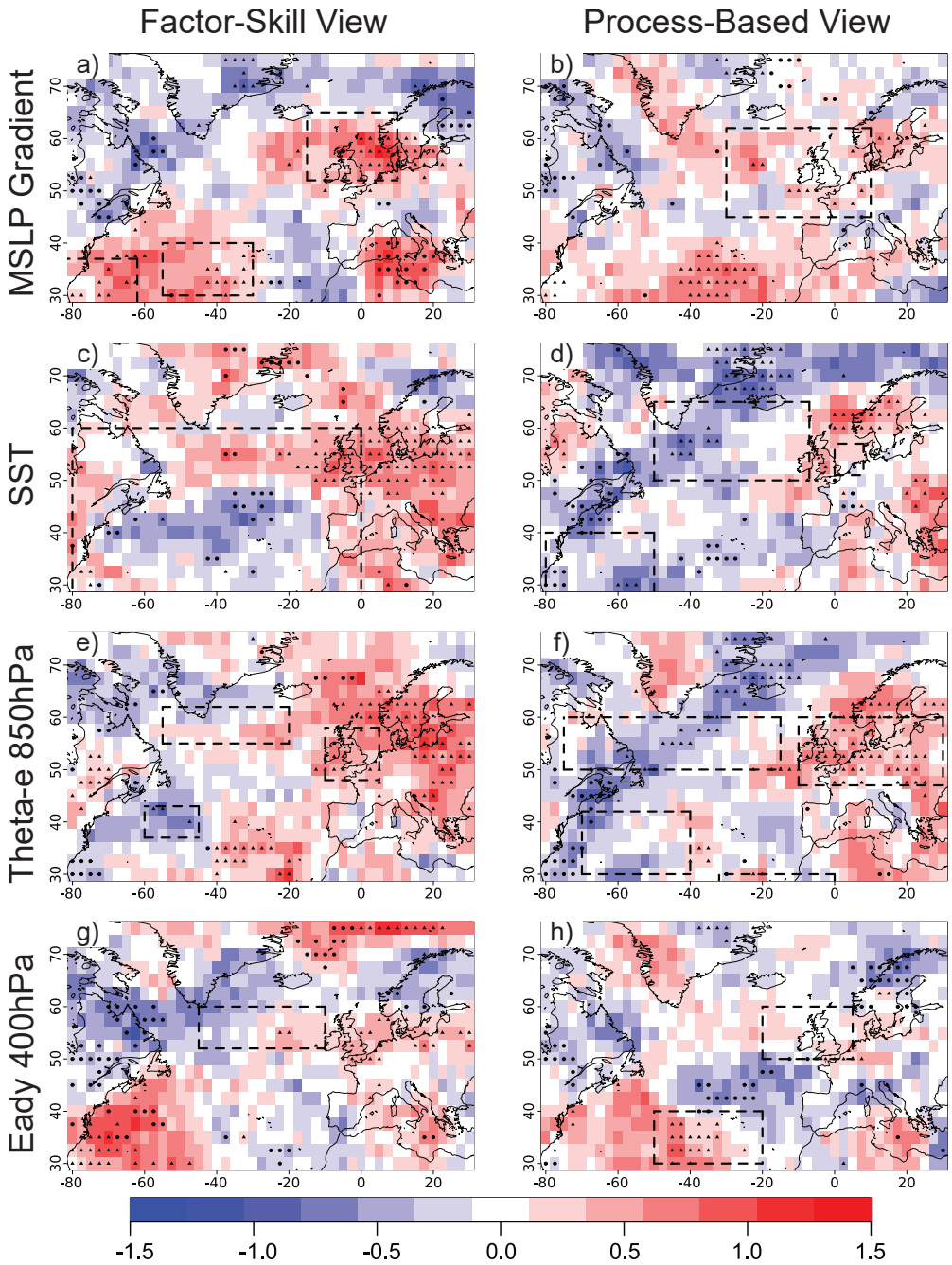


Figure 6.6 Difference of windstorm forecast skill (as Kendall correlation) with seasonal separation by the Factor-skill-view (left column) and the Process-based-view (right column). The separation is based on spatial averages over the shown boxes from Fig. 6.5 for the left column and Fig. 6.3 for the right column, individually per factor. Dotted grid cells showed significant storm forecast skill in the bad predicted season, Triangle in the well predicted sub-sample.

for years which are overall well predicted minus storm forecast skill of bad predicted years. It can be concluded that for years in which the MSLP Gradient in these three regions is well predicted, these years show an increase in storm prediction skill over parts of the North Atlantic, British Isles and Scandinavia. The second view, separated by centres of action in the composite anomalies (Fig. 6.3a), shows a less strong increase in storm forecast skill for the selected region of MSLP Gradient but still a slight increase in skill over Scandinavia. This shows the difference between the 2 separations. The process-based view (centre of action) improves the windstorm forecast skill less, and the regions already skilful in the factor forecast have more influence on the windstorm forecast skill. As SST was overall well predicted (Fig. 6.5b), the whole North Atlantic region was used to identify well and badly predicted SST-seasons. When SSTs over the North Atlantic are well predicted, the total storm prediction skill over Europe increases. The Northern European part shows the well-predicted years have a significant value on these grid cells, but the bad predicted not (indicated by the dots). The process-based-view for SST uses the four centres of action defined from the composite analysis (Fig. 6.3c) in the North Atlantic. A good forecast in these four centres of action leads to an increase in windstorm forecast skill over Europe as well. The Θ_e relevance for windstorms is tested by using three regions of skilful Θ_e-forecast (Fig. 6.5c). When all these three regions are well predicted, the windstorm forecast over Europe, especially Scandinavia and East Europe, is increasing (Fig. 6.6e). This means that the model needs a well-predicted Θ_e-value in these three Θ_e-poles to create a skilful or even improved windstorms forecast. As Θ_e and SST seem to have a similar link to windstorms (composite patterns in Fig. 6.3), as potentially higher SSTs result in more convection, hence, more moisture in 850hPa and a higher Θ_e, these factors show very similar centres of action and the Θ_e process-based-view has similar four boxes selected as for SST. A good forecast of Θ_e in these four boxes leads to an overall increase in seasonal windstorm forecast over Europe. This increase is higher and covers a bigger area than the increase by well-predicted SST regions, which might be because Θ_e is a primary factor, hence, influencing cyclonic systems directly, and SST is a

secondary factor which is more an atmospheric state surrounding the cyclonic systems. The relevant signal from the factor-skill-view is not as strong for EGR in 400hPa (Fig. 6.6g,h) as for the previous three factors in the same view, but still is a well-predicted region related to an increase of storm forecast skill downstream the box and over the British Isles. As well as the factor-skill-view, the process-based-view shows less increase in windstorm forecast skill for EGR compared to the previous three factors in this view. The remaining factors can be found in Fig. C.5, appendix. EGR in the lower troposphere (700hPa) has two very similar boxes in both views and hence, almost the same increase in windstorm forecast skill over Europe. Factors like MSLP meridional Gradient, precipitation and divergence show the skill-dependent selected regions are increasing the windstorm forecast skill over Europe significantly. The process-based-view is showing increasing signals for factors like precipitation and PV 350K.

6.1.5 Discussion and Conclusion

This study investigates the connection between primary and secondary atmospheric dynamical factors and the forecast performance of seasonal winter windstorm predictions. As skilful seasonal prediction for tracked windstorms (Befort et al., 2019; Degenhardt et al., 2022) was recently shown, this study aimed to explain the forecast skill further. A dependency of windstorms and their forecast skill on large-scale patterns, like NAO, SCA or EA, has previously been established (Degenhardt et al., 2022). Here, a more in-depth analysis of the mechanics of forecast skill generation is presented, and consequently, 10 dynamically important factors were selected and tested in multiple settings concerning their impact on the seasonal forecast skill of windstorm frequency (see Tab. 6.1). To reflect on the main contribution of those individual processes to the complex development of extra-tropical cyclones and storms, it has been differentiated between primary (small- and short-scale) or secondary (large- and long-scale) factors. These factors are investigated in a 3-step approach: first, validation of

the relevance of the factor to winter windstorms. Second, the forecast skill of the individual factor itself on a seasonal scale. And third, the relevance of the factor's forecast for the overall winter windstorm frequency forecast skill.

The strong link between windstorms and factors seen in the ERA5 composite anomalies of the four focus factors, MSLP gradient, SST, Θ_e and EGR (400hPa), are important because these four factors are knowingly the most driving factors for storm and cyclones (e.g., Pinto et al., 2008). The relation to windstorms for all these important factors is well simulated in the seasonal forecast suite, GloSea5. The SST shows the known horseshoe anomaly pattern (Nie et al., 2019a) and a clear connection is identified with a positive SST and Θ_e signal over Europe (Northern Sea and Baltic Sea): leading to stronger storm seasons as stronger SSTs may enhance Θ_e, leading to more baroclinic instability e.g., in the lower troposphere in favour of baroclinic wave development and thus for windstorms. The lower tropospheric EGR (700hPa) agrees with this concept in ERA5, as the stronger EGR (700hPa) reaches over the North Atlantic until central Europe but lacks in spatial dimension in GloSea5. The SST composites in GloSea5 show similar three centres of action (positive - east of America, negative – south of Iceland and positive – North Sea), but a more extended negative SST composite anomaly in GloSea5 further south over the North Atlantic is in line with the recently found SST bias south of Greenland in CMIP6 models causing a bias in cyclone tracks (Priestley et al., 2023). The Θ_e composite anomalies of GloSea5 show a slightly different pattern over the North Atlantic, with a more extended positive signal reaching from southwest to northeast than in ERA5. This is in line with the results from the factor precipitation, where in GloSea5, the North Atlantic precipitation is simulated further west. Studies like Fink et al. (2009) and Pinto et al. (2008) investigated storms from a Lagrangian perspective, but some of their characteristics can also be seen in the here presented Eulerian view. E.g., the dry pole in the northwest of the Atlantic is in line with studies like Fink et al. (2009), which show the general atmospheric state around an extreme cyclone and that a strong cyclone leaves dry air behind. The composites of EGR (400hPa) in ERA5 and

GloSea5 show a strong link of EGR just upstream of the target area (extended region of British Isles). Especially the pattern of GloSea5 is in line with the knowledge that EGR affects strong cyclones in a west-east band through their centre (Pinto et al., 2008) and the cyclone centre is located north of the windstorm field (cf. Leckebusch et al., 2008b), which explains the strong EGR influence north of the North Atlantic windstorm track.

With mostly agreeing on the physical connection between windstorms and individual factors within the observational and model data, these connections may enhance model forecast performance when the individual factors are well forecasted. The individual forecast skills of these factors show high and significant skill in windstorm-relevant regions over the North Atlantic but also some gaps. The forecast skill of the upper tropospheric EGR is significant at the north-easterly end of the Atlantic storm track, an important area for intensifying strong cyclones before making landfall in Europe. Even the forecast skill of the MSLP gradient, SST and Θ_e show significant skill around the British Isles, but the area around 50° N and 40° to 50° W is a gap for these factors. This reduction in forecast skill may link to previous studies, e.g. Scaife et al. (2011), which identify large SST biases in model data.

After the factors have been verified as having the same physical link in observations and models and the model shows forecast skill for important regions of the factor, the third step is connecting the factor performance to windstorm forecast skill. It has been found that all main factors increase the forecast skill of winter windstorms over the British Isles and the North Sea by increasing the forecast skill of relevant factors in storm-relevant regions. SST and Θ_e additionally improve the windstorm forecast skill over Central Europe and Southern Scandinavia. The process-based-view, sub-sampling based on the centre of action from the composite analysis (step 1), is less conclusive. But the factors SST and Θ_e present four centres of action helping to increase the windstorm prediction over Europe when these regions are well predicted.

The overall conclusion from this three-step approach leads to a well-represented connection between the four focused dynamical factors and winter windstorm forecast skill. All four

factors (MSLP gradient, SST, Θ_e & EGR 400hPa) show an agreement in the physical link, as composite analysis and in the stricter correlation maps, suggesting the model does include the physical link overall correctly. The model provides positive forecast skill within relevant regions for all four factors, which means the model performance for the individual factor is positive and well-predicted seasons in these regions support skilful windstorm forecasts.

In addition, the further investigated factors (cf. appendix) show similar results. Well-predicted regions of precipitation and divergence over the tropics and sub-tropics have a positive influence on storm predictability over Europe. For precipitation, this is in line with Scaife et al. (2017b), which found that tropical Atlantic precipitation is an influencing factor for European predictability of atmospheric patterns. Further crucial factors (not shown) in this study were, e.g., the Rossby Wave Source (RWS), SST gradient (total and meridional component) or the North-America/North-Atlantic temperature gradient identified by Wild et al. (2015). The RWS factor did not show a clear pattern or relation. The ERA5 composite is very scattered, but the GloSea5 mean shows at least a pattern agreeing with the conceptional idea of the tropical North Atlantic precipitation triggering convective rising, which triggers the RWS further North (Scaife et al., 2017b). A similar scattered result results from all approach steps for the SST gradients. The temperature dipole from Wild et al. (2015) has been tested, as a connection between North American surface temperature and North Atlantic sea surface temperature anomalies are linked to windstorms over Europe. But the results in this study are not conclusive, probably because the storm target region differs in both studies.

This study concludes that the existing windstorm forecast skill in GloSea5 can be explained by different dynamical atmospheric factors, primary or secondary, connected to cyclones and windstorms. Thus, the model predicts the winter storm season well for the correct reasons, increasing forecast confidence. Secondary and large-scale factors like the MSLP gradient or SST strongly relate to windstorms in the observational and model data sets. Their seasonal forecast skill is high, and seasons which are well predicted have a positive

influence on windstorm forecasts. The same is found for primary factors like Θ_e in 850hPa and EGR in the upper (400hPa) troposphere. This approach results in a new understanding of dynamical factors and covers multiple perspectives, which implies new knowledge about where the windstorms forecast skill might originate and where additional efforts, besides the also for windstorms existing signal-to-noise paradox (Degenhardt et al., 2022), are needed to potentially improve windstorm forecast skill over the downstream end of the North-Atlantic storm track.

6.2 Additional Outcome: Different Views on Assessing the Connection between Dynamical Factors and Windstorms

The initial investigation of windstorms and dynamical factors had multiple perspectives before the above-shown results were summarised in the publication of chapter 6.1. This chapter will present and discuss additional not-published results. Nonetheless, these results lead to a better understanding of the methods and relationships between factors and windstorms. Within publication 2, chapter 6.1, only one perspective is shown on investigating the physical link between windstorms and dynamical factors. Different potential perspectives to investigate the connection between seasonal winter windstorms and dynamical atmospheric factors have been assessed:

(a) **Grid-to-grid:** This perspective aims to understand how storms and dynamical factors locally, per individual grid cell, are connected. This view is an investigation between storm frequency and a dynamical factor on a grid-to-grid basis, which means each grid cell of the storm frequency is compared to the same grid cell of the dynamical factor. The correlation between the respective storm parameter and factor is done for each grid cell individually. A study on a grid cell basis can be rather variable.

(b) **Grid-factor to impact region:** This view is evaluating the spatial distribution of dynamical factors, in regards to a target impact storm region. This investigates from where the targeted storms are influenced by the dynamical factor. This view uses the dynamical factors per grid cell and links them to a storm impact region, e.g., UK (50-60°N, -25-5°E). Therefore, the storm parameters are averaged for this region and correlated to each individual grid cell of the factor. This results in a map of correlations which present the parts over the North Atlantic of the factors that influences this specific windstorm area.

(c) **Grid-storm to factor region:** This perspective is the inverted view of (b), with the same aim of a spatial distribution of the connection between dynamical factor and windstorms, but with regards to a target factor region. This investigates where the targeted factors influence the windstorms. Grid cell-based factors will be averaged in 10°x10° boxes over the whole study region (-100°-40°E & 80-0°N), and these region-averaged factors are correlated to windstorm characteristics per individual grid cell.

(d) **Box-to-box:** This perspective is a combined view to connect an impact storm region with averaged but homogeneously distributed factor regions. This perspective reduces the noise due to an individual grid-cell analysis, compared to view a). And is a combination of view b) and c) as both the target regions of storms and dynamical factors are used. This view is to understand which region of a factor is mostly linked to storms in a target region but without being restricted to individual grid points. A storm impact region is defined, e.g., the UK, and each factor is used as time series of a 10°x10° box average. This view is included in the 2^{nd} publication (chapter 6.1, Fig. 6.4 & C.3).

The additional chapter will focus on the views a) to c). Two example factors, EGR 400hPa (Fig. 6.7) and MSLP gradient (Fig. 6.8), are shown to represent this method, to understand the link between dynamic factors and windstorm frequency and assess the robustness of the physical connection.

6.2.1 Eady Growth Rate (400hPa)

The EGR (400hPa) is one of the primary dynamical factors that chapter 6.1 focuses on. Fig. 6.7 shows views a) to c) for EGR 400hPa. The grid-to-grid view strongly connects EGR and storm frequency over most of the North Atlantic. The maximum correlation at around 35° N and 50° W aligns with many publications showing this area as the storm genesis region. The upper tropospheric EGR is a significant driver of that genesis (Hoskins & Valdes, 1990). The view b) as the grid-factor perspective is used to understand where the link between

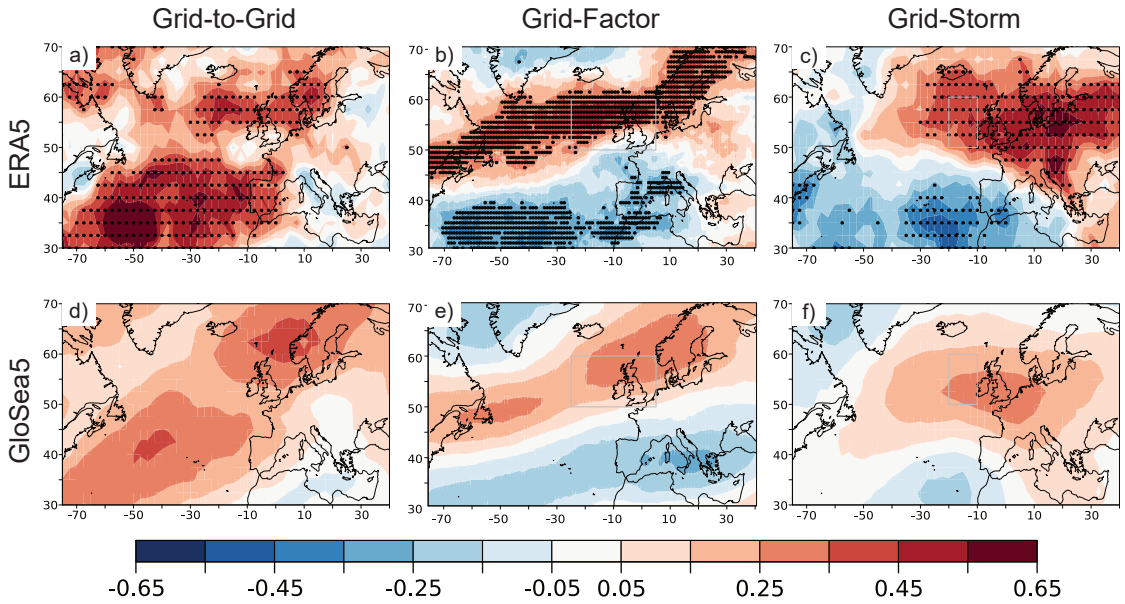


Figure 6.7 Different perspectives on the analysis of the connection between windstorm frequency and EGR in 400hPa, in ERA5 (top) and GloSea5 mean over all member-individual correlations (bottom). The significance correlation-test is only calculated for ERA5 (top row), as GloSea5 (bottom row) is an average over all individual members.

EGR over the North Atlantic is the strongest to windstorms over the British Isles (and the region with the highest forecast skill, Fig. 5.1). This is seen as a whole band from Newfoundland, north of the North Atlantic, over Scotland to Scandinavia. While regions west of the target region are an indicator for causing storms, regions east of the target box might be more consequence of the storms. The opposite perspective of this is view c) which takes storm-related regions of EGR and represents the areas where this factor is linked the most to storms. Here the region of 10° to 20°W and 50° to 60°N is shown as an example (Fig. 6.7). This concludes clearly that EGR (400hPa) affects mostly storms downstream, as highlighted by Pinto et al. (2008). The ensemble mean is expected to be lower, but similar regions show the link between storms and EGR 400hPa in the model.

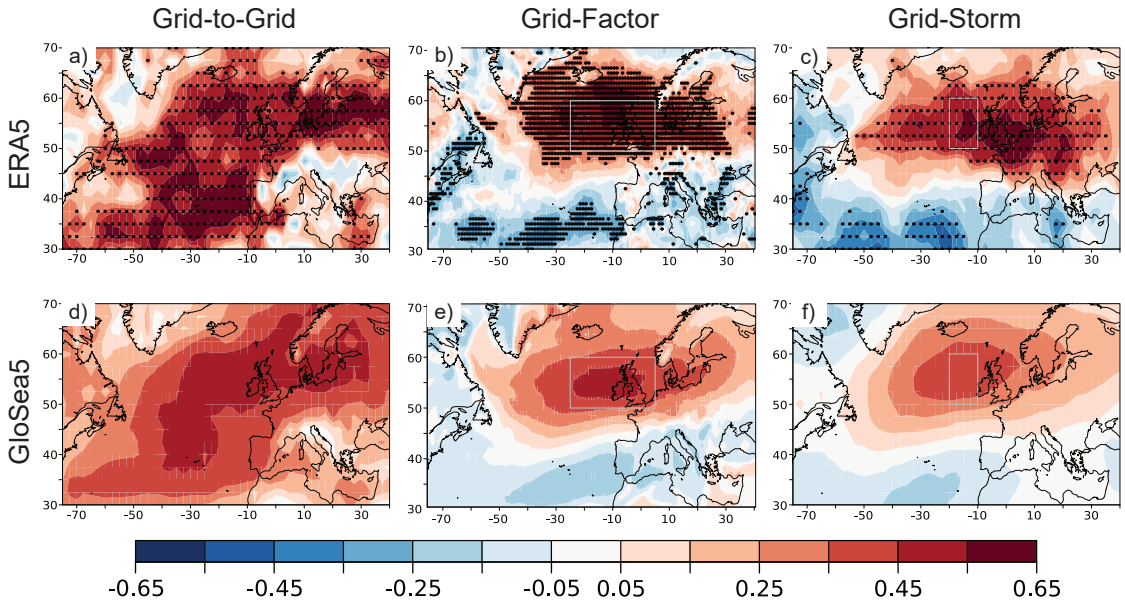


Figure 6.8 Different perspectives on analysing the connection between windstorm frequency and Mean Sea-Level Pressure Gradient, in ERA5 (top) and GloSea5 mean over all member-individual correlations (bottom). The significance correlation-test is only calculated for ERA5 (top row), as GloSea5 (bottom row) is an average over all individual members.

6.2.2 Mean Sea-Level Pressure Gradient

The MSLP gradient is an example of a secondary dynamical factor from chapter 6.1. The grid-to-grid view in ERA5 shows over the whole North Atlantic region a significant link between MSLP gradient and windstorm frequency (Fig. 6.8). MSLP gradient north of the target region are mostly linked to storms over the UK (see grid-factor). This agrees well with existing knowledge that extreme wind speeds are mainly south of the ETCs hitting Europe (Leckebusch et al., 2008b). For the grid-storm perspective (view c), the same factor region as for EGR has been selected. The MSLP gradient in this region is strongly linked to windstorms downstream the box, hence, the British Isles and Central Europe.

Similar to the EGR 400hPa example, the GloSea5 results in Fig. 6.8 show a weaker link than the ERA5 results, but the patterns are similar in all three views. While the grid-to-grid view is generally strongly linked over the North Atlantic, the other two views show slightly shifted centres compared to ERA5. The strongest link appears in the grid-factors view,

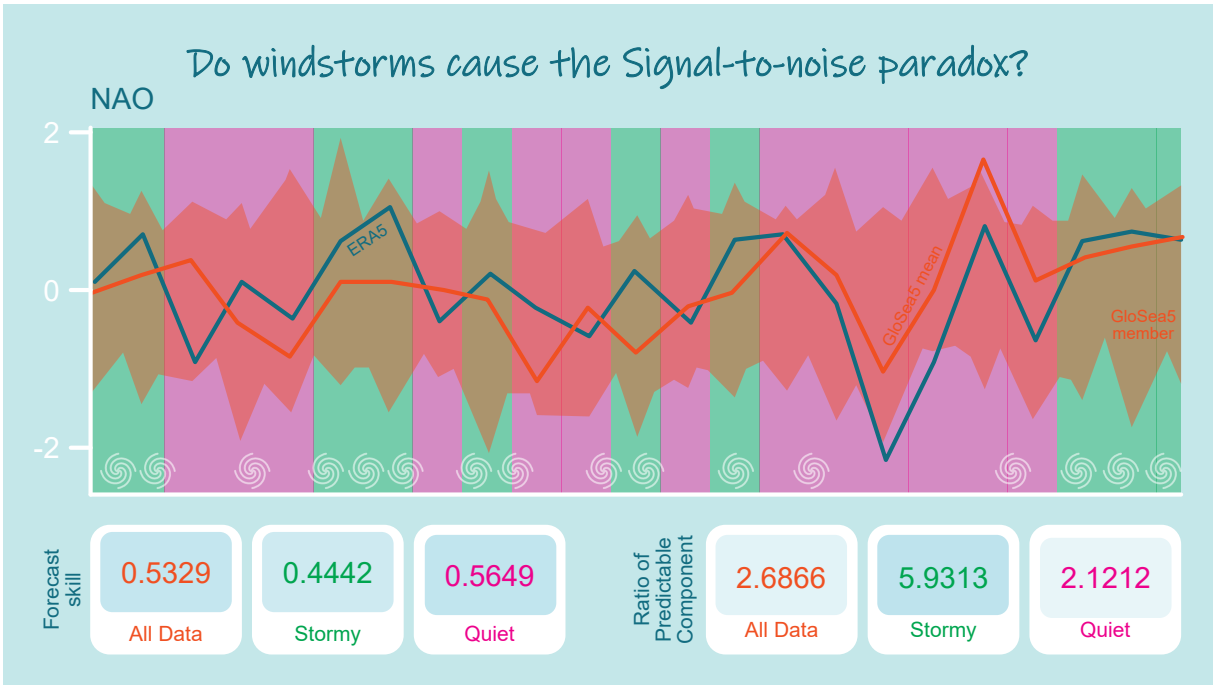
and the grid-storm perspective is more centred on the target regions. This could mean the physical connection between windstorms and MSLP gradient is slightly shifted in GloSea5 compared to ERA5. Nonetheless, chapter 6.1 showed the MSLP gradient as an essential factor for improving windstorm forecast skill, even with that shift.

