



Call: H2020-SC5-2014-two-stage

Topic: SC5-01-2014

PRIMAVERA

Grant Agreement 641727



PRocess-based climate sIMulation: AdVances in high resolution modelling and European climate Risk Assessment

Deliverable D10.3

***Comparison of physics of selected meteorological events
in CMIP5, CORDEX and in PRIMAVERA models***

Deliverable Title	<i>Comparison of physics of selected meteorological events in CMIP5, CORDEX and in PRIMAVERA models</i>	
Brief Description	<p><i>Assess the model biases and representation of physics of selected extreme events drawing from results of WP 1, 2, 3 as simulated by PRIMAVERA high-resolution models (Stream 1 and 2) and compare them to the existing CMIP5 and CORDEX simulations. Focus will be on mechanisms controlling storm development and their tracks: tropical (intense rain, storm surge) and extra-tropical (intense rain, wind), as well as extra-tropical transition along changing storm tracks, particularly in the North Atlantic sector.</i></p>	
WP number	10	
Lead Beneficiary	<i>Alexander Baker, UREAD</i>	
Contributors	<i>Gustav Strandberg, SMHI David Brayshaw, UREAD Paula Gonzalez, UREAD Galia Guentchev, UKMO Julia Lockwood, UKMO Alexander Baker, UREAD Panos Athanasiadis, CMCC Reinhard Schiemann, UREAD Ségolène Berthou, UKMO Erika Palin, UKMO Rein Haarsma, KNMI Marie-Estelle Demory, ETHZ Malcolm Roberts, UKMO Pier Luigi Vidale, UREAD</i>	
Creation Date	22/10/2019	
Version Number	4	
Version Date	30/01/2020	
Deliverable Due Date	31/01/2020	
Actual Delivery Date	31/01/2020	
Nature of the Deliverable	X	<i>R - Report P - Prototype D - Demonstrator O - Other</i>
Dissemination Level/ Audience	X	<i>PU - Public PP - Restricted to other programme participants, including the Commission services RE - Restricted to a group specified by the consortium, including the Commission services CO - Confidential, only for members of the consortium, including the Commission services</i>

Version	Date	Modified by	Comments
1	30/10/2019	Alexander Baker	Outline
2	18/12/2019	Alexander Baker	First draft
3	24/01/2019	All contributors	Revised draft
4	30/01/2019	Alexander Baker	Final draft for submission

Table of Contents

1. Executive Summary.....	5
2. Project Objectives.....	10
3. Detailed Report.....	11
3.1 The North Atlantic midlatitude storm track.....	11
3.1.1 Comparison of physics of extreme events in PRIMAVERA, CMIP5 and CORDEX models.....	11
3.1.2 The ability of PRIMAVERA models to simulate extreme surface wind speeds.....	25
3.2 Atmospheric blocking.....	32
3.2.1 Blocking evaluation overview.....	32
3.2.2 Impact of blocking on low wind events and its representation by high-resolution GCMs: An energy perspective.....	36
3.3 The North Atlantic eddy-driven jet.....	49
3.3.1 Representation of the North Atlantic eddy-driven jet in PRIMAVERA simulations.....	49
3.4 Precipitation over Europe.....	56
3.4.1 Comparing simulated daily precipitation distribution between PRIMAVERA and CORDEX ensembles.....	56
3.4.2 The importance of model resolution of simulated precipitation in Europe.....	64
3.4.3 The impact of atmospheric resolution on simulated extreme European precipitation..	74
3.5 Potential emerging hazards for Europe.....	78
3.5.1 Post-tropical cyclones impacting Europe.....	78
3.6 Temperature extremes.....	84
3.6.1 Temperature trends over Europe.....	84
4. Lessons Learnt.....	99
5. Links Built.....	99
5.1 Outside PRIMAVERA.....	99
5.2 Between PRIMAVERA Work Packages.....	100
6. References.....	100

1. Executive Summary

1.1 Scope

This deliverable—D10.3—concerns model biases and the representation of extreme events, as simulated by low- and high-resolution models participating in PRIMAVERA and CMIP6 HighResMIP. It draws from results of PRIMAVERA Work Packages (WP) 1, 2, 3, 10 and 11, focussing on Stream 1 integrations and the role of atmospheric horizontal resolution. Comparisons with existing (global) CMIP5 and (regional) EURO-CORDEX simulations were performed to establish where high-resolution global simulations add value to both the simulation and our understanding of extreme events, focussing here across Europe and the North Atlantic sector. While a constrained geographic region is the focus of this report, it should be borne in mind that extreme events affecting European nations' overseas territories, particularly tropical cyclones, as well as events affecting regions where European sectors have commercial interests, are relevant to the overall European risk portfolio. Future deliverables, particularly D10.4, will further explore the value of global modelling to address risk across borders.

This deliverable considers a range of topics unified by a key underlining research question: whether or not, under the HighResMIP protocol, increased model resolution (atmosphere or ocean) increases simulation fidelity and/or whether processes emerge at high resolution. The occurrence of synoptic-scale extremes and large-scale atmospheric variability and circulation are directly linked. Two key advantages of PRIMAVERA's experimental design are exploited by the research summarised in this report. Firstly, results are based on multi-model analysis and, in many cases, conclusions are drawn from ensemble-to-ensemble comparisons. Secondly, PRIMAVERA allows for interaction across spatial scales, with smaller ('weather') scales potentially feeding back onto large-scale variability and impacting the simulate climate. As such, this deliverable considers (i) midlatitude cyclones and the North Atlantic storm track, including wind hazards; (ii) large-scale North Atlantic blocking and eddy-driven jet stream variability, including impacts on low-wind events; (iii) extreme precipitation and temperature occurrences. As research has progressed, the scope of D10.3 has changed since the Description of Work: less emphasis is placed on the extratropical transition of tropical cyclones in the midlatitudes and the scope expanded to incorporate temperature extremes.

1.2 Storm tracks and midlatitude cyclones

As part of WP11, interviews were conducted with eight participants from the insurance industry (see D11.6 and D10.1) to determine the potential benefit from high-resolution global climate models. The sector's key concerns were windstorms and flooding, which are the highest-loss hazards affecting insurance policies. Due to the lack of sufficient observational data, climate models can be used to augment observational datasets to estimate long return period losses. PRIMAVERA models could prove very useful for this purpose if it can be shown that they realistically simulate windstorms and flooding, which are primarily related to the occurrence of extra-tropical cyclones (ETCs). Here, ETC characteristics between

reanalysis data, PRIMAVERA and CMIP5 simulations were compared with establish what added value could support insurance sector interests.

The characteristics of ETCs in PRIMAVERA models compare well to reanalyses. A significant improvement compared with CMIP5 is seen in the distribution of ETC intensities, as measured by minimum SLP and maximum vorticity, as PRIMAVERA models better simulate more extreme ETCs. Track density biases are smaller across PRIMAVERA models compared with CMIP5 models.

A use case for the insurance/finance industry outlined in D10.1 concerns using PRIMAVERA model data to estimate the present-day risk of European windstorm damage. Windstorm damage is typically quantified by a loss or severity index, of which there are many definitions, but most are proportional to the cubic exceedance of maximum winds or gusts over a threshold—an approach taken herein. Large biases in such a severity index, particularly over high-altitude regions, are seen in low-resolution PRIMAVERA models, which could result from different parameters (e.g., roughness length) used in the models to estimate 10m wind speeds, but this requires further investigation. Overall, most models, particularly MOHC, MPI-M, EC-Earth, show an improvement in representation of a severity index with increasing resolution, although some models show an increase in bias (CMCC, CNRM-CERFACS). More investigation is needed to establish whether these changes with resolution are caused by changes in storm frequency, storm intensity or another cause. The correct representation of extreme wind speeds is important for estimating insured losses due to storm damage, and major biases will need to be corrected for a model to give realistic estimations of storm loss.

Upon completion, PRIMAVERA will have considered the whole annual cycle of storm hazards. Windstorms, which are typical during the cold season, have been a key focus across WPs one and 10. Ongoing work is investigating summer storms with tropical origins—termed post-tropical cyclones. Model improvements seen in the representation of wintertime ETCs do not hold true for post-tropical systems, which establishes a priority for further research.

1.3 Large-scale atmospheric circulation variability

The representation of blocking is key to capturing extreme event occurrence because blocking events coincide with a range of weather extremes, including temperature extremes prevailing over blocked areas and storm activity and heavy precipitation over regions adjacent to blocks. An evaluation of blocking has contributed to IPCC WG1 AR6. PRIMAVERA and CMIP6 HighResMIP models exhibit improvement (compared with reanalysis data) over CMIP5 models in simulated blocking, which is seen most clearly over the North Sea in winter. There is also improvement with resolution in the PRIMAVERA atmosphere-land-only and coupled ensembles, both in the Atlantic and Pacific in winter and summer. This improvement is seen more clearly in the coupled simulations than in the forced simulations. Moreover, in coupled simulations, increased resolution improves eddy-driven jet latitude variability compared with reanalysis data. A key conclusion, therefore, is that horizontal resolution, as investigated within PRIMAVERA, is one of the factors important

for simulating atmospheric blocking and jet position, yet an increase in resolution to about 25 km alone does not fully remedy blocking biases in climate models.

It has been shown that blocking events over Europe have a significant impact on the occurrence and duration of low wind speeds at the country level, which is of direct relevance to the energy sector. Low-wind events are more frequent and more persistent under blocking conditions over large areas of Europe. In general, both effects are captured by most of the PRIMAVERA models, revealing that, under highresSST-present forcing, models that simulate blocking reasonably well capture the basic dynamical connection with wind anomalies. Nonetheless, that deficiencies in simulated weather conditions exist introduces biases in the properties of the events and their joint occurrence. Such model errors depend on the metric employed, country, and resolution, but some spatially consistent bias patterns are found (e.g., north-south dipolar structures). Despite the overall improvement in Euro-Atlantic blocking statistics found in PRIMAVERA simulations, it has proven difficult to identify robust improvements in the corresponding impact of blocking on persistent low-surface-wind events. This indicates that caution should be exercised in the use of energy system simulations based on GCM output. In particular, although model wind speed mean biases may be corrected, errors in the frequency or duration of weather events are less easily overcome and may introduce errors in wind power and energy demand simulations. In general, however, high-resolution global simulations, such as those delivered by PRIMAVERA, offer potential benefits as higher-fidelity driving models for regional climate downscaling. Their improved representation of large-scale phenomena, such as blocking and storm track processes, goes some way to mitigating systematic errors in regional outputs inherited from coarse-resolution models' comparably poor representation of these drivers of extreme event occurrence.

1.4 Extreme precipitation

Extreme precipitation events across the Euro-Atlantic region, particularly during winter, are primarily related to midlatitude storm occurrence, indicating that global models, which capture storm track processes and variability, provide information that cannot be captured by regional models. In PRIMAVERA, efforts have concentrated on comparison between PRIMAVERA and EURO-CORDEX (EUR-44 and EUR-11) and on applying multiple extreme precipitation metrics.

CORDEX and PRIMAVERA differ most for the most intense precipitation rates over all regions. EUR-44 overestimates intense precipitation, but PRIMAVERA models are generally in better agreement with observations. When using reduced CORDEX and PRIMAVERA ensembles (i.e., those models common between the two), the ensembles exhibit greater similarity, implying the majority of the precipitation rate distribution across most European sub-regions depends on model formulation and physics rather than on the downscaling method. The largest difference is found for the most intense precipitation rates in most seasons and regions. PRIMAVERA appears to provide better performance in general than the full EUR-44 ensemble when compared with raw observations, but EUR-44 is closer to a synthetic observational dataset in which an average 20% precipitation under-catch error is applied. PRIMAVERA and CORDEX (both EUR-11 or EUR-44) should therefore be considered equally credible, depending on the users' needs. The added value of RCMs over

CMIP5 GCMs emphasizes the importance of a well-designed, well-evaluated model chain when using dynamical downscaling as a method to obtain higher resolution climate data. We show in this report that considering climate information from various sources is crucial.

Two further studies investigate the importance of model resolution on the simulated precipitation in Europe, aiming to investigate the differences between models and model ensembles and evaluate models' agreement. It is clear that the type of model has a large effect on precipitation, mostly on more extreme precipitation. For example, the number of precipitation days does not depend significantly on resolution, and instead depends primarily, at least for annual precipitation, on large-scale weather patterns rather than local topography and convection. For extreme precipitation events that are more local and short-lived, model resolution is more important. A high-resolution model better resolves such events and distinguishes better between different sub-regions. Thus, extreme precipitation is more extreme and more frequent in high-resolution models compared with their low-resolution counterparts. Given the same frequency of wet days, precipitation intensifies such that wet days are wetter.

Employing generalised extreme value metrics reveals improved representation of 'typical' extreme precipitation values compared with observations when resolution is increased from low to high, but the representation of year-to-year variability in extremes is degraded. Extreme precipitation increases across the downstream region of the North Atlantic storm track, particularly over ocean, in most PRIMAVERA models at high resolution. Crucially, improved model performance—demonstrated by reduced root-mean-square error—is also found in most models over this region. Spatially widespread inter-model agreement in both of these results is found, which is key in establishing robustness and indicates that, at CMIP5-like resolutions, extreme precipitation over the Euro-Atlantic region is underestimated. These results are consistent with Baker et al. (2019a) and will form the basis of a model evaluation manuscript currently in preparation.

Resolution sensitivity of extreme European precipitation is complex. Higher percentiles are impacted by resolution increase more than low percentiles for all studied indices. Increasing resolution has similar same effects in both global PRIMAVERA and regional CORDEX simulations. Furthermore, GCMs and RCMs of comparable resolution simulate comparable precipitation climates. Increasing resolution from low (~100 km) to moderate (~60 km) tends has the largest effect of increasing precipitation; increasing from moderate to high (~20 km) has a comparably small effect. This does not, however, mean that incremental increases in resolution become less and less worthwhile; once the models reach convection-permitting (not shown herein) resolutions (~3 km), resolution increases explicitly resolve mesoscale precipitation-generating processes.

1.5 Extreme temperature events

Low- and high-resolution PRIMAVERA models were compared over 1970–2014 to gridded and homogenised daily series of observed land surface temperatures. This analysis was performed focusing on mean and trend biases in mean values of winter minimum temperatures and summer maximum temperatures. Other indices considered here include the number of days with minimum temperatures below the 10th percentile of winter values

('cold nights') and (ii) the number of days with maximum temperatures exceeding the 90th percentile of summer values ('warm days').

Common spatial patterns were found among models, such as an underestimation of winter minimum temperatures over Italy and Norway and an overestimation in the north of Sweden and Finland, which may be related to a lack of snow coverage simulated by the models. The models share a common north-south gradient in maximum summer temperature biases, with warmer values along the European coasts of the Mediterranean. This may be related to excessive moisture in Northern Europe and a lack of moisture in Southern Europe. During summer, high-resolution models overestimate observed trends over Northern Europe and underestimate those observed over Southern Europe.

Trends in extreme values show common spatial patterns among the models. Models underestimate trends in wintertime night temperatures simulated over Eastern Europe and underestimate the percentage of cold days—thus, warmer trends—over Southern Europe. At the same time, models underestimate the percentage of warm days, indicating colder trends than observed, consistent with the findings of Min et al. (2013) for CMIP5. The most serious multi-model discrepancy is a substantial underestimation of the observed increasing trend of warm extremes over the Mediterranean. Considering the high economic and societal vulnerability of these areas to very-warm events during summer, as well as the importance of the prediction of heatwave intensity and frequency for the next decades, it is fundamental to improve the simulation of these phenomena and of their projections to future decades. Model development beyond HighResMIP may therefore be needed to address these shortcomings.

2. Project Objectives

With this deliverable, the project has contributed to the achievement of the following objectives (DOA, Part B Section 1.1) WP numbers are in brackets:

No.	Objective	Yes	No
A	To develop a new generation of global high-resolution climate models. (3, 4, 6)		X
B	To develop new strategies and tools for evaluating global high-resolution climate models at a process level, and for quantifying the uncertainties in the predictions of regional climate. (1, 2, 5, 9, 10)	X	
C	To provide new high-resolution protocols and flagship simulations for the World Climate Research Programme (WCRP)'s Coupled Model Intercomparison Project (CMIP6) project, to inform the Intergovernmental Panel on Climate Change (IPCC) assessments and in support of emerging Climate Services. (4, 6, 9)		X
D	To explore the scientific and technological frontiers of capability in global climate modelling to provide guidance for the development of future generations of prediction systems, global climate and Earth System models (informing post-CMIP6 and beyond). (3, 4)	X	
E	To advance understanding of past and future, natural and anthropogenic, drivers of variability and changes in European climate, including high impact events, by exploiting new capabilities in high-resolution global climate modelling. (1, 2, 5)		X
F	To produce new, more robust and trustworthy projections of European climate for the next few decades based on improved global models and advances in process understanding. (2, 3, 5, 6, 10)	X	
G	To engage with targeted end-user groups in key European economic sectors to strengthen their competitiveness, growth, resilience and ability by exploiting new scientific progress. (10, 11)	X	
H	To establish cooperation between science and policy actions at European and international level, to support the development of effective climate change policies, optimize public decision making and increase capability to manage climate risks. (5, 8, 10)		X

3. Detailed Report

3.1 The North Atlantic midlatitude storm track

3.1.1 Comparison of physics of extreme events in PRIMAVERA, CMIP5 and CORDEX models

Galina Guentchev, Julia Lockwood, Erika Palin (UKMO)

3.1.1.1 Background

As part of WP11, interviews were conducted with 8 participants from the insurance industry, to discover how they might benefit from high resolution global climate models (see PRIMAVERA deliverables 11.6 and 10.1). Their main concerns were windstorms and flooding, since these are the highest loss hazards covered by property insurance policies. Due to the lack of sufficient observational data, climate models can be used to augment observational datasets to estimate long return period losses. PRIMAVERA models could prove very useful for this purpose if it can be shown that they realistically simulate windstorms and flooding. Since extra-tropical cyclones (ETCs) are the main cause of winter European windstorms, as well as being associated with flooding in this document we compared ETC characteristics between re-analysis dataset, the PRIMAVERA models, and CMIP5 models to see if PRIMAVERA shows any improvement.

A brief summary of some results already presented in D10.2 is given below for the sake of completeness. Present day AMIP runs at the highest resolution for the set of PRIMAVERA models were compared with their counterpart models with lower resolution from CMIP5 (Table 3.1) and to 2 reanalysis datasets (ERA Interim and MERRA2) in terms of extratropical cyclones' (ETCs) characteristics. The comparisons were performed over the winter season (DJF) for the period 1979/80-2007/08. A brief summary of the results is provided in the following section.

3.1.1.2 D10.2 summary

The characteristics of ETCs in PRIMAVERA models compare well to the MERRA2 and the ERA Interim reanalyses. A large improvement compared with CMIP5 is seen in the distribution of ETC intensities as measured by minimum MSLP and maximum vorticity, as PRIMAVERA models better simulate more extreme ETCs. Regarding track densities, the PRIMAVERA models are characterised by overall smaller biases compared with the CMIP5 models.

3.1.1.2.1 ETC track density

The results were mixed: the ECEARTH and the MPI models showed large areas of reduced bias in PRIMAVERA, but the CMCC model showed an increased bias in northern Europe

(although the very high resolution CMCC model was not tracked yet at the time of the report – D10.2). Overall, the PRIMAVERA models were characterised by smaller biases compared with the CMIP5 models. One exception was the CMCC model which had a higher resolution in the CMIP5 set of models at the time of the report (D10.2).

3.1.1.2.2 ETC variability

All models had similar standard deviations in the ETC numbers and compare well to ERA Interim. The PRIMAVERA high-resolution models show somewhat higher numbers of storms entering the European domain compared with the MERRA and ERA Interim reanalyses. The CMIP5 models compare favourably to the MERRA reanalysis-based storm frequencies and standard deviation values.

3.1.1.2.3 ETCs intensity—measured in terms of Minimum Sea Level Pressure (MSLP)

The CMIP5 models tend to underestimate the number of extreme ETCs in Europe (with minimum MSLP < 970hPa). This bias is reduced in PRIMAVERA models, although some PRIMAVERA models underestimate the number of weaker ETCs. In general, the differences are small. The CMCC PRIMAVERA model seems to be an outlier.

3.1.1.2.4 ETCs intensity—measured in terms of maximum vorticity

The CMIP5 models seem to underestimate the frequency of more extreme storms with higher vorticity, while overestimating the frequency of lower vorticity storms. These biases are largely reduced in the PRIMAVERA models especially regarding the underestimation of the stronger storms with higher vorticity.

3.1.1.2.5 ETCs intensity—measured in terms of maximum 925hPa winds

This analysis is based solely on PRIMAVERA models, since there was no data about max 925hPa winds for the CMIP5 models. Several PRIMAVERA models overestimate the upper section of the wind distribution vs MERRA – indicating overestimation of the frequency of storms with higher wind speeds. Some models underestimate the lower end of the distribution – representing lower frequencies of storms with lower wind speeds at 925hPa level. The biases are much smaller vs the MERRA reanalysis compared with the biases evident when the PRIMAVERA models are compared with the ERA Interim.

Modelling centre	PRIMAVERA model analysed	CMIP5 model analysed
CMCC	CMCC-CM2-VHR4 (18km)	CMCC-CM (70km)
CNRM	CNRM-CM6-1-HR (50km)	CNRM-CM5 (100km)
ECEARTH	ECEARTH3-HR (36km)	ECEARTH (80km)
MOHC	HadGEM3-GC31-HM (25km)	HadGEM2-A (90km)
MPI	MPIESM-1-2-XR (34km)	MPI-ESM-MR (130km)
ECMWF	ECMWF-IFS-HR (25km)	Not available for analysis

Table 3.1. Summary of models analysed with approximate native horizontal resolutions at mid latitudes (50°N) for the PRIMAVERA models. CMIP5 resolutions from <https://portal.enes.org/data/enes-model-data/cmip5/resolution>

3.1.1.3 Analysis of temporal clustering of the winter storms as represented by the CMIP5 and the PRIMAVERA counterpart models

Temporal clustering was calculated for the ETCs passing through the European region (-15°E to 25°E, 35° to 70°N). Storms lasting 2 days and passing within 6.3° radius from a template (2.5° x 2.5°) of grid points were counted. The choice of the 6.3° radius follows the method by Economou et al. (2015) where they explain that: “6.3° (700 km) radius circle is considered following the approach of Pinto et al. (2013). This choice of radius is within the range of effective radius for extratropical cyclones (600 – 1000 km; Rudeva and Gulev, 2007) and corresponds to a plateau of quasi-constant values of a particular dispersion statistic (section 2.4) over most of the study area”. The temporal clustering is defined by the dispersion (ratio of variance to mean) of the December–February counts of North Atlantic storms, following Economou et al. (2015), where positive (negative) values of the dispersion statistic indicate clustering (regularity). Only one simulation per model was included in the analyses from all of the models including those that were run in ensemble-mode. Figures 3.1 and 3.2 show the temporal clustering as represented by the ERAI and the CMIP5, as well as by the PRIMAVERA high resolution models (Table 3.1). A typical distribution of high clustering of storms is evident in the ERAI map towards Iceland and northern Europe and towards the British Isles and over central Europe (shown in yellow/orange/red colours, Figure 3.1, upper left corner). Most of the CMIP5 models represent lower levels of clustering in these areas as indicated by the maps of the differences.

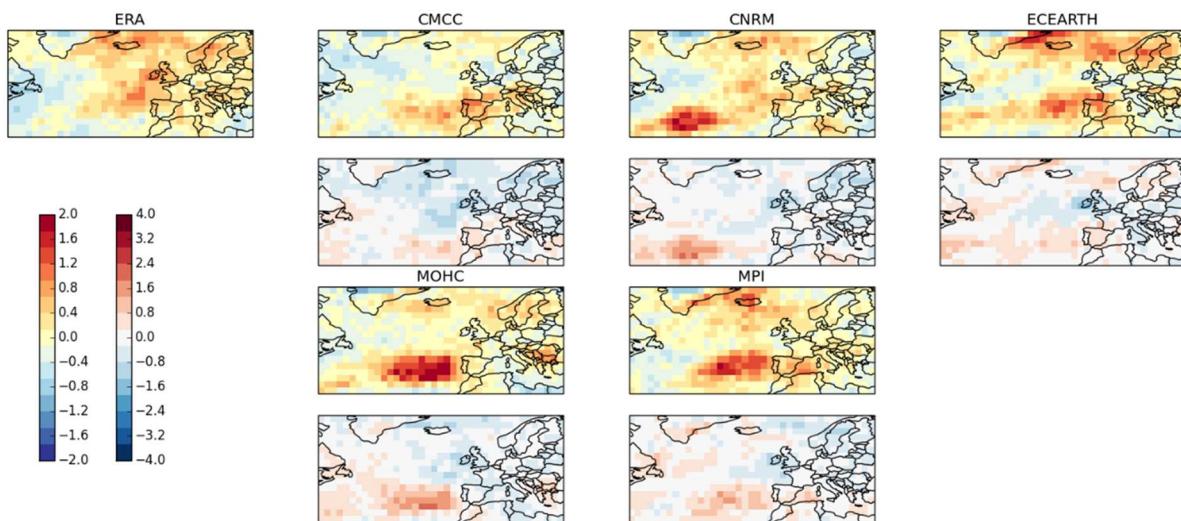


Figure 3.1. Temporal clustering as represented by the dispersion statistic calculated from ERAI, and the CMIP5 models (maps in rainbow colours, positive (negative) values of the dispersion statistic indicate clustering (regularity).) and differences (biases) between the CMIP5 models and ERAI (in red and blue colours).

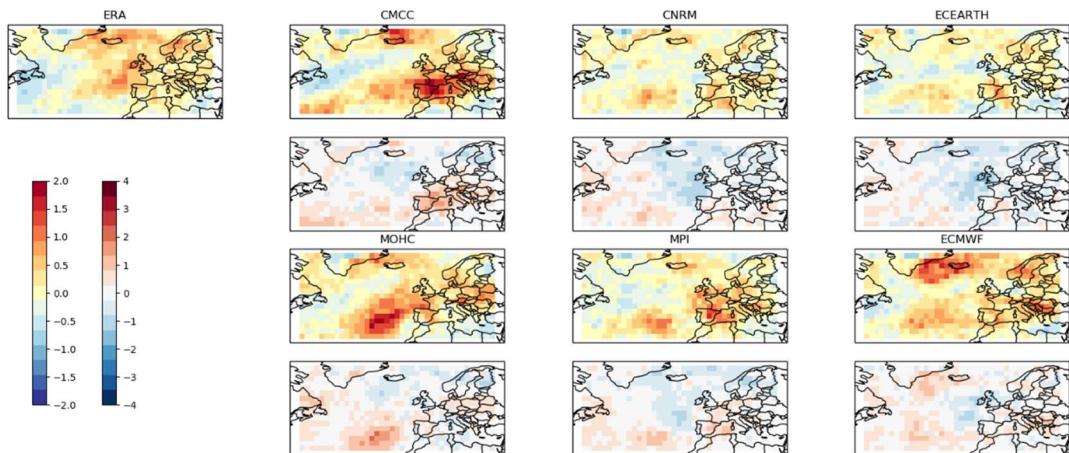


Figure 3.2. Temporal clustering as represented by the dispersion statistic calculated from the ERA Interim, and the PRIMAVERA high resolution models (maps in rainbow colours, positive (negative) values of the dispersion statistic indicate clustering (regularity).) and differences (biases) between the PRIMAVERA models and ERAI (in red and blue colours).

While overall the PRIMAVERA models are still representing somewhat lower clustering towards Iceland and northern Europe and/or the British Isles, several models represent somewhat higher clustering over central Europe – CMCC, MOHC, MPI, ECMWF (Figure 3.2).

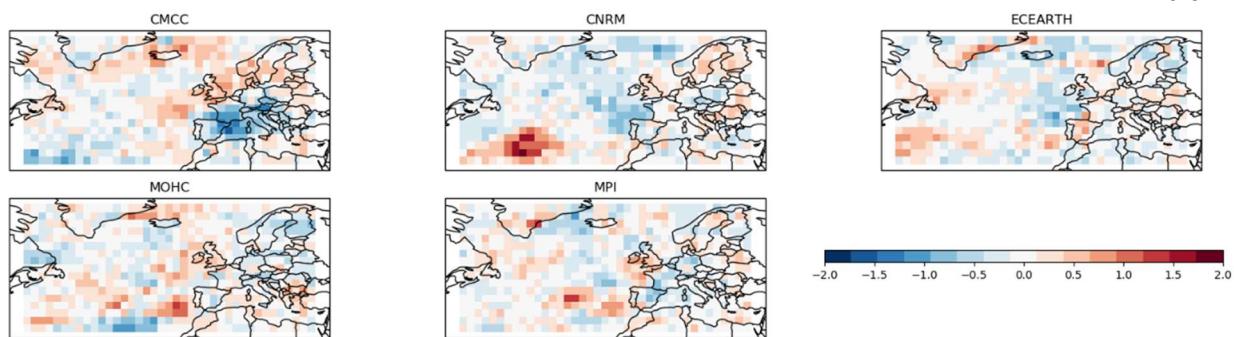


Figure 3.3. Change in temporal clustering bias between CMIP5 and PRIMAVERA ($|CMIP5 \text{ bias}| - |PRIMAVERA \text{ bias}|$), as compared with ERAI. Red areas (reduction in bias) show where there is improvement in PRIMAVERA models vs. CMIP5 models.

Generally, an improvement is evident in the CMCC model in the area around Iceland, in northern Europe, and towards the British Isles, and in the MOHC model towards the British Isles and north of Iceland; The rest of the models either do not show definite improvement in this characteristic (like the MPI model) or indicate definite increase in bias within the PRIMAVERA models (CNRM and ECEARTH) in representing the temporal clustering within the northeast Atlantic (towards Iceland and the British Isles).

3.1.1.4 Analyses of the track density, winter storm variability and storm intensity measures as represented by PRIMAVERA high- and low-resolution models

The PRIMAVERA high- and low-resolution models (Table 3.2) were compared in terms of the same characteristics used in the CMIP5 comparisons. The PRIMAVERA models simulations analysed here are the present day AMIP runs at the highest resolution currently available as well as at the available lower resolution. Only one ensemble member has been analysed from each modelling centre. TRACK (Hodges 1995) using standard settings has been used to track all the ETCs. The TRACK algorithm tracks maxima in the 850hPa relative vorticity field filtered to T42 resolution. Tracks are retained for ETCs which last at least 2 days, travel $>1000\text{km}$ and have a maximum relative vorticity $>10^{-5} \text{ s}^{-1}$. ERA-Interim (Dee et al 2011) as well as MERRA 2 (Gelaro et al. 2017) re-analysis datasets have been tracked in the same way (tracks from the ERA Interim data kindly provided by Kevin Hodges and Robert Lee and from the MERRA2 data by Malcolm Roberts). All the results presented here are for winter (DJF) ETCs over the period common to all datasets (1980/81 – 2007/08).

3.1.1.4.1 Track density

Track densities from each model were calculated by counting the number of storms each month passing within a 6.3° radius of each grid point within the region of interest, as in Economou et al. (2015). The storm track density based on the ERA Interim data is represented in Figure 3.4 (upper left map). The main North Atlantic storm track is clearly evident extending from northeast North America across the Atlantic towards northern Europe. Higher storm track density is evident over the Mediterranean as well.

Modelling centre	PRIMAVERA high-resolution models	PRIMAVERA low-resolution models
CMCC	CMCC-CM2-VHR4 (18km)	CMCC-CM2-HR4 (64km)
CNRM	CNRM-CM6-1-HR (50km)	CNRM-CM6-1 (142km)
ECEARTH	ECEARTH3-HR (36km)	ECEARTH3 (71km)
MOHC	HadGEM3-GC31-HM (25km)	HadGEM3-GC31-LM (135km)
MPI	MPIESM-1-2-XR (34km)	MPIESM-1-2-HR (67km)
ECMWF	ECMWF-IFS-HR (25km)	ECMWF-IFS-LR (50km)

Table 3.2. Summary of models analysed with approximate native horizontal resolutions at mid latitudes (50°N).

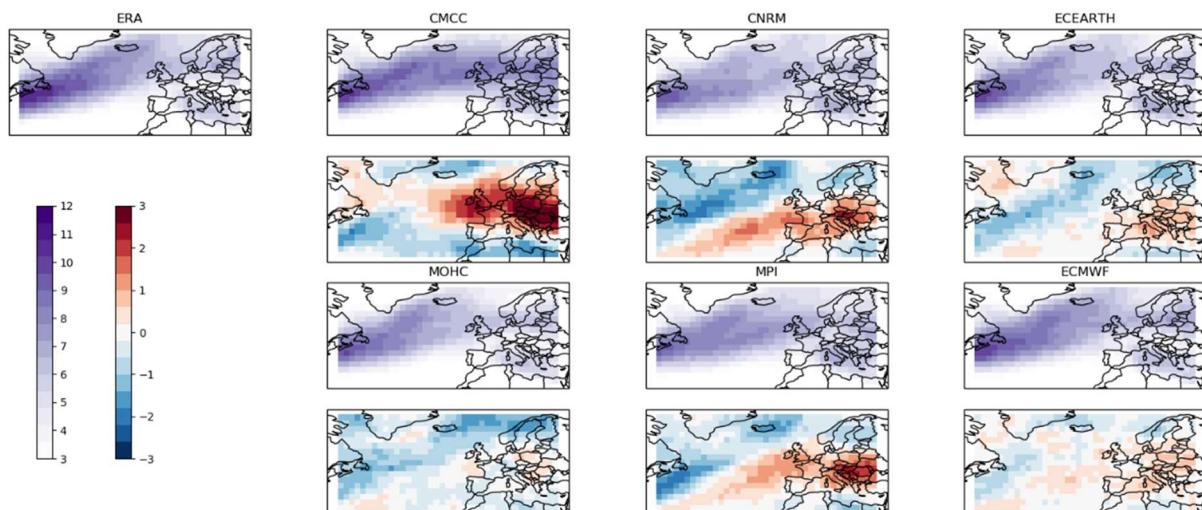


Figure 3.4. Track density as represented by ERA Interim and PRIMAVERA low resolution models (maps in purple colours) and differences (biases) between the PRIMAVERA models and ERAI (in red and blue colours).

Most of the PRIMAVERA low-resolution models underestimate the frequency of storms within the main North Atlantic storm track. The CMCC model shows a notable large positive bias in storm frequency over the UK and northern Europe. This is probably related to the bias in jet position and frequency seen in this model, described in section 3.3 (figs 3.37 and 3.38). Furthermore, some of the models simulate either a southerly displaced storm track (CNRM, MPI) or a storm track that is too zonal (CMCC) where too many storms hit central Europe, as was found for CMIP5 models (Zappa et al. 2013). In addition, the CMCC model shows a notable large positive bias in storm frequency over the UK and northern

Europe. This is probably related to the bias in jet position and frequency seen in this model, described in section 3.3 (figs 3.37 and 3.38). Some overestimation is evident within the Mediterranean region in most models (Figure 3.4).

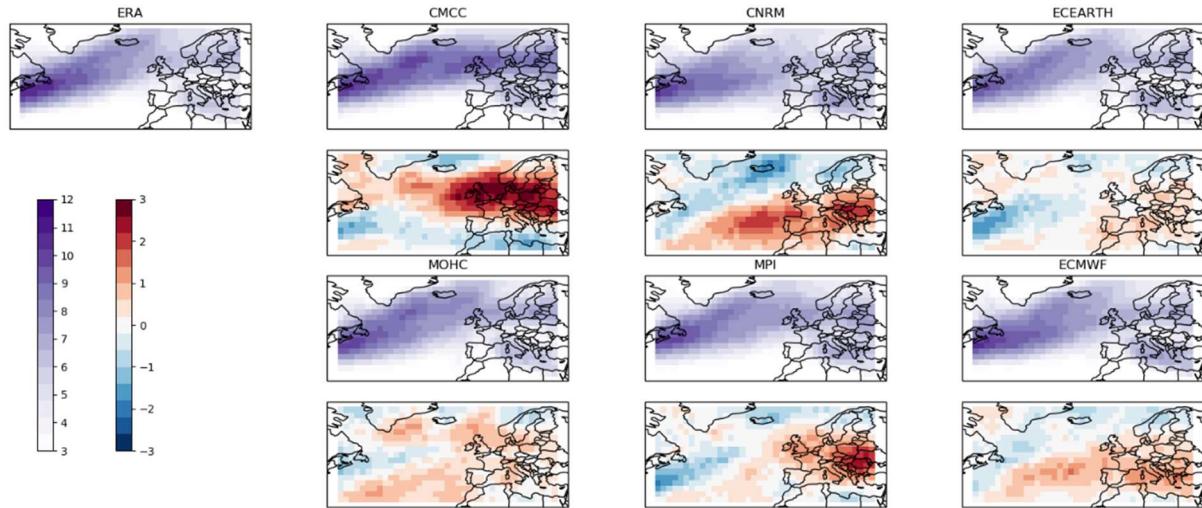


Figure 3.5. Track density as represented by ERA Interim and PRIMAVERA high resolution models (maps in purple colours) and differences (biases) between the PRIMAVERA models and ERAI (in red and blue colours).

The underestimation of the number of storms passing along the main North Atlantic storm track is decreased in all of the high-resolution versions of the PRIMAVERA models, although the slight overestimation over the Mediterranean region is still evident (Figure 3.5). There seems to be a general increase in storm frequency for all models rather than changing the tilt or position of the storm track.

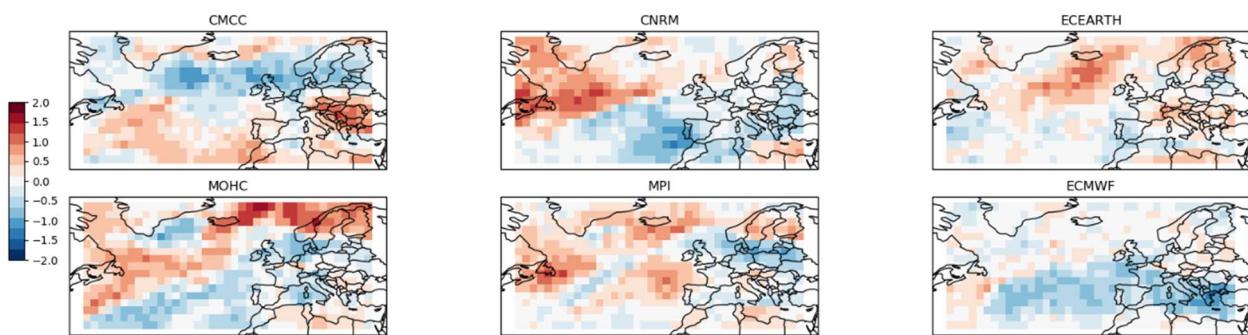


Figure 3.6. Change in track density bias between PRIMAVERA low and PRIMAVERA high resolution models ($| \text{PRIMAVERA low res bias} | - | \text{PRIMAVERA high res bias} |$), as compared with ERAI. Red areas (reduction in bias) show where there is improvement in PRIMAVERA high resolution vs. low resolution models.

Figure 3.6 demonstrates whether there is an improvement in the PRIMAVERA high-resolution models in terms of track density. It is clearly evident that the high-resolution models present an improvement over the main North Atlantic storm track for almost all models; a clear exception is the CMCC model which shows an improvement within the southern sections of the main track, and lack of such over the northern sections and Northern Europe. The results are mixed regarding the track density over the Mediterranean region.

3.1.1.4.2 ETC variability

Figures 3.7 and 3.8 show the PRIMAVERA high- and low-resolution models as compared with ERAI in terms of variability of the number of ETCs. The storms are counted within the European region (-15°E to 25°E, 35° to 70°N) by checking whether the coordinates of the storm track fall within the region boundaries.