6.2.3 Conclusion

This chapter investigates different ways of connecting windstorm frequency and dynamical factors in ERA5 and GloSea5. Three additional views have been presented for two example factors (MSLP gradient and EGR 400hPa). These different views are additional information to the published view in chapter 6.1. Each view focuses on a slightly different aspect, but all investigate the physical connection between windstorms and the individual dynamical factor. View a) grid-to-grid is the most complicated perspective for the model as it investigates each grid cell individually, which means the results can be quite scattered. View b) grid-factor and c) grid-storm are contradicting perspectives in terms of the impact. One is showing in which region the factor is influencing storms over a target region (b) grid-factor) and the other is the opposite of which storms are influenced by a factor in a fixed area (c) grid-storms). And all show an agreement of the physical link in ERA5 and GloSea5, which means the results, published in chapter 6.1, Fig. 6.4, are robust even with different perspectives.

Chapter Seven

Signal-to-Noise paradox in seasonal winter windstorm predictions



The third part of this thesis focuses on the signal-to-noise paradox (Eade et al., 2014). The first result chapter revealed that the signal-to-noise paradox also exists in seasonal forecasts of winter windstorms (chapter 5.1.4, Fig. 5.8). It is still the question of where the paradox originates. The first part of this chapter is written as a Research Letter and will be submitted for publication to Atmospheric Science Letters. The idea and concept of this part are from

Lisa Degenhardt, Gregor C. Leckebusch and Adam A. Scaife. A research visit at the Met Office in May 2023 led to cooperation and scientific meetings with Doug Smith and Steve Hardiman. Latter assisted with the eddy feedback calculation and provided his scripts for this part of the study.

7.1 Signal-to-noise paradox dependent on winter windstorm counts over Europe

Abstract

The signal-to-noise paradox exists in different seasonal and decadal ensemble forecasts. The causes for its existence are not fully understood yet but are known to be a limitation for the predictability of forecasts on seasonal to decadal time scales. Multiple studies detect the paradox in forecasting different parameters, especially for the North Atlantic Oscillation (NAO). This study focuses on the fact that the seasonal forecast of severe winter windstorm characteristics is skilful in different forecast suites. However, the signal-to-noise paradox is also present in these predictions. The seasonal multi-model data set from Copernicus (C3S) focusing on the UK Met Office model version 5, GloSea5, is used to investigate if severe winter windstorms impact the known NAO signal-to-noise paradox. The ratio of predictable components (RPC) measures the signal-to-noise paradox, while it exists when the RPC is greater than 1. A sub-sampling approach shows that stormy seasons lead to a higher RPC for the NAO than quiet storm seasons, thus resulting in a stronger signal-to-noise paradox. This is mainly due to a drop in the ensemble mean standard deviation of the NAO. Other models of the C3S ensemble are not revealing the same clear result but show a tendency towards it. The eddy feedback is one potential cause of the signal-to-noise paradox but cannot be identified as the cause of significantly higher RPC in stormy seasons than in quiet seasons for the NAO forecast in GloSea5.

7.1.1 Introduction

The signal-to-noise paradox is a limitation in mid to long-range ensemble predictions. First investigations in understanding and quantifying this paradox are published by Eade et al. (2014) and Scaife and Smith (2018). They established the ratio of predictable components (RPC), which compares the forecast skill of the model (as the correlation between observation and ensemble mean) to the averaged prediction of the model itself (as the averaged correlation between each ensemble member and the ensemble mean). If the RPC is higher than 1, the model can predict the reality better than its ensemble members (Eade et al., 2014). Multiple studies investigated the existence of the signal-to-noise paradox in different model suites and for different parameters, like for the North Atlantic Oscillation (NAO), mean-sea-level pressure (MSLP; Smith et al., 2020) and the Eddy-driven Jet (Marcheggiani et al., 2023) in decadal data, the surface atmospheric temperature in CMIP5 data (Sévellec & Drijfhout, 2019) or the Arctic Oscillation (Hardiman et al., 2022) and East-Atlantic Pattern (Thornton et al., 2023) in seasonal forecast suites. In addition to these diagnostic studies, first attempts are published to understand the signal-to-noise paradox and where it originates. Williams et al. (2023) found that El-Nino-Southern-Oscillation (ENSO) is important for extratropical predictability but is unlikely the direct cause for the signal-to-noise paradox. Other studies analysed new model simulations to understand the signal-to-noise paradox, like Knight et al. (2022). They reduced the error and dynamics in the tropics but still found the signal-to-noise paradox for the extra-tropics, even with relaxed tropics. One investigation implies a connection between the signal-to-noise paradox in the NAO forecast and an eddy feedback parameter (EFP) (Hardiman et al., 2022). They investigated six different seasonal model suites of the RPC and their relation to an EFP. They summarise that higher eddy feedback (as the correlation between eddy forcing and the averaged zonal wind) is connected to a lower signal-to-noise paradox and that multiple models underestimate the eddy feedback. Correcting the eddy feedback in the seasonal forecast could double the seasonal forecast skill

of the Atlantic Oscillation for extratropical regions (Hardiman et al., 2022). The EFP as used in Hardiman et al. (2022) is the correlation of how much energy is transported between eddies (meaning disturbances of the zonal mean) and the zonal mean itself (as described in Lorenz, 1955).

Synoptic-scale eddies in the extra-tropics are disruptions in the jet stream, causing synoptical circulations like cyclones and hence, extreme wind events. Recent studies showed that cyclones (Befort et al., 2019) and windstorm frequencies and intensities (Befort et al., 2019; Degenhardt et al., 2022) are predictable on a seasonal scale. The latter showed that windstorm frequency and intensity also have an RPC greater than one (Degenhardt et al., 2022) with a higher RPC, up to 5, for two different intensity measures over the British Isles.

This study aims to investigate how big the potential influence of windstorms on the signal-to-noise paradox is because of the assumption that eddies have a huge influence on the signal-to-noise paradox and the knowledge that a strong paradox exists in the storm forecast. Therefore, the NAO forecast is used as a reference study, as it is known to have a high RPC (Smith et al., 2020). The signal-to-noise paradox of NAO will be investigated for different sub-samples to distinguish the influence of windstorms. Seasons with many storms represent the sub-sample *stormy seasons*, and with few storms, the sample *quiet seasons*.

This introduction will be followed by a short Data and Method section (section 7.1.2 & 7.1.3). The results are separated into a multi-model investigation of the NAO-RPC in section 7.1.4 and an analysis to assess the found signal (section 7.1.5). This section finishes with a short Discussion and Conclusion (section 7.1.6).

7.1.2 Data

This study investigates seasonal forecast ensemble models compared to the observational data set, ERA5 from the ECMWF (Hersbach et al., 2019). The UK Met Office's Global Seasonal Forecasting System version 5 (GloSea5) is mainly used (MacLachlan et al., 2015), but mul-

tiple European seasonal forecast suites have been assessed (not all shown). All data sets are available on the Copernicus Climate Change Service (C3S; Copernicus Climate Change Service, 2023c), German weather service (DWD; Fröhlich et al., 2021), European Centre for Medium-Range Weather Forecasts (ECMWF; Johnson et al., 2019), Euro-Mediterranean Center on Climate Change (CMCC; Gualdi et al., 2020), and Météo France (MF; Batté et al., 2021). The main focus of this study is on GloSea5 because it is known to have a high RPC for NAO (Scaife & Smith, 2018) and a skilful windstorm prediction over Europe (Degenhardt et al., 2022). Model suites like DWD and CMCC will be used for a multi-model approach because they show, besides GloSea5, high forecast skill for the NAO (close to 90% significance). Models like the ECMWF and MF have been investigated but are not presented in this section as they have very low forecast skills for the NAO.

The following parameters have been used for all data sets: the mean sea-level pressure (MSLP) for the NAO-index and further investigations and wind components in 500hPa for the EFP. The windstorm tracking used to separate into stormy and quiet seasons is only calculated for ERA5 with 10m wind speed. All data sets have been used on their original spatial resolution. Analysis has been done for the core winter (December-January-February) and the period from 1993 to 2016 (23 seasons).

7.1.3 Method

The NAO can be calculated in multiple ways (Leckebusch et al., 2008a). The Empirical Orthogonal Function (EOF) followed by a Principle Component Analysis (PCA) has been used on the observational data, each ensemble member and a total model data set (overall members) to obtain the individual NAO indices. As the EOF patterns of ERA5 and GloSea5 are very similar, the GloSea5 data is projected onto the ERA5 pattern in the PCA for a better comparison. The NAO index is calculated with monthly MSLP data, but the analysis is performed with the obtained seasonal mean. The NAO index is per model member indivi–

Table 7.1 List of storm counts in different regions over North Atlantic and Europe and the allocated sub-sample, incl. NAO-separation.

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
-15°-10°E & 48°-60°N	6	6	4	4	4	5	5	4	6	2	1	6	3
Stormy													
Quiet													
-15°-20°E & 48°-64°N	8	6	5	6	5	5	6	4	8	2	3	7	3
Stormy													
Quiet													
-40°-20°E & 48°-64°N	13	8	8	8	6	9	7	7	11	8	5	8	3
Stormy													
Quiet													
50°-10° W & 40°-60°N	14	6	8	6	8	10	6	10	11	10	4	7	5
Stormy													
Quiet													
NAO	0.094	0.672	-0.954	0.116	-0.358	0.606	1.032	-0.437	0.190	-0.241	-0.599	0.241	-0.450

Continue of Tab. 7.1.

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
	8	3	4	3	3	4	3	14	6	8
S										
Q										
	8	5	4	3	4	4	3	14	6	8
S										
Q										
	12	9	6	5	10	6	5	18	8	9
S										
Q										
	12	8	9	12	12	6	10	21	9	10
S										
Q										
	0.589	0.704	-0.175	-2.135	-0.925	0.773	-0.645	0.612	0.707	0.581

dually standardised, meaning each NAO time series has been subtracted by its individual mean and divided by its standard deviation before being used in any further calculations.

This study separates the whole data set into stormy and quiet seasons. This is based on ERA5 storm tracks, tracked with an impact-based algorithm (Leckebusch et al., 2008b). The sub-samples are defined over the number of storms crossing a pre-defined region. Four regions with different reasons have been used in this study for the sub-sampling. Fig. 7.1 shows the defined regions, and Tab. 7.1 summarises the storm counts and separation for each region. Region 1 focused on the UK (-15-10°E & 48-60°N) includes the centre of skilful seasonal windstorm frequency predictions

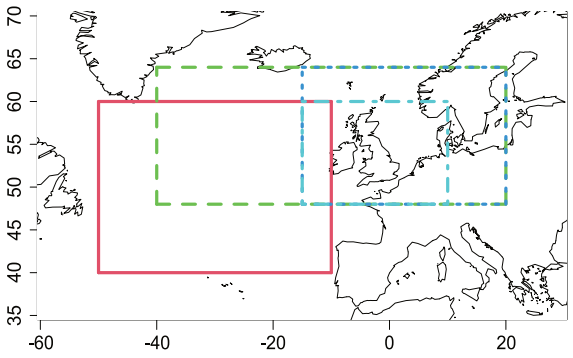


Figure 7.1 Regions for different separation per storm counts. Region 1 - -15-10°E & 48-60°N (light blue), region 2 - -15-20°E & 48-64°N (dark blue), region 3 - -40-20°E & 48-64°N (green), region 4 - 50-10°W & 40-60°N(red).

found by Degenhardt et al. (2022). Region 2 (-15-20°E & 48-64°N) is a similar box but slightly extended to southern Scandinavia, which is also skilful in windstorm predictions. Regions 3 (-40-20°E & 48-64°N) and 4 (50-10°W & 40-60°N) are selected with the reason of the Atlantic storm tracks and represent areas of more storms. Regions 3 and 4 have been investigated but do not show as clear and significant results as the separation with storms from regions 1 and 2, hence, only these two regions are shown in the following chapter.

The ratio of predictable components (RPC) is used to identify the signal-to-noise paradox. In this study, a similar equation (see eq. 7.1) as in Eade et al. (2014), Scaife and Smith (2018), Smith et al. (2020), and Hardiman et al. (2022) is used.

$$RPC = \frac{r_{mo}}{r_{mm}} \simeq \frac{r_{mo}}{\frac{\sigma_{ensmean}}{\sigma_{total}}} \tag{7.1}$$

All the different components will be presented in the results, r_{mo} as the correlation between observations and model ensemble mean, r_{mm} as the averaged correlation of each ensemble member to the ensemble mean, $\sigma_{ensmean}$ as the standard deviation of the ensemble means time series and σ_{total} as the standard deviation over all members, hence the total model. The used measures, Pearson's correlation and standard deviation, are used as uncentred versions. This means that while using sub-samples (stormy, quiet, NAO+ or NAO–), the statistical measure is not centred on the sub-samples mean but on the mean of all data. This is done to ensure a clear comparison of the sub-sample to the results of all data. Eq. 7.2 & 7.3 indicate the used equations, commonly used standard deviation and Pearson's correlation equations with different means of the individual data sets.

$$sd = \sqrt{\frac{\sum(x_{i\ subsample} - \overline{x}_{all\ data})^2}{n-1}} \tag{7.2}$$

$$cor = \frac{\sum(x_{i\ subsample} - \overline{x}_{all\ data})(y_{i\ subsample} - \overline{y}_{all\ data})}{\sqrt{(x_{i\ subsample} - \overline{x}_{all\ data})^2 (y_{i\ subsample} - \overline{y}_{all\ data})^2}} \tag{7.3}$$

Further investigations in this study investigate the EFP. The EFP is calculated as in Hardiman et al. (2022) and Scaife et al. (2019a). They used the horizontal Eliassen–Palm-flux divergence (Eq. 7.4, with ρ is density, a the radius of the earth, ϕ is the latitude, (u, v) are the zonal and meridional wind components with $'$ being the difference to the respective zonal mean) as a seasonal average and correlated it with the seasonal zonal mean zonal wind $\bar{u}$.

$$\frac{\nabla \cdot \mathbf{F}_H}{\rho a cos\phi} = \frac{-1}{a cos^2\phi} \frac{d(\overline{u'v'} cos^2\phi)}{d\phi} \tag{7.4}$$

This parameter represents the connection between the zonal Eddy forcing and the general zonal wind. As higher this parameter's value, as stronger the energy exchange between the Eddy Kinetic Energy and the zonal mean kinetic energy in the Lorenz energy cycle (Lorenz, 1955).

7.1.4 Influence of windstorms on NAO paradox

To understand the influence of severe winter windstorms on the signal-to-noise paradox, Fig. 7.2 shows the RPC and all related components for the NAO in different sub-samples in different seasonal forecast models. The results in Fig 7.2 used region 1 as the basis for the sub-samples. Fig. 7.4 shows an example for region 2. The bottom row of Fig. 7.2a (GloSea5) represents all components of the RPC in the total data set of 23 seasons within the UKMO model, GloSea5. The RPC of 2.69 for the NAO forecast is similar to values that have been found in previous studies for that model (e.g., Scaife & Smith, 2018).

To understand the influence of storms on the signal-to-noise paradox, the same RPC, incl. all components, has been calculated for the two sub-samples, stormy and quiet seasons. This means the 10 seasons with the most storms in the predefined region create the sub-sample of stormy seasons. The sub-sample of quiet seasons, in this case, has 13 seasons because ERA5 storm counts are integers, and the ranks 8 to 13 have the same amount of storms (see Tab. 7.1). These two sub-samples show that the NAO-RPC for stormy seasons is significantly higher than a random sampling (significance tested by bootstrapping with 10.000 random

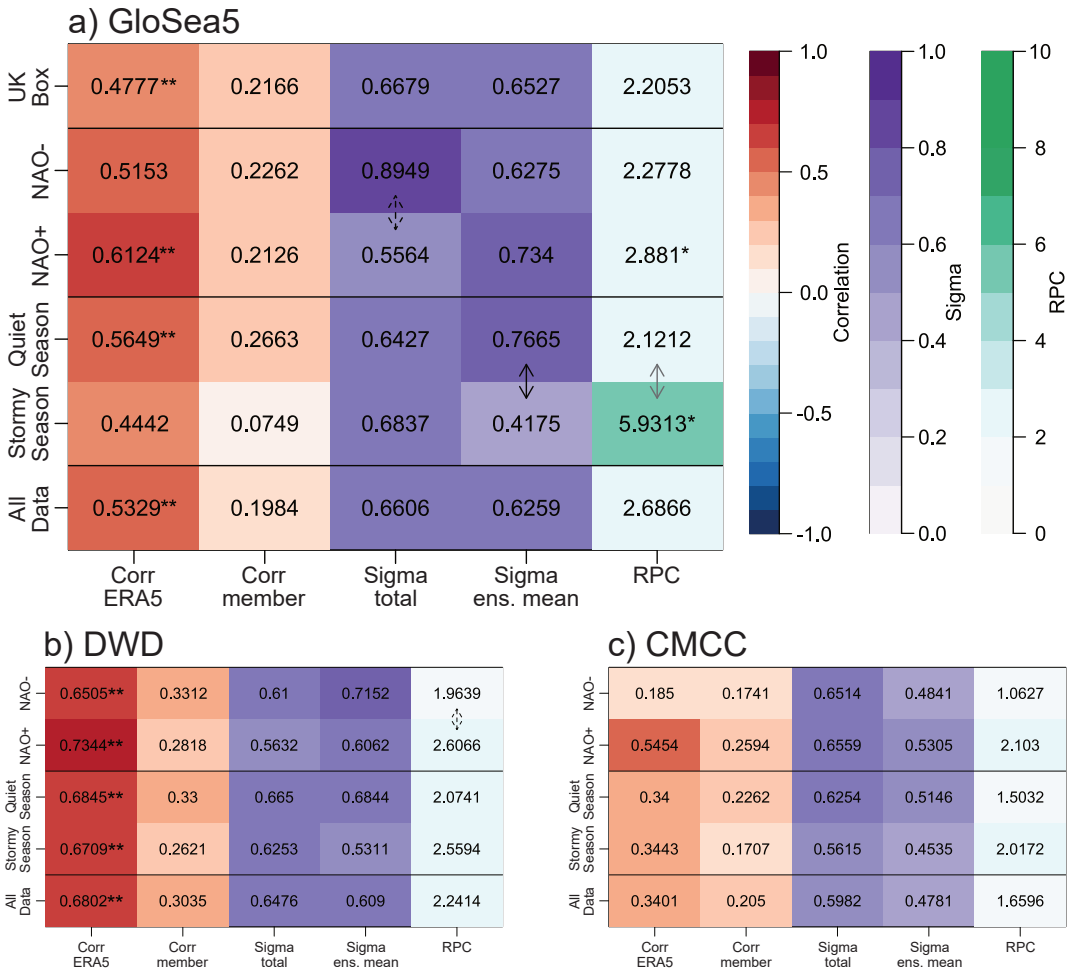


Figure 7.2 Ratio of Predictable component (RPC - green legend) and the individual components (r_{mo} & r_{mm} - blue-to-orange legend, σ_{total} & $\sigma_{ensmean}$ - purple legend) for NAO in the whole data set, in stormy, quiet, NAO+ and NAO– seasons. For selected C3S data sets. Separation for region 1 (-15°-10°E & 48°-60°N). Significance tested with correlation test for correlation columns, rest by 10.000 bootstrapping, * is 90%, ** is 95% significant. (Dashed) Arrows mark when two values are more than 2σ (1.7σ) apart. Grey arrows represent the test after removing 5% outliers in bootstrap distribution.

resamplings). The RPC for quiet seasons is similar to the RPC with all data. Due to the different components of RPC, it can be concluded that this change in RPC is resulting from a drop of the r_{mm}, which is related to the lower $\sigma_{ensmean}$ for stormy seasons compared to quiet seasons. The arrow between stormy and quiet seasons shows that these values are 2 standard deviations from each other (dotted arrow means 1.7 standard deviations $\sim 90^{th}$ percentile, Pukelsheim, 1994). This aligns with the idea that stormy seasons have a higher

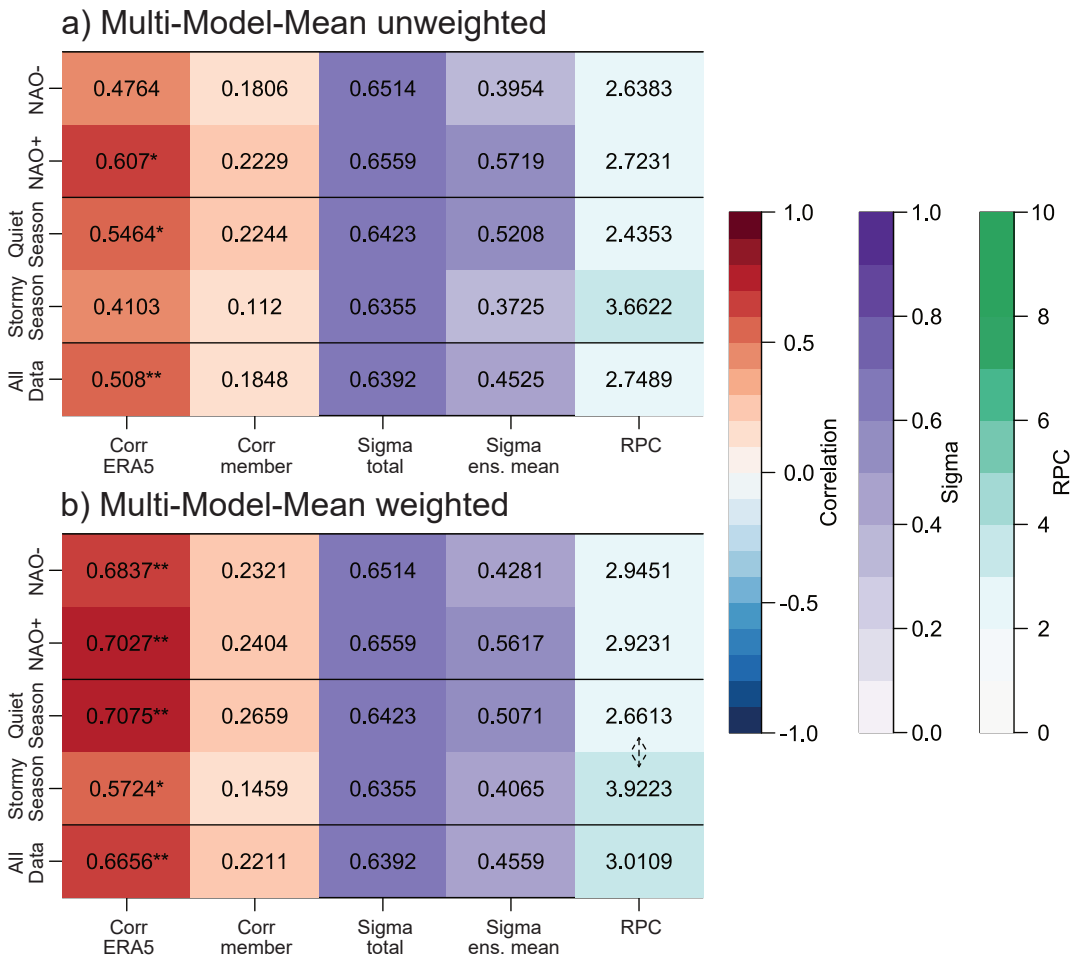


Figure 7.3 As Fig. 7.2 for Multi-Model-Mean approaches in region 1 (-15°-10°E & 48°-60°N).

RPC, hence, a higher signal-to-noise paradox. It is known that stormy seasons tend to be in NAO+ phases (Donat et al., 2010). Hence, a sub-sampling into NAO+ and NAO– has additionally calculated (see separation in Tab. 7.1). 8 out of 10 seasons are in sub-sample stormy seasons, and NAO+ and all NAO– years are within the sub-sample of quiet seasons. Still, Fig. 7.2a shows that even the RPC for NAO+ seasons is slightly increasing but is only significantly higher on a 90% level than random sampling. The difference in the RPC for NAO+ and NAO– seasons has another reason than the increase for the stormy season RPC. The NAO+ seasons show a higher forecast skill (r_{mo}) than the NAO– seasons. This is due to a higher (1.7σ) σ_{total} in the NAO– seasons.

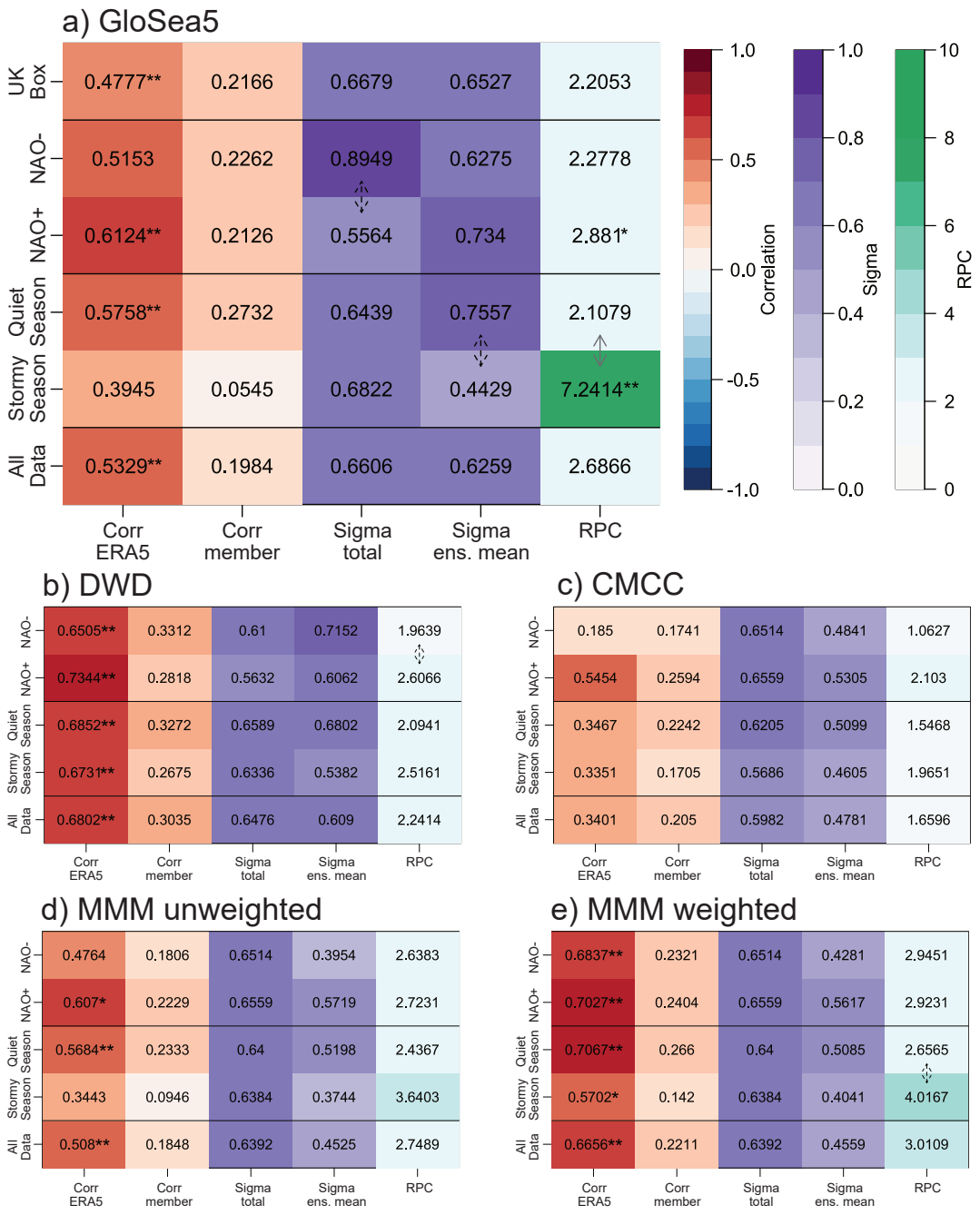


Figure 7.4 As Fig. 7.2 & Fig. 7.3 for region 2 (-15°-20°E & 48°-64°N).

Another approach to identifying if severe winter windstorms impact the NAO predictability is to remove data points from the data set that include storms. This has been assessed with multiple definitions of a storm data point, all storms over a specific region (North Atlantic, UK) or the strongest 30/50 storms in a region or the longest 30/50 storms in a region. Only

one set is presented in Fig. 7.2a as the results are all similar. The example set is calculated by removing all time steps that include a storm that affects the British Isles. This new data set without UK storms has been used to calculate the NAO index, and the top row of Fig. 7.2a shows the RPC results for this index (only for GloSea5). The RPC for the NAO without UK storms is not significantly lower than usual, and the skill (r_{mo}) is still significant for the NAO. The RPC drops slightly due to a slight increase in the averaged member correlation (r_{mm}). This led to the result that the signal-to-noise paradox in the NAO forecast is hardly influenced by the storms, as there is no difference in the RPC if no storms over the UK exist. The GloSea5 results show that the strongest and most significant signal appears with a sub-sampling of stormy and quiet seasons. Hence, this has been done for more available seasonal forecast models in the C3S data collection. Fig. 7.2 includes the results for GloSea5, CMCC and DWD, as these are the only three model suites with a forecast skill of NAO close to 90% significance. Hence, these three systems are used for a multi-model-mean (MMM) approach (Fig. 7.3). The MMM has been performed with two different approaches, firstly "unweighted", means all ensemble members have been used, GloSea5 (63 members), DWD (30 members) and CMCC (40 members), and secondly "weighted", means that each model suite was averaged to an individual model mean before used for the MMM. The same pattern of a stronger RPC in stormy seasons appears in both individual model suites (CMCC and DWD, Fig. 7.2b & c) and both MMM-approaches (Fig. 7.3a & b), but are all not significant on a 90% level. The stormy RPC values in the MMM (Fig. 7.3) are above the 70% percentile of a random sampling distribution. The stormy and quiet RPCs of the MMM weighted approach differ on a 90% level. DWD (Fig. 7.2c) shows that the sub-sampled RPCs for NAO+ and NAO– prediction are 1.7 standard deviations from each other, means separated by roughly 90% of the data, but individually not significant on a 90% level.

Fig. 7.4 includes the individual model and MMM results for a second example region. Region 2 is an extended version of region 1. Hence, a similar result was expected. It is seen that the RPC difference between stormy and quiet seasons in GloSea5 (Fig. 7.4a) is persistent

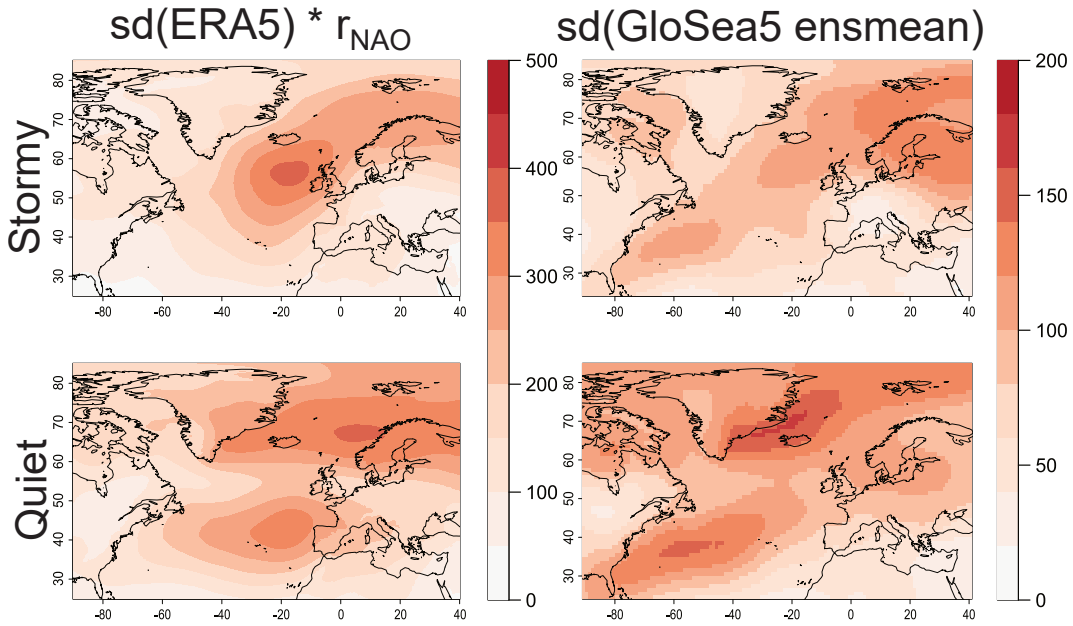


Figure 7.5 Mean sea level pressure predictable component of the observations as standard deviation of ERA5 multiplied with the forecast skill (left) and the predictable component of the model ensemble mean as standard deviation of the ensemble mean (right) for stormy (top) and quiet (bottom) seasons in GloSea5 for region 1 (-15°-10°E & 48°-60°N).

and even increases, the same for the weighted MMM (Fig. 7.4e).

7.1.5 Potential reasons

Eddies are one of the expected causes of the signal-to-noise paradox, with the argument that the eddy feedback is underestimated in ensemble models (Hardiman et al., 2022). Synoptical, near-surface representations of eddies are cyclones and storms. Hence, it is assumed that stormy seasons have more eddies and might be more affected by a potential misrepresentation (underestimation, Hardiman et al., 2022), which causes the signal-to-noise paradox. Fig. 7.5 and Tab. 7.2 represent calculations along this hypothesis for the storm separation by region 1. The MSLP predictable components (Fig. 7.5) are used as a spatial representative for the NAO. It shows ERA5 has a clear difference between the stormy seasons and the quiet seasons, while GloSea5 is not reproducing this variability as clearly. In stormy seasons GloSea5 (Fig.

7.5b) cannot reproduce the signal of a strong pole downstream the British Isles as in ERA5 (Fig. 7.5a). In the quiet season, the pattern is similar, but the two poles are further east in ERA5 (Fig. 7.5c) than in GloSea5 (Fig. 7.5d).

The EFP (Tab. 7.2) is a parameter to represent the strength of the connection between the eddy forcing and the general zonal wind. This parameter describes one component in the Lorenz energy cycle (Lorenz, 1955). A lack of this process could explain the signal-to-noise paradox (Hardiman et al., 2022). The values for ERA5 and GloSea5 in different sub-samples are listed in Tab. 7.2. The GloSea5 correlations are lower than the ERA5 EFPs. More important, the EFPs in the stormy seasons are higher in both data sets, ERA5 and GloSea5. The EFP was expected to show a stronger error in stormy seasons than in quiet seasons between ERA5 and GloSea5, which is not the case for these regions.

Table 7.2 Eddy feedback parameter (like Hardiman et al., 2022) for all 23 seasons and as sub-samples from stormy and quiet seasons in ERA5 and GloSea5 ensemble mean, incl. model member error as $2\sigma_{member}/\sqrt{n_{member}}$, for all tested regions.

	ERA5	GloSea5	Error
All	0.323	0.253	0.014
-15°-10°E & 48°-60°N			
Stormy	0.351	0.301	0.022
Quiet	0.309	0.261	0.020
-15°-20°E & 48°-64°N			
Stormy	0.320	0.296	0.020
Quiet	0.349	0.262	0.021

7.1.6 Discussion and Conclusion

This study aims to understand further the cause of the signal-to-noise paradox (e.g., Eade et al., 2014), which is found in seasonal and decadal ensemble forecasts. Previous investigations presented that the winter windstorm seasonal prediction includes the signal-to-noise paradox in seasonally averaged storm amount and intensity measures (Degenhardt et al., 2022). An especially high RPC was found for the latter parameter. Other studies investigated the cause of the signal-to-noise paradox and found the EFP is linked to the paradox and might be a potential reason (Hardiman et al., 2022). As windstorms are eddies from the mean zonal

flow, this could mean that windstorms are part of the reason for the signal-to-noise paradox. This study analysis if severe winter windstorms influence the strength of the paradox within the NAO seasonal prediction.

This study investigates five seasonal forecast models (GloSea5, ECWMF, DWD, CMCC & MF). However, only three (GloSea5, DWD & CMCC) have a general NAO forecast skill close to 90% significance. Hence, they are chosen to be presented in this study. All models show an RPC greater than one for the NAO forecast from 1993 to 2015. The UK model, GloSea5, shows the highest RPC with 2.6866. CMCC and DWD show an NAO forecast skill close to or over the 90% significance and also show an RPC greater one for 23 seasons. These three seasonal models show combined high forecast skill in the multi-model-mean approach but also result in a high RPC, 2.7489 (unweighted) and 3.0109 (weighted). This study investigates if different atmospheric configurations influence the RPC based on respective sub-samples of seasons. The stormy sub-sample, with the 10 most stormy seasons in reality, shows a significantly higher RPC only for GloSea5. Originally the 10.000-bootstrap did not show a significant difference between the stormy and quiet RPC for GloSea5. Still, analysis of the bootstrap distribution with whisker-box plots shows a high rate of outliers (Aguinis et al., 2013). Removing the 5% highest and lowest outliers for the distribution of bootstrapped RPC reduces the standard deviation and shows that the stormy & quiet RPCs are significantly different. In the MMM approaches, this pattern is reproduced but with a lower and less significantly higher RPC for stormy seasons. This could be justified by the fact that only GloSea5 is so far known to be skilful in storm frequency forecasts and has a high signal-to-noise paradox in windstorm forecasts (Degenhardt et al., 2022). The highest total NAO-RPC from the here presented individual models exists in the GloSea5 model. The reduced RPC for DWD or CMCC could be due to the definition of regions separating into stormy and quiet seasons. This is based on the GloSea5 forecast skill and not the individual model regions with skilful storm forecasts. However, the weighted MMM show a significant difference between the stormy RPC and quiet RPC on a 90% level (1.7σ

difference, Pukelsheim, 1994) in both regions.

The separation into stormy and quiet seasons is based on four different regions as references, chosen for different reasons. The regions presented in this section are the two that show the clearest increase for the stormy sub-sample in NAO-RPC but are simultaneously the regions based on the found forecast skill in windstorm characteristics. This means the influence of windstorms on the NAO signal-to-noise paradox is the strongest in regions that show a significant forecast skill in windstorm predictions and the highest signal-to-noise paradox in windstorm forecasts.

This study shows a significant increase in RPC for stormy seasons in the GloSea5 seasonal model. It is robust in similar regions but insignificant in other seasonal models or a multi-model approach. The higher RPC in stormy seasons results from a lower standard deviation of the ensemble mean, which reduces the forecast skill slightly but mainly reduces the average of member correlations to the ensemble mean (significant with 1.7σ difference). This might be in relation to the difference in the seasonal MSLP pattern over the North Atlantic. The predictable component of MSLP is higher in ERA5 than the standard deviation of the GloSea5 ensemble means. Besides the strength of the pattern, the general structure of the MSLP is not well represented in GloSea5. The strong pole at the end of the Atlantic storm track in ERA5 for stormy seasons is not reproduced in GloSea5.

The eddy feedback has been assessed as a potential cause for the increased RPC in stormy seasons. Hardiman et al. (2022) investigated the eddy feedback in relation to the RPC in multiple seasonal models. They found an underestimation of the EFP in the different models and that a potential improvement could reduce the signal-to-noise paradox (Hardiman et al., 2022). The same EFP was used in this study, resulting in similar values for ERA5 and GloSea5 for all data as in Hardiman et al. (2022). However, the EFPs for stormy and quiet seasons do not confirm the assumption that the eddy feedback might be even stronger underestimated for stormy than for quiet seasons and hence, causing the higher signal-to-noise paradox. However, GloSea5 underestimates the EFP compared to ERA5, in general.

The initial objective of this study was to investigate the influence of windstorm activity on the signal-to-noise paradox. Therefore, the NAO forecast has been used and investigated for stormy and quiet sub-samples. This study concludes that storms might influence the signal-to-noise paradox but are not the full explanation, given that removing storm time steps has not decreased the RPC and the signal is not coherent between different models. However, the strongly increased RPC in stormy seasons in GloSea5 is robust within slightly different regions. Hence, the windstorm activity has a potential influence, although only significant in GloSea5.

7.2 Additional Outcome: The Signal-to-noise Paradox in Seasonal Predictions of Storm-related Dynamical Factors

The RPC has been investigated for different storm characteristics (see chapter 5.1) and for the NAO and different sub-samples of the NAO (see chapter 7.1). However, the established dynamical forcing factors (see chapter 6.1) can show signs of the signal-to-noise problem within their forecast, which would make it more difficult to reach the actually possible predictability. As these factors influence storm predictions, a potential signal-to-noise paradox within the factors could be a reason for a reduced skill in storm predictions. This additional chapter investigates the signal-to-noise problem in the dynamical factors analysed in chapter 6.1, with the focus on the two secondary factors, MSLP gradient and SST, and the two primary factors, Θ_e and EGR (400hPa). The RPC maps for all other factors are in Appendix D, Fig. D.1. More details and further explanations about the signal-to-noise paradox can be found in the Literature overview (chapter 2.5) and more about the RPC-calculation in the Methods (chapter 4.3).

The major studies published so far on the signal-to-noise paradox and the RPC are investigating mainly the NAO forecast (e.g., Eade et al., 2014; Scaife & Smith, 2018). Chapter 5.1 shows the first time the RPC for storm characteristics (see Fig. 5.8). An RPC greater than 1 indicates the signal-to-noise paradox to exist in the individual forecast. For storm frequencies, the RPC is >1 downstream of the North Atlantic storm track, over the British Isles, the North Sea and Scandinavia (chapter 5.1, Fig. 5.8a). The storm intensity measures show strong RPC-values over the British Isles with values higher than three close to Ireland (chapter 5.1, Fig. 5.8b-c).

The second part of this thesis connects primary and secondary dynamical factors to windstorms and the windstorm seasonal forecast skill. The four main factors investigated, MSLP

gradient, SST, Θ_e and EGR, show an increase in windstorm forecast skill when specific regions (skilful predicted regions) are well predicted (see Fig. 6.6 and respective discussion in chapter 6.1). The areas for the MSLP gradient, which are regions with significant MSLP gradient forecast skill, are a combination of a box over the British Isles and the North Sea, the southern North Atlantic and close to America's east coast. SST is well predicted over the whole North Atlantic, and in seasons where this region is well predicted, the storm forecast skill is improving over Europe. Θ_e showed a substantial increase in European seasonal windstorm forecast skill when a region over the UK, south of Greenland, and south of Newfoundland is well predicted in Θ_e. The EGR for the upper troposphere (400hPa) was only skilful predicted in a small region over the North Atlantic between 50° and 60°N. However, correct EGR forecasts in this region have a small but positive effect on the windstorm forecast skill over Europe.

Fig. 7.6 shows the RPC per grid cell for the four most important factors for seasonal storm predictions (MSLP gradient, SST, Θ_e & EGR). An RPC higher than one implies the existence of a signal-to-noise paradox.

Fig. 7.6a shows the RPC for the MSLP gradient. The signal-to-noise paradox (RPC>1) exists as a coherent region over the northern North Atlantic and the British Isles, even with values higher than three around Scotland. SST predictions (Fig. 7.6b) have very low RPC values, with only a few values above one around the British Isles, which means that the found forecast skill for SST is robust and can be assumed as being the real skill (Fig. 6.5b). The primary dynamical factor, Θ_e, shows a similar region of high RPC values like MSLP gradient (Fig. 7.6c). The RPC values show a signal-to-noise paradox for an extensive part of the northern North Atlantic, including the British Isles. GloSea5 predictions of EGR 400hPa have a high RPC mainly over the northern parts of the North Atlantic, with scattered grid cells of an RPC over three.

The information gained from chapter 5.1 - Fig. 5.8 and from chapter 6.1 - Fig. 6.6 are nicely connected due to the results from this additional result chapter and Fig. 7.6. The areas of

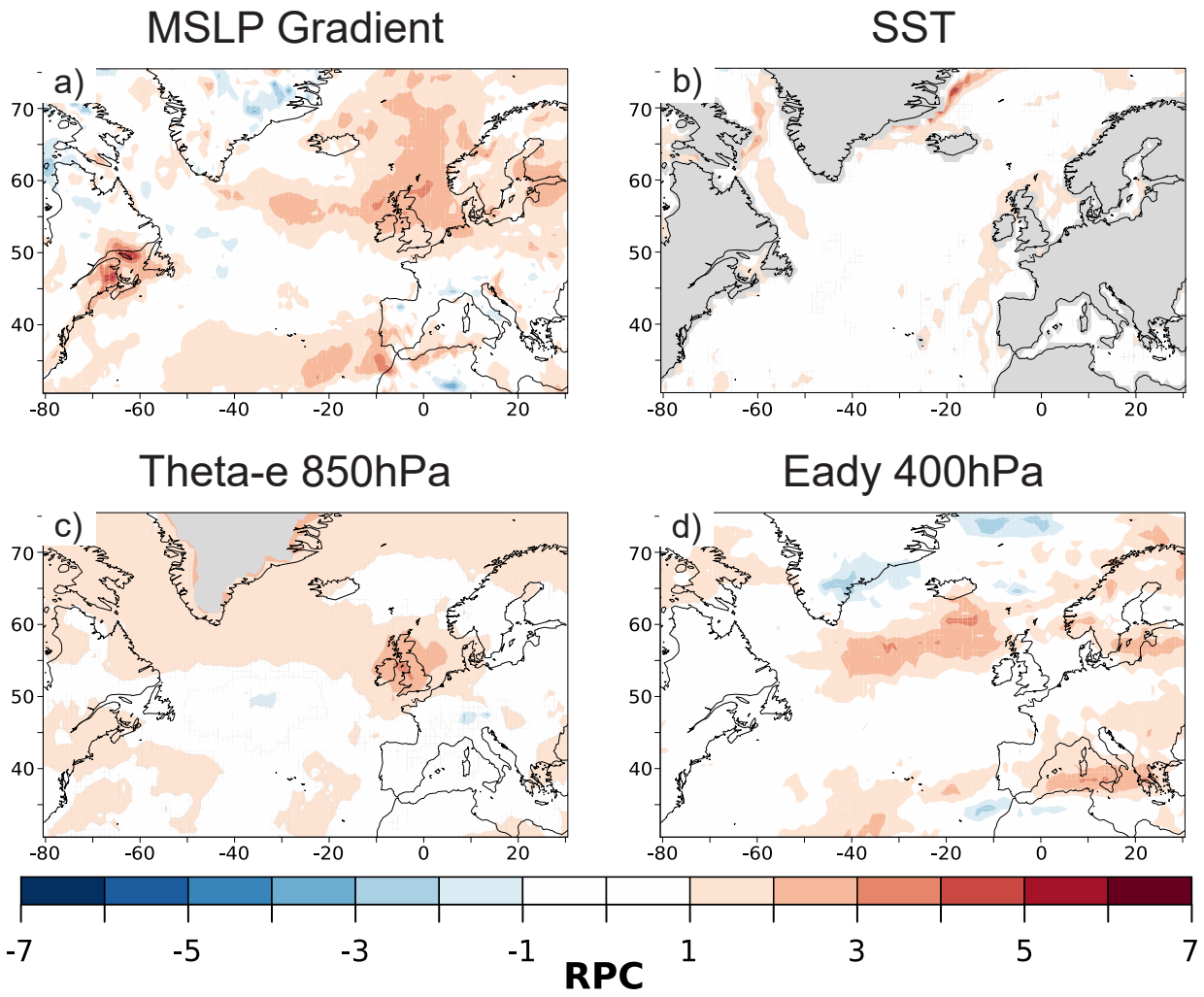
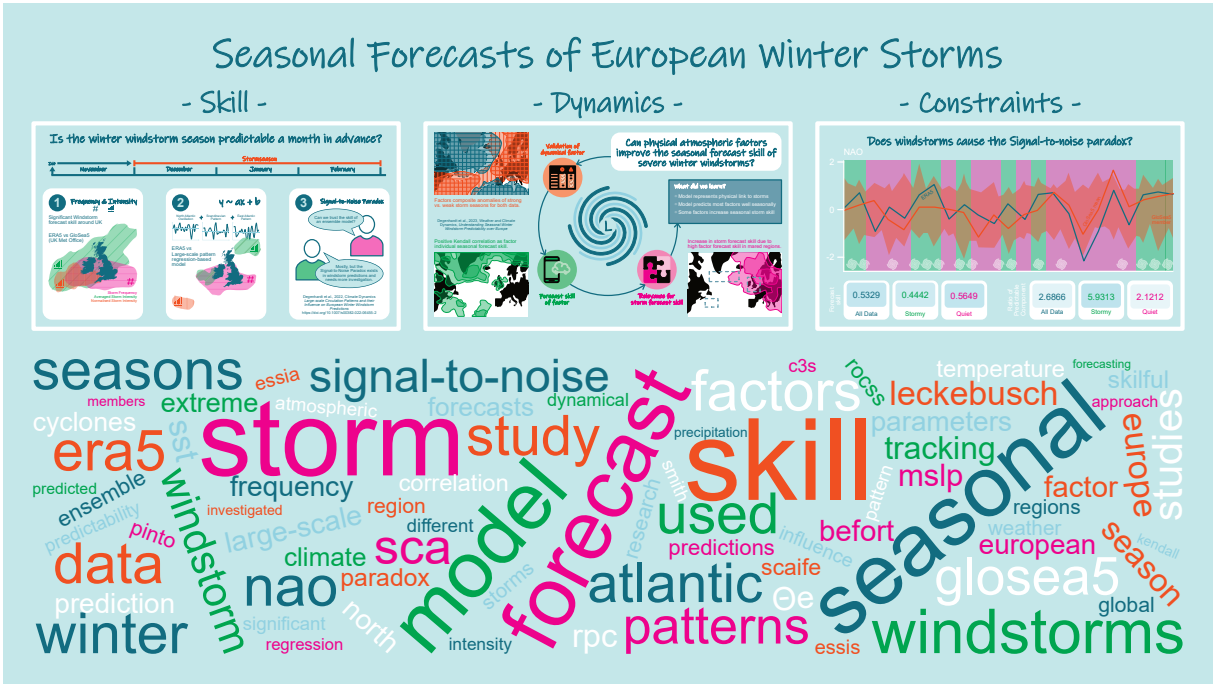


Figure 7.6 Ratio of Predictable component (PRC) for the four main influencing dynamical factors, only values higher than one are coloured.

high RPC in MSLP gradient, Θ_e and EGR 400hPa are similar to the areas of high RPC in storms, which agrees to the idea that these factors are mainly influencing winter windstorms and also have a high impact on the forecast skill of seasonal windstorm characteristics. But these regions are also the same as found in chapter 6.1, Fig. 6.6, influencing the forecast skill of winter windstorms. The same regions have a high RPC/the signal-to-noise paradox and are essential to improve the windstorm forecast skill. This means the factor-own forecast skill has the potential to be higher which could enhance the storm forecast skill, as shown in chapter 6.1.