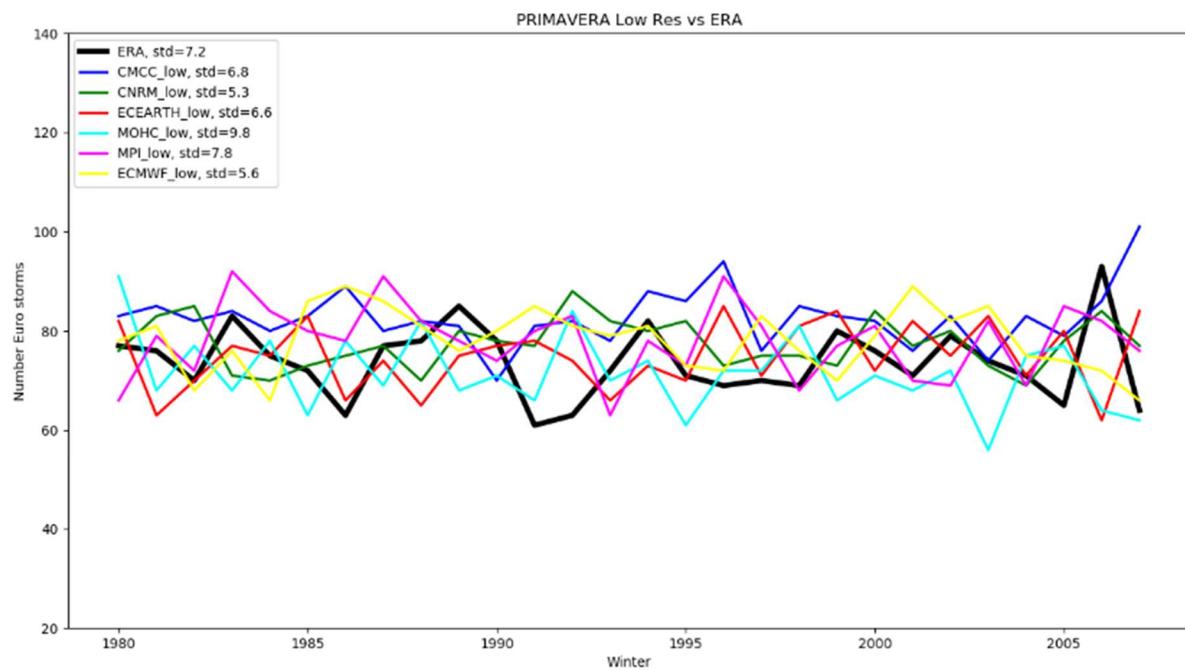


Figure 3.7. Numbers of ETCs entering Europe each winter for PRIMAVERA low-resolution models. ERAI storm numbers are given with thick black line.

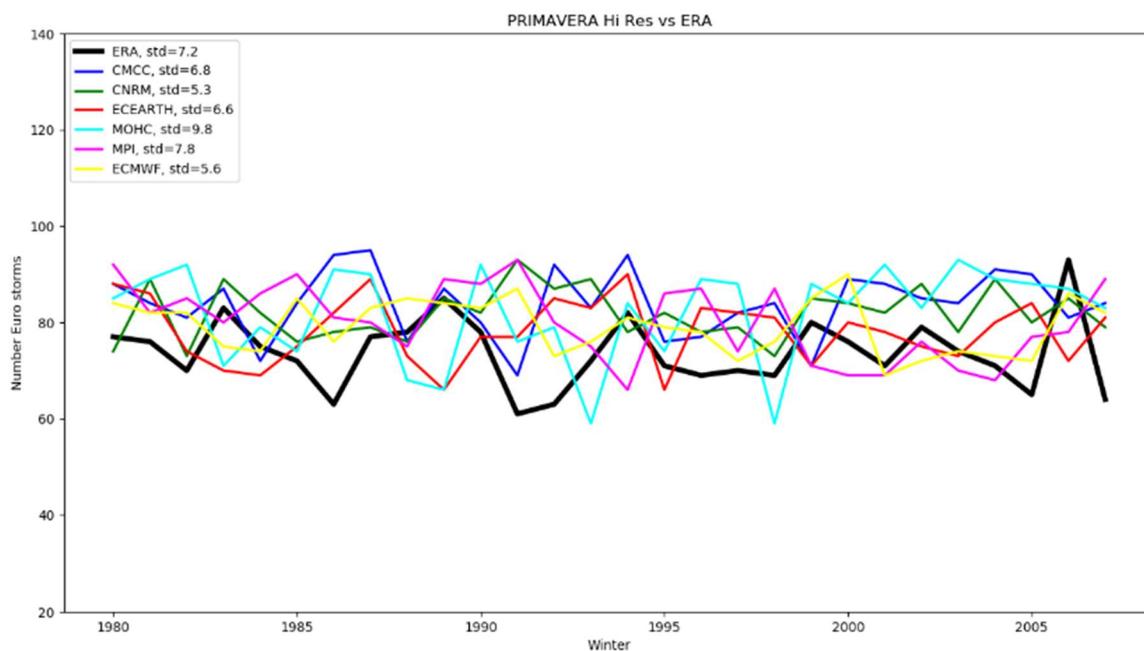


Figure 3.8. Numbers of ETCs entering Europe each winter for PRIMAVERA high-resolution models. ERAI storm numbers are given with thick black line.

Overall the standard deviations of the storm time series are similar. The numbers of winter storms from year to year are slightly higher compared with ERAI in the high-resolution models.

3.1.1.4.3 ETC intensity—measured by minimum SLP

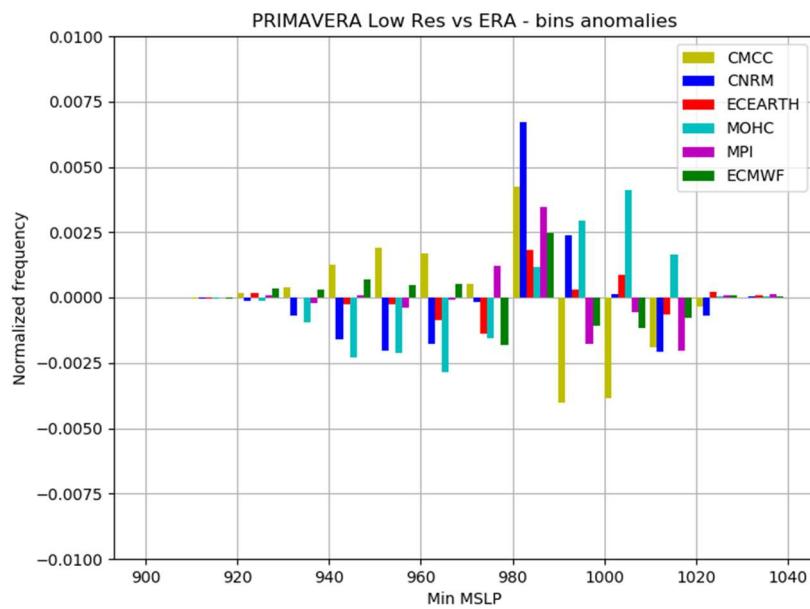
Figure 3.9 (a and b) shows frequency anomalies of the storms entering Europe (-15°E to 25°E , 35° to 70°N) with minimum MSLP within a given range, for PRIMAVERA high- and low-resolution models compared with the ERAI reanalysis. The bins within the figure span the range 900–1040 hPa and have a width of 10 hPa.

Most of the high-resolution PRIMAVERA models demonstrate lower biases compared with ERAI when looking at more extreme storms with min MSLP below 980 hPa (one exception is the CMCC model), while half of the low-resolution models show larger biases for the more extreme storms. The results are mostly similar for less intense storms with higher minimum MSLP.

3.1.1.4.4 ETC intensity—measured by maximum vorticity

Figure 3.10 (a and b) shows the frequency anomalies of winter storms with specific maximum vorticity characteristics within the European region as compared with ERAI. The bins span vorticity from 0– 21 s^{-1} and the width of the bins is 3 s^{-1} .

a)



b)

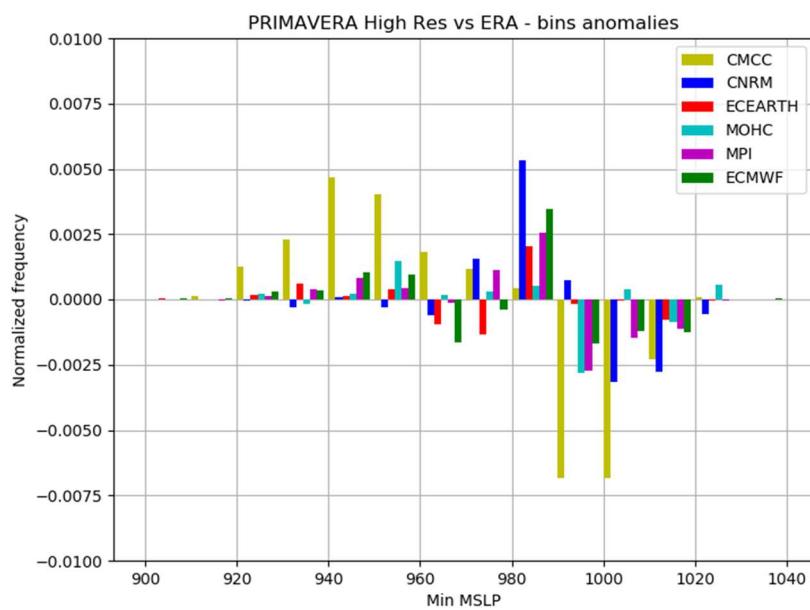
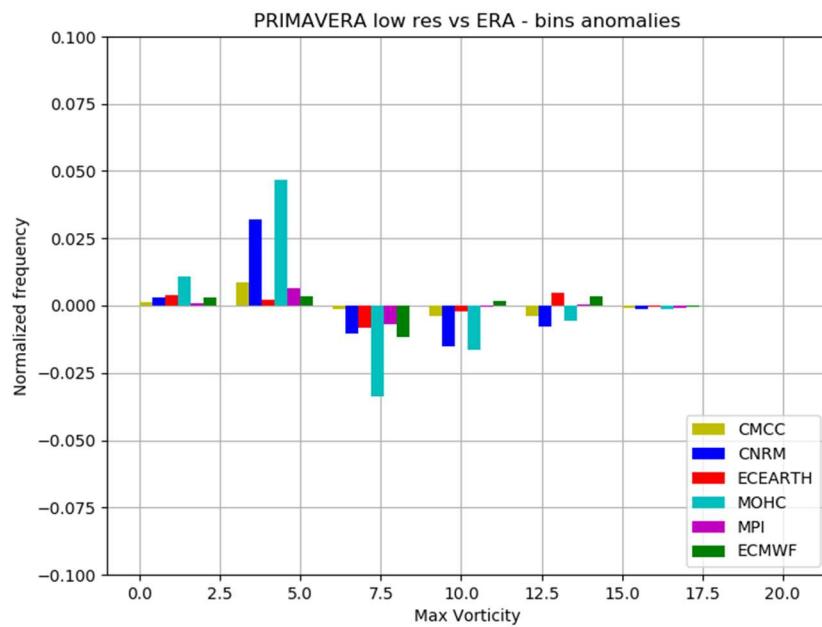


Figure 3.9. Frequency anomalies based on the minimum MSLP of the storms entering Europe for PRIMAVERA low (a) and high (b) resolution models vs ERAI.

a)



b)

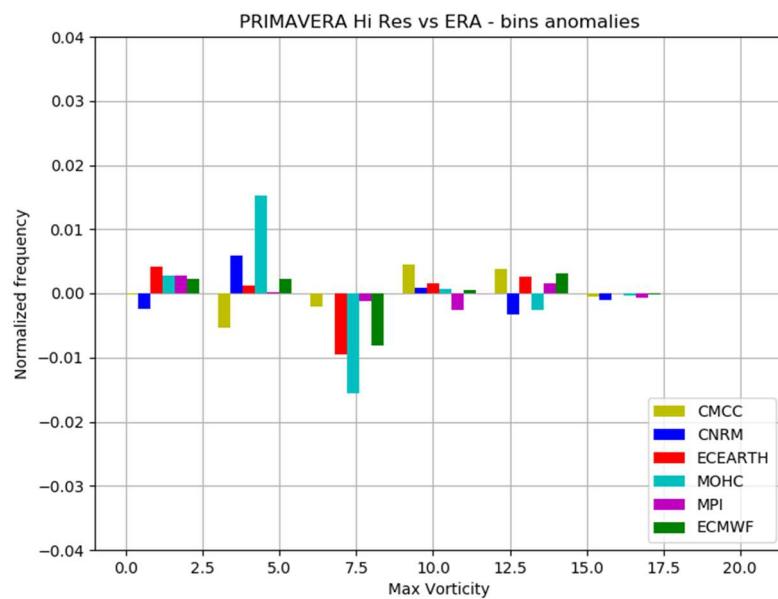


Figure 3.10. Frequency anomalies based on the maximum vorticity of the storms entering Europe for PRIMAVERA low (a) and high (b) resolution models vs ERAI.

There is not a large difference in the frequency of storms with given maximum vorticity as represented by the high- and low-resolution models. Overall the biases from ERAI are low, much lower compared with the CMIP5 models as mentioned above.

3.1.1.4.5 ETC intensity—measured by maximum 925hPa wind speed

Figure 3.11 (a and b) shows the frequency anomalies of winter storms with maximum winds at 925hPa level within a given range within the European region as compared with ERAI. The bins span the range 5-65 m/s and the width of the bins is 5 m/s.

The results for the low-resolution models are mixed but the differences are smaller compared with the high-resolution models which show a clear underestimation of the frequency of storms with lower winds and overestimation of the frequency of storms with higher winds (above 30m/s) as compared with the ERAI reanalysis. The results from the MERRA2 comparisons differ from the results shown above with the high-resolution models showing somewhat smaller biases compared with the low-resolution models. This is also due to the two reanalyses being distinct from one another in terms of frequency of storms with winds within a certain range at the 925hPa level (Figure 3.12).

3.1.1.5 Summary of D10.3 results

3.1.1.5.1 Temporal clustering – CMIP5 vs PRIMAVERA high-resolution models

Generally, an improvement is evident in the CMCC model and in the MOHC model towards the British Isles and around Iceland; The rest of the models do not show definite improvement in this characteristic compared with the CMIP5 models.

3.1.1.5.2 Track density, winter storm variability and storm intensity measures in high- vs low-resolution PRIMAVERA models

a) Track density

The high-resolution models present an improvement over the main North Atlantic storm track for almost all models (an exception is the CMCC model). The results are mixed over the Mediterranean region.

b) Storm variability

Overall the standard deviations of the storm time series are similar. The numbers of winter storms from year to year are slightly higher compared with ERAI in the high-resolution models.

c) Storm intensity (minimum MSLP)

Most of the high-resolution PRIMAVERA models demonstrate lower biases compared with ERAI when looking at more extreme storms with minimum MSLP below 980 hPa (one exception is the CMCC model).

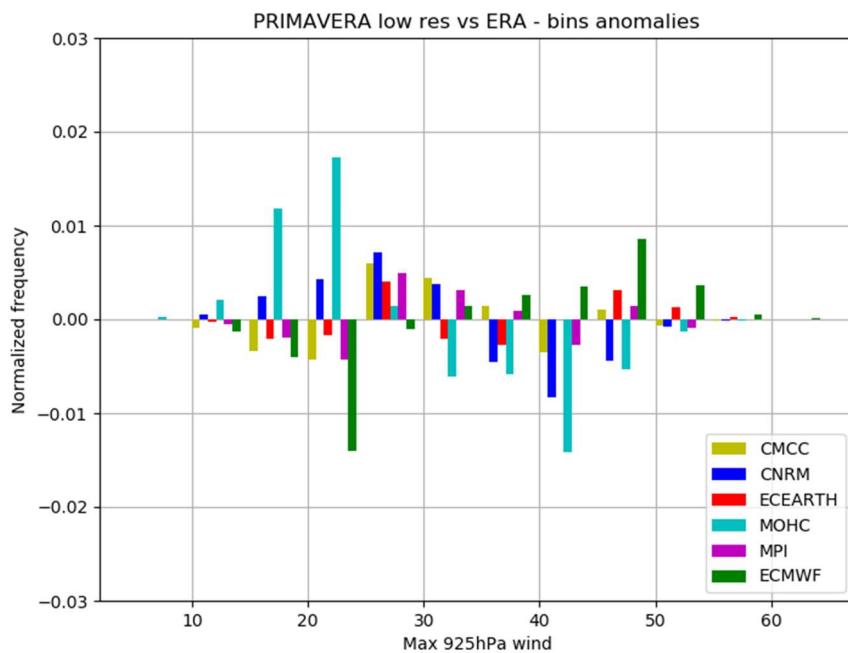
d) Storm intensity (maximum vorticity)

There is not a large difference in the frequency of storms with given maximum vorticity as represented by the high- and low-resolution models.

e) Storm intensity (maximum 925hPa wind)

The results for the low-resolution models are mixed but the differences are smaller compared with the high-resolution models which show a clear underestimation of the frequency of storms with lower winds and overestimation of the frequency of storms with higher winds (above 30m/s) as compared with the ERAI reanalysis.

a)



b)

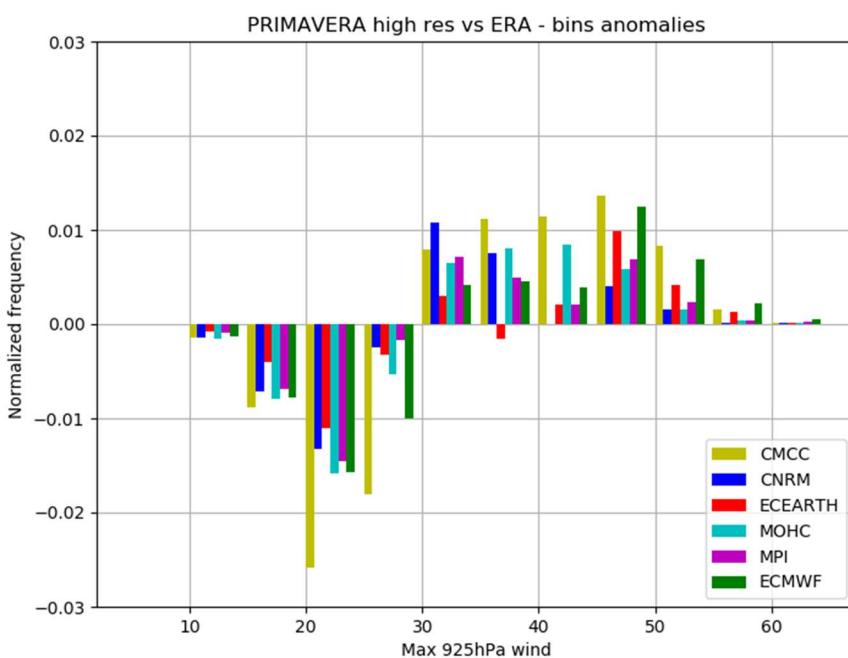


Figure 3.11. Frequency anomalies based on the maximum 925hPa winds of the storms entering Europe for PRIMAVERA low (a) and high (b) resolution models vs ERAI.

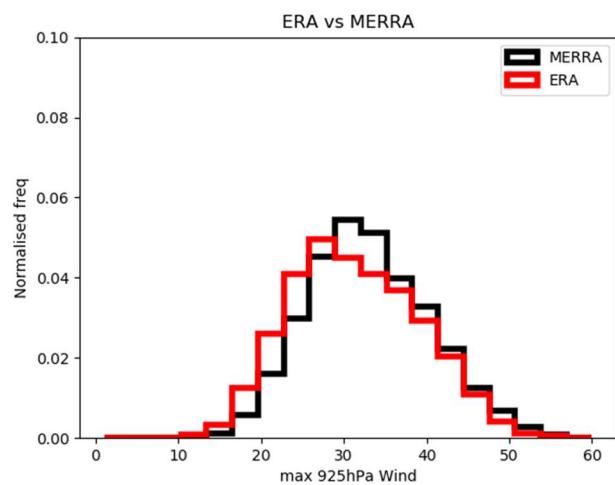


Figure 3.12. Frequency histograms for storms with maximum 925hPa winds entering Europe for ERA and MERRA2 reanalyses.

3.1.2 The ability of PRIMAVERA models to simulate extreme surface wind speeds

Julia Lockwood (UKMO)

3.1.2.1 Introduction, method and definitions

One of the use cases for the insurance/finance industry outlined in D10.1 was using PRIMAVERA model data to estimate the present-day risk of European windstorm damage. Windstorm damage is usually quantified by a loss or severity index, of which there are many definitions, but most are proportional to the cubic exceedance of maximum winds or gusts over a threshold (e.g. Lamb 1991, Klawa and Ulbrich 2003, Pinto et al 2007). The threshold can be an absolute value (e.g., 15 m/s for winds or 20-25m/s for gusts, which are the thresholds at which winds/gusts become damaging), or a percentile, with the 98th percentile being frequently used (Klawa and Ulbrich 2003). The loss index is calculated either per storm or per day, and then total losses for a season are estimated by summing these values. For PRIMAVERA, we will base our analysis on the loss index defined in Klawa and Ulbrich (2003):

$$\text{Loss index} \propto \sum_{i,d} p_i \times \left(\frac{v_{i,d}}{v_{i,98}} - 1 \right)^3$$

Where p_i is the population density at grid point i , $v_{i,d}$ is the maximum wind speed on day d at grid point i , and $v_{i,98}$ is the 98th percentile value of daily maximum wind speed at grid point i . The loss index is usually summed over the winter half-year (October- March), when the majority of European windstorm damage occurs (Pinto et al 2007). This loss index was chosen because it has been found to correlate well with insured losses over Germany

($r=0.96$ Klapwijk and Ulbrich 2003), and since it is based on a percentile threshold it should be less dependent on model biases.