Chapter Eight

Synthesis



8.1 Overall Summary

This thesis investigates seasonal forecasts of severe winter windstorms over Europe. The three main result chapters present new investigations and gained knowledge about forecast skill (chapter 5), the physical dynamics leading to the skill (chapter 6) and constraints of the seasonal ensemble forecast model (chapter 7).

The first investigation (chapter 5.1) analyses the seasonal forecast model of the UKMO in regards to seasonal predictability of severe winter windstorm frequency and intensity. This part aims to understand the link between storm characteristics and large-scale patterns over Europe. A persisting seasonal forecast skill over the British Isles and southern Scandinavia has been found for winter windstorm frequencies in the model's updated version, like Befort et al. (2019). But also, for the first time, a skilful prediction of similar regions for seasonally averaged storm intensities has been discovered. These results show that the seasonal model can predict objectively tracked windstorms as a seasonal average. These include the 2% strongest storm events over Europe as the threshold for tracking is to exceed the local 98^{th} percentile per grid cell. A known connection between windstorms and the large-scale pattern NAO has been found for not only frequency but also intensities. Further investigations show a positive connection between windstorm characteristics and additional patterns, SCA and EA. These have been used as predictors to build a statistical model predicting storm characteristics (frequency and intensity measures) per season. These multi-linear regressions can explain up to 80% of the variance for the storm frequency and up to 60% of storm intensities. However, the indirect-regression approach is not as skilful as the direct model approach, using the dynamical simulations of windstorms in the numerical model. However, adding SCA and EA in the multi-linear regression has improved the statistical forecast approach. Both patterns explain storm variance over the North Atlantic region, where previous studies found gaps using only NAO (Befort et al., 2019). Further investigations regarding the large-scale patterns influencing the storm's forecast skill (chapter 5.2) showed that the seasonal skill in all storm characteristics mainly developed from the neutral seasons of the large-scale patterns.

The knowledge gained about skilful storm characteristics (frequency and intensity measures) on a seasonal average leads to the question of which physical mechanisms are well represented in the model, hence, support the storm forecast skill or might need to be improved to improve the existing storm skill. Ten dynamical atmospheric factors have been assessed in a newly

developed 3-step approach to investigate whether the ensemble model is skilful on a seasonal scale due to the correct physical reasons (see chapter 6). The developed approach for this study investigates firstly if the link between storm frequency and the dynamical factor exists in the model in the same way as known in the reanalysis. The second step investigates if the dynamical factors which show a well-represented connection are well-predicted on a seasonal scale. And the third step analysis which regions with higher factor forecast skill are a necessary criterion to improve the seasonal forecast of windstorms. The four main focus factors (two primary factors - MSLP gradient & SST - and two secondary factors - Θ_e & EGR - see chapter 6.1) represent the physical link between factor and storm known from observations in the model (step 1). Additional perspectives on the connection between storms and factors have been evaluated and conclude a coherent result for the main primary and secondary factors, upper tropospheric EGR and MSLP gradient. Primary and secondary factors showed in the second step storm-related regions of positive and significant seasonal forecast skill. In the third step, some of these regions were identified as fundamental for improving the forecast skill of severe winter windstorms over large parts of Central Europe and southern Scandinavia/North Sea.

The third part of this study investigates the signal-to-noise paradox for extreme windstorms over Europe and the North Atlantic. This paradox has been found in multiple model suites, from seasonal to decadal scales (see chapter 2.5 for more details). This study shows that the paradox exists for the seasonal winter windstorm predictions for both storm frequency and intensity (see chapter 5.1). A stronger paradox exists in the model for storm intensities (RPC > 3) than for frequency (RPC > 1) over the British Isles. This study looks into the main influencing factors for storm predictions. Hence, these factors were also assessed for the existence of the signal-to-noise paradox. The primary factors, Θ_e and EGR 400hPa, and the secondary factor, MSLP gradient, show a high RPC in the same regions as the storm frequency. Hence, this could also be an additional limitation in forecasting severe winter windstorms over Europe, as the dynamical factors responsible for the seasonal storm

forecast skill are limited in their predictability due to the signal-to-noise paradox. Chapter 7.1 summarises if storms could be influencing or causing the signal-to-noise paradox in models. Previous studies have connected the eddy feedback, which describes the energy transport between small-scale eddies and the zonal mean flow, to the paradox. The hypothesis is that the underrepresented eddy feedback causes the signal-to-noise paradox (Hardiman et al., 2022). The third part investigates how the RPC of the NAO index changes with different data sub-samples dependent on the season's storm activity. Seasons with many storms have a significantly high RPC, as the signal in the model decreases while the noise and the skill in the NAO persist the same. This agrees with the idea that the eddy feedback influences the signal-to-noise paradox (Hardiman et al., 2022), as the EFP is higher in stormy than in quiet seasons.

8.2 Discussion

This study presents three investigations regarding seasonal forecasts of severe winter windstorms over Europe. Firstly, the general skill of these seasonal forecasts has been investigated. Significant skill, as Kendall correlation, has been found over the British Isles and Southern Scandinavia for windstorm frequencies in the UKMO model, GloSea5. This aligns with previous investigations, finding similar regions of significant skill in GloSea5 for storm frequency (Befort et al., 2019). The same study also investigated two ECMWF seasonal models, which show less strong forecast skill. Renggli et al. (2011) found forecast skill for storm frequencies in other data sets with similar skill measures as tested in this study, ROC and RPSS. This study did not find a clear signal with the RPSS skill measure, but the same results as Renggli et al. (2011) in the ROC analysis, that the extreme terciles (low and high) are better predicted. All these studies, including this thesis, used an objective tracking algorithm (Leckebusch et al., 2008b), which tracks the 2% most extreme storm events. These individually tracked storms are investigated on a seasonal average, either as storm frequency

or intensity. The latter is a newly investigated measure in this study, as previous studies mostly focused on the frequency/counts (Befort et al., 2019; Walz et al., 2018a). Other studies use a different definition of storms or storminess, like Scaife et al. (2014) uses the 10^{th} percentile of MSLP minima, but find high and coherent forecast skill for this storminess measure. The wind speed, in general, is underconfident in the seasonal forecast (Clark et al., 2017). A previously mentioned study that used the impact-based tracking algorithm analysed the connection of windstorm frequency to the NAO (Befort et al., 2019). They found a clear positive link between storms and NAO north of 45°N and a negative link south of 42°N. The assumption for this study was that NAO explains most of the storm variance but not all, so maybe other dominant large-scale patterns describe the remaining variance. The "map of drivers" from Walz et al. (2018a) identifies the SCA and EA patterns as the most dominant and important patterns over Central Europe and the North Atlantic after the NAO. A multi-linear regression approach in this study agreed with the assumption and showed that SCA and EA increase the total explained variance for storm frequency to up to 80%. This also leads to a potential explanation for why the seasonal forecast of windstorms over the Central North Atlantic is not skilful. SCA and EA mostly explain it but might not be as well represented in the model as the NAO. The extra-tropical northern hemispheric winter (DJF) is, in general, a more complex challenge for the seasonal forecast suites than the tropics (cf. Kirtman et al., 2013; Scaife et al., 2017b). First attempts to predict the atmospheric conditions of the extra-tropics by predicting baroclinic waves show that these are not well predicted yet in seasonal forecast suites (Zhang et al., 2021a). They found the essential SST-forcing relation exists, which makes a skilful prediction on a seasonal scale possible, but also, the predictable component in the models needs improvements. Most of the general seasonal predictability originates from the tropics, with large-scale patterns like ENSO, QBO or MJO (UK Met Office, 2023c). The UKMO Hadley Centre publishes studies on the predictability of different variables for European winter seasons (season 2015/16 - Scaife et al. (2017a); 2016/17 - Dunstone et al. (2018a); 2017/18 & 2018/19 - Knight et al.

(2021); 2019/20 - Hardiman et al. (2020) and 2020/21 - Lockwood et al. (2022)). These studies analyse each winter season individually and their forecast performance explicitly, focusing on general patterns, like the NAO, MSLP, and potential reasons for the individual forecast performance in the particular season. Seasons like 2015/16 were well predicted compared to observations because of a strong El Niño and westerly QBO which influence the North Atlantic pressure patterns (Scaife et al., 2017a). The hot and wet winter in 2019/20 in the UK (Hardiman et al., 2020) was caused by a strong NAO and was well predicted, even on a monthly scale, because of a strong IOD which caused Rossby wave trains from the tropics into the extra-tropical North Atlantic. Other seasons, like the most recent of that list, 2020/21, show more uncertainty in the seasonal prediction, as the right NAO sign was just represented in the model's ensemble mean after the MJO and the SSW as overall important teleconnections captured the proper signal (Lockwood et al., 2022).

The second part of the thesis aims to understand the dynamical reasons of seasonal forecast skill of windstorm frequency. Ten dynamical factors with the knowledge of influencing cyclones or windstorms have been investigated in a newly developed 3-step approach. The main factors are, first, well represented in their connection between dynamical factors and storm in reanalysis and model. Second, show skilful regions which are influential for windstorms, and third, skilful and storm-relevant regions have a potential to improve seasonal storm forecasts over wider parts of Europe, incl. British Isles, North Sea and southern Scandinavia. These factors are, for example, the Eady Growth Rate in the upper troposphere (400hPa) and equivalent potential temperature in 850hPa. The same factors have been investigated and linked to storms by Pinto et al. (2008). Other well-represented but more general large-scale factors are, for example, the MSLP gradient or SST (Shaw et al., 2016). Both show coherent results in the approach, including improved storm forecast skill. Other factors analysed have not shown a clear and coherent result, as some might be linked on different spatial or temporal scales. This study utilised all factors as seasonal averages, but dynamical factors like PV might influence windstorms or cyclones on smaller time scales. Hoskins and Hodges (2019)

linked cyclone track density to vorticity maxima or a 2-6-day bandpass-filtered vorticity in 250hPa. This study used PV in 350K as used in Hoskins (2015) and found moderate results in the physical link between storms and PV but hardly any forecast skill. A bandpass-filtered PV (2-8 days) was too scattered in the grid-cell-based approach. Similar scattered results have been produced for the dynamical factor, Rossby Wave Source. This has been included in this study to include the theory of Scaife et al. (2017b), that tropical precipitation triggers Rossby Wave trains towards the extra-tropics and increases predictability over Europe. RWS has been used in previous studies to show this teleconnection (e.g., Beverley et al., 2019). They used the RWS to link Indian summer monsoons to European summer forecasts and found a too weak representation of the RWS in their seasonal model. In this study, the RWS did not show a clear connection to seasonal storm frequency. An assumption might be that RWS connects on other temporal scales, not on a seasonal average, as Beverley et al. (2019) used monthly values.

The third investigation concerns the signal-to-noise paradox, described by Eade et al. (2014) and Scaife and Smith (2018). This study showed, for the first time, that the signal-to-noise paradox not only exists in the NAO forecast in multiple seasonal forecast models (Scaife & Smith, 2018; Smith et al., 2020) but also in the windstorm forecast. The RPC, the measurement for the signal-to-noise paradox, indicates the paradox if the value is greater than 1. This is the case in the storm frequency forecasts and even higher (RPC>3) for intensity over the British Isles. Previous studies suggested that eddies could cause the paradox within the model as seasonal models underestimate the positive eddy feedback (Hardiman et al., 2022). This was only seen in GloSea5 in this study, as only GloSea5 showed a significantly higher NAO-RPC for stormy seasons.

This study is mainly investigated with correlations as a statistical measure. They are frequently used in previous studies to present skill (Dunstone et al., 2016; Höschel et al., 2019; Nie et al., 2019b; Weisheimer et al., 2017). Here especially, the Kendall-τ_b correlation is used, as the data (storm frequency or intensity) cannot be assumed to be normally distributed. A

wrongly used Pearson's correlation can result in an overprediction (Armstrong, 2019), hence, a ranked measure without assuming the normal distribution is an adequate choice. Kendall correlation is also used in the study of Befort et al. (2019), which makes these two studies easier comparable. Other common predictability measures have been assessed, like ROC curves and RPSS (e.g. used in Renggli et al., 2011; Hansen et al., 2019; Weisheimer et al., 2017). The latter measure did not show a clear signal of predictability in this study, similar to Renggli et al. (2011). They used this measure and could not find skill for storm intensities. Besides the skill measures, correlation is also used to investigate connections between storms and dynamic factors. Therefore, different perspectives on connecting these two parameters have been investigated, but only one has been published (chapter 6.1). The rest is presented in an additional chapter (chapter 6.2). All the different perspectives agree on the connections between windstorms and individual dynamical factors, which makes this a robust result. The third part of this study uses the RPC as a measure for the signal-to-noise paradox. This measure has been used in multiple previous studies investigating this paradox (Eade et al., 2014; Scaife & Smith, 2018; Smith et al., 2020; Hardiman et al., 2022). A known issue with the RPC measure is the dependence on the ensemble size (Scaife & Smith, 2018; Baker et al., 2018). This study uses 63 members of the seasonal model, which should reduce this issue. All statistics have been tested on significance either with an individual test, like the Kendall test (Kendall, 1945), a standard Students-T-test for differences or bootstrapping with more than 10.000 iterations. These tests are used to understand which results are robust.

8.3 Conclusion & Outlook

This study concludes and summarises the skill, dynamics and constraints of the UK Met Office seasonal forecast model, GloSea5, regarding its winter windstorm forecasts. It has been shown for the first time that the average intensity of tracked storm events can be predicted on a seasonal scale. For the first time, it has been presented that SCA and EA

are strongly linked to windstorm characteristics and are important for a statistical model of seasonal storm characteristics. Combining the three most dominant large-scale patterns over Europe reaches almost the same predictability as the direct model forecast. The model forecasts windstorm characteristics from the right physical connections between storms and dynamic atmospheric factors. But the signal-to-noise paradox is also existing in the storm predictions. Hence, there is still a limitation on the forecast performance.

The results presented here for the GloSea5 model lead to higher confidence in the seasonal windstorm predictions. It could lead to more general use of this information, e.g. for end users like insurance or the general public to inform about the upcoming winter season and what to expect. Besides peer-reviewed scientific studies (e.g., Lockwood et al., 2022; Dunstone et al., 2018a), UKMO also releases a 3-month seasonal outlook for the public each month. The focus was on temperature and precipitation for many years as these are in great demand by the government and the public. Since the winter season of 2022, wind has been included in these monthly summaries (UK Met Office, 2022). These summaries show the potential outcome of the next three months by categorising them into higher, lower or near average. Additionally, an overview over the last ten years for the same forecast is presented, and a comparison of the predicted distribution to an average distribution. Regarding the new additional parameter wind, a quick and simple check can be done by comparing the number of windstorms in each season to the predicted wind forecast issued in November. UKMO only names a storm when it has a severe impact on the UK (UK Met Office, 2023d). This happened only once in the last winter season, December 2022 to February 2023, with Storm "*Otto*" on the 16^{th} of February. The 3-month outlook said this season would be near-average regarding wind (UK Met Office, 2022). The number of named storms and the average wind over the season is officially not directly linked, but Lockwood et al. (2023) found a strong correlation between average wind and storminess of the seasons, which can be seen as similar. This indicator can be used for previous seasons to see that the 3-month outlook predicted a *windy* season 2019/2020, which had at the end

five named storms affecting the UK. The UKMO seasonal outlook has not yet a windstorm prediction included, but having the general wind predicted on a confident level and presented to the public every month is a significant improvement. This study would help to build the confidence to include windstorm frequencies or intensities. This next step could look similar to released hurricane predictions from NOAA (National Weather Service Climate Prediction Center, 2023). The NOAA hurricane prediction is separated into the three affected basins (Atlantic, East Pacific and Central Pacific). The seasonal outlook gives a percentage of how likely the upcoming season is above, near or below normal and a range of expected storm counts in different categories (named storms, hurricanes or major hurricanes). This public outlook is possible because the tropical cyclone (TC) forecasts are very advanced already. The forecast error has dropped from 300 nautical miles in 1990 to only 100 nautical miles in 2016 (Cangialosi & Franklin, 2017) and might have even reached the theoretically possible maximum of predictability (Landsea & Cangialosi, 2018). To investigate if this is the case for ETCs and mid-latitude windstorms would be a research question of another study. This study showed forecast skill in windstorms over Europe, but the skill has some limitations (signal-to-noise paradox) that must be understood. This study leads to more confidence in the seasonal windstorm forecast of the UKMO which could now be used to extend the 3-month outlook of the UKMO with windstorm predictions. This would make the public aware of the expected upcoming season and could help the government and companies with potential preparations.

Each result chapter could be extended with further investigations. In general, the whole approach of this thesis could be tested in other available forecast suites. Previous studies showed lower skill in windstorm frequency in other models (Befort et al., 2019). Reasons for this could be of interest, but also, the physical link in each model can be different, and the signal-to-noise paradox has a different influence on the forecast skill. The latter, in general, can be investigated in more depth, as the assumption remains that the disturbance of storms is a cause for the paradox.

References

Aguinis, Herman, Ryan K. Gottfredson, and Harry Joo (2013). "Best-Practice Recommendations for Defining, Identifying, and Handling Outliers". In: *Organizational Research Methods* 16.2. DOI: 10.1177/1094428112470848.

Ambaum, Maarten HP, Brian J Hoskins, and David B Stephenson (2001). "Arctic oscillation or North Atlantic oscillation?" In: *Journal of Climate* 14.16. DOI: 10.1175/1520-0442(2001)014<3495:aoonao>2.0.co;2.

Angus, M. and G. C. Leckebusch (2020). "On the Dependency of Atlantic Hurricane and European Windstorm Hazards". In: *Geophysical Research Letters* 47.23. DOI: https://doi.org/10.1029/2020GL090446.

Armstrong, Richard A (2019). "Should Pearson's correlation coefficient be avoided?" In: *Ophthalmic and Physiological Optics* 39.5. DOI: https://doi.org/10.1111/opo.12636.

Athanasiadis, Panos J., Alessio Bellucci, Leon Hermanson, Adam A. Scaife, Craig MacLachlan, Alberto Arribas, Stefano Materia, Andrea Borrelli, and Silvio Gualdi (2014). "The Representation of Atmospheric Blocking and the Associated Low-Frequency Variability in Two Seasonal Prediction Systems". In: *Journal of Climate* 27.24. DOI: 10.1175/jcli-d-14-00291.1.

Athanasiadis, Panos J., Alessio Bellucci, Adam A. Scaife, Leon Hermanson, Stefano Materia, Antonella Sanna, Andrea Borrelli, Craig MacLachlan, and Silvio Gualdi (2017). "A Multisystem View of Wintertime NAO Seasonal Predictions". In: *Journal of Climate* 30.4. DOI: 10.1175/jcli-d-16-0153.1.

AXA Agentur Koller & Baur (2022). *kollerbaur - Hohe Schäden durch Wintersturm-Serie im Februar 2022*. URL: https://www.koller-baur.de/blog/hohe-schaeden-durch-wintersturm-serie-im-februar-2022#:~:text=Der%5C%20Gesamtverband%5C%20der%5C%20Deutschen%5C%20Versicherungswirtschaft,auf%5C%201%5C%2C4%5C%20Milliarden%5C%20Euro. (visited on 05/08/2023).

Baker, L. H., L. C. Shaffrey, R. T. Sutton, Antje Weisheimer, and Adam A. Scaife (2018). "An Intercomparison of Skill and Overconfidence/Underconfidence of the Wintertime North Atlantic Oscillation in Multimodel Seasonal Forecasts". In: *Geophysical Research Letters* 45.15. DOI: 10.1029/2018gl078838.

Barnston, Anthony G and Robert E Livezey (1987). "Classification, seasonality and persistence of low-frequency atmospheric circulation patterns". In: *Monthly weather review* 115.6. DOI: 10.1175/1520-0493(1987)115<1083:csapol>2.0.co;2.

Barnston, Anthony G., Michael K. Tippett, Michelle L. L'Heureux, Shuhua Li, and David G. DeWitt (2012). "Skill of Real-Time Seasonal ENSO Model Predictions during 2002–11: Is

Our Capability Increasing?" In: *Bulletin of the American Meteorological Society* 93.5. DOI: https://doi.org/10.1175/BAMS-D-11-00111.1.

Batté, L, L Dorel, C Ardilouze, and JF Guérémy (2021). "Documentation of the METEO-FRANCE Seasonal Forecasting System 8". In.

BBC (2022). *Storm Eunice reaches the UK*. URL: https://www.bbc.co.uk/news/in-pictures-60428222 (visited on 05/08/2023).

Befort, Daniel J., Simon Wild, Jeff R. Knight, Julia F. Lockwood, Hazel E. Thornton, Leon Hermanson, Philip E. Bett, Antje Weisheimer, and Gregor C. Leckebusch (2019). "Seasonal forecast skill for extratropical cyclones and windstorms". In: *Quarterly Journal of the Royal Meteorological Society*. DOI: 10.1002/qj.3406.

Berliner Wetterkarte e.V. (1998). *Berliner Wetterkarte e.V.* URL: https://berliner-wetterkarte.de/ (visited on 05/08/2023).

Beverley, Jonathan D, Steven J Woolnough, Laura H Baker, Stephanie J Johnson, and Antje Weisheimer (2019). "The northern hemisphere circumglobal teleconnection in a seasonal forecast model and its relationship to European summer forecast skill". In: *Climate dynamics* 52.5. DOI: https://doi.org/10.1007/s00382-018-4371-4.

Board, Ocean Studies, Engineering National Academies of Sciences, and Medicine (2016). *Next generation earth system prediction: strategies for subseasonal to seasonal forecasts*. National Academies Press. ISBN: 030938883X. DOI: 10.17226/21873.

Bolton, David (1980). "The Computation of Equivalent Potential Temperature". In: *Monthly Weather Review* 108.7. DOI: https://doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2.

Bueh, Cholaw and Hisashi Nakamura (2007). "Scandinavian pattern and its climatic impact". In: *Quarterly Journal of the Royal Meteorological Society* 133.629. DOI: 10.1002/qj.173.

Büeler, Dominik and Stephan Pfahl (2017). "Potential Vorticity Diagnostics to Quantify Effects of Latent Heating in Extratropical Cyclones. Part I: Methodology". In: *Journal of the Atmospheric Sciences* 74.11. DOI: 10.1175/JAS-D-17-0041.1.

Butterworth, Stephen (1930). "On the theory of filter amplifiers". In: *Wireless Engineer* 7.6.

Cangialosi, J. P. and J. L. Franklin (2017). *National Hurricane Center forecast verification report: 2016 hurricane season*. National Hurricane Center Rep., p. 72.

Catto, Jennifer L., Duncan Ackerley, James F. Booth, Adrian J. Champion, Brian A. Colle, Stephan Pfahl, Joaquim G. Pinto, Julian F. Quinting, and Christian Seiler (2019). "The Future of Midlatitude Cyclones". In: *Current Climate Change Reports* 5.4. DOI: 10.1007/s40641-019-00149-4.

Chang, CB, DJ Pepkey, and CW Kreitzberg (1984). "Latent heat induced energy transformations during cyclogenesis". In: *Monthly weather review* 112.2. DOI: https://doi.org/10.1175/1520-0493(1984)112<0357:LHIETD>2.0.CO;2.

Clark, Robin T, Philip E Bett, Hazel E Thornton, and Adam A Scaife (2017). "Skilful seasonal predictions for the European energy industry". In: *Environmental Research Letters* 12.2. DOI: 10.1088/1748-9326/aa57ab.

Copernicus Climate Change Service (2023a). *Description of HadGEM3-GC3.2-v20200929 C3S contribution*. URL: https://confluence.ecmwf.int/display/CKB/Description+of+HadGEM3-GC3.2-v20200929+C3S+contribution (visited on 05/08/2023).

– (2023b). *Description of the C3S seasonal multi-system*. URL: https://confluence.ecmwf.int/display/CKB/Description+of+the+C3S+seasonal+multi-system (visited on 05/08/2023).
– (2023c). *Description of the C3S seasonal multi-system*. URL: https://confluence.ecmwf.int/display/CKB/Description+of+the+C3S+seasonal+multi-system (visited on 05/08/2023).
Dawson, Andrew (2016). "Windspharm: A high-level library for global wind field computations using spherical harmonics". In: *Journal of Open Research Software* 4.1. DOI: 10.5334/jors.129.
Dee, Dick P, S M Uppala, Adrian J Simmons, Paul Berrisford, Paul Poli, Shinya Kobayashi, U Andrae, MA Balmaseda, G Balsamo, and d P Bauer (2011). "The ERA-Interim reanalysis: Configuration and performance of the data assimilation system". In: *Quarterly Journal of the royal meteorological society* 137.656.
Degenhardt, Lisa, Gregor Leckebusch, and Adam Scaife (2020). "Assessing reasons for skilful predictions of winter windstorms over the Atlantic/European region". In: *EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-10991*. DOI: 10.5194/egusphere-egu2020-10991.
Degenhardt, Lisa, Gregor C. Leckebusch, and Adam A. Scaife (2022). "Large-scale circulation patterns and their influence on European winter windstorm predictions". In: *Climate Dynamics*. DOI: 10.1007/s00382-022-06455-2.
Degenhardt, Lisa and Haraldur Ólafsson (2019). "Persistence of observed air temperatures in Iceland". In: *International Journal of Climatology* 39.3. DOI: https://doi.org/10.1002/joc.5875.
Déqué, Michel (2011). "Deterministic forecasts of continuous variables". In: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. DOI: 10.1002/9781119960003.ch5.
Déqué, Michel, Christine Dreveton, Alain Braun, and Daniel Cariolle (1994). "The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling". In: *Climate Dynamics* 10. DOI: https://doi.org/10.1007/BF00208992.
Domeisen, Daniela (2020). "Origins of variability and predictability in the North Atlantic region". In: *EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-7815*. DOI: 10.5194/egusphere-egu2020-7815.
Domeisen, Daniela I. V., Christopher J. White, Hilla Afargan-Gerstman, Ángel G. Muñoz, Matthew A. Janiga, Frédéric Vitart, C. Ole Wulff, Salomé Antoine, Constantin Ardilouze, Lauriane Batté, Hannah C. Bloomfield, David J. Brayshaw, Suzana J. Camargo, Andrew Charlton-Pérez, Dan Collins, Tim Cowan, Maria del Mar Chaves, Laura Ferranti, Rosario Gómez, Paula L. M. González, Carmen González Romero, Johnna M. Infanti, Stelios Karozis, Hera Kim, Erik W. Kolstad, Emerson LaJoie, Llorenç Lledó, Linus Magnusson, Piero Malguzzi, Andrea Manrique-Suñén, Daniele Mastrangelo, Stefano Materia, Hanoi Medina, Lluís Palma, Luis E. Pineda, Athanasios Sfetsos, Seok-Woo Son, Albert Soret, Sarah Strazzo, and Di Tian (2022). "Advances in the Subseasonal Prediction of Extreme Events: Relevant Case Studies across the Globe". In: *Bulletin of the American Meteorological Society* 103.6. DOI: 10.1175/BAMS-D-20-0221.1.
Domeisen, Daniela IV, Gualtiero Badin, and Inga M Koszalka (2018). "How predictable are the Arctic and North Atlantic Oscillations? Exploring the variability and predictability of

the Northern Hemisphere". In: *Journal of Climate* 31.3. DOI: https://doi.org/10.1175/JCLI-D-17-0226.1.

Domeisen, Daniela IV, Amy H Butler, Kristina Fröhlich, Matthias Bittner, Wolfgang A Müller, and Johanna Baehr (2015). "Seasonal predictability over Europe arising from El Niño and stratospheric variability in the MPI-ESM seasonal prediction system". In: *Journal of Climate* 28.1. DOI: 10.1175/jcli-d-14-00207.1.

Donat, Markus G, Gregor C Leckebusch, Joaquim G Pinto, and Uwe Ulbrich (2010). "Examination of wind storms over Central Europe with respect to circulation weather types and NAO phases". In: *International Journal of Climatology* 30.9. DOI: 10.1002/joc.1982.

Donat, Markus G., Gregor C. Leckebusch, Simon Wild, and Uwe Ulbrich (2011a). "Future changes in European winter storm losses and extreme wind speeds inferred from GCM and RCM multi-model simulations". In: *Natural Hazards and Earth System Sciences* 11.5. DOI: 10.5194/nhess-11-1351-2011.

Donat, Markus G., D. Renggli, Simon Wild, L. V. Alexander, Gregor C. Leckebusch, and Uwe Ulbrich (2011b). "Reanalysis suggests long-term upward trends in European storminess since 1871". In: *Geophysical Research Letters* 38.14. DOI: 10.1029/2011gl047995.

Dunstone, Nick, Adam A Scaife, Craig MacLachlan, Jeff Knight, Sarah Ineson, Doug Smith, Hazel Thornton, Margaret Gordon, Peter McLean, and Erika Palin (2018a). "Predictability of European winter 2016/2017". In: *Atmospheric Science Letters* 19.12. DOI: https://doi.org/10.1002/asl.868.

Dunstone, Nick, Doug Smith, Adam Scaife, Leon Hermanson, Rosie Eade, Niall Robinson, Martin Andrews, and Jeff Knight (2016). "Skilful predictions of the winter North Atlantic Oscillation one year ahead". In: *Nature Geoscience* 9.11. DOI: 10.1038/ngeo2824.

Dunstone, Nick, Doug Smith, Adam Scaife, Leon Hermanson, David Fereday, Chris O'Reilly, Alison Stirling, Rosie Eade, Margaret Gordon, and Craig MacLachlan (2018b). "Skilful seasonal predictions of summer European rainfall". In: *Geophysical Research Letters* 45.7. DOI: https://doi.org/10.1002/2017GL076337.

Eade, R., D. Smith, Adam A. Scaife, E. Wallace, N. Dunstone, L. Hermanson, and N. Robinson (2014). "Do seasonal-to-decadal climate predictions underestimate the predictability of the real world?" In: *Geophys Res Lett* 41.15. DOI: 10.1002/2014gl061146.

Eady, Eric T (1949). "Long waves and cyclone waves". In: *Tellus* 1.3. DOI: https://doi.org/10.3402/tellusa.v1i3.8507.

ECMWF (2023a). *Data assimilation*. URL: https://www.ecmwf.int/en/research/data-assimilation#:~:text=Data%20assimilation%20is%20typically%20a,is%20initiated%2C%20and%20so%20on. (visited on 09/14/2023).

– (2023b). *Modelling and Prediction*. URL: https://www.ecmwf.int/en/research/modelling-and-prediction (visited on 09/14/2023).

Fereday, DR, A Maidens, A Arribas, AA Scaife, and JR Knight (2012). "Seasonal forecasts of Northern Hemisphere winter 2009/10". In: *Environmental Research Letters* 7.3. DOI: 10.1088/1748-9326/7/3/034031.

Feser, F., M. Barcikowska, O. Krueger, F. Schenk, R. Weisse, and L. Xia (2015). "Storminess over the North Atlantic and northwestern Europe-A review". In: *Quarterly Journal of the Royal Meteorological Society* 141.687. DOI: 10.1002/qj.2364.

Fink, Andreas H, Tim Brücher, Volker Ermert, Andreas Krüger, and Joaquim G Pinto (2009). "The European storm Kyrill in January 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change". In: *Natural Hazards and Earth System Sciences* 9.2. DOI: https://doi.org/10.5194/nhess-9-405-2009.

Folland, C. K., A. A. Scaife, J. Lindesay, and D. B. Stephenson (2012). "How potentially predictable is northern European winter climate a season ahead?" In: *International Journal of Climatology* 32.6. DOI: https://doi.org/10.1002/joc.2314.

Fröhlich, Kristina, Mikhail Dobrynin, Katharina Isensee, Claudia Gessner, Andreas Paxian, Holger Pohlmann, Helmuth Haak, Sebastian Brune, Barbara Früh, and Johanna Baehr (2021). "The German Climate Forecast System: GCFS". In: *Journal of Advances in Modeling Earth Systems* 13.2. DOI: https://doi.org/10.1029/2020MS002101.

Gualdi, S, A Borrelli, A Cantelli, G Davoli, M del Mar Chavesmontero, S Masina, A Navarra, A Sanna, and S Tibaldi (2020). "The new CMCC operational seasonal prediction system". In: *CMCC Research Paper* RP0288. DOI: https://doi.org/10.25424/CMCC/SPS3.5.

Hansen, F., Tim Kruschke, R. J. Greatbatch, and Antje Weisheimer (2019). "Factors Influencing the Seasonal Predictability of Northern Hemisphere Severe Winter Storms". In: *Geophysical Research Letters* 46.1. DOI: 10.1029/2018gl079415.

Hardiman, Steven C, Nick J Dunstone, Adam A Scaife, Doug M Smith, Jeff R Knight, Paul Davies, Martin Claus, and Richard J Greatbatch (2020). "Predictability of European winter 2019/20: Indian Ocean dipole impacts on the NAO". In: *Atmospheric Science Letters*. DOI: https://doi.org/10.1002/asl.1005.

Hardiman, Steven C., Nick J. Dunstone, Adam A. Scaife, Doug M. Smith, Ruth Comer, Yu Nie, and Hong-Li Ren (2022). "Missing eddy feedback may explain weak signal-to-noise ratios in climate predictions". In: *npj Climate and Atmospheric Science* 5.1. DOI: 10.1038/s41612-022-00280-4.

Hemri, Stephan, Jonas Bhend, Mark A Liniger, Rodrigo Manzanas, Stefan Siegert, David B Stephenson, José M Gutiérrez, Anca Brookshaw, and Francisco J Doblas-Reyes (2020). "How to create an operational multi-model of seasonal forecasts?" In: *Climate Dynamics* 55.5. DOI: https://doi.org/10.1007/s00382-020-05314-2.

Hersbach, Hans, Bill Bell, Paul Berrisford, András Horányi, Joaquín Muñoz Sabater, Julien Nicolas, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, and Dick Dee (2019). "Global reanalysis: goodbye ERA-Interim, hello ERA5". In: *ECMWF Newsletter* No. 159. DOI: 10.21957/vf291hehd7.

Hoffmann, Lars, Gebhard Günther, Dan Li, Olaf Stein, Xue Wu, Sabine Griessbach, Yi Heng, Paul Konopka, Rolf Müller, and Bärbel Vogel (2019). "From ERA-Interim to ERA5: the considerable impact of ECMWF's next-generation reanalysis on Lagrangian transport simulations". In: *Atmospheric Chemistry and Physics* 19.5. DOI: 10.5194/acp-19-3097-2019.

Holton, JR (2004). *An introduction to dynamic meteorology*. New York, London: Elsevier Academic Press.

Höschel, Ines, Sebastian Illing, Jens Grieger, Uwe Ulbrich, and Ulrich Cubasch (2019). "On skillful decadal predictions of the subpolar North Atlantic". In: *Meteorologische Zeitschrift* 28.5. DOI: 10.1127/metz/2019/0957.

Hoskins, BJ and KI Hodges (2019). "The annual cycle of Northern Hemisphere storm tracks. Part I: Seasons". In: *Journal of Climate* 32.6. DOI: https://doi.org/10.1175/JCLI-D-17-0870.1.

Hoskins, Brian (2015). "Potential vorticity and the PV perspective". In: *Advances in Atmospheric Sciences* 32.1. DOI: https://doi.org/10.1007/s00376-014-0007-8.

Hoskins, Brian J and Kevin I Hodges (2002). "New perspectives on the Northern Hemisphere winter storm tracks". In: *Journal of the Atmospheric Sciences* 59.6. DOI: https://doi.org/10.1175/1520-0469(2002)059<1041:NPOTNH>2.0.CO;2.

Hoskins, Brian J, Michael E McIntyre, and Andrew W Robertson (1985). "On the use and significance of isentropic potential vorticity maps". In: *Quarterly Journal of the Royal Meteorological Society* 111.470. DOI: https://doi.org/10.1002/qj.49711147002.

Hoskins, Brian J and Paul J Valdes (1990). "On the existence of storm-tracks". In: *Journal of Atmospheric Sciences* 47.15. DOI: https://doi.org/10.1175/1520-0469(1990)047<1854:OTEOST>2.0.CO;2.

Hoskins, Brian J., Ian N. James, and Glenn H. White (1983). "The Shape, Propagation and Mean-Flow Interaction of Large-Scale Weather Systems". In: *Journal of Atmospheric Sciences* 40.7. DOI: https://doi.org/10.1175/1520-0469(1983)040<1595:TSPAMF>2.0.CO;2.

Hurrell, James W (1995). "Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation". In: *Science* 269.5224. DOI: 10.1126/science.269.5224.676.

Hurrell, James W, Yochanan Kushnir, Geir Ottersen, and Martin Visbeck (2003). "An overview of the North Atlantic oscillation". In: *Geophysical Monograph-American Geophysical Union* 134. DOI: 10.1029/134gm01.

Hurrell, James W, Yochanan Kushnir, and Martin Visbeck (2001). "The north Atlantic oscillation". In: *Science* 291.5504. DOI: 10.1126/science.1058761.

Jin, Emilia K., James L. Kinter, B. Wang, C. K. Park, I. S. Kang, B. P. Kirtman, J. S. Kug, A. Kumar, J. J. Luo, J. Schemm, J. Shukla, and T. Yamagata (2008). "Current status of ENSO prediction skill in coupled ocean–atmosphere models". In: *Climate Dynamics* 31.6. DOI: 10.1007/s00382-008-0397-3.

Jin, Yishuai (2022). "The Signal-to-Noise Paradox in ENSO Prediction: Role of ENSO Growth Rate and Period". In: *Geophysical Research Letters* 49.12. DOI: https://doi.org/10.1029/2022GL097965.

Johnson, S. J., T. N. Stockdale, L. Ferranti, M. A. Balmaseda, F. Molteni, L. Magnusson, S. Tietsche, D. Decremer, A. Weisheimer, G. Balsamo, S. P. E. Keeley, K. Mogensen, H. Zuo, and B. M. Monge-Sanz (2019). "SEAS5: the new ECMWF seasonal forecast system". In: *Geosci. Model Dev.* 12.3. DOI: 10.5194/gmd-12-1087-2019.

Karremann, MK and JG Pinto (2020). "2 Clustering von Zyklonen". In: *promet: Außertropische Zyklonen*.

Kendall, M. G. (1945). "The Treatment of Ties in Ranking Problems". In: *Biometrika* 33.3. DOI: 10.2307/2332303.

Kendon, Mike (2022). *Storms Dudley, Eunice and Franklin, February 2022*. URL: https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/weather/learn-about/uk-past-events/interesting/2022/2022_01_storms_dudley_eunice_franklin_r1.pdf (visited on 05/08/2023).

Kharin, Viatcheslav V, Qiaobin Teng, Francis W Zwiers, George J Boer, Jacques Derome, and Juan Sebastian Fontecilla (2009). "Skill assessment of seasonal hindcasts from the Canadian Historical Forecast Project". In: *Atmosphere-ocean* 47.3. DOI: https://doi.org/10.3137/AO1101.2009.

Kharin, Viatcheslav V and Francis W Zwiers (2003). "On the ROC score of probability forecasts". In: *Journal of Climate* 16.24. DOI: https://doi.org/10.1175/1520-0442(2003)016<4145:OTRSOP>2.0.CO;2.

Kharin, VV and FW Zwiers (2001). "Skill as a function of time scale in ensembles of seasonal hindcasts". In: *Climate Dynamics* 17.2. DOI: 10.1007/s003820000102.

Kharin, VV, FW Zwiers, and N Gagnon (2001). "Skill of seasonal hindcasts as a function of the ensemble size". In: *Climate Dynamics* 17.11. DOI: https://doi.org/10.1007/s003820100149.

Kim, Hye-Mi, Peter J Webster, and Judith A Curry (2012). "Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter". In: *Climate Dynamics* 39.12. DOI: https://doi.org/10.1007/s00382-012-1364-6.

Kirtman, Ben, Scott B Power, Akintayo John Adedoyin, George J Boer, Roxana Bojariu, Ines Camilloni, Francisco Doblas-Reyes, Arlene M Fiore, Masahide Kimoto, and Gerald Meehl (2013). "Near-term climate change: projections and predictability". In.

Klawa, M and Uwe Ulbrich (2003). "A model for the estimation of storm losses and the identification of severe winter storms in Germany". In: *Natural Hazards and Earth System Sciences* 3. DOI: https://doi.org/10.5194/nhess-3-725-2003.

Knight, Jeff, Adam Scaife, Philip E Bett, Tamara Collier, Nick Dunstone, Margaret Gordon, Steven Hardiman, Leon Hermanson, Sarah Ineson, and Gillian Kay (2021). "Predictability of European Winters 2017/2018 and 2018/2019: Contrasting influences from the Tropics and stratosphere". In: *Atmospheric Science Letters* 22.1. DOI: https://doi.org/10.1002/asl.1009.

Knight, Jeff R., Adam A. Scaife, and Anna Maidens (2022). "An Extratropical Contribution to the Signal-To-Noise Paradox in Seasonal Climate Prediction". In: *Geophysical Research Letters* 49.23. DOI: https://doi.org/10.1029/2022GL100471.

Kruschke, Tim (2015). "Winter wind storms: identifcation, verification of decadal predictions, and regionalization". Thesis. Frei Universität Berlin.

Kruschke, Tim, Henning W. Rust, Christopher Kadow, Gregor C. Leckebusch, and Uwe Ulbrich (2014). "Evaluating decadal predictions of northern hemispheric cyclone frequencies". In: *Tellus A: Dynamic Meteorology and Oceanography* 66.1. DOI: 10.3402/tellusa.v66.22830.

Kruschke, Tim, Henning W. Rust, Christopher Kadow, Wolfgang A. Müller, Holger Pohlmann, Gregor C. Leckebusch, and Uwe Ulbrich (2016). "Probabilistic evaluation of decadal prediction skill regarding Northern Hemisphere winter storms". In: *Meteorologische Zeitschrift* 25.6. DOI: 10.1127/metz/2015/0641.

Kurz, Manfred (1990). *Synoptische Meteorologie; Mit 189 häufig mehrteiligen Abb. im Text.* Selbstverl. des Dt. Wetterdienstes. ISBN: 3881482628.

Landsea, Christopher W and John P Cangialosi (2018). "Have we reached the limits of predictability for tropical cyclone track forecasting?" In: *Bulletin of the American Meteorological Society* 99.11. DOI: https://doi.org/10.1175/BAMS-D-17-0136.1.

Leckebusch, Gregor C, Alice Kapala, Herman Mächel, Joaquim G Pinto, and Mark Reyers (2008a). "Indizes der Nordatlantischen und Arktischen Oszillation". In: *Promet* 34.3-4.

Leckebusch, Gregor C, Dominik Renggli, and Uwe Ulbrich (2008b). "Development and application of an objective storm severity measure for the Northeast Atlantic region". In: *Meteorologische Zeitschrift* 17.5. DOI: 10.1127/0941-2948/2008/0323.

Leckebusch, Gregor C and Uwe Ulbrich (2004). "On the relationship between cyclones and extreme windstorm events over Europe under climate change". In: *Global and planetary change* 44.1-4. DOI: https://doi.org/10.1016/j.gloplacha.2004.06.011.

Lee, Simon H., Zachary D. Lawrence, Amy H. Butler, and Alexey Y. Karpechko (2020). "Seasonal Forecasts of the Exceptional Northern Hemisphere Winter of 2020". In: *Geophysical Research Letters* 47.21. DOI: https://doi.org/10.1029/2020GL090328.

Lin, Hai, Zhiyong Huang, Harry Hendon, and Gilbert Brunet (2021). "NAO Influence on the MJO and its Prediction Skill in the Subseasonal-to-Seasonal Prediction Models". In: *Journal of Climate* 34.23. DOI: https://doi.org/10.1175/JCLI-D-21-0153.1.

Liu, Chunlei, Keith Haines, Alan Iwi, and Doug Smith (2012). "Comparing the UK Met Office Climate Prediction System DePreSys with idealized predictability in the HadCM3 model". In: *Quarterly Journal of the Royal Meteorological Society* 138.662. DOI: 10.1002/qj.904.

Lledó, Llorenç, Irene Cionni, Verónica Torralba, Pierre-Antoine Bretonnière, and Margarida Samsó (2020). "Seasonal prediction of Euro-Atlantic teleconnections from multiple systems". In: *Environmental Research Letters*. DOI: 10.1088/1748-9326/ab87d2.

Lledó, Llorenç, Jaume Ramon, Albert Soret, and Francisco-Javier Doblas-Reyes (2022). "Seasonal prediction of renewable energy generation in Europe based on four teleconnection indices". In: *Renewable Energy* 186. DOI: https://doi.org/10.1016/j.renene.2021.12.130.

Lockwood, Julia, Nicky Stringer, Katie Hodge, Philip Bett, Jeff Knight, Doug Smith, Adam Scaife, Matthew Patterson, Nick Dunstone, and Hazel Thornton (2023). *Seasonal prediction of UK mean and extreme winds*. Report. Copernicus Meetings. DOI: https://doi.org/10.5194/egusphere-egu23-13639.

Lockwood, Julia F, Nicky Stringer, Hazel E Thornton, Adam A Scaife, Philip E Bett, Tamara Collier, Ruth Comer, Nick Dunstone, Margaret Gordon, and Leon Hermanson (2022). "Predictability of European winter 2020/2021: Influence of a mid-winter sudden stratospheric warming". In: *Atmospheric Science Letters* 23.12. DOI: https://doi.org/10.1002/asl.1126.

Lorenz, Edward N. (1955). "Available Potential Energy and the Maintenance of the General Circulation". In: *Tellus* 7.2. DOI: https://doi.org/10.1111/j.2153-3490.1955.tb01148.x.

– (1963). "Deterministic Nonperiodic Flow". In: *Journal of Atmospheric Sciences* 20.2. DOI: 10.1175/1520-0469(1963)020<0130:Dnf>2.0.Co;2.

MacLachlan, C., A. Arribas, K. A. Peterson, A. Maidens, D. Fereday, Adam A. Scaife, M. Gordon, M. Vellinga, A. Williams, R. E. Comer, J. Camp, P. Xavier, and G. Madec (2015). "Global Seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system". In: *Quarterly Journal of the Royal Meteorological Society* 141.689. DOI: 10.1002/qj.2396.

Manzanas, R., V. Torralba, Ll Lledó, and P. A. Bretonnière (2022). "On the Reliability of Global Seasonal Forecasts: Sensitivity to Ensemble Size, Hindcast Length and Re-

gion Definition". In: *Geophysical Research Letters* 49.17. DOI: https://doi.org/10.1029/2021GL094662.

Marcheggiani, Andrea, Jon Robson, Paul-Arthur Monerie, Thomas J. Bracegirdle, and Doug Smith (2023). "Decadal Predictability of the North Atlantic Eddy-Driven Jet in Winter". In: *Geophysical Research Letters* 50.8. DOI: https://doi.org/10.1029/2022GL102071.

Marshall, Andrew G., Paul A. Gregory, Catherine O. de Burgh-Day, and Morwenna Griffiths (2022). "Subseasonal drivers of extreme fire weather in Australia and its prediction in ACCESS-S1 during spring and summer". In: *Climate Dynamics* 58.1. DOI: 10.1007/s00382-021-05920-8.

Mason, Simon J and David B Stephenson (2008). *How do we know whether seasonal climate forecasts are any good?* Report 1402069901. Springer.

Mason, Simon J., Christopher A. T. Ferro, and Willem A. Landman (2021). "Forecasts of "normal"". In: *Quarterly Journal of the Royal Meteorological Society* 147.735. DOI: https://doi.org/10.1002/qj.3968.

Mayer, Björn, André Düsterhus, and Johanna Baehr (2021). "When Does the Lorenz 1963 Model Exhibit the Signal-To-Noise Paradox?" In: *Geophysical Research Letters* 48.4. DOI: https://doi.org/10.1029/2020GL089283.

Meehl, Gerald A., Lisa Goddard, George Boer, Robert Burgman, Grant Branstator, Christophe Cassou, Susanna Corti, Gokhan Danabasoglu, Francisco Doblas-Reyes, Ed Hawkins, Alicia Karspeck, Masahide Kimoto, Arun Kumar, Daniela Matei, Juliette Mignot, Rym Msadek, Antonio Navarra, Holger Pohlmann, Michele Rienecker, and Tony Rosati (2014). "Decadal Climate Prediction: An Update from the Trenches". In: *Bulletin of the American Meteorological Society*. 95.2. DOI: 10.1175/bams-d-12-00241.1.

Merryfield, William J, Woo-Sung Lee, George J Boer, Viatcheslav V Kharin, John F Scinocca, Gregory M Flato, RS Ajayamohan, John C Fyfe, Youmin Tang, and Saroja Polavarapu (2013). "The Canadian seasonal to interannual prediction system. Part I: Models and initialization". In: *Monthly weather review* 141.8. DOI: https://doi.org/10.1175/MWR-D-12-00216.1.

Mishra, Niti, Chloé Prodhomme, and Virginie Guemas (2019). "Multi-model skill assessment of seasonal temperature and precipitation forecasts over Europe". In: *Climate dynamics* 52.7-8. DOI: 10.1007/s00382-018-4404-z.

Moulds, S., L. J. Slater, N. J. Dunstone, and D. M. Smith (2023). "Skillful Decadal Flood Prediction". In: *Geophysical Research Letters* 50.3. DOI: https://doi.org/10.1029/2022GL100650.