This loss index is highly dependent on the distribution of daily maximum wind speeds above the 98th percentile threshold. Therefore, in order for a model to give reliable estimations of windstorm loss, the distribution of extreme wind speeds needs to be realistically simulated.

In this section we present maps of the quantity $x = (\frac{v_{i,99.5}}{v_{i,98}} - 1)^3$, where $v_{i,99.5}$ is the 99.5th percentile of the daily maximum wind speed, for ERA5 (used to represent the ‘true’ distribution; Copernicus Climate Change Service 2017), CMIP5 and PRIMAVERA models, to compare the ‘weight’ of the tail of extreme daily maximum wind speeds, and examine its spatial distribution. The 99.5th percentile of daily maximum wind speeds was chosen rather than the maximum (of maximum wind speeds) because it reveals the systematic differences between the daily maximum wind speed distributions rather than being dependent upon a single storm.

The values of $v_{i,98}$ and $v_{i,99.5}$ are calculated empirically from one ensemble member from each model for present day, atmosphere-only simulations. The whole data period is used for each model (1979/80-2017/18 for ERA5; 1979/80-2007/8 for CMIP5; 1950/51-2013/14 for PRIMAVERA). To test the difference model resolution has on large scales and for a fairer comparison between models and, all daily maximum wind speed data is regressed to N96 resolution before percentiles are calculated.

3.1.2.2 Results

Figures 3.13-3.17 show the maps of $x = (\frac{v_{i,99.5}}{v_{i,98}} - 1)^3$ for the observations, CMIP5 and PRIMAVERA models, from the modelling centres MOHC, MPI-M, EC-Earth Consortium, CNRM-CERFACS and CMCC. The models from ECMWF are not shown because they do not output daily maximum winds.

From ERA5, we see that x is rather uniform over the Atlantic, and has higher values over land, increasing at higher elevations. The model results are discussed in more detail below.

3.1.2.2.1 MOHC

Areas of high elevation tend to show a large positive bias, meaning that the loss estimation from these models would be too high in these regions. Interestingly, the bias increases from the CMIP5 generation model to the low-resolution PRIMAVERA model. The bias over high elevation improves in the PRIMAVERA models as resolution increases, but a positive bias is introduced over the Atlantic. An example of the difference in daily maximum wind speed distribution between the CMIP5 model and PRIMAVERA N96 model for a single point in Norway can be seen clearly in Figure 3.18. Here the PRIMAVERA model is biased towards low wind speeds, but is more skewed overall, with a heavier tail. A similar result is seen for points in Spain. A possible reason for this is differences in the parametrisation schemes between the models, but this requires further investigation.

3.1.2.2.2 MPI-M

The two CMIP5 models both show a negative bias in x over land, which would result in lower estimated losses. This bias seems to be eliminated in the PRIMAVERA models, although the lower resolution PRIMAVERA model shows a positive bias over the Iberian Peninsula.

3.1.2.2.3 EC-Earth

The PRIMAVERA EC-Earth models show a positive bias in x over Eastern Europe and Russia, which would lead to disproportionately high losses in those regions. This bias is reduced in the high-resolution model. No CMIP5 daily maximum wind speed data was available from this modelling centre.

3.1.2.2.4 CNRM-CERFACS

The CMIP5 model shows a positive bias over the Atlantic, the UK and Norway. These biases are reduced in the PRIMAVERA low-resolution model, but there is a slight negative bias over land. The negative land bias is reduced the high-resolution PRIMAVERA model, but a positive bias is re-introduced over the Atlantic.

3.1.2.2.5 CMCC

There is a negative bias in x over central Europe, and a positive bias over Norway and the Iberian Peninsula. The PRIMAVERA models show a moderate negative bias over land, and positive bias in the far North Atlantic, and the Atlantic off the coast off west Africa.

3.1.2.3 Discussion

The large bias in x over high altitude seen in the PRIMAVERA MOHC models, particularly at N96 and N216 resolution, could be a result of different parameters (such as roughness length) used in the models to estimate 10m wind speeds, but this requires further investigation.

The cause of notable increases in x seen with resolution in the MOHC, MPI-M and CNRM-CERFACS models is more difficult to ascertain. Changes in the storm frequency have been linked to the distribution of extreme wind speeds (Walz et al 2017), but the effect on the weight of the tail is difficult to predict and would depend on the severity of the storms, and possibly whether or not storms are already rare or frequent events in each location.

Figures 3.19 and 3.20 compare the storm track density bias and bias in x for the MOHC PRIMAVERA N216 and N512 models, and CNRM-CERFACS low- and high-resolution models. The track densities were calculated using the TRACK algorithm (Hodges 1995), as described in section 3.1.1.3. The track density biases were unavailable for the Oct-Mar period so are calculated for Dec-Feb (DJF) only, so for consistency in figs 3.19 and 3.20 the x bias is also calculated for DJF. The DJF x biases are very similar to those for the full Oct-Mar winter shown in figs 3.13 and 3.15.

In the case of the MOHC models (Figure 3.19), storm frequency increases from N216 to N512 over large areas of the domain, although it is difficult to know whether this is associated with the increase in x . The same is true for the CNRM-CERFACS models (Figure

3.20), although a similar negative bias in storm frequency at high latitudes is seen at both resolutions. This therefore cannot account for the increase in x seen around these latitudes for this model. More investigation is needed to see whether the changes in x seen with resolution are caused by changes in storm frequency, storm intensity or another cause.

The correct representation of extreme wind speeds is important for estimating insured losses due to storm damage, and major biases will need to be corrected for a model to give realistic estimations of storm loss. The MOHC N96 PRIMAVERA model, for example, estimates storm losses over the Iberian Peninsula to be far higher than those over the UK, which is not seen in reality. This may preclude the use of this model in estimating present day windstorm risk.

Overall, most models (MOHC, MPI-M, EC-Earth) show an improvement in representation of the quantity $(\frac{v_{i,99.5}}{v_{i,98}} - 1)^3$ with increasing resolution, although some models show an increase in bias (CMCC, CNRM-CERFACS).

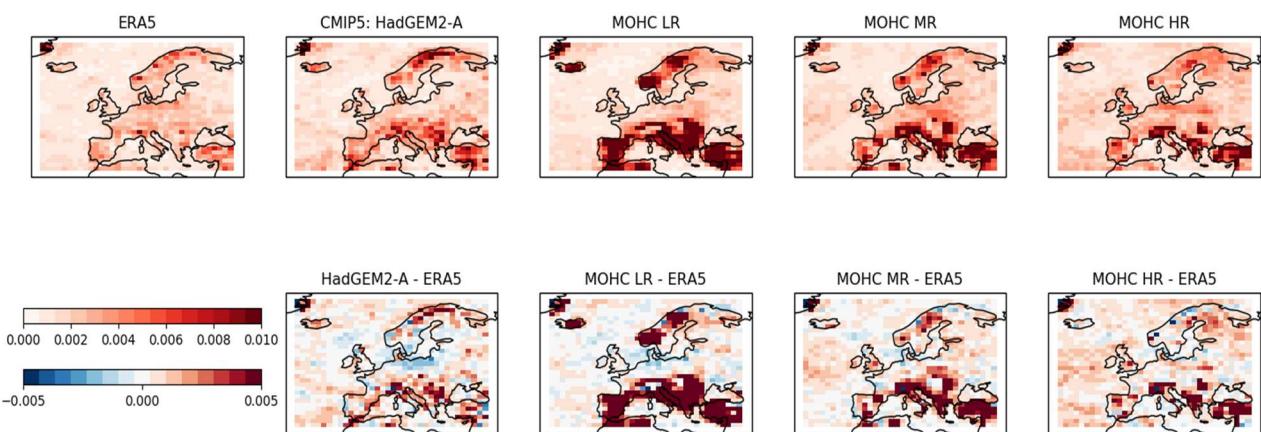


Figure 3.13. Top row: Maps of the quantity $(\frac{v_{i,99.5}}{v_{i,98}} - 1)^3$ for ERA5 and the MOHC models for CMIP5 and PRIMAVERA. The MOHC models are (left to right): HadGEM2-A (CMIP5, N96 resolution), and HadGEM3-GC3.1 (PRIMAVERA) with native resolutions N96, N216 and N512. Bottom row: Bias in $(\frac{v_{i,99.5}}{v_{i,98}} - 1)^3$ (model – ERA5). All data has been regressed to N96 before analysis.

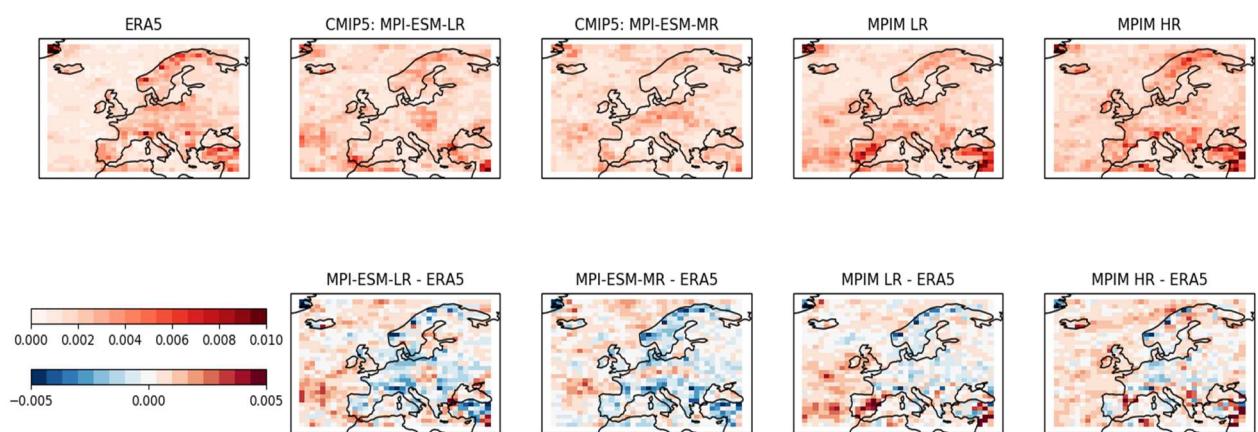


Figure 3.14. As for figure 3.13 but for MPI-M models. The models are (left to right): MPI-ESM-LR (CMIP5, native resolution T63), MPI-ESM-MR (CMIP5, native resolution T63), and PRIMAVERA MPI-ESM1-2-HR (native resolution T127), MPI-ESM-1-2-XR (native resolution T255). The CMIP5 models MPI-ESM-LR and MPI-ESM-MR differ in their number of vertical levels in the atmosphere (47 and 95 respectively).

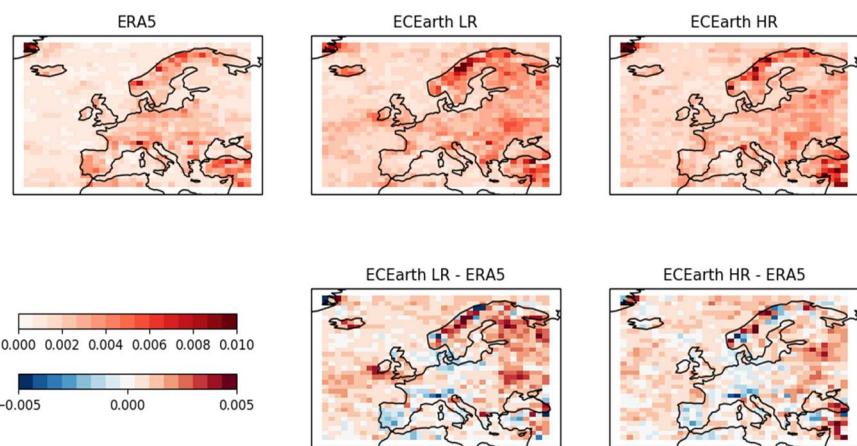


Figure 3.15. As for figure 3.13 but for EC-Earth Consortium models. The models are PRIMAVERA EC-Earth3 at native resolutions T1255 and T1511. No CMIP5 EC-Earth daily sfcWindmax data was available at the time of analysis.

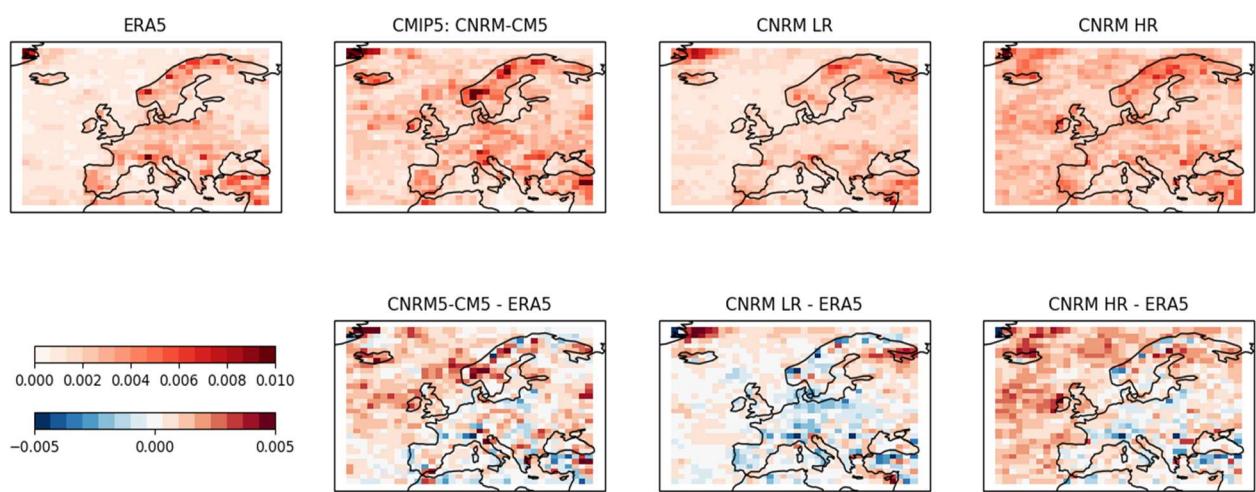


Figure 3.16. As for figure 3.13 but for CNRM-CERFACS models. The models are CNRM-CM5 (CMIP5, native resolution TI127), and PRIMAVERA CNRM-CM6 at native resolutions TI127 and TI359.

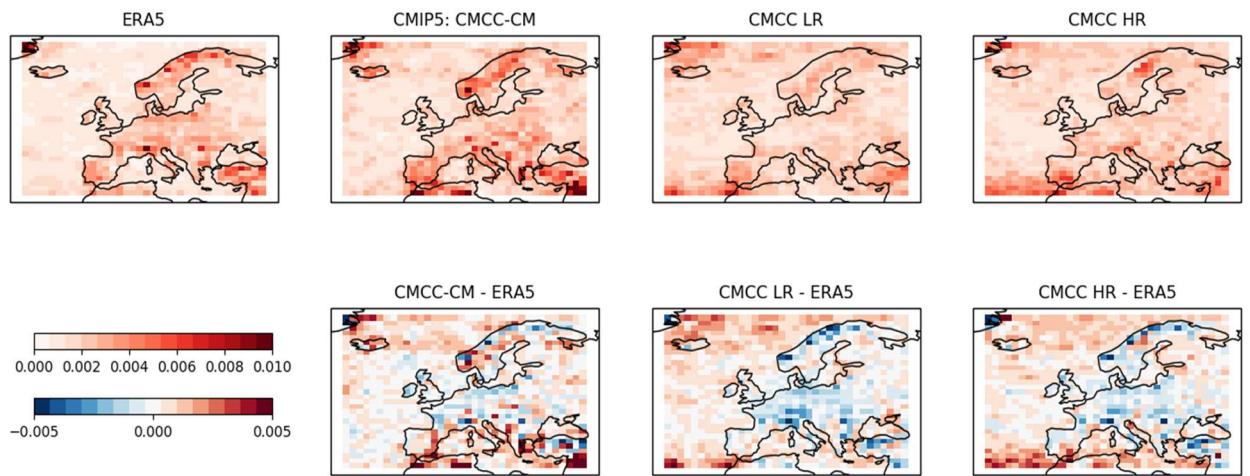


Figure 3.17. As for figure 3.13 but for CMCC models. The models are CMCC-CM (CMIP5, native resolution $2^\circ \times 2^\circ$), and PRIMAVERA CMCC-CM2 at native resolutions $1^\circ \times 1^\circ$ and $0.25^\circ \times 0.25^\circ$.

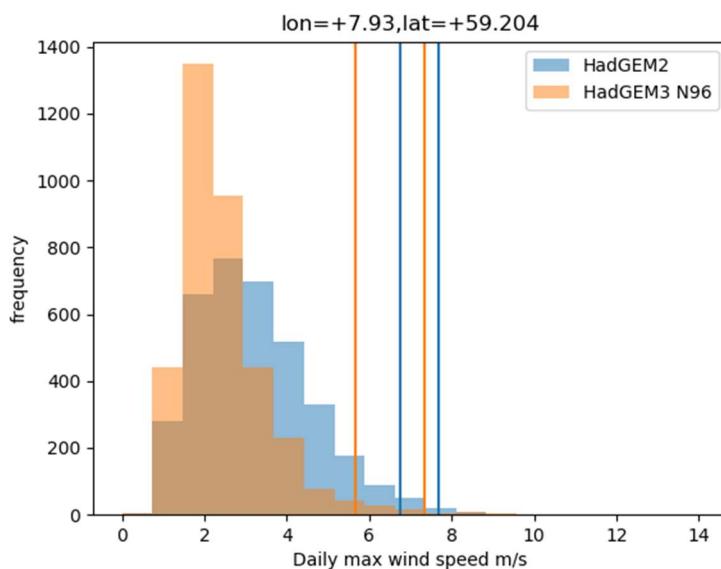


Figure 3.18. Histograms of daily maximum wind speed at a point in Norway ($+7.93^{\circ}\text{E}$, $+59.204^{\circ}\text{N}$) from the MOHC CMIP5 model (HadGEM2) and PRIMAVERA low-resolution model (HadGEM3 at N96). Histograms are calculated from 10 years of daily data (1979-1988) for each model. The 98th and 99.5th percentiles for each model are marked with vertical lines (6.77 and 7.68 m/s for HadGEM2 and 5.68 and 7.36 m/s for HadGEM3).

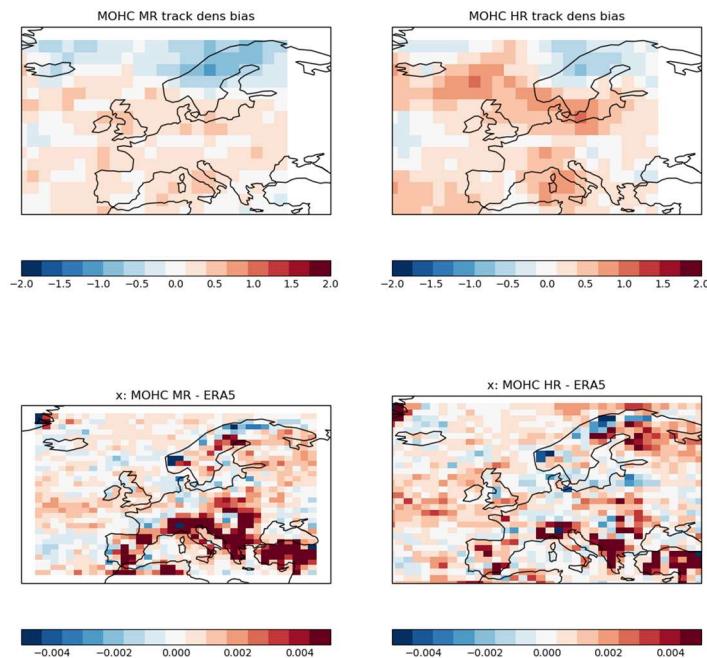


Figure 3.19. Top row: track density bias (for DJF) for the PRIMAVERA MOHC N216 (left) and N512 (right) models. Bottom row: bias in x (for DJF) for the same models.