MunichRE (2010). "Topics Geo 2009: Natural Catastrophes: Analyses, assessments, positions". In.

MunichRe (2019). *NatCatSERVICE - World Weather-Related Natural Catastrophes By Peril, 1980-2018*. URL: https://www.iii.org/graph-archive/218207 (visited on 06/26/2023).

– (2023a). *Climate change and La Niña driving losses: the natural disaster figures for 2022*. URL: https://www.munichre.com/en/company/media-relations/media-information-and-corporate-news/media-information/2023/natural-disaster-figures-2022.html (visited on 05/08/2023).

– (2023b). *Winter storms and blizzards - A risk to entire continents*. URL: https://www.munichre.com/en/risks/natural-disasters/winter-storms.html (visited on 05/08/2023).

Murphy, Allan H. (1969). “On the "Ranked Probability Score"”. In: *Journal of Applied Meteorology* 8. DOI: https://doi.org/10.1175/1520-0450(1969)008<0988:OTPS>2.0.CO;2.

Murray, Ross J and Ian Simmonds (1991). “A numerical scheme for tracking cyclone centres from digital data. Part I: Development and operation of the scheme”. In: *Aust. Meteor. Mag* 39.3.

Namias, Jerome (1968). “Long range weather forecasting—history, current status and outlook”. In: *Bulletin of the American Meteorological Society* 49.5-1. DOI: https://doi.org/10.1175/1520-0477-49.5.438.

National Weather Service Climate Prediction Center (2005). *Teleconnection Pattern Calculation Precedures*. URL: https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/history/method.shtml (visited on 07/04/2023).

– (2023). *NOAA 2023 Atlantic Hurricane Season Outlook*. URL: https://www.cpc.ncep.noaa.gov/products/outlooks/hurricane.shtml (visited on 06/08/2023).

Nesterov, E. S. (2009). “East Atlantic oscillation of the atmospheric circulation”. In: *Russian Meteorology and Hydrology* 34.12. DOI: 10.3103/s1068373909120048.

Neu, Urs, Mirseid G. Akperov, Nina Bellenbaum, Rasmus Benestad, Richard Blender, Rodrigo Caballero, Angela Cocozza, Helen F. Dacre, Yang Feng, Klaus Fraedrich, Jens Grieger, Sergey Gulev, John Hanley, Tim Hewson, Masaru Inatsu, Kevin Keay, Sarah F. Kew, Ina Kindem, Gregor C. Leckebusch, Margarida L. R. Liberato, Piero Lionello, Igor I. Mokhov, Joaquim G. Pinto, Christoph C. Raible, Marco Reale, Irina Rudeva, Mareike Schuster, Ian Simmonds, Mark Sinclair, Michael Sprenger, Natalia D. Tilinina, Isabel F. Trigo, Sven Ulbrich, Uwe Ulbrich, Xiaolan L. Wang, and Heini Wernli (2013). “IMILAST: A Community Effort to Intercompare Extratropical Cyclone Detection and Tracking Algorithms”. In: *Bulletin of the American Meteorological Society* 94.4. DOI: 10.1175/bams-d-11-00154.1.

Ng, K. S. and G. C. Leckebusch (2021). “A new view on the risk of typhoon occurrence in the western North Pacific”. In: *Nat. Hazards Earth Syst. Sci.* 21.2. DOI: 10.5194/nhess-21-663-2021.

Nie, Yu, Hong-Li Ren, and Yang Zhang (2019a). “The role of extratropical air–sea interaction in the autumn subseasonal variability of the North Atlantic Oscillation”. In: *Journal of Climate* 32.22. DOI: https://doi.org/10.1175/JCLI-D-19-0060.1.

Nie, Yu, Adam A. Scaife, Hong-Li Ren, Ruth E. Comer, Martin B. Andrews, Philip Davis, and Nicola Martin (2019b). “Stratospheric initial conditions provide seasonal predictability of the North Atlantic and Arctic Oscillations”. In: *Environmental Research Letters* 14.3. DOI: 10.1088/1748-9326/ab0385.

Nissen, Katrin M., Gregor C. Leckebusch, Joaquim G. Pinto, and Uwe Ulbrich (2013). “Mediterranean cyclones and windstorms in a changing climate”. In: *Regional Environmental Change* 14.5. DOI: 10.1007/s10113-012-0400-8.

Ogimet (2005). *Professional information about meteorological conditions in the world*. URL: http://ogimet.com/home.phtml.en (visited on 05/08/2023).

Olauson, Jon (2018). “ERA5: The new champion of wind power modelling?” In: *Renewable Energy* 126. DOI: 10.1016/j.renene.2018.03.056.

Palmer, Tim N, A Alessandri, U Andersen, P Cantelaube, M Davey, Pascale Delécluse, Michel Déqué, E Diez, Francisco J Doblas-Reyes, and Henrik Feddersen (2004). “Devel-

opment of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER)". In: *Bulletin of the American Meteorological Society* 85.6. DOI: https://doi.org/10.1175/BAMS-85-6-853.

Parker, Tess, Tim Woollings, Antje Weisheimer, Chris O'Reilly, Laura Baker, and Len Shaffrey (2019). "Seasonal predictability of the winter North Atlantic Oscillation from a jet stream perspective". In: *Geophysical Research Letters* 46.16. DOI: https://doi.org/10.1029/2019GL084402.

Pearson, Karl and Alice Lee (1903). "On the Laws of Inheritance in Man: I. Inheritance of Physical Characters". In: *Biometrika* 2.4. DOI: 10.2307/2331507.

Pinto, Joaquim G., Iñigo Gómara, Giacomo Masato, Helen F. Dacre, Tim Woollings, and Rodrigo Caballero (2014). "Large-scale dynamics associated with clustering of extratropical cyclones affecting Western Europe". In: *Journal of Geophysical Research: Atmospheres* 119.24. DOI: https://doi.org/10.1002/2014JD022305.

Pinto, Joaquim G., Stefan Zacharias, Andreas H. Fink, Gregor C. Leckebusch, and Uwe Ulbrich (2008). "Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO". In: *Climate Dynamics* 32.5. DOI: 10.1007/s00382-008-0396-4.

Portal, Alice, Paolo Ruggieri, Froila M. Palmeiro, Javier García-Serrano, Daniela I. V. Domeisen, and Silvio Gualdi (2022). "Seasonal prediction of the boreal winter stratosphere". In: *Climate Dynamics* 58.7. DOI: 10.1007/s00382-021-05787-9.

Postel, Gregory A. and Matthew H. Hitchman (1999). "A Climatology of Rossby Wave Breaking along the Subtropical Tropopause". In: *Journal of the Atmospheric Sciences* 56.3. DOI: https://doi.org/10.1175/1520-0469(1999)056<0359:ACORWB>2.0.CO;2.

Priestley, M. D. K., H. F. Dacre, L. C. Shaffrey, K. I. Hodges, and J. G. Pinto (2018). "The role of serial European windstorm clustering for extreme seasonal losses as determined from multi-centennial simulations of high-resolution global climate model data". In: *Nat. Hazards Earth Syst. Sci.* 18.11. DOI: 10.5194/nhess-18-2991-2018.

Priestley, Matthew D. K., Duncan Ackerley, Jennifer L. Catto, Kevin I. Hodges, Ruth E. McDonald, and Robert W. Lee (2020). "An Overview of the Extratropical Storm Tracks in CMIP6 Historical Simulations". In: *Journal of Climate* 33.15. DOI: 10.1175/jcli-d-19-0928.1.

Priestley, Matthew D. K., Joaquim G. Pinto, Helen F. Dacre, and Len C. Shaffrey (2017a). "Rossby wave breaking, the upper level jet, and serial clustering of extratropical cyclones in western Europe". In: *Geophysical Research Letters* 44.1. DOI: https://doi.org/10.1002/2016GL071277.

– (2017b). "The role of cyclone clustering during the stormy winter of 2013/2014". In: *Weather* 72.7. DOI: https://doi.org/10.1002/wea.3025.

Priestley, Matthew DK, Duncan Ackerley, Jennifer L Catto, and Kevin I Hodges (2023). "Drivers of biases in the CMIP6 extratropical storm tracks. Part I: Northern Hemisphere". In: *Journal of Climate* 36.5. DOI: https://doi.org/10.1175/JCLI-D-20-0976.1.

Pukelsheim, Friedrich (1994). "The three sigma rule". In: *The American Statistician* 48.2. DOI: 10.1080/00031305.1994.10476030.

Raymond, DJ (1992). "Nonlinear balance and PV thinking at large Rossby number". In: *Quart. J. Roy. Meteor. Soc* 118.

Ren, Hong-Li, Adam A Scaife, Nick Dunstone, Ben Tian, Ying Liu, Sarah Ineson, June-Yi Lee, Doug Smith, Changzheng Liu, and Vikki Thompson (2019). "Seasonal predictability of winter ENSO types in operational dynamical model predictions". In: *Climate Dynamics* 52. DOI: https://doi.org/10.1007/s00382-018-4366-1.

Renggli, Dominik, Gregor C. Leckebusch, Uwe Ulbrich, Stephanie N. Gleixner, and Eberhard Faust (2011). "The Skill of Seasonal Ensemble Prediction Systems to Forecast Wintertime Windstorm Frequency over the North Atlantic and Europe". In: *Monthly Weather Review* 139.9. DOI: 10.1175/2011mwr3518.1.

Roberts, J. F., A. J. Champion, L. C. Dawkins, K. I. Hodges, L. C. Shaffrey, D. B. Stephenson, M. A. Stringer, H. E. Thornton, and B. D. Youngman (2014). "The XWS open access catalogue of extreme European windstorms from 1979 to 2012". In: *Natural Hazards and Earth System Sciences* 14.9. DOI: 10.5194/nhess-14-2487-2014.

Saha, S, S Nadiga, C Thiaw, J Wang, W Wang, Q Zhang, HM Van den Dool, H-L Pan, S Moorthi, and D Behringer (2006). "The NCEP climate forecast system". In: *Journal of Climate* 19.15. DOI: https://doi.org/10.1175/JCLI3812.1.

Saha, Suranjana, Shrinivas Moorthi, Hua-Lu Pan, Xingren Wu, Jiande Wang, Sudhir Nadiga, Patrick Tripp, Robert Kistler, John Woollen, David Behringer, Haixia Liu, Diane Stokes, Robert Grumbine, George Gayno, Jun Wang, Yu-Tai Hou, Hui-ya Chuang, Hann-Ming H. Juang, Joe Sela, Mark Iredell, Russ Treadon, Daryl Kleist, Paul Van Delst, Dennis Keyser, John Derber, Michael Ek, Jesse Meng, Helin Wei, Rongqian Yang, Stephen Lord, Huug van den Dool, Arun Kumar, Wanqiu Wang, Craig Long, Muthuvel Chelliah, Yan Xue, Boyin Huang, Jae-Kyung Schemm, Wesley Ebisuzaki, Roger Lin, Pingping Xie, Mingyue Chen, Shuntai Zhou, Wayne Higgins, Cheng-Zhi Zou, Quanhua Liu, Yong Chen, Yong Han, Lidia Cucurull, Richard W. Reynolds, Glenn Rutledge, and Mitch Goldberg (2010). "The NCEP Climate Forecast System Reanalysis". In: *Bulletin of the American Meteorological Society* 91.8. DOI: 10.1175/2010BAMS3001.1.

Saha, Suranjana, Shrinivas Moorthi, Xingren Wu, Jiande Wang, Sudhir Nadiga, Patrick Tripp, David Behringer, Yu-Tai Hou, Hui-ya Chuang, Mark Iredell, Michael Ek, Jesse Meng, Rongqian Yang, Malaquías Peña Mendez, Huug van den Dool, Qin Zhang, Wanqiu Wang, Mingyue Chen, and Emily Becker (2014). "The NCEP Climate Forecast System Version 2". In: *Journal of Climate* 27.6. DOI: https://doi.org/10.1175/JCLI-D-12-00823.1.

Saito, Naoaki, Shuhei Maeda, Tosiyuki Nakaegawa, Yuhei Takaya, Yukiko Imada, and Chihiro Matsukawa (2017). "Seasonal predictability of the North Atlantic Oscillation and zonal mean fields associated with stratospheric influence in JMA/MRI-CPS2". In: *Sola* 13. DOI: https://doi.org/10.2151/sola.2017-038.

Sanna, A, A Borrelli, P Athanasiadis, S Materia, A Storto, A Navarra, S Tibaldi, and S Gualdi (2016). "CMCC-SPS3: the CMCC Seasonal Prediction System 3". In: *CMCC Research Paper* RP0285.

Scaife, A. A., A. Arribas, E. Blockley, A. Brookshaw, R. T. Clark, N. Dunstone, R. Eade, D. Fereday, C. K. Folland, M. Gordon, L. Hermanson, J. R. Knight, D. J. Lea, C. MacLachlan, A. Maidens, M. Martin, A. K. Peterson, D. Smith, M. Vellinga, E. Wallace, J. Waters, and A. Williams (2014). "Skillful long-range prediction of European and North American winters". In: *Geophysical Research Letters* 41.7. DOI: 10.1002/2014gl059637.

Scaife, Adam A., Joanne Camp, Ruth Comer, Philip Davis, Nick Dunstone, Margaret Gordon, Craig MacLachlan, Nicola Martin, Yu Nie, Hong-Li Ren, Malcolm Roberts, Walter Robinson, Doug Smith, and Pier Luigi Vidale (2019a). "Does increased atmospheric resolution improve seasonal climate predictions?" In: *Atmospheric Science Letters* 20.8. DOI: 10.1002/asl.922.

Scaife, Adam A., Ruth Comer, Nick Dunstone, David Fereday, Chris Folland, Elizabeth Good, Margaret Gordon, Leon Hermanson, Sarah Ineson, Alexey Karpechko, Jeff Knight, Craig MacLachlan, Anna Maidens, K. Andrew Peterson, Doug Smith, Julia Slingo, and Brent Walker (2017a). "Predictability of European winter 2015/2016". In: *Atmospheric Science Letters* 18.2. DOI: 10.1002/asl.721.

Scaife, Adam A., Ruth E. Comer, Nick J. Dunstone, Jeff R. Knight, Doug M. Smith, Craig MacLachlan, Nicola Martin, K. Andrew Peterson, Dan Rowlands, Edward B. Carroll, Stephen Belcher, and Julia Slingo (2017b). "Tropical rainfall, Rossby waves and regional winter climate predictions". In: *Quarterly Journal of the Royal Meteorological Society* 143.702. DOI: 10.1002/qj.2910.

Scaife, Adam A., Dan Copsey, Chris Gordon, Chris Harris, Tim Hinton, Sarah Keeley, Alan O'Neill, Malcolm Roberts, and Keith Williams (2011). "Improved Atlantic winter blocking in a climate model". In: *Geophysical Research Letters* 38.23. DOI: https://doi.org/10.1029/2011GL049573.

Scaife, Adam A., Laura Ferranti, Oscar Alves, Panos Athanasiadis, Johanna Baehr, Michel Dequé, Tina Dippe, Nick Dunstone, David Fereday, Richard G. Gudgel, Richard J. Greatbatch, Leon Hermanson, Yukiko Imada, Shipra Jain, Arun Kumar, Craig MacLachlan, William Merryfield, Wolfgang A. Müller, Hong-Li Ren, Doug Smith, Yuhei Takaya, Gabriel Vecchi, and Xiaosong Yang (2019b). "Tropical rainfall predictions from multiple seasonal forecast systems". In: *International Journal of Climatology* 39.2. DOI: 10.1002/joc.5855.

Scaife, Adam A. and Doug Smith (2018). "A signal-to-noise paradox in climate science". In: *npj Climate and Atmospheric Science* 1.1. DOI: 10.1038/s41612-018-0038-4.

Schulzweida, Uwe, Luis Kornblueh, and Ralf Quast (2019). *CDO User Guide (Version 1.9.6)*. Report.

Schuster, Mareike, Jens Grieger, Andy Richling, Thomas Schartner, Sebastian Illing, Christopher Kadow, Wolfgang A. Müller, Holger Pohlmann, Stephan Pfahl, and Uwe Ulbrich (2019). "Improvement in the decadal prediction skill of the northern hemisphere extratropical winter circulation through increased model resolution". In: *Earth System Dynamics Discussions*. DOI: 10.5194/esd-2019-18.

Sévellec, F. and S. S. Drijfhout (2019). "The Signal-to-Noise Paradox for Interannual Surface Atmospheric Temperature Predictions". In: *Geophysical Research Letters* 46.15. DOI: https://doi.org/10.1029/2019GL083855.

Shaw, T. A., M. Baldwin, E. A. Barnes, R. Caballero, C. I. Garfinkel, Y. T. Hwang, C. Li, P. A. O'Gorman, G. Rivière, I. R. Simpson, and A. Voigt (2016). "Storm track processes and the opposing influences of climate change". In: *Nature Geoscience* 9.9. DOI: 10.1038/ngeo2783.

Simon, Horst, Thomas Zacharia, and Rick Stevens (2007). "Modeling and Simulation at the Exascale for Energy and the Environment". In: *Department of Energy Technical Report*.

Smith, D. M., A. A. Scaife, R. Eade, P. Athanasiadis, A. Bellucci, I. Bethke, R. Bilbao, L. F. Borchert, L. P. Caron, F. Counillon, G. Danabasoglu, T. Delworth, F. J. Doblas-Reyes, N. J. Dunstone, V. Estella-Perez, S. Flavoni, L. Hermanson, N. Keenlyside, V. Kharin, M. Kimoto, W. J. Merryfield, J. Mignot, T. Mochizuki, K. Modali, P. A. Monerie, W. A. Müller, D. Nicolí, P. Ortega, K. Pankatz, H. Pohlmann, J. Robson, P. Ruggieri, R. Sospedra-Alfonso, D. Swingedouw, Y. Wang, S. Wild, S. Yeager, X. Yang, and L. Zhang (2020). "North Atlantic climate far more predictable than models imply". In: *Nature* 583.7818. DOI: 10.1038/s41586-020-2525-0.

Smith, Doug M., Stephen Cusack, Andrew W. Colman, Chris K. Folland, Glen R. Harris, and James M. Murphy (2007). "Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model". In: *Science* 317.5839. DOI: 10.1126/science.1139540.

Smith, Doug. M., R. Eade, A. A. Scaife, L. P. Caron, G. Danabasoglu, T. M. DelSole, T. Delworth, F. J. Doblas-Reyes, N. J. Dunstone, L. Hermanson, V. Kharin, M. Kimoto, W. J. Merryfield, T. Mochizuki, W. A. Müller, H. Pohlmann, S. Yeager, and X. Yang (2019). "Robust skill of decadal climate predictions". In: *npj Climate and Atmospheric Science* 2.1. DOI: 10.1038/s41612-019-0071-y.

Stockdale, TN (2000). "An overview of techniques for seasonal forecasting". In: *Stochastic Environmental Research and Risk Assessment* 14. DOI: https://doi.org/10.1007/s004770000053.

SwissRe (2023). *Financial Report 2022*. Report.

Takaya, Yuhei, Shoji Hirahara, Tamaki Yasuda, Satoko Matsueda, Takahiro Toyoda, Yosuke Fujii, Hiroyuki Sugimoto, Chihiro Matsukawa, Ichiro Ishikawa, Hirotoshi Mori, Ryoji Nagasawa, Yutaro Kubo, Noriyuki Adachi, Goro Yamanaka, Tsurane Kuragano, Akihiko Shimpo, Shuhei Maeda, and Tomoaki Ose (2018). "Japan Meteorological Agency/ Meteorological Research Institute-Coupled Prediction System version 2 (JMA/MRI-CPS2): atmosphere-land-ocean-sea ice coupled prediction system for operational seasonal forecasting". In: *Climate Dynamics* 50.3. DOI: 10.1007/s00382-017-3638-5.

Tang, Youmin, Hai Lin, Jacques Derome, and Michael K. Tippett (2007). "A Predictability Measure Applied to Seasonal Predictions of the Arctic Oscillation". In: *Journal of Climate* 20.18. DOI: 10.1175/jcli4276.1.

Thornton, H. E., D. M. Smith, A. A. Scaife, and N. J. Dunstone (2023). "Seasonal Predictability of the East Atlantic Pattern in Late Autumn and Early Winter". In: *Geophysical Research Letters* 50.1. DOI: https://doi.org/10.1029/2022GL100712.

Toreti, Andrea, Franco Desiato, Guido Fioravanti, and Walter Perconti (2010). "Seasonal temperatures over Italy and their relationship with low-frequency atmospheric circulation patterns". In: *Climatic Change* 99.1. DOI: https://doi.org/10.1007/s10584-009-9640-0.

Trigo, Isabel Franco (2006). "Climatology and interannual variability of storm-tracks in the Euro-Atlantic sector: a comparison between ERA-40 and NCEP/NCAR reanalyses". In: *Climate Dynamics* 26.2. DOI: 10.1007/s00382-005-0065-9.

Troccoli, Alberto, Mike Harrison, David LT Anderson, and Simon J Mason (2008). *Seasonal climate: forecasting and managing risk*. Vol. 82. Springer Science & Business Media. ISBN: 1402069901.

UK Met Office (2022). *UK 3-month Outlook - Period: December 2022 - February 2023*. Report.

– (2023a). *National Severe Weather Warnings Archive - National Meteorological Library & Archive*. URL: https://library.metoffice.gov.uk/Portal/Default/en-GB/Search/SimpleSearch (visited on 06/26/2023).
– (2023b). *Numerical weather prediction models*. URL: https://www.metoffice.gov.uk/research/approach/modelling-systems/unified-model/weather-forecasting (visited on 09/14/2023).
– (2023c). *Seasonal forecasts and climate drivers resources*. URL: https://www.metoffice.gov.uk/services/government/contingency-planners/seasonal-forecasts-and-climate-drivers-resources (visited on 06/09/2023).
– (2023d). *UK Storm Centre*. URL: https://www.metoffice.gov.uk/weather/warnings-and-advice/uk-storm-centre/index (visited on 06/07/2023).

Ulbrich, U and Michael Christoph (1999). "A shift of the NAO and increasing storm track activity over Europe due to anthropogenic greenhouse gas forcing". In: *Climate dynamics* 15.7. DOI: 10.1007/s003820050299.

Ulbrich, U, AH Fink, M Klawa, and Joaquim G Pinto (2001). "Three extreme storms over Europe in December 1999". In: *Weather* 56.3. DOI: http://dx.doi.org/10.1002/j.1477-8696.2001.tb06540.x.

Ulbrich, U, GC Leckebusch, and Joaquim G Pinto (2009). "Extra-tropical cyclones in the present and future climate: a review". In: *Theoretical and Applied Climatology* 96.1. DOI: https://doi.org/10.1007/s00704-008-0083-8.

Volonté, Ambrogio, Suzanne L. Gray, Peter A. Clark, Oscar Martínez-Alvarado, and Duncan Ackerley (2023a). "Strong surface winds in Storm Eunice. Part 1: storm overview and indications of sting jet activity from observations and model data". In: *Weather* n/a.n/a. DOI: https://doi.org/10.1002/wea.4402.
– (2023b). "Strong surface winds in Storm Eunice. Part 2: airstream analysis". In: *Weather* n/a.n/a. DOI: https://doi.org/10.1002/wea.4401.

Von Storch, Hans and Francis W Zwiers (2001). *Statistical analysis in climate research*. Cambridge university press. ISBN: 0521012309.

Wallace, John M. and David S. Gutzler (1981). "Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter". In: *Monthly Weather Review* 109.4. DOI: 10.1175/1520-0493(1981)109<0784:titghf>2.0.co;2.

Walz, Michael A., Daniel J. Befort, N. O. Kirchner-Bossi, Uwe Ulbrich, and Gregor C. Leckebusch (2018a). "Modelling serial clustering and inter-annual variability of European winter windstorms based on large-scale drivers". In: *Int J Climatol* 38.7. DOI: 10.1002/joc.5481.

Walz, Michael A., Markus G. Donat, and Gregor C. Leckebusch (2018b). "Large-Scale Drivers and Seasonal Predictability of Extreme Wind Speeds Over the North Atlantic and Europe". In: *Journal of Geophysical Research: Atmospheres* 123.20. DOI: 10.1029/2017jd027958.

Walz, Michael A. and Gregor C. Leckebusch (2019). "Loss potentials based on an ensemble forecast: How likely are winter windstorm losses similar to 1990?" In: *Atmospheric Science Letters* 20.4. DOI: 10.1002/asl.891.

Weigel, Andreas P, Mark A Liniger, and Christof Appenzeller (2007). "The discrete Brier and ranked probability skill scores". In: *Monthly Weather Review* 135.1. DOI: https://doi.org/10.1175/MWR3280.1.

Weisheimer, A, FJ Doblas-Reyes, TN Palmer, A Alessandri, A Arribas, M Déqué, Noel Keenlyside, M MacVean, A Navarra, and P Rogel (2009). "ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions—Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs". In: *Geophysical research letters* 36.21. DOI: https://doi.org/10.1029/2009GL040896.

Weisheimer, Antje, Daniel J. Befort, Dave MacLeod, Tim Palmer, Chris O'Reilly, and Kristian Strømmen (2020). "Seasonal forecasts of the 20th Century". In: *Bulletin of the American Meteorological Society*. DOI: 10.1175/bams-d-19-0019.1.

Weisheimer, Antje, Damien Decremer, David MacLeod, Christopher O'Reilly, Tim N. Stockdale, Stephanie Johnson, and Tim N. Palmer (2019). "How confident are predictability estimates of the winter North Atlantic Oscillation?" In: *Quarterly Journal of the Royal Meteorological Society* 145.S1. DOI: 10.1002/qj.3446.

Weisheimer, Antje, N. Schaller, C. O'Reilly, D. A. MacLeod, and T. Palmer (2017). "Atmospheric seasonal forecasts of the twentieth century: multi-decadal variability in predictive skill of the winter North Atlantic Oscillation (NAO) and their potential value for extreme event attribution". In: *Q J R Meteorol Soc* 143.703. DOI: 10.1002/qj.2976.

Wetter3.de (2003). *The World in Weather Charts*. URL: https://wetter3.de/index_dt.html (visited on 05/08/2023).

Wild, Simon, Daniel J Befort, and Gregor C Leckebusch (2015). "Was the extreme storm season in winter 2013/14 over the North Atlantic and the United Kingdom triggered by changes in the West Pacific warm pool?" In: *Bulletin of the American Meteorological Society* 96.12. DOI: https://doi.org/10.1175/BAMS-D-15-00118.1.

Wilks, Daniel S (2011). *Statistical methods in the atmospheric sciences*. 3rd ed. Vol. 100. Oxford, Waltham, MA: Oxford, Waltham, MA : Academic Press. ISBN: 0123850223.

Williams, K. D., D. Copsey, E. W. Blockley, A. Bodas-Salcedo, D. Calvert, R. Comer, P. Davis, T. Graham, H. T. Hewitt, R. Hill, P. Hyder, S. Ineson, T. C. Johns, A. B. Keen, R. W. Lee, A. Megann, S. F. Milton, J. G. L. Rae, M. J. Roberts, A. A. Scaife, R. Schiemann, D. Storkey, L. Thorpe, I. G. Watterson, D. N. Walters, A. West, R. A. Wood, T. Woollings, and P. K. Xavier (2018). "The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations". In: *Journal of Advances in Modeling Earth Systems* 10.2. DOI: https://doi.org/10.1002/2017MS001115.

Williams, N. C., A. A. Scaife, and J. A. Screen (2023). "Underpredicted ENSO Teleconnections in Seasonal Forecasts". In: *Geophysical Research Letters* 50.5. DOI: https://doi.org/10.1029/2022GL101689.

WMO (2023). *Mean Square Skill Score*. URL: http://www.bom.gov.au/wmo/lrfvs/msss.shtml (visited on 06/26/2023).

Yang, Dejian, Xiu-Qun Yang, Dan Ye, Xuguang Sun, Jiabei Fang, Cuijiao Chu, Tao Feng, Yiquan Jiang, Jin Liang, Xuejuan Ren, Yaocun Zhang, and Youmin Tang (2018). "On the Relationship Between Probabilistic and Deterministic Skills in Dynamical Seasonal Climate Prediction". In: *Journal of Geophysical Research: Atmospheres* 123.10. DOI: https://doi.org/10.1029/2017JD028002.

Zhang, Gan, Hiroyuki Murakami, William F Cooke, Zhuo Wang, Liwei Jia, Feiyu Lu, Xiaosong Yang, Thomas L Delworth, Andrew T Wittenberg, and Matthew J Harrison

(2021a). "Seasonal predictability of baroclinic wave activity". In: *npj Climate and Atmospheric Science* 4.1. DOI: https://doi.org/10.1038/s41612-021-00209-3.

Zhang, Wei and Ben Kirtman (2019). "Understanding the Signal-to-Noise Paradox with a Simple Markov Model". In: *Geophysical Research Letters* 46.22. DOI: https://doi.org/10.1029/2019GL085159.

Zhang, Wei, Ben Kirtman, Leo Siqueira, Amy Clement, and Junfei Xia (2021b). "Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model". In: *Climate Dynamics* 56.9. DOI: 10.1007/s00382-020-05621-8.

Zubiate, Laura, Frank McDermott, Conor Sweeney, and Mark O'Malley (2017). "Spatial variability in winter NAO–wind speed relationships in western Europe linked to concomitant states of the East Atlantic and Scandinavian patterns". In: *Quarterly Journal of the Royal Meteorological Society* 143.702. DOI: 10.1002/qj.2943.

Appendix A

For Chapter 4

This Appendix includes a summary table for used skill score measures in literature and is reference in chapter 4.

Table A.1 Skill Score list of skill score measure in different forecast skill publications.

Paper	Skill Score	Used Parameter	Significance Test
Athanasiadis et al. (2017)	ACC	MSLP for NAO; DJF	t-test, against the null hypothesis of nonpositive correlation
Athanasiadis et al. (2014)	Correlation coefficient	Blockings	t test
Baker et al. (2018)	ACC, RPC	MSLP for NAO; DJF	/
Befort et al. (2019)	Kendall's τ_b rank correlation coefficient ACC	Wind storm and cyclone frequency (density) DJF	obtained under the null hypothesis of no association, and as our data contains ties
Déqué (2011)			
Domeisen (2020)	ACC		
Domeisen et al. (2015)	ACC	GPH	1000 bootstrap samples

Dunstone et al. (2018b)	uncentered Pearson's correlation coefficient	precipitation; JJA	two-sided and 1-sided Student's t test
Dunstone et al. (2016)	correlation, RPC	NAO out of MSLP	Student t test , Durbin and Watson test
Eade et al. (2014)	RPC MSSS	Surface Temperature, MSLP, precipitation DJF	nonparametric block bootstrap method
Hansen et al. (2019)	RPSS	winter storm frequency; DJF	two-sided t test
Höschel et al. (2019)	Pearson Correlation MSESS	SST NAI (NA SST Index)	bootstrap
Kruschke et al. (2016)			
Kruschke (2015)	RPSS Lead-time-dependent bias	Wind storm frequency ONDJFM	fold overlapping-blockbootstrapping
Kruschke et al. (2014)	cross-validated bias corrections	Wind storms and cyclones frequency	bootstrapping by 1000-fold re-sampling the 41 hindcasts

	RPSS	ONDJFM	one-sided test
Liu et al. (2012)	standard deviation of anomaly errors (anomaly differences with observations)	Surface Air Temperature (SAT)	significant differences with F-test
MacLachlan et al. (2015)	ACC Reliability diagram, ROC	compares GloSea5/4 (ENSO, MSLP)	
Meehl et al. (2014)			
Mishra et al. (2019)	ACC Continuous RPSS (CRPSS) Fair CRPSS	temperature precipitation	one-sided student t-test & controlling False Discovery Rate standard deviation of skill score
Nie et al. (2019b)	Correlation between initial wind and the monthly-mean zonal wind anomaly	Monthly MSLP (NAO), zonal velocity	one-tailed test

	ACC	DJF	
Renggli et al. (2011)	RPSS ROC	wind storm frequency DJF and JFMA	bootstrap method (Weigel et al. 2008)
Scaife et al. (2019a)	correlation?	ENSO, tropical cyclone track frequency, NAO; DJF and JJA	/
Scaife et al. (2019b)	correlation of time series	/	
Scaife and Smith (2018)	RPC average correlation score	MSLP (NAO) DJF	statistically significantly different from 1
Scaife et al. (2017b)	correlations between predicted and observed winter	MSLP (NAO), rainfall, 200hPa geopotential height; DJF	two tailed t-test
Scaife et al. (2017a)	/	MSLP, SST (ENSO), stratospheric (QBO, SSW); Winter 15/16	/

Scaife et al. (2014)	correlation	windstorm frequency, temperature adn wind speed	student t-test
Schuster et al. (2019)	ACC	Stormtrack, Blocking Cyclone, Windstorm Frequency	1000 times resampling
Smith et al. (2019)	RPC	Signal-to-Noise Paradox Temperature, Precipitation, Pressure	greater 1
Smith et al. (2007)			
Tang et al. (2007)			
Walz et al. (2018a)	VIF AIC dispersion (D) score	wind storm frequency large scale drivers DJF	10-fold cross-validation method χ^2 test (for Drivers)
Walz et al. (2018b)	AIC Entropy	MSLP (EOF), wind DJF	T-Test

	Predictive power (PP)		
Weisheimer et al. (2020)	ACC, RPC	SST, GPH	
Weisheimer et al. (2019)	RPC, ACC, ACP Spread/standard deviation RMSE	500hPa geopotential height DJF	/
Weisheimer et al. (2017)	ACC ROCSS	GPH 500hPa DJF	t-test Mann-Whitney U-test; bootstrap

Appendix B

For Chapter 5

This is the Appendix for paper No. 1 from chapter 5 and already published.

The Appendix includes additional material for this paper. Starting with a more detailed explanation of the storm parameter calculation, the description of the two different regression methods, the used skill measures and the signal-to-noise paradox. This will be followed by further additional figures that help understanding some minor processes of the study.

Storm Parameters and large-scale patterns (for Chapter 5.1.3)

The tracking output includes 3 different views of the storm for either the whole storm or each individual time step of the storm (Leckebusch et al. 2008b). Firstly, the storm's footprint, a map that includes each grid cell the storm effects. Secondly, a storm track which is the storm location (longitude/latitude) per time step as the mass-weighted average of the cluster. Thirdly, the Storm Severity Index (SSI, as introduced for gridded data by Leckebusch et al. (2008b), an objective measure of storm intensity which is calculated either for each time step separately or as event-integrated storm severity.

The large-scale patterns are calculated using the principal component analysis (PCA). This extracts time series of indices from the EOF patterns which will be used for further calculations. This is done by projecting the anomalies of ERA5 or GloSea5 onto the eigenvector values of the EOF analysis of ERA5.

Regression Methods (for Chapter 5.1.3)

The ERA5-based regression leads to a study that analyses the influence of NAO, SCA and EA respectively on the individual windstorm parameters for realistic conditions. The following step of including model large-scale patterns into the calculated regression equation would show how well this real-life connection can be represented with modelled large-scale mode indices. Hence, the forecast skill of windstorms would solely be based on the modelled large-scale patterns forecast skill.

A second way regressions can be calculated is to use the model climate only. A comparison to the results from the ERA5-based regression can reveal how the real-world link is represented in the forecast model. A statistical calculation of windstorms using this regression equation would suggest that the GloSea5 NAO, SCA and EA is very close to reality. The resulting windstorm forecast skill would be a result of the large-scale mode performance of the model and the models climatological link to windstorm parameters.

The validation is done with the determination coefficient to validate the reliability of the total regression and the individual percentage of explained variance per predictor (NAO, SCA and EA) is used to understand the role of the three different large-scale patterns and their main area of influence. Both validations can be calculated from an ANOVA (analysis of variance) testing (Von Storch and Zwiers 2001).

Skill Measures (for Chapter 5.1.3)

Like in previous studies e.g., Befort et al. (2019), Kendall-τ_b-ranked correlation coefficient is used as a skill measure. It has no assumption of normally distributed data as for the Pearson's correlation and hence, can be used for a wider range of data (Wilks 2011). The significance is tested by a two-sided T-test asking for the dependence of the two-time series. The Kendall correlation is a skill measure that compares the directionality of both data sets. It is neither explaining variabilities in the amplitude of increase or decrease nor analysing

the correct prediction of different classifications in the data sets.

The ROC curve is used for probabilistic and discrete forecasts and is based on validations of the contingency table (Wilks 2011). The ROC curves show the Hit (right forecast of an occurred event) and False Alarm (wrong forecast that the event occurs) Rate of a forecast. In the context of the seasonal windstorm predictions, the windstorm parameters are categories into 3 groups: lowest, middle and highest tercile of frequency or intensity measure. The area under this curve is defined as the ROC skill score (ROCSS: Wilks 2011) and is used in similar studies e.g., in Renggli et al. (2011). The results of ROCSS and individual ROC curves will reveal the forecast skill of the individual categories; lowest, middle and high storm frequency or intensity.

Additional Figures

Fig. B.1 shows the EOF patterns from ERA5 and GloSea5, which are patterns-wise very similar and the PCA could be calculated by just using the ERA5 pattern to make the indices more comparable. (Relates to chapter 5.1.3)

Fig. B.2 & B.3 are the additional regression verification figures for the GloSea5-based regression analysis. These figures show the regression slopes and variance per grid cell per large-scale pattern. (Relates to chapter 5.1.4)

Fig. B.4 shows the test run of using the leave-one-out method for the regressions and show no significant differences to the method used in the main paper. (Relates to chapter 5.1.4)

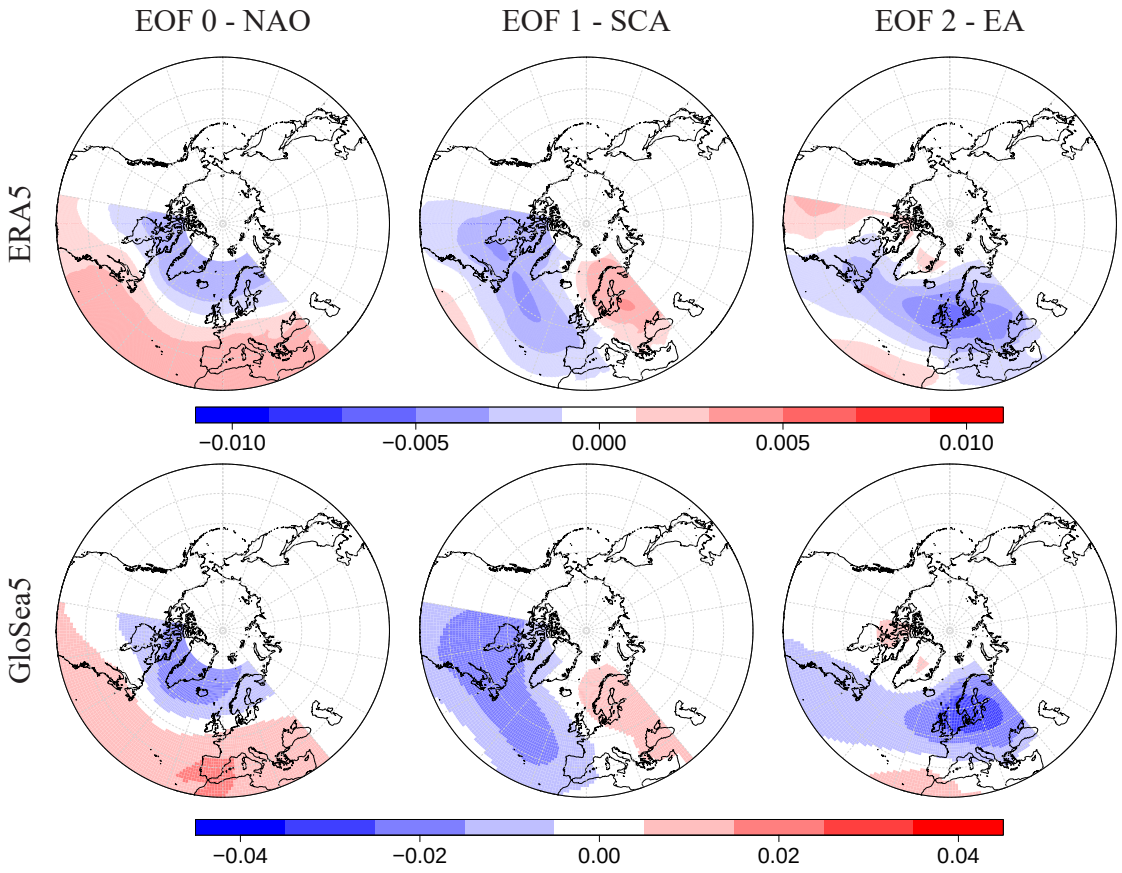


Figure B.1 EOF pattern of ERA5 and GloSea5 for EOF 1 to 3, correspond to NAO, SCA and EA for North Atlantic, DJF, 1993-2015.

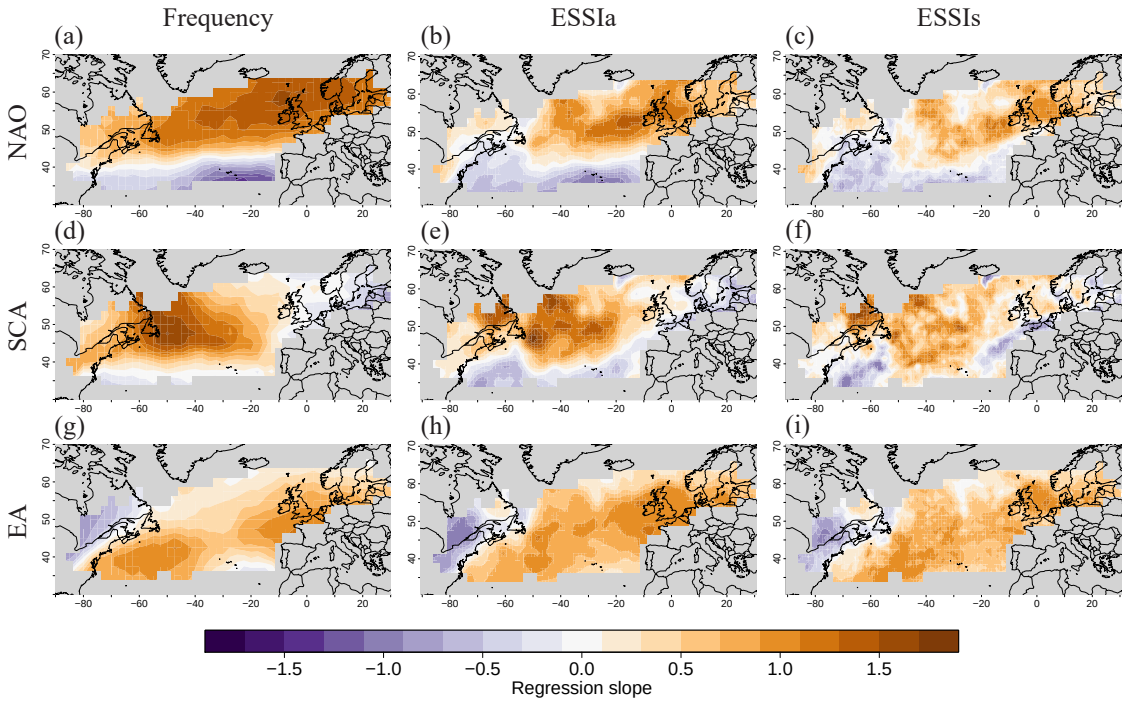


Figure B.2 Regression slopes as in Fig. 5.5 but based on GloSea5.

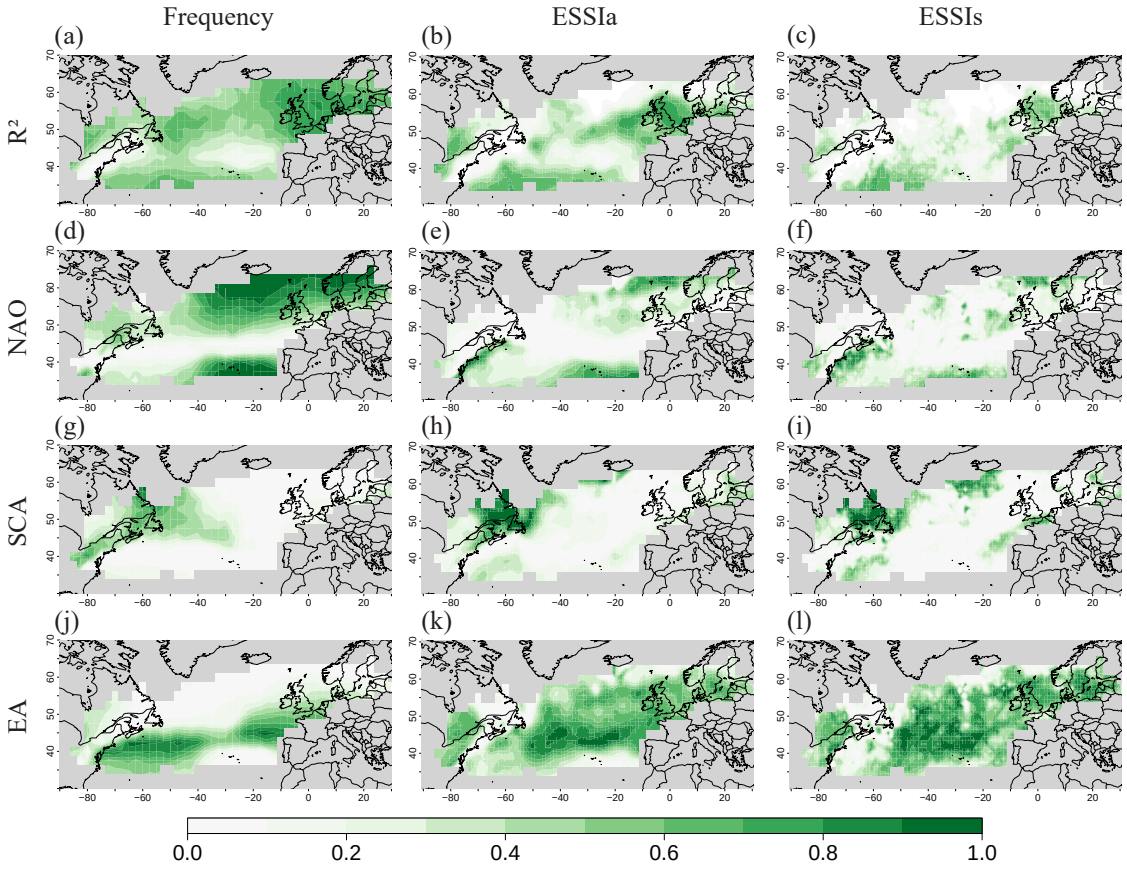


Figure B.3 Regression model variance as in Fig. 5.6 but with GloSea5 regression.

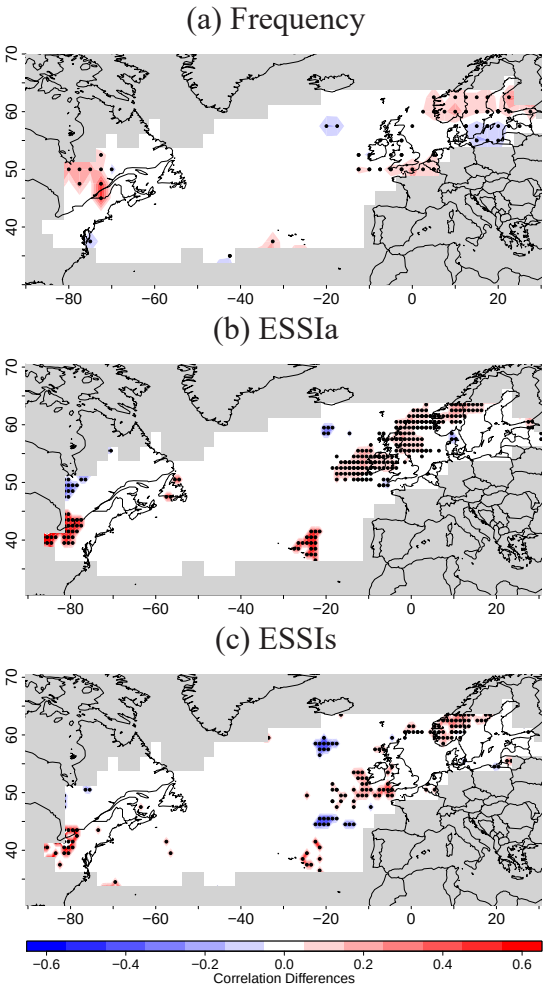


Figure B.4 Kendall Correlation Difference of direct and indirect approach for ERA5-based regression with the leave-one-out method.

Appendix C

For Chapter 6

This is the Appendix for paper No. 2 from chapter 6 and already published.

This part of the Appendix include additional information about the method and calculation for the dynamical factors, hence for chapter 6.1.3.

MSLP and SST represent more general information about environmental conditions. Their respective gradients are calculated using the NCL (NCAR Command Language) implemented function (grad_latlon_cfd) and computed the absolute value of the gradient vectors. The Climate Data Operator (CDO; Schulzweida et al., 2019) has an implemented function (uv2dv) to calculate the respective wind divergence from both wind components (u & v). For calculating the Rossby Wave Source (RWS), the python package windspharm (Dawson, 2016) was used as an example script from GitHub. This script is based on the RWS equation used, e.g., by Beverley et al. (2019) and Dunstone et al. (2018b). Studies like Parker et al. (2019) investigated the jet stream on its seasonal predictability and connection to the NAO. This study follows their calculation of jet location and speed but for 200hPa rather than 850hPa. The jet is defined over a 9-day running mean of the zonal average of the wind; both only the u-component and the total wind were tested. The jet location is defined here as the latitude at which the maximum wind (respectively u or total wind) is found, and as jet speed, the respective wind is used. An investigation from Wild et al. (2015) analysed how temperature anomalies over North America and the North Atlantic can influence the winter windstorm season over Europe. They created a Temperature-Dipole index which uses surface temperature at two regions, one over North America (105° - 80°W, 38° - 55°N) and one over

the western North Atlantic (85° - 50°W, 15° - 35°N). The difference in the respective anomalies creates the so-called temperature index. The PV (Hoskins, 2015; Hoskins et al., 1985) is calculated using two implemented NCL-functions (pot_vort_isobaric & int2p_n_Wrap). Therefore first, the pressure level data are used to calculate PV on pressure levels, and second, these values are interpolated onto *Theta* levels. The 350K level is later used in this study. The PV Advection is calculated from the pressure-level data and then advected by both (u & v) wind components. Θ_e as an individual factor on 850hPa (Chang et al., 1984), is calculated with the NCL-function, pot_temp_equiv. The Eady Growth Rate (EGR) is calculated with an implemented NCL function (eady_growth_rate) which uses the 2-layer approach. This means whenever it is referred to EGR at 400hPa, it is calculated using data from 300hPa and 500hPa, and for EGR at ∼700hPa, it is 500 & 850hPa. PV and EGR are additionally analysed in this study after post-processing with a bandpass filter. This bandpass filter was run with an R-implemented function using the Butterworth filter (Butterworth, 1930), with filter characteristics of 2 to 8 days for PV and 2 to 4 days for EGR. The filter was performed for each GloSea5 member individually. Because of data storage and computational times, the filtering was only executed for a region -100° to 40°E and 30° to 75°N. The total precipitation is used as in Scaife et al. (2017b) to investigate the link between tropical precipitation and the predictability of European climate conditions, like geopotential height. To be not restricted to the four used tropical regions used in Scaife et al. (2017b) and for a better comparison to the other used factors, the seasonal precipitation mean is investigated on the grid-cell level.