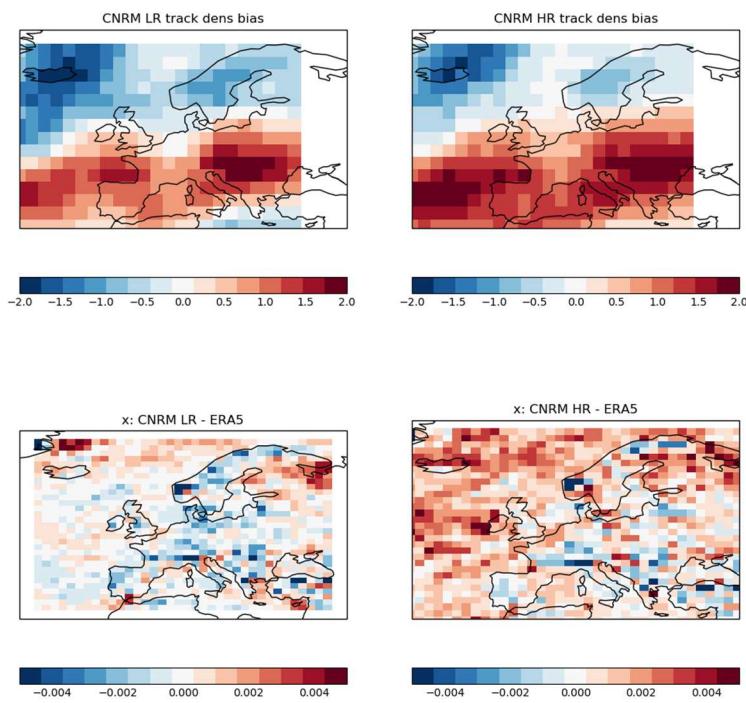


Figure 3.20. As for Figure 3.19 but for the PRIMAVERA CNRM-CERFACS low- and high-resolution models.

3.2 Atmospheric blocking

3.2.1 Blocking evaluation overview

Reinhard Schiemann (UREAD) and Panos Athanasiadis (CMCC)

We compare the representation of Northern Hemisphere atmospheric blocking in the CMIP5, CMIP6 DECK, and PRIMAVERA Stream 1 ensembles. To this end, multi-model-mean biases in the Absolute Geopotential Height (AGP) blocking index (Scherrer et al. 2006) with respect to a 50-year reanalysis climatology (Schiemann et al. 2017) are shown in Figures 3.21 and 3.22. The six panels comprise four PRIMAVERA sub-ensembles (low-resolution forced, low-resolution coupled, high-resolution forced, and high-resolution coupled simulations) as well as one panel each for CMIP5 and CMIP6 simulations.

Long-standing model biases in blocking can be identified in all these ensembles such as an underestimation of blocking occurrence especially over the North Atlantic/Northern Europe, and an underestimation across high-latitude Eurasia and over Alaska in summer. The magnitude of these biases varies between the different ensembles. There is an improvement in CMIP6 over CMIP5 seen most clearly over the North Sea in DJF. There is also an improvement with resolution in the PRIMAVERA ensembles, both in the Atlantic and Pacific and both in winter and summer. This improvement is more clearly seen in the coupled simulations than in the SST-forced simulations. We conclude that horizontal resolution, as explored in PRIMAVERA, is one of the factors important for simulating atmospheric blocking, yet an increase in resolution, to about 25 km, alone does not fully remedy blocking biases in

climate models. We also caution that our results should be considered to be conservative given that PRIMAVERA models have not been re-tuned at the higher resolutions and the known sensitivity of simulated blocking to the mean state (Scaife et al. 2010, Woollings et al. 2010).

Further discussion and analyses are available as a discussion paper that is currently under review (Schiemann et al. 2020)

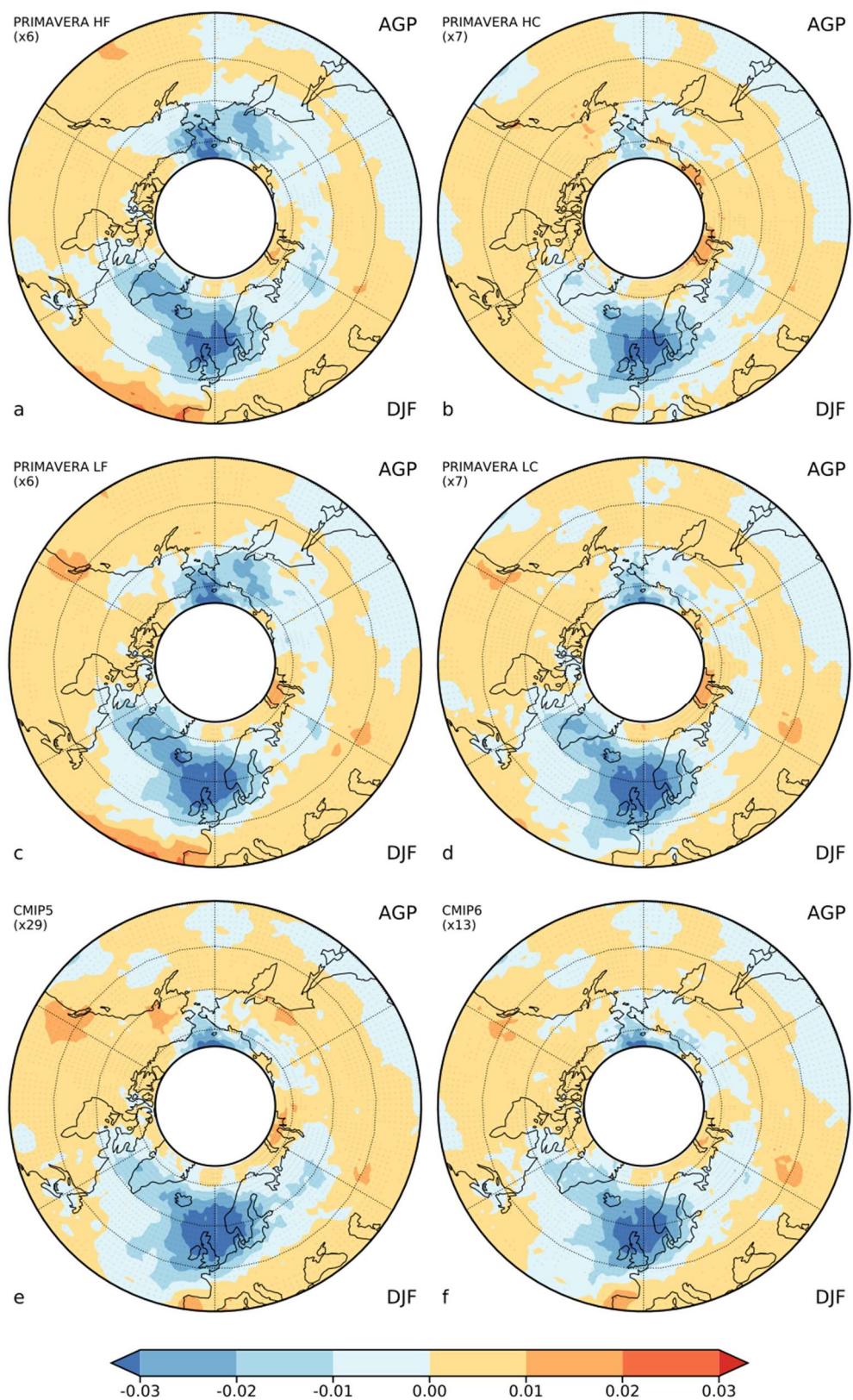


Figure 3.21. Winter (DJF) multi-model mean biases (versus ERA-40 and -Interim) in blocking frequency for PRIMAVERA (a) high-resolution highresSST-present, (b) hist-1950, (c) low-resolution highresSST-present, (d) low-resolution hist-1950, (e) CMIP5, and (f) CMIP6. Number of models is given in parentheses.

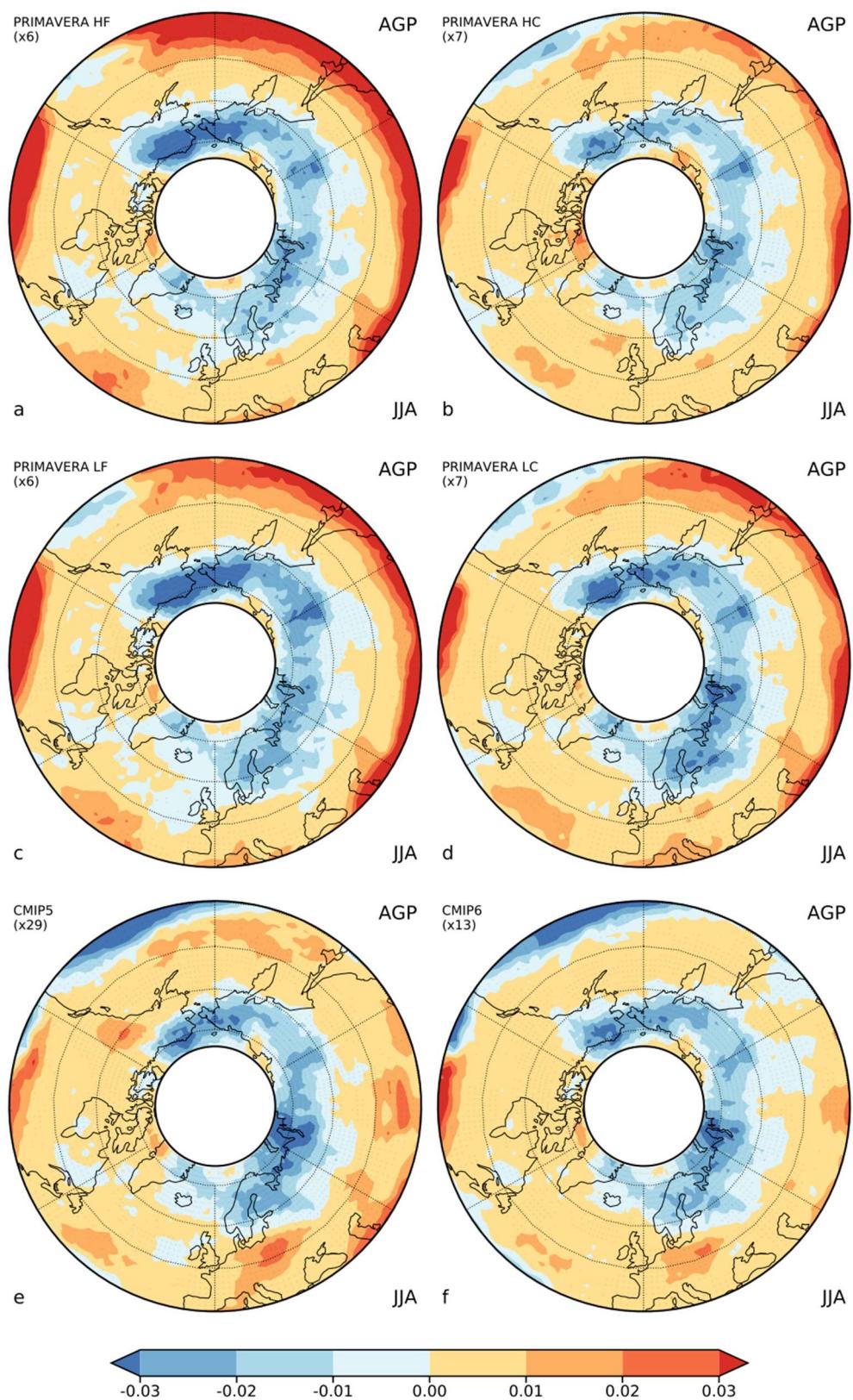


Figure 3.22. Summer (JJA) multi-model mean biases (versus ERA-40 and -Interim) in blocking frequency for PRIMAVERA (a) high-resolution highresSST-present, (b) hist-1950, (c) low-resolution highresSST-present, (d) low-resolution hist-1950, (e) CMIP5, and (f) CMIP6. Number of models is given in parentheses.

3.2.2 Impact of blocking on low wind events and its representation by high-resolution GCMs: An energy perspective

Paula Gonzalez, David Brayshaw and Reinhard Schiemann (UREAD)

3.2.2.1 Introduction

With higher penetration of renewable energies and the effort to decarbonise power production, there is a strong interest in the objective characterization of wind resource. Over Europe, wind power accounts for around 17% of total power capacity and almost 30% of renewable capacity and is the overall second largest form of generation capacity after gas.

In addition to the description of mean capacity factors, there is a need to characterise extremes. Low wind events and persistent low wind events are of particular interest because, during these events, the energy system needs to rely on ‘backup’ sources, such as gas, coal and nuclear. Over the United Kingdom and other parts of Europe, these events are often linked to the occurrence of blocking (e.g., Cannon et al. 2015, Grams et al. 2017), which is the initial focus of this study. Additionally, blocking events have an impact on near-surface temperatures over Europe, which implies an effect in weather-dependent energy demand.

This study focuses on the impacts of blocking conditions on low wind events and their persistence, and the representation of these effects on the PRIMAVERA models. As will be shown, blocking is strongly associated with low wind events. The PRIMAVERA model simulations are able to capture this basic relationship, though biases in frequency and duration of low wind events remain.

3.2.2.2. Datasets

The blocking index used here was calculated as in Schiemann et al. 2016 and is a 2-D extension of the Scherrer et al. (2006) absolute geopotential height (AGP) index, requiring:

- a reversal of the meridional gradient of the 500-hPa geopotential south of the point
- westerly flow to the north of the point
- a persistence of the previous conditions for at least 5 days

To calculate this, geopotential heights were re-gridded to a common $1.8758^\circ \times 1.258^\circ$ grid prior to the calculation of blocked days.

In addition to the blocking index, daily mean 10m wind speeds and 2m temperatures were considered from the ERA Interim reanalysis dataset (ERA1, Dee et al. 2011) and PRIMAVERA models (Table 3.3).

Institution	MOHC	MPI-M	CMCC	EC-Earth	ECMWF
Model Name	HadGEM3-GC3.1	MPI-ESM-1-2	CMCC-CM2	EC-Earth3	ECMWF-IFS
Model Versions	MM, HM	HR, XR	HR4, VHR4	, HR	LR, HR
Atmos grid	N216, N512	T127, T255	1x1, 0.25x0.25	TI255, TI511	Tco199, Tco399
Atmos res @ 50N	60km, 25km	67km, 34km	128km, 64km	71km, 36km	50km, 25km

Table 3.3. Main features of the PRIMAVERA models considered in the study.

Simulations corresponding to the experiment highresSST-present for 1950-2014 were considered, and the output was interpolated to the coarser grid of the blocking index.

3.2.2.3 Results: blocking and low wind events in ERA-Interim

As a first step, this study evaluates the impact of blocking events on low wind conditions over Europe using the ERAI reanalysis. For the purpose of this analysis, magnitudes were evaluated at the country level. A day was considered as ‘blocked’ for the whole country when at least 50% of the grid points within the country polygon were blocked. This simple criterion ensures that the resulting country-aggregate average frequency was very close to the country-mean frequency derived from the grid-point blocking index (Figure 3.23, black versus red dots). Analogously, country-average daily mean 10m wind speeds were considered. Low wind (LW) days were defined as those below the country-average monthly-mean 20th percentile.

Prior to the study of their connection, this study presents some basic statistics of blocking and low wind events (hereafter referred to as BE and LWE, respectively). Events are defined as any number of consecutive days on which the conditions are met — i.e., no lower threshold for the duration of an event is imposed.

3.2.2.3.1 Blocking events

Figure 3.23 presents the annual mean percent of days that qualified as ‘blocked’ for each country according to the criteria described above. It shows that the overall frequencies are very small in every country (less than 10%) and that they decrease from North to South. This also reveals that blocked days are by themselves rarer than LW days, which by definition account for 20% of the record.

Figure 3.24a presents the annual mean frequency of blocked events (BE) per year. This frequency decreases from north to south and tends to be smaller than one event per year for most countries. In northern Europe, however, it can be more than 2. Figure 3.24b shows the annual mean average duration of BE in ERAI. Durations take values between 4.5 and 7 days. A lower threshold close to 5 is to be expected given the definition of the blocking index. The shorter BE are observed in the Mediterranean countries, whereas the longest events occur in central and eastern Europe.

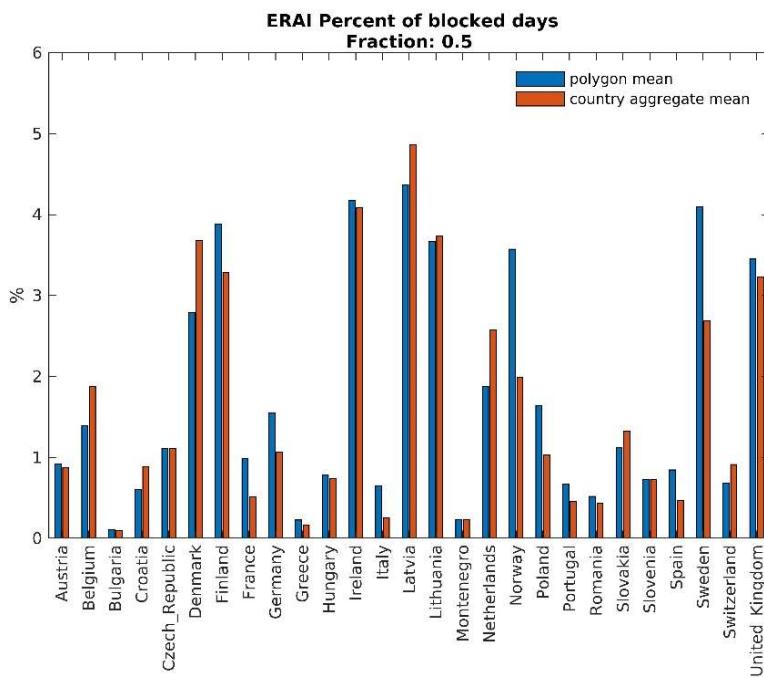


Figure 3.23. ERAI percent of blocked days in each country considering the full year. Black dots correspond to the frequency obtained averaging frequencies for grid points inside the country polygon and red dots correspond to the percent obtained from the country-average blocking index defined with a threshold of 50% of the grid points.

3.2.2.3.2 Low wind events

Figure 3.25a shows the average annual frequency of LWE. Values range from 34 to 46 events per year, and they tend to be more frequent in central and eastern Europe. Their average duration is presented in Figure 3.25b and it shows that LWE are not a long-lasting feature, with average durations ranging between 1.5 and 2.2 days. The longest LWE are observed in western Europe, showing to some extent an opposing picture to the one seen for frequency (i.e., countries with frequent LWE, observe short ones and vice versa).

3.2.2.3.3 Impact of blocking on wind

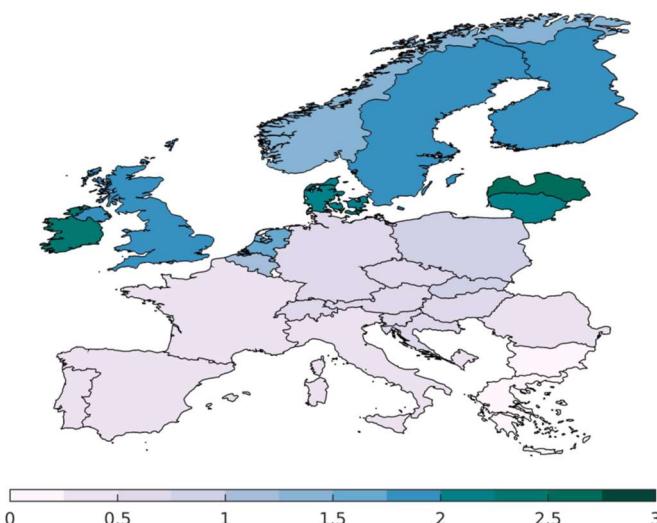
As a starting point for the evaluation of the impact of blocking conditions on near-surface wind over Europe, a comparison was performed between the country-average daily 10m wind speed PDFs and the ones corresponding to blocked days only. Here, focus is placed on results for four countries: France (FR), Germany (DE), Spain (ES) and the United Kingdom (UK), chosen for being large, populous countries and with a significant amount of wind power generation. Figure 3.26 reveals a clear shift of the country-average speeds towards calmer conditions during blocked days. The shift is such that the interquartile range of the blocked days' PDF does not include the mean value for the full PDF. Nonetheless, it must be noted that the sample sizes are very different between sets and a subsampling test of the full dataset should be performed to assess statistical significance.

Another way to evaluate the impact of blocking is to assess the probability of observing LW days during blocked days. This percent of occurrence is presented in Figure 3.27 and can be compared with the climatological probability of observing LW days, which is 20%. All

countries have an increased probability of observing LE days during blocking, with the highest increases observed in central Europe. France and Germany, for example, are 3 times more likely to experience LW days during blocking events than the climatology.

a)

ERA1: average frequency of BE [events/year]
Full year - 1979-2011



b)

ERA1: average duration of BE
Full year - 1979-2011

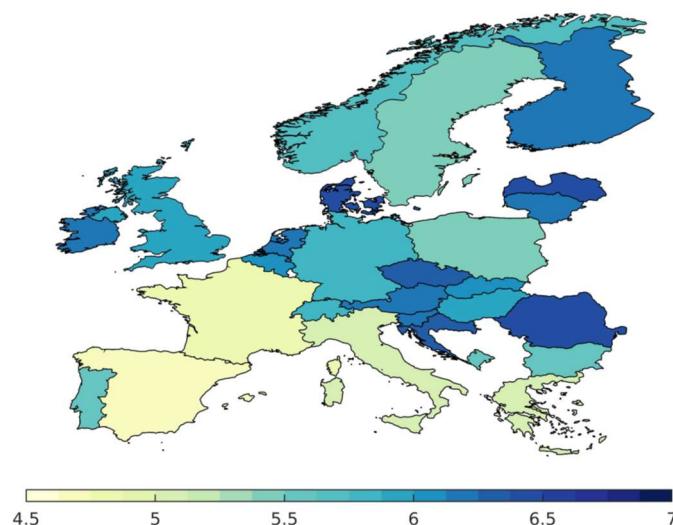
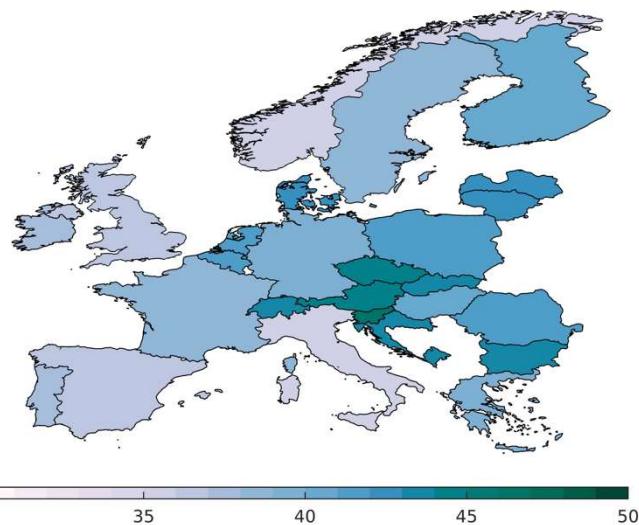


Figure 3.24. ERAI blocking events statistics. a) Annual mean frequency of blocked events [events/year]. b) Annual mean duration of the events [days].

a)

ERA1: average frequency of LWE [events/year]
Full year - 1979-2011



b)

ERA1: average duration of LWE
Full year - 1979-2011

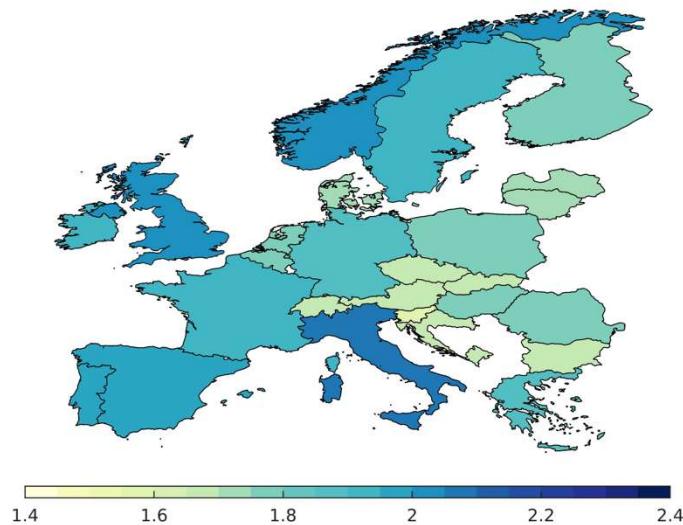


Figure 3.25. ERAI low wind events statistics. a) Annual mean frequency of low wind events [events/year]. b) Annual mean duration of the events [days].

A significant impact of blocking on near-surface winds was also observed on the persistence of low-wind events. To explore these changes a comparison was performed between the probability of obtaining subsequent LW days considering all events and considering only those for which the start day (day 0) was also blocked. It is important to remark that the

requirement for the LW day to also be blocked is only imposed on day 0. Figure 3.28 presents such comparison for two of the four countries, but results are analogous on the remainder. Having a blocked day 0 results in an increased probability of observing subsequent LW days for several days (up to seven days, in some cases like Spain and the United Kingdom).

To explore the spatial distribution of the effect, Figure 3.29 presents maps showing for each country, the change in the probability of obtaining a LW day 3 days later. This is the difference between the red and green bars in Figure 3.28. For day +3, most countries show increases in the probabilities, with higher values in central and southern Europe. For days +5 and +7 (not shown), the pattern becomes less homogeneous and only some countries like the UK, Spain, Sweden and Italy continue to show increases.

The results above show that in the ERAI reanalysis dataset blocking conditions have a significant effect on near-surface speeds over Europe. In addition to fostering lower wind speeds, BE are associated with more frequent and more persistent LWE.

3.2.2.4 Blocking and low wind events in the PRIMAVERA GCMs

The study now presents results that assess the representation of the processes described above by the PRIMAVERA GCMs. A focus is also placed in analysing whether the increase in the horizontal resolution of the models resulted in a more accurate representation.

3.2.2.4.1 Representation of blocking events

A complete analysis of the representation of blocking events in the PRIMAVERA models is not within the scope of this study and can rather be found in Schiemann and Athanasiadis (2018) and Schiemann et al. (2020 and under review).

We first compared the annual mean percent of days that qualified as ‘blocked’ for each country in ERAI and in the PRIMAVERA models considered here. In general, the PRIMAVERA GCMs capture the north-south decrease quite well (not shown), but depending on the region, they have some significant biases. Figure 3.30 shows that GCMs consistently tend to underestimate the number of blocked days on northern Europe and overestimate it on central and southern Europe. No clear improvements are observed with increased resolution, except for the EC-Earth3 model.

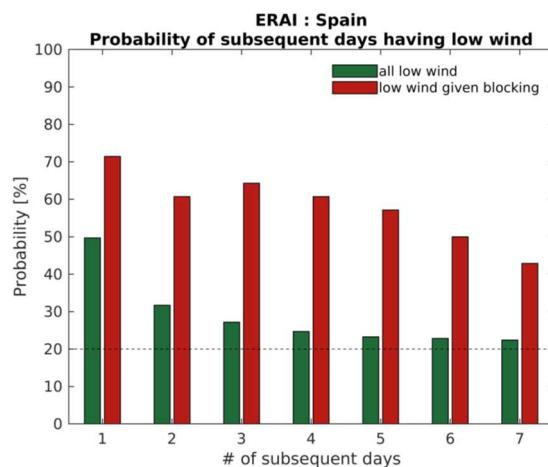
Figure 3.31 compares the error in the annual mean average duration of BE. The main bias in the GCMs is an overestimation of the duration in Iberia and France and an underestimation in the north-eastern sector. In some models like CMCC-CM2, the errors decrease with the increase in resolution, but this is not a robust feature across models and countries.

3.2.2.4.2 Representation of low wind events

Country-average LW days were defined as those below the country-average monthly-mean 20th percentile. Using each dataset’s monthly mean threshold in the definition allows to ‘correct’ any potential biases in mean speeds and their seasonality in the GCMs. Figure 3.32

compares the relative error in the average frequency of the events with respect to ERAI, which reveals that the models have biases of up to +/- 20%, with some models generally underestimating the frequency, like CMCC-CM2 and some others like ECMWF-IFS generally overestimating the duration, and some in between, such as HadGEM3-GC3.1. The opposite biases are seen from the comparison of LWE durations in Figure 3.33, which implies that models that simulate too frequent LWE simulate short durations, and vice versa. There are no improvements with increased resolution.

a)



b)

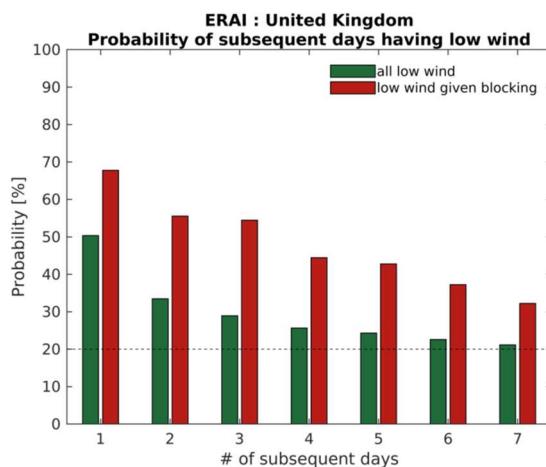


Figure 3.28. Comparison of the probability of observing subsequent low-wind days given all low wind days (green) and the probabilities obtained when day 0 was also a blocked day (red) for a) Spain and b) the United Kingdom. The dashed horizontal line corresponds to 20%, which is the probability of any given day having low wind, by definition.

**Increase in prob of LW in subsequent days - Day + 3
ERA1 - 1979-2011**

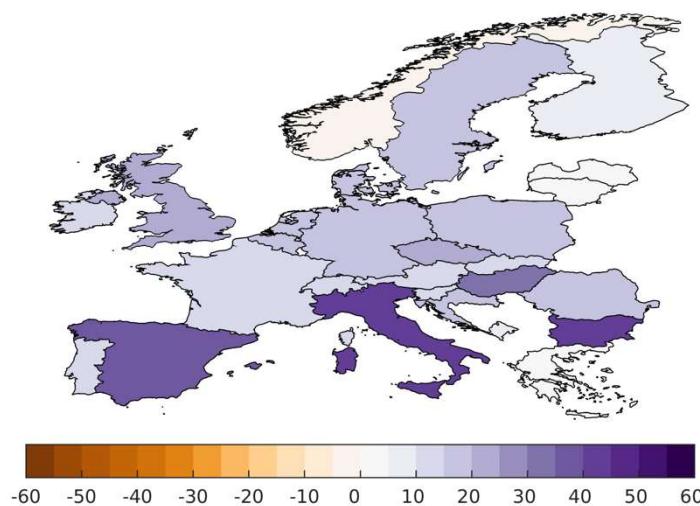


Figure 3.29. ERAI change in the probability of observing subsequent low wind at day +3 given that day 0 is also blocked. The requirement for a blocked day is only applied to day 0.

3.2.2.4.3 Representation of the impact of blocking on wind

Figure 3.34 compares the impact of blocking conditions on the wind speed PDFs. It shows that, with the exception of CMCC-CM2-HR4 over the UK, all models show a clear shift of the country-average speeds towards calmer conditions during blocked days. Regarding the effect of increased resolution in the individual models, some show improvements in the representations of both the full speed PDF and its shift: for example, MPI over Spain (Figure 3.34, left panel) and CMCC-CM2 and EC-Earth3 over the UK (Figure 3.34, right panel), but some perform worse too (e.g.; HadGEM3 and EC-Earth3 over Spain).

Regarding the joint occurrence of blocking and LW days, Figure 3.35 shows that with the exception of ECMWF-IFS, the GCMs capture reasonably well the pattern of increased probability of LW days, but there is a tendency to underestimate the frequency in most countries (with the exception of CMCC-CM2). Some models like HadGEM3-GC31 and MPI-ESM1-2 indicate an improvement with increased resolution, but not robustly across countries.

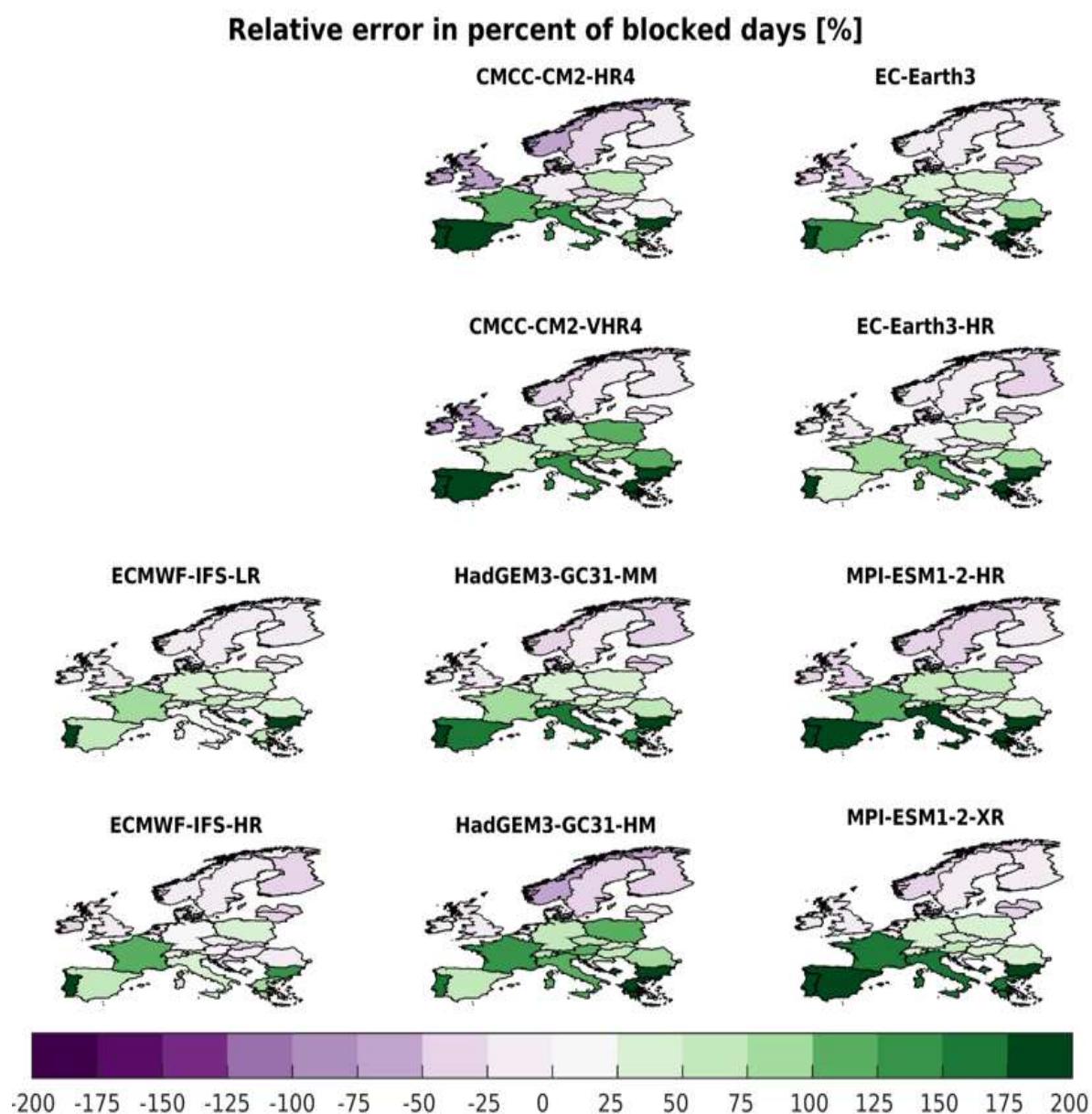


Figure 3.30. Percent relative error with respect to ERAI in the percent of blocked days in each country considering the full year [%].

A comparison of the probability of subsequent LW days for the four focus countries (not shown) revealed that, despite the aforementioned biases, the impact of blocking on LWE persistence is also present in the PRIMAVERA GCMs, though the magnitude of the impact suffers from biases. The comparison of the spatial distribution of the effect is presented in Figure 3.36. For day +3, most models and resolutions capture the widespread increase in probability of LW days, and the main bias is in not showing higher probabilities around the Mediterranean region, regardless of the resolution. CMCC-CM2-VHR4 is the model that captures more closely the probability change, despite not being amongst the finest in resolution. For higher lags (not shown), the change in probabilities in ERAI is less widespread than in the models and the sign differences between the countries are not well captured by the GCMs.

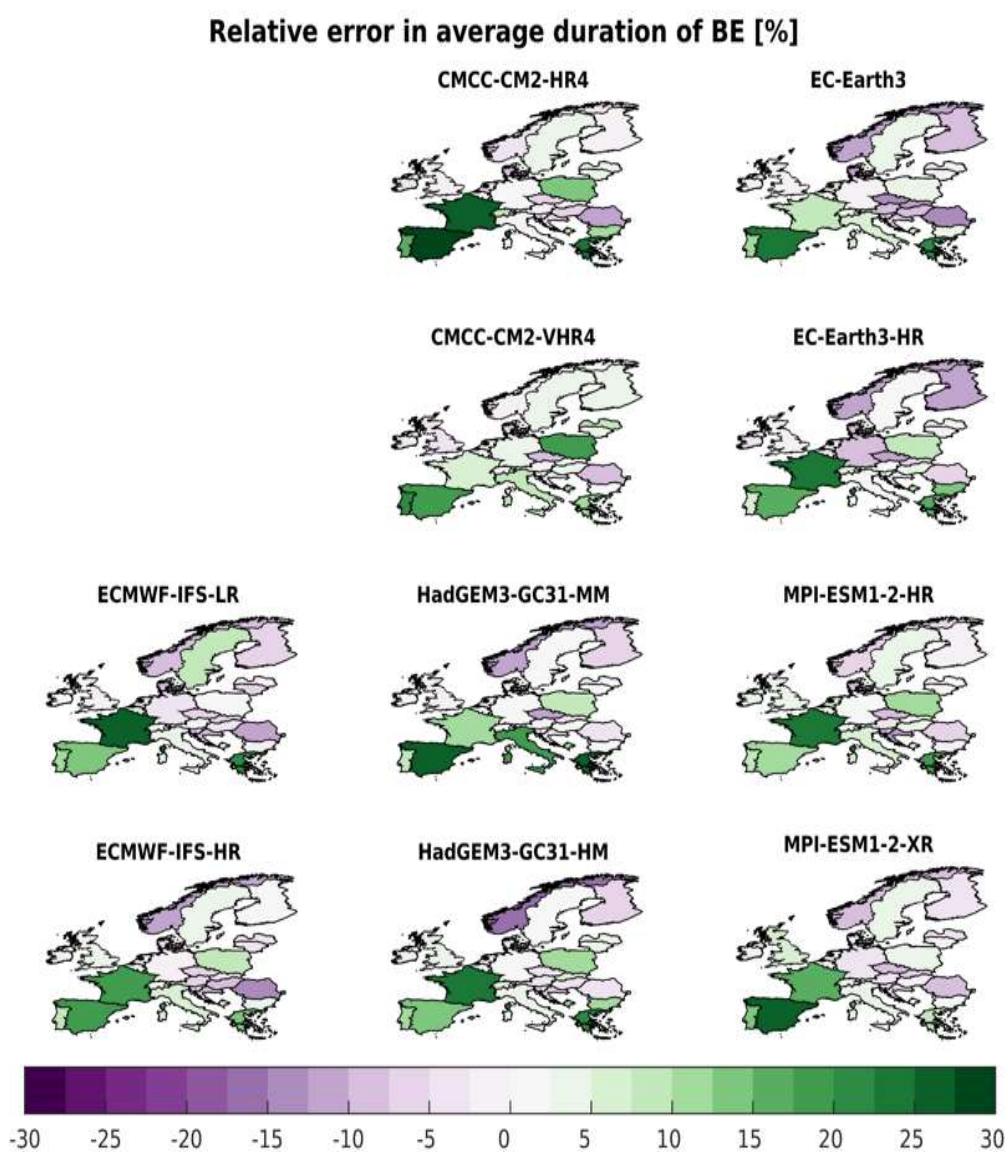


Figure 3.31. Percent relative error with respect to ERAI in average duration of blocking events [%].