This part of the Appendix includes the results for the remaining tested dynamical factors. Therefore, it belongs to the Result chapter 6.1.4.

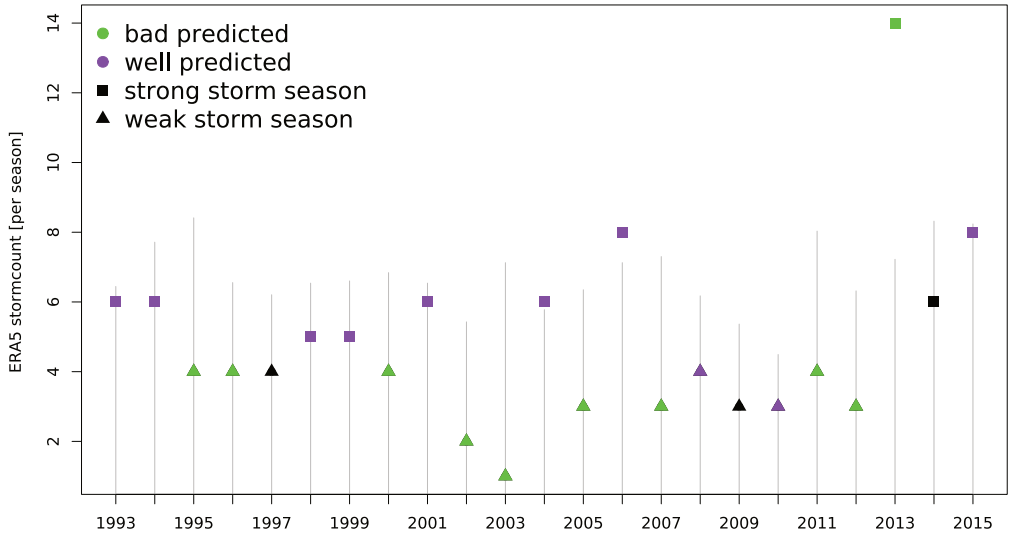


Figure C.1 ERA5 UK storm counts as dots and GloSea5 ensemble mean counts as bars. Bad predicted seasons (green), well predicted seasons (purple), weak ERA5 seasons (triangles) and strong ERA5 seasons (squares).

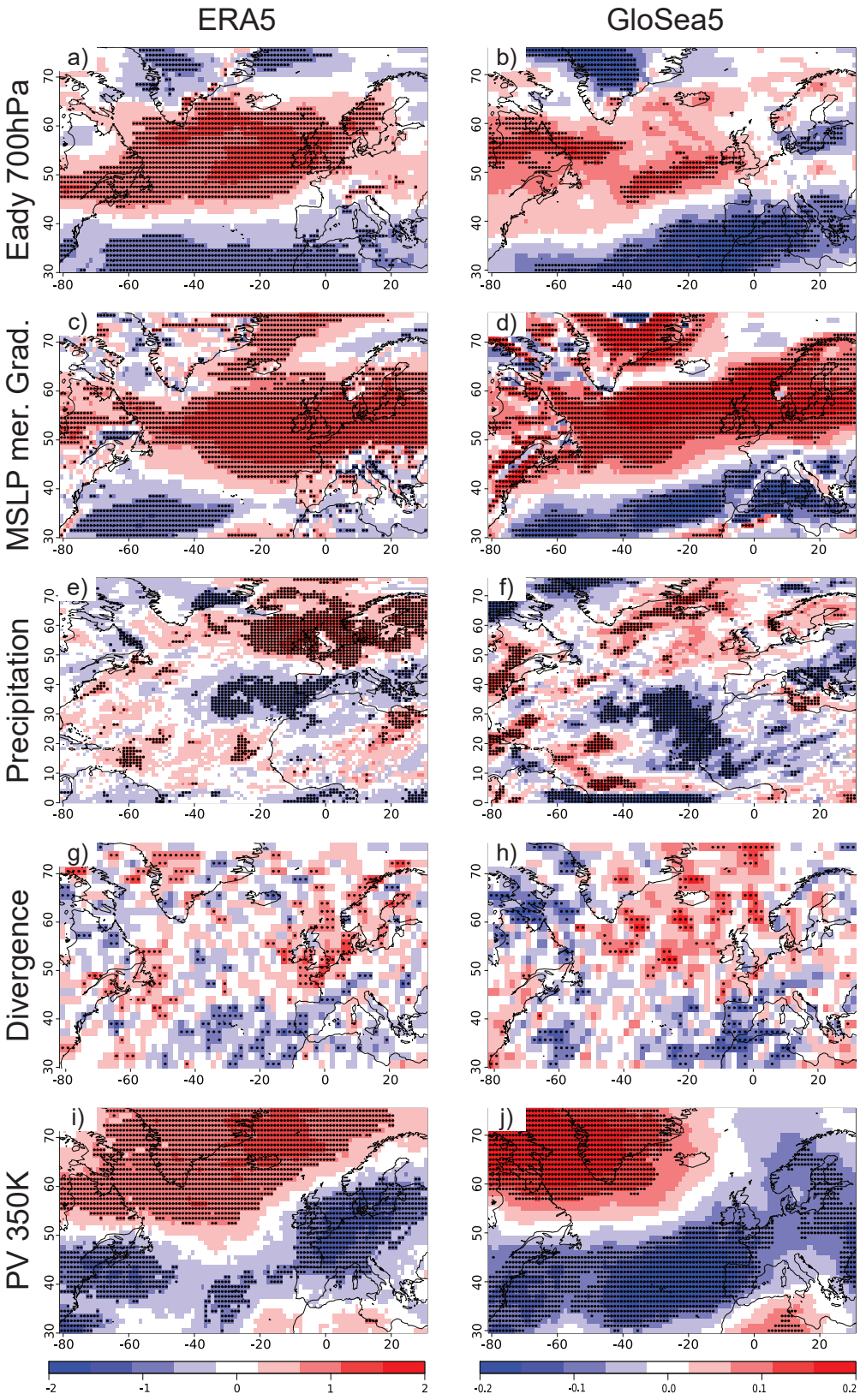


Figure C.2 As Fig. 6.3 for remaining primary and secondary factors.

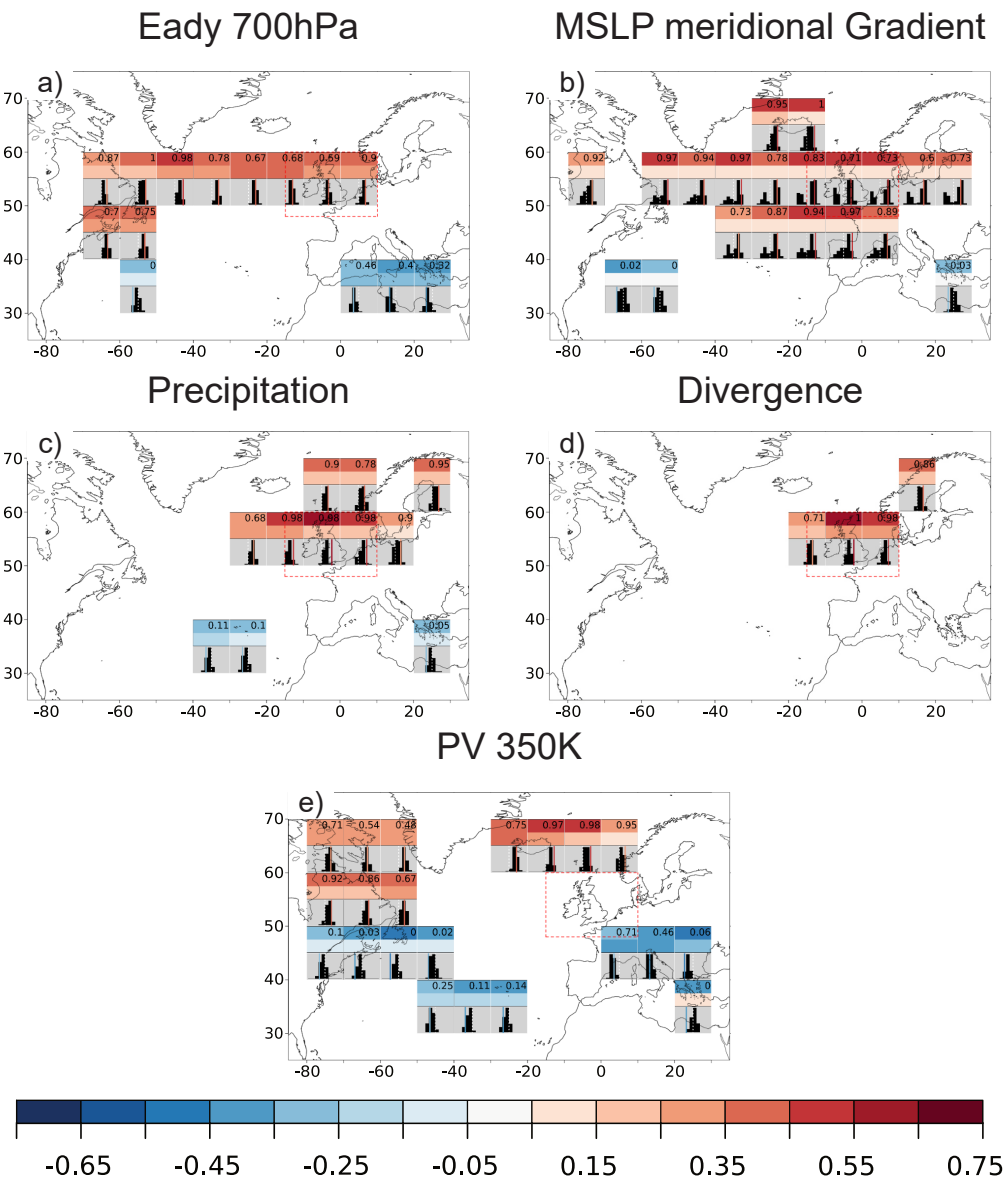


Figure C.3 As Fig. 6.4 for remaining primary and secondary factors.

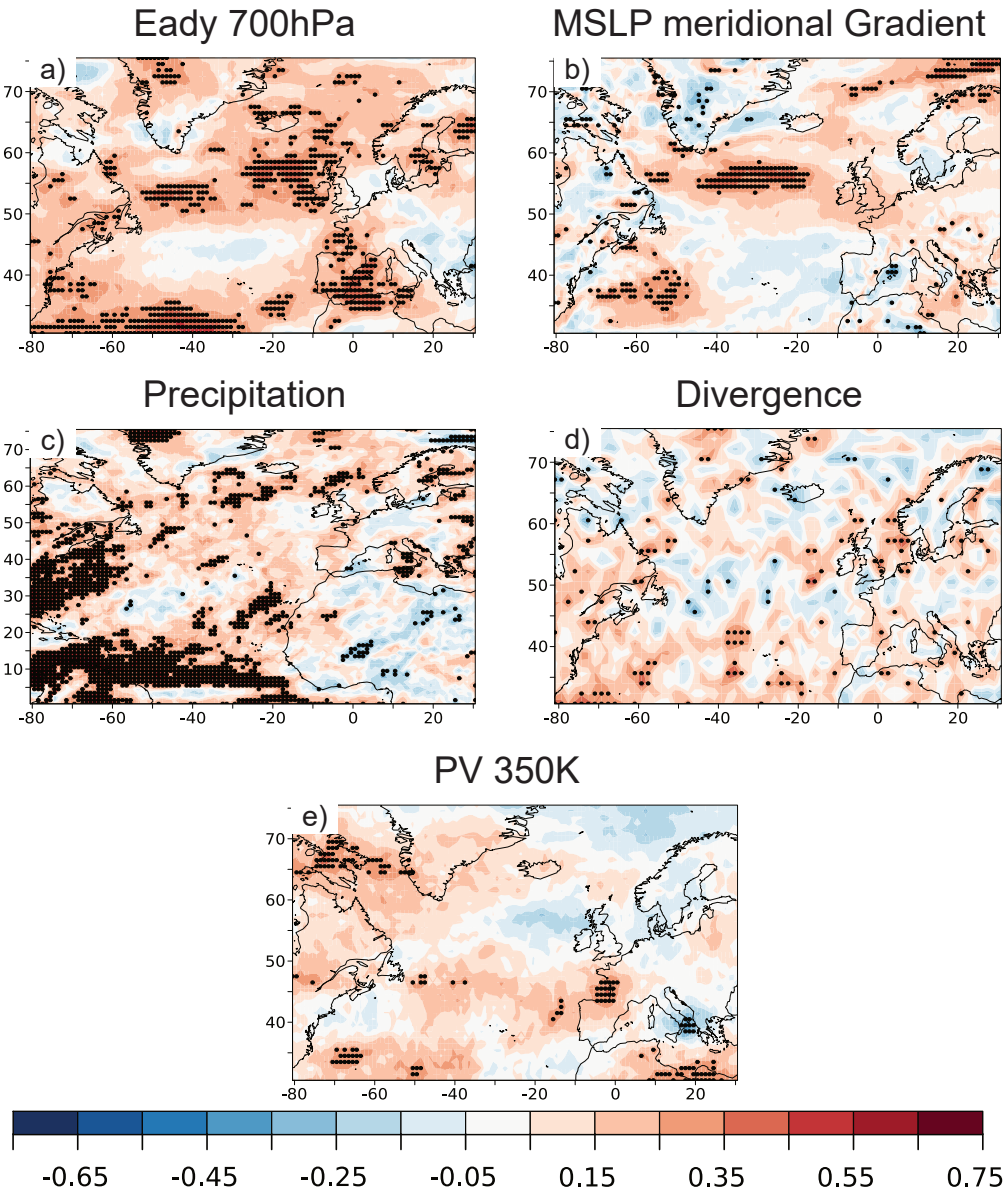


Figure C.4 As Fig. 6.5 for remaining primary and secondary factors.

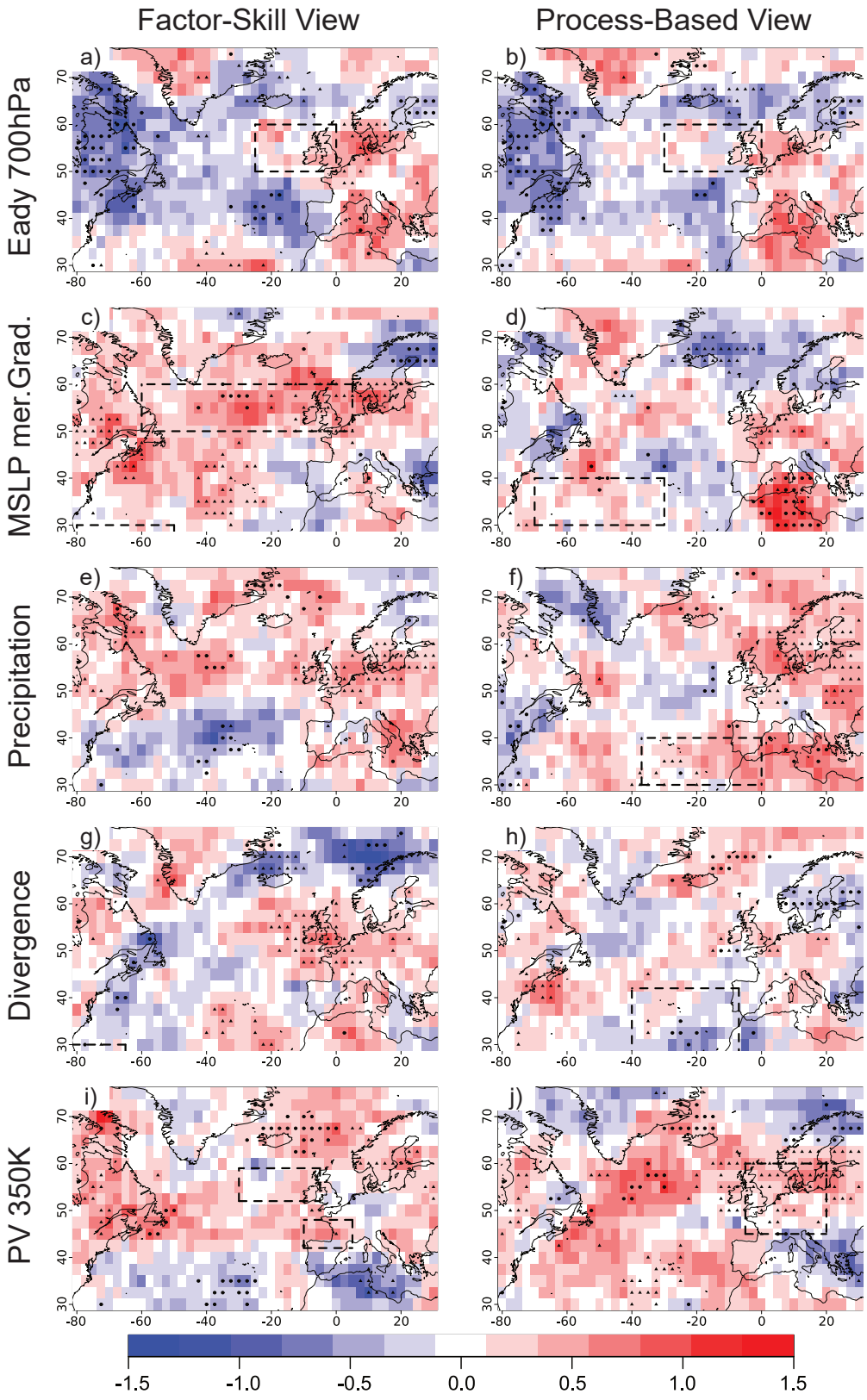


Figure C.5 As Fig. 6.6 for remaining primary and secondary factors.

Table C.1 Tested skilful regions of factor forecast skill.

Factor	Box 1	Box 2	Box 3	Boxmean
Mean Sea-Level Pressure Gradient	-15° - 10°E 52° - 65°N	-55° - -30°E 30° - 40°N	-85° - -62°E 27° - 37°N	Box 1-3
Sea-Surface Temperature	-35° - -5°E 40° - 64°N	-80° - -45°E 20° - 35°N	-80° - 0°E 10° - 60°N	Box 1&2
Equivalent Potential Temperature Θ_e	-10° - 5°E 48° - 58°N	-60° - -45°E 37° - 43°N	-55° - -20°E 55° - 62°N	Box 1-3
Eady Growth Rate 400hPa	-45° - -10°E 52° - 60°N			
Eady Growth Rate 700hPa	-25° - 0°E 50° - 60°N			
Mean Sea-Level Pressure Meridional Gradient	-60° - 5°E 50° - 60°N	-80° - -50°E 10° - 30°N		Box 1&2
Total Precipitation	-85° - -15°E 5° - 20°N	-90° - -55°E 20° - 45°N		Box 1&2
Divergence	-90° - -65°E 20° - 30°N			
Potential Vorticity 350K	-30° - -5°E 52° - 59°N	-10° - 5°E 42° - 48°N	-30° - -10°E 12° - 24°N	Box 1-3

Table C.2 Tested relevant regions of ERA5 & GloSea5 composite anomalies.

Factor	Box 1	Box 2	Box 3	Box 4	Boxmean
Mean Sea-Level Pressure Gradient	-40° - 0°E 30° - 40°N	-30° - 10°E 45° - 62°N	-40° - 0°E 15° - 30°N		Box 1-3
Sea-Surface Temperature	-80° - -50°E 27° - 40°N	-50° - -7°E 50° - 65°N	0° - 8°E 51° - 57°N	-20° - -10°E 21° - 27°N	Box 1-4
Equivalent Potential Temperature Θ_e	-70° - -40°E 30° - 42°N	-32° - 0°E 25° - 30°N	-10° - 30°E 47° - 60°N	-75° - -15°E 50° - 60°N	Box 1-4
Eady Growth Rate 400hPa	-50° - -20°E 30° - 40°N	-20° - 5°E 50° - 60°N			Box 1&2
Eady Growth Rate 700hPa	-70° - 10°E 25° - 35°N	-30° - 0°E 50° - 60°N			Box 1&2
Mean Sea-Level Pressure Meridional Gradient	-70° - -30°E 30° - 40°N	-30° - 10°E 45° - 57°N	-40° - 0°E 15° - 30°N		Box 1-3
Total Precipitation	-37° - 0°E 30° - 40°N	-25° - 10°E 50° - 62°N			Box 1&2
Divergence	-40° - -7°E 42° - 27°N	-15° - 7°E 45° - 63°N			Box 1&2
Potential Vorticity 350K	-5° - 20°E 45° - 60°N	-80° - -52°E 15° - 23°N			Box 1&2

Appendix D

For Chapter 7

These are additional dynamical forcing factors tested in the additional investigation (chapter 7.2) on the signal-to-noise paradox.

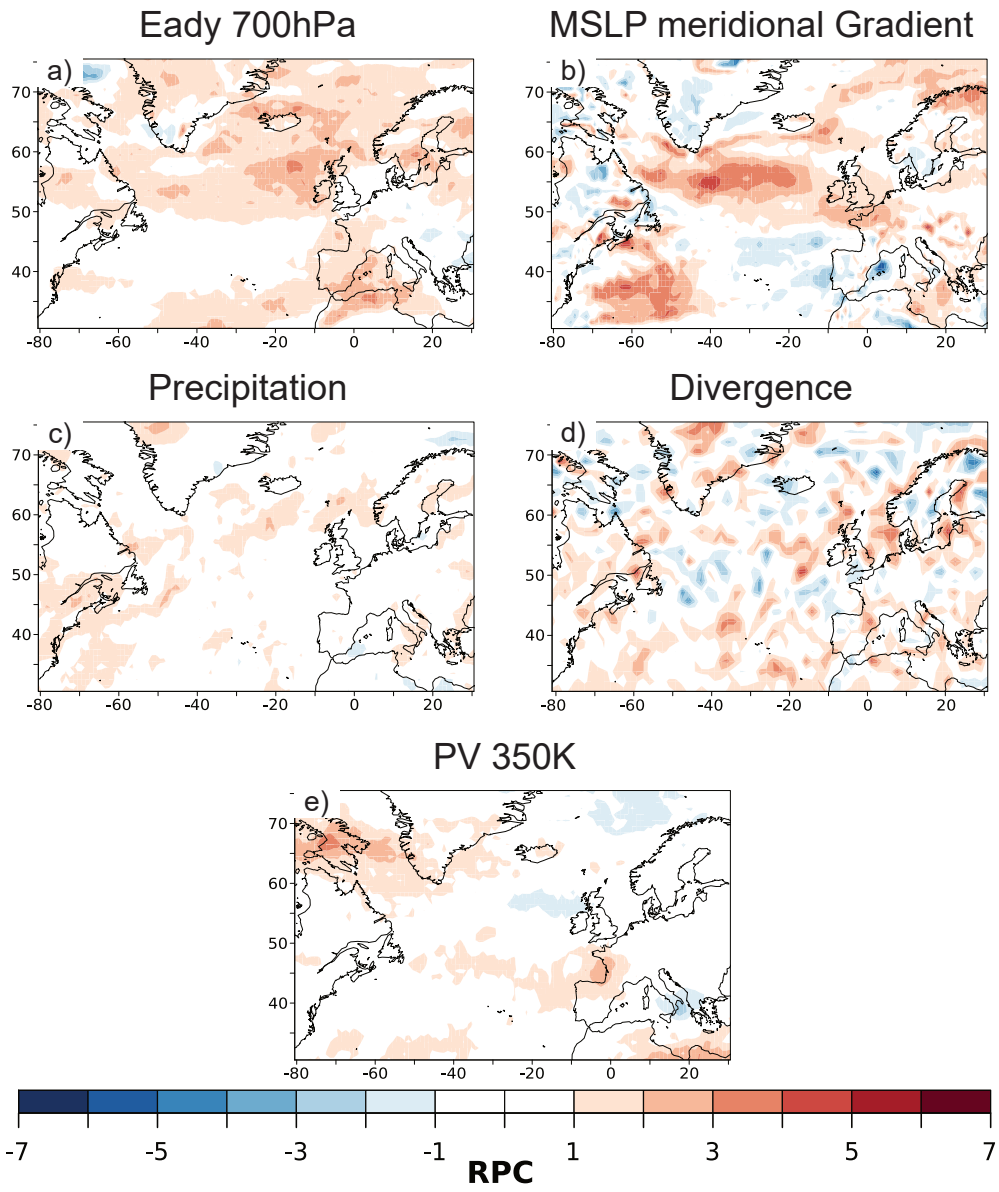


Figure D.1 Ratio of Predictable component (PRC) for the remaining primary and secondary dynamical factors.