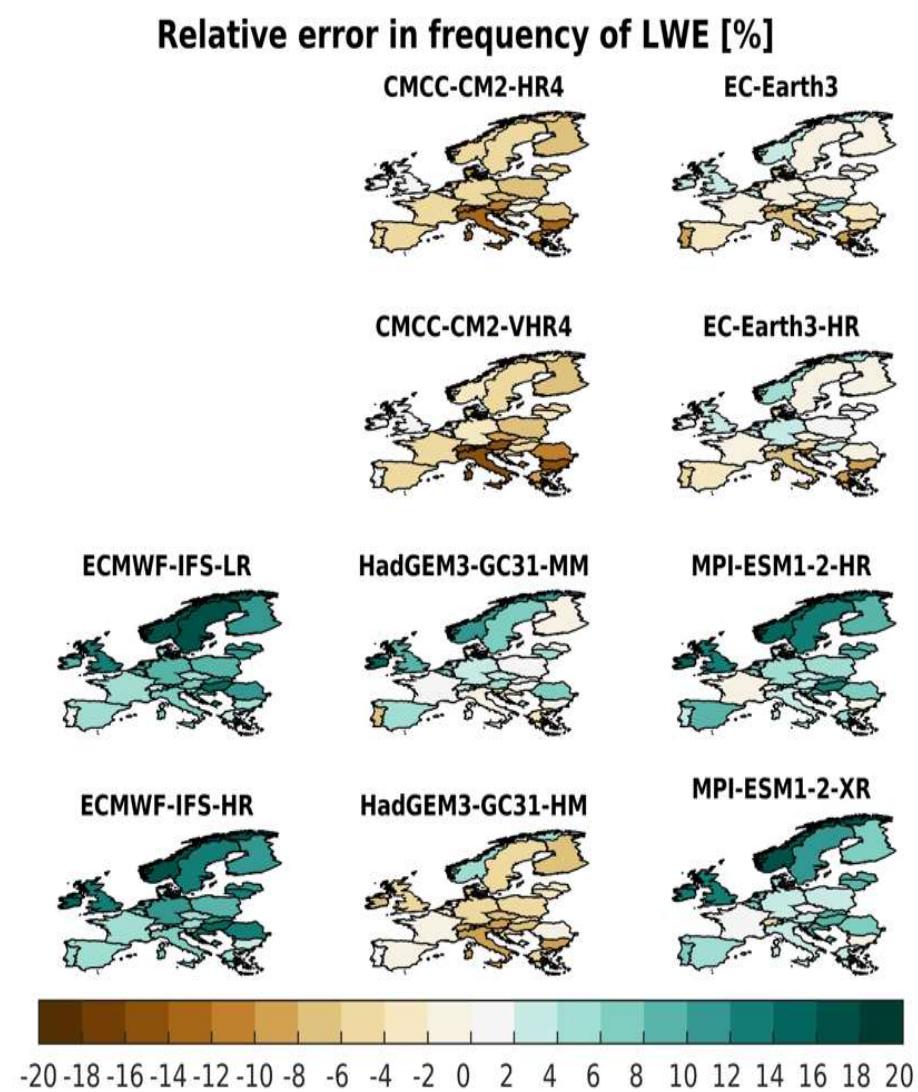


Figure 3.32. Percent relative error with respect to ERAI in average frequency of low wind events [%].

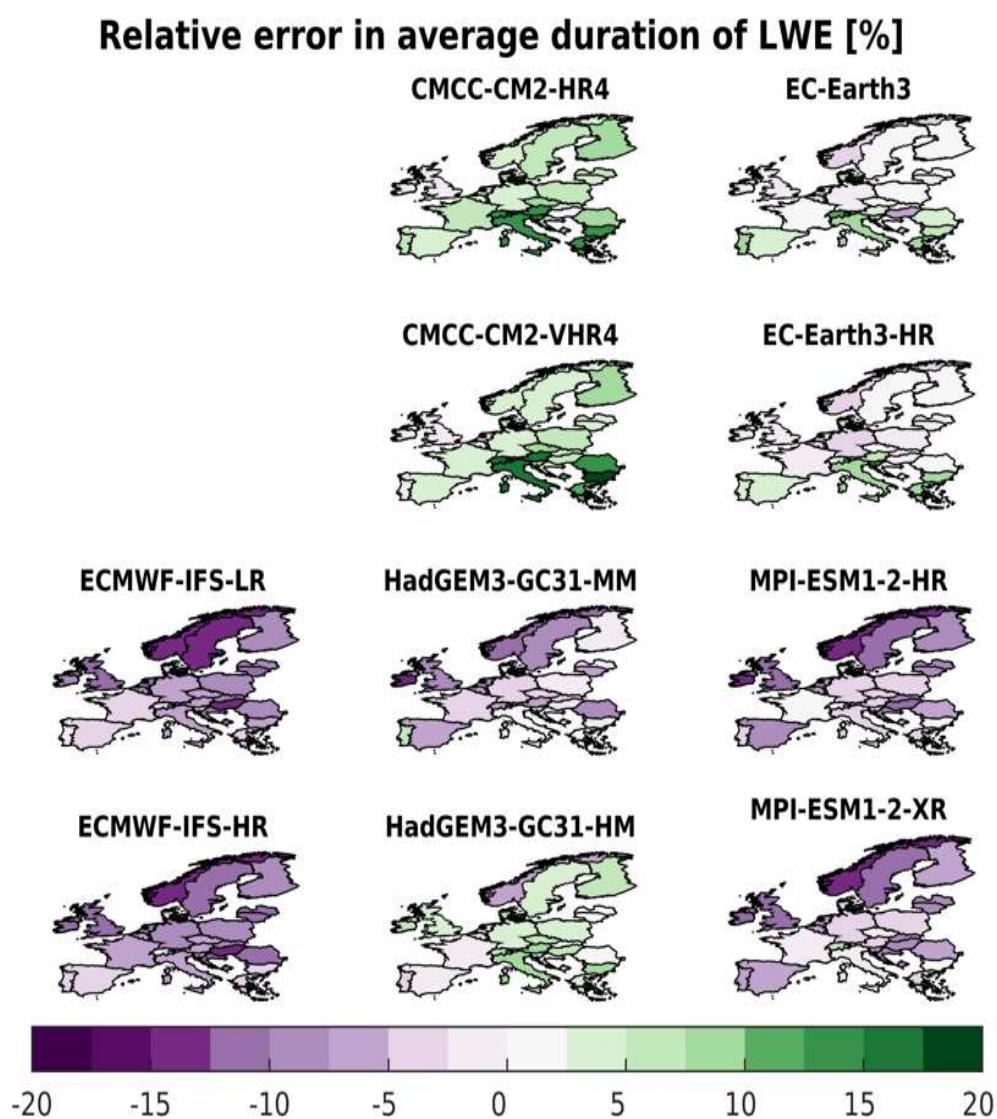
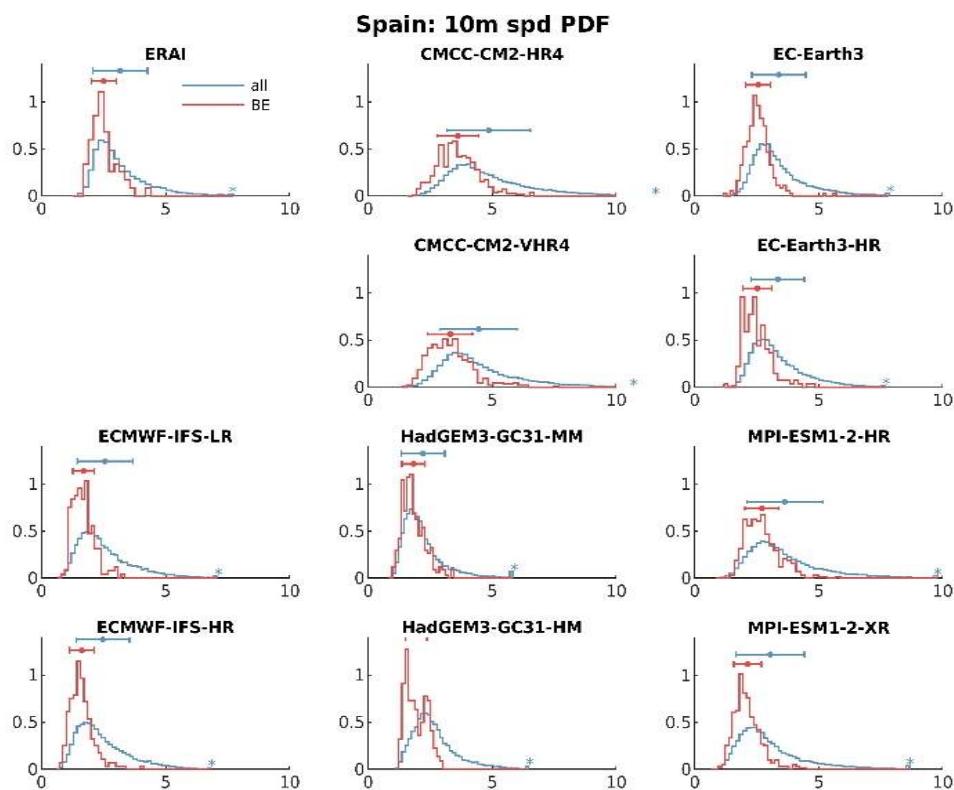


Figure 3.33. Percent relative error with respect to ERAI in average duration of low wind events [%].

a)



b)

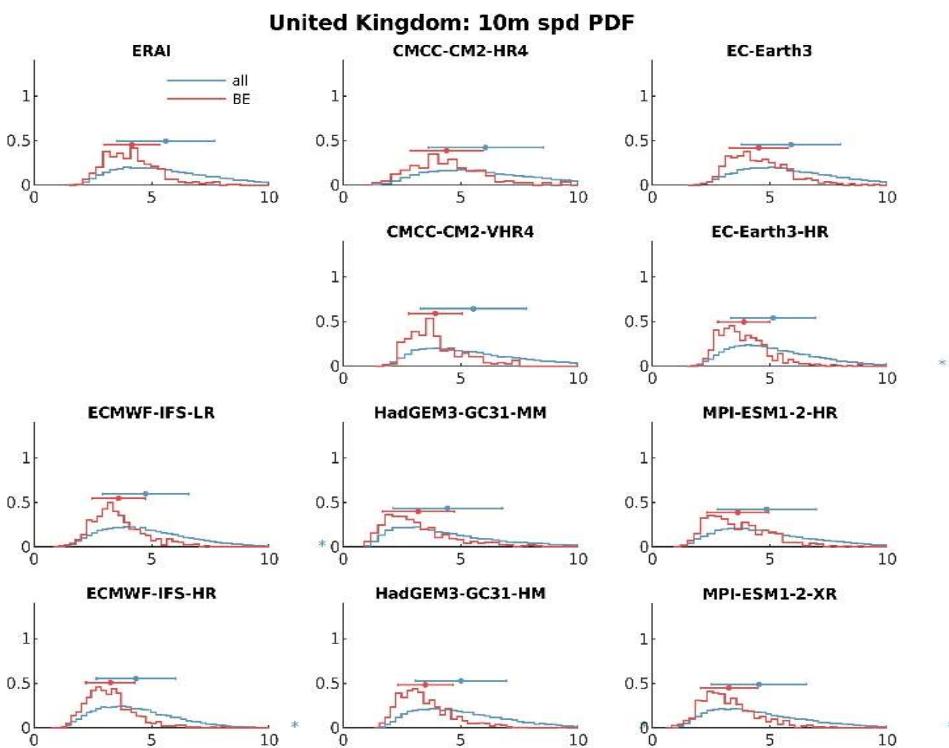


Figure 3.34. Comparison between the full-year country-average 10m speed PDF (in blue) and that corresponding only to wind speeds during blocked days (in red), for a) Spain and B) United Kingdom.

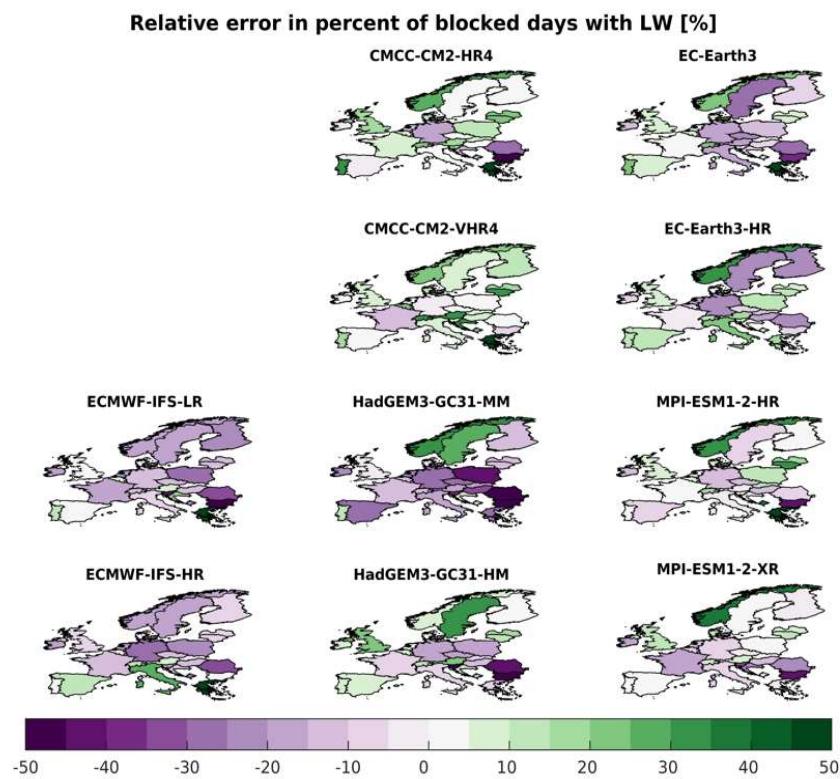


Figure 3.35. Percent relative error with respect to ERAI in the proportion of blocked days with low wind [%].

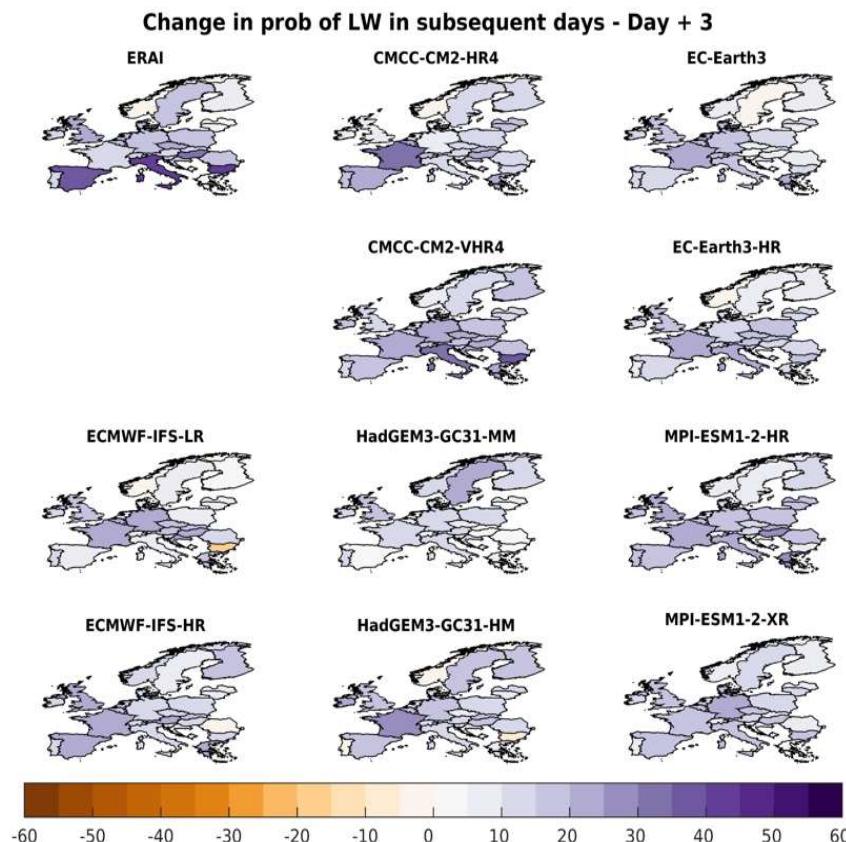


Figure 3.36. Change in the probability of observing subsequent low wind at day +3 given that day 0 is also blocked. The requirement for a blocked day is only applied to day 0.

3.2.2.5 Summary and discussion

It has been shown that blocking events over Europe have a significant impact on the occurrence and duration of low wind speeds at the country level, which is of direct relevance to the energy sector. In addition to becoming more frequent, LWE are also more persistent under blocking conditions over large areas of Europe. In general, both effects are captured by most of the PRIMAVERA GCMs analysed here, revealing that models which simulate BE reasonably under highresSST-present forcing also capture the basic dynamical connection with wind anomalies. Nonetheless, the fact that the simulated weather conditions have deficiencies introduces biases in the properties of the events and their joint occurrence. The errors in the models depend on the statistic, the country and the resolution, but some consistent bias patterns can be observed (e.g., north-south dipolar structures). No robust improvements in the representation of these effects were observed in the high-resolution versions of the PRIMAVERA models, nor where the highest resolution runs consistently outperforming coarser simulations. It is noted that CMCC-CM2-HR4 is the worst-performing simulation and also that with the coarsest resolution.

The limited length of the record (in particular for ERAI) constrains the robustness of this analysis, in particular in the case of compound events, which occur very rarely. Other limitations of the study arise from the size of the model ensemble.

Given the relevance the impacts discussed here have for the energy sector, the results could be taken as a note of caution for energy system simulations that employ this type of GCM simulations. It follows that, although some wind speed mean biases in the models could be easily corrected, errors in the frequency or duration of weather events are not easily dealt with and will introduce errors in wind power and energy demand simulations.

These results will be further improved and complemented by developing two aspects: in the first place, rather than working on blocked conditions at the country level, a set of timeseries will be used to represent conditions such as Scandinavian or Greenland blocking, that are known to have the most significant surface impacts over Europe. Additionally, to account for the full impact of blocking conditions on the energy systems, the effect on temperature will be included as a proxy for the impacts to electricity demand.

3.3 The North Atlantic eddy-driven jet

3.3.1 Representation of the North Atlantic eddy-driven jet in PRIMAVERA simulations *Panos Athanasiadis (CMCC) and Alex Baker (UREAD)*

3.3.1.1 Introduction

Climate models exhibit biases in the representation of the North Atlantic eddy-driven jet, particularly its climatological mean position and variability (e.g. Iqbal et al., 2018). Despite

improvements over previous model generations (Hannachi et al., 2013), an accurate representation of the statistics of the eddy-driven jet variability (meridional shifts and pulsing) remains a challenge, upon which also depends the representation of important aspects of European climate, including weather extremes and the frequency of severe prolonged anomalies such as cold spells. Interestingly, however, state-of-the-art climate models developed in PRIMAVERA and contributing to CMIP6 exhibit significantly reduced biases comparing to previous model generations (referring to CMIP5 and CMIP3) that may be partly attributed to better resolving oceanic eddies (and thus ocean–atmosphere interactions) at the Gulf Stream extension region.

Here, we evaluate the midlatitude atmospheric circulation over the North Atlantic using historical simulations for the period 1950–2014, both coupled (hist-1950) and AMIP-like with prescribed observed daily SSTs (highresSST-present). We also compare low-resolution (LR) and high-resolution (HR) simulations. For brevity, we herein focus on winter (DJF) and summer (JJA) seasons only.

3.3.1.2 Methodology

To identify the eddy-driven jet, daily mean u-wind field at 850 hPa was extracted from each simulation and interpolated from native model grids to a $2.5^\circ \times 2.5^\circ$ regular grid. Following Woollings et al. (2010) and Woollings et al. (2018) and applying an additional orography mask (to account for the 850 hPa isobaric level being underground over most of Greenland). Jet latitude and jet speed are defined over the domain 0–60W and the respective distributions are determined for each simulation and the NCEP/NCAR reanalysis, binned, respectively, at 2.5° latitude and 1.0 m s^{-1} speed and smoothed by a PDF kernel (Silverman, 1986). The results are largely insensitive to the kernel estimation method, the practical effect of which is smoothing.

3.3.1.3 Key results

In Figure 3.37, considerable inter-model variability is evident in the representation of the North Atlantic jet in wintertime. EC-EARTH3 and ECMWF-IFS models, which share the same dynamical core, exhibit the smallest biases in the distribution of the jet. HadGEM3-GC31, MPIESM-1-2-HR and CNRM-CM6-1 exhibit moderate biases, mainly in the representation of the trimodal character of the jet latitude distribution, whilst CMCC-CM2 appears to fail in reproducing the observed trimodality. Ignoring some details, these results generally hold true for both the AMIP-like (upper panel) and the coupled (lower panel) simulations. For most models, coupling and increasing the model resolution (alone and together) seem to increase the south-jet occurrences, indicating (Madonna et al., 2017) increased frequencies of Greenland blocking and the NAO(–) circulation regime. Coupling also brings an equatorward shift of the mean jet position, as indicated by the coloured markers in the two panels. This result is corroborated by the respective PRIMAVERA diagnostics for North Atlantic blocking (Reinhard et al., 2020) and weather regimes (Fabiano et al., – in review) and is consistent with recent studies (e.g. Haarsma et al., 2019) pointing to the importance of fine-scale coupled processes for the realistic representation of low-frequency variability in the North Atlantic sector. Considering also the respective jet speed distributions in Figure 3.38, coupling (upper to lower panel) and increasing the model

resolution (solid to dashed lines) seem to slightly lower the mean jet speed in most models. Overall, significant improvements are seen comparing to CMIP5 and CMIP3 models, specifically in capturing the trimodal character of the jet latitude distribution, which is directly related to the existence of distinct Euro-Atlantic circulation regimes.

Key results for summer: jet speeds (Figure 3.40) get slightly reduced with coupling, while, for most models, jet latitude distributions (Figure 3.39) get slightly closer observations with coupling. No robust changes are found with increasing resolution. For the transitional seasons (spring and autumn, results not shown) models generally represent the observed jet latitude and jet speed distributions better than in winter and exhibit larger intra-ensemble spread when coupling is introduced.

The present analysis is based on single realizations (Stream-1 simulations). Therefore, it is fair to wonder whether the above-discussed findings would hold for a different realization. Results from a recent study (Kwon et al., 2018) using large ensembles, referring to the Large Ensemble Simulations run with the Community Earth System Model (CESM-LENS), provide evidence that the jet statistics over a similar historical period (1951–2005) exhibit small intra-ensemble spread. This boosts our confidence that single-member analyses are a viable option for similar studies, yet the presented results will be expanded to include more available members for the models that this is applicable (incorporating Stream-2 simulations).

3.3.1.4 *Forthcoming research*

- Two-dimensional assessment of the eddy-driven jet position and strength.
- Evaluation of the role of SST biases for the jet biases over the North Atlantic.
- Analysis of future simulations (ongoing)—follow-up on Baker et al., 2019a.

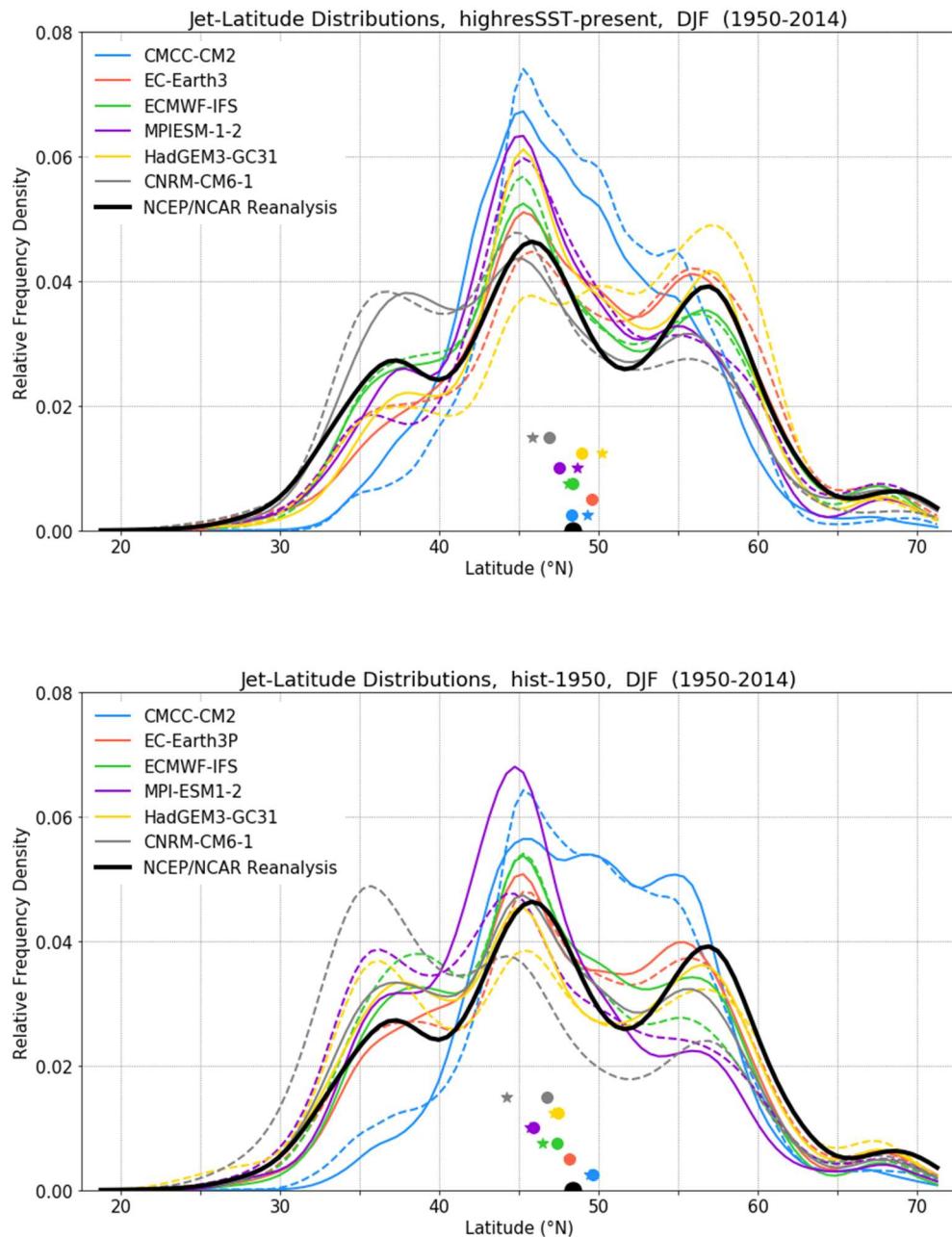


Figure 3.37. Model (coloured lines) and observed (thick black line) distributions of jet latitude for the North Atlantic eddy-driven jet in winter. All distributions are computed for the same period (1950–2014) and are estimated by a PDF kernel (see text for details). Upper panel for AMIP-like simulations (highresSST-present) and lower panel for the respective coupled (hist-1950) simulations. The dashed lines correspond to the high-resolution version of each model. Coloured markers (circles for LR and asterisks for HR) indicate the respective mean jet latitude.

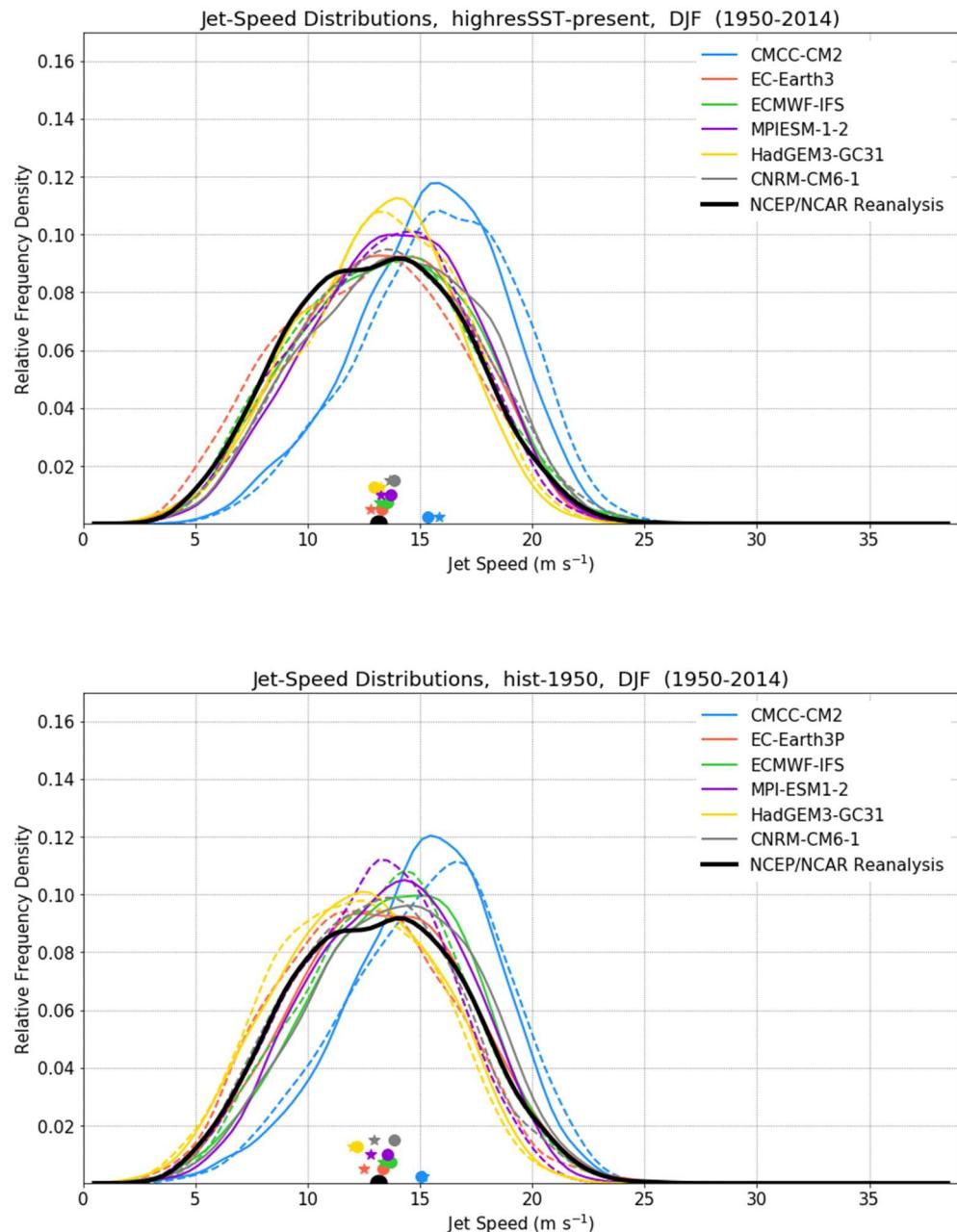


Figure 3.38. Model (coloured lines) and observed (thick black line) distributions of jet speed for the North Atlantic eddy-driven jet in winter. All distributions are computed for the same period (1950–2014) and are estimated by a PDF kernel (see text for details). Upper panel for AMIP-like simulations (highresSST-present) and lower panel for the respective coupled (hist-1950) simulations. The dashed lines correspond to the high-resolution version of each model. Coloured markers (circles for LR and asterisks for HR) indicate the respective mean jet latitude.

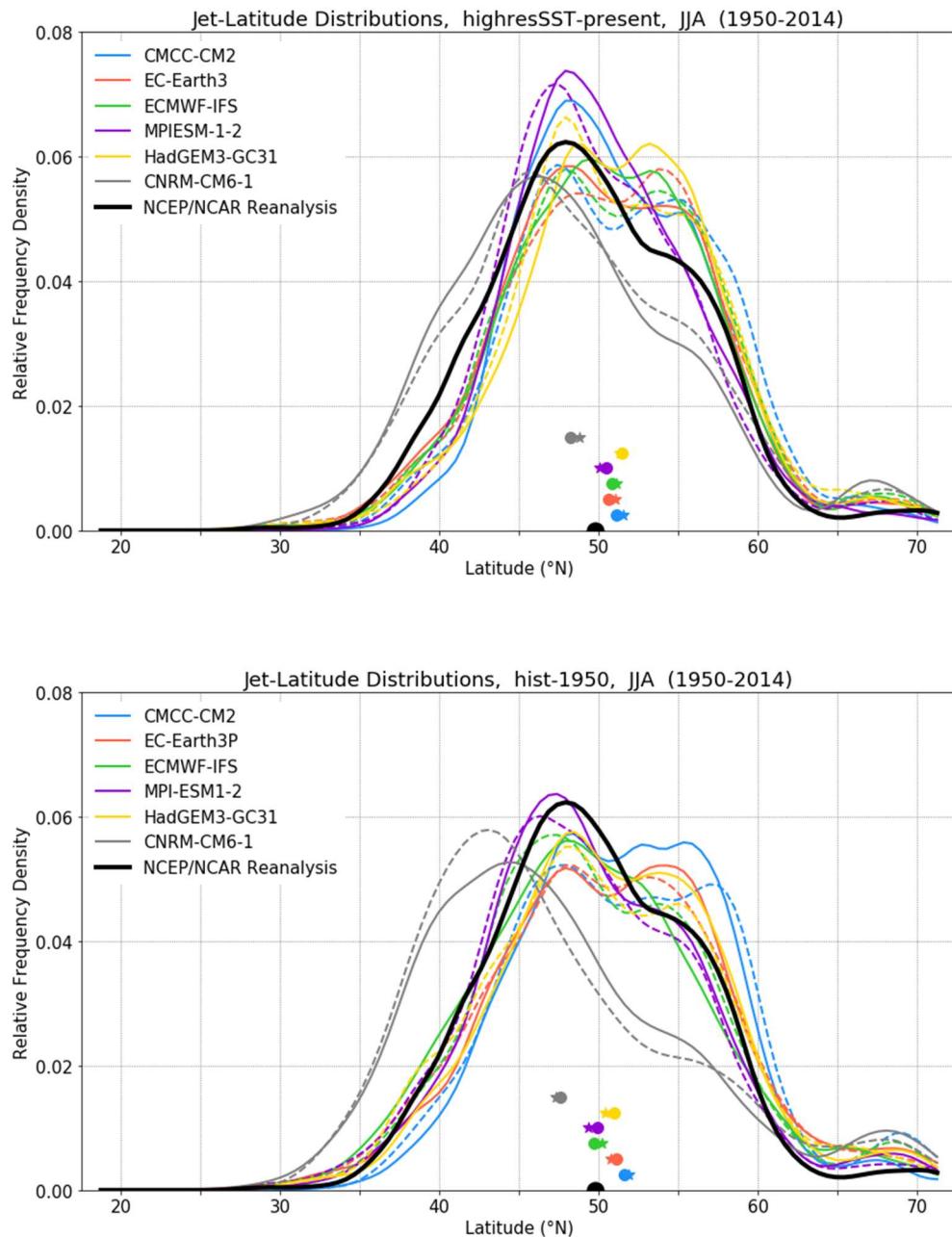


Figure 3.39. Model (coloured lines) and observed (thick black line) distributions of jet latitude for the North Atlantic eddy-driven jet in summer. All distributions are computed for the same period (1950–2014) and are estimated by a PDF kernel (see text for details). Upper panel for AMIP-like simulations (highresSST-present) and lower panel for the respective coupled (hist-1950) simulations. The dashed lines correspond to the high-resolution version of each model. Coloured markers (circles for LR and asterisks for HR) indicate the respective mean jet latitude.

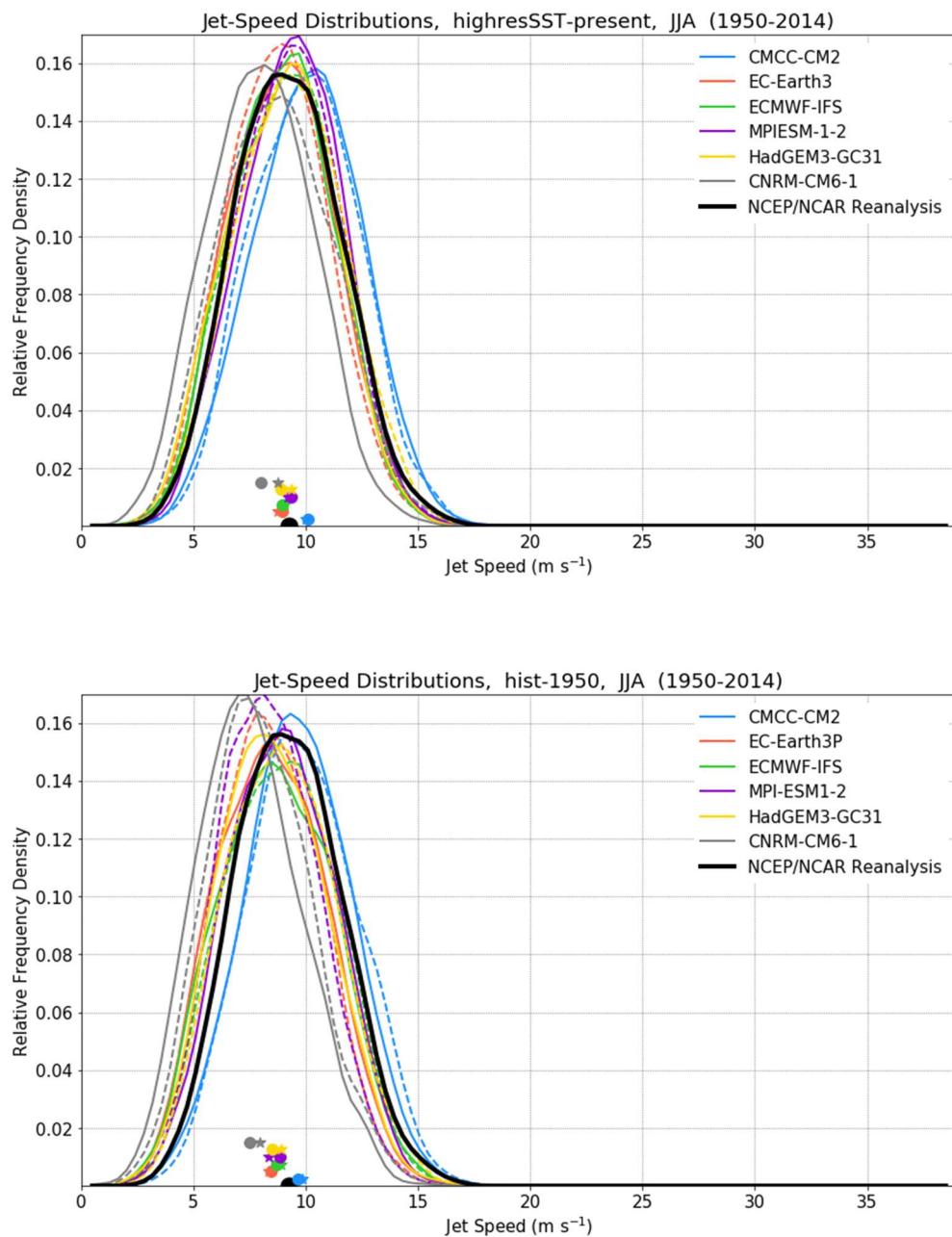


Figure 3.40. Model (coloured lines) and observed (thick black line) distributions of jet speed for the North Atlantic eddy-driven jet in summer. All distributions are computed for the same period (1950–2014 and are estimated by a PDF kernel (see text for details). Upper panel for AMIP-like simulations (highresSST-present) and lower panel for the respective coupled (hist-1950) simulations. The dashed lines correspond to the high-resolution version of each model. Coloured markers (circles for LR and asterisks for HR) indicate the respective mean jet latitude.

3.4 Precipitation over Europe

3.4.1 Comparing simulated daily precipitation distribution between PRIMAVERA and CORDEX ensembles

Marie-Estelle Demory (ETHZ) and Ségolène Berthou (UKMO)

3.4.1.1 *Introduction*

In this study, we perform an evaluation of PRIMAVERA high-resolution (25-50 km) Global Climate Models (GCMs) (Table 3.4) relative to CORDEX Regional Climate Models (RCMs) (Table 3.5) over Europe (12-50 km resolutions). It is the first time such assessment is performed for regional climate information using ensembles of GCMs and RCMs at similar horizontal resolutions. We perform this exercise for the distribution of daily precipitation contributions to rainfall bins over Europe under current climate conditions. Full details are described in Demory et al. (GMDD, submitted).

3.4.1.2 *Datasets*

3.4.1.2.1 PRIMAVERA GCMs

We use the ocean-atmosphere coupled GCMs developed and run within the EU-Horizon 2020 PRIMAVERA project (<https://www.primavera-h2020.eu>), which is a European contribution to HighResMIP. PRIMAVERA uses the HighResMIP protocol (Haarsma et al., 2016), which is different from CMIP (e.g. different aerosols; refer to Haarsma et al., 2016, for details). As PRIMAVERA simulations are still running, we use the ones which were available at the time of the study. So far, PRIMAVERA simulations consist of 6 GCMs (Table 3.4). Most high-resolution simulations include one member only, but in case there are more (such as the IFS-HR that provides 6 members), we consider one per model in order to apply equal weights to each model.

3.4.1.2.2 CORDEX RCMs

Over Europe, we use the CMIP5-driven EUR-44 and EUR-11 CORDEX simulations (please refer to the EURO-CORDEX simulation list here: <https://eurocordex.net/imperia/md/content/csc/cordex/20180130-eurocordex-simulations.pdf>) run at 0.44° (~50 km) and 0.11° (~12 km) resolution. Daily precipitation model data have been extracted from the Earth System Grid Federation (ESGF; <https://esgf.llnl.gov>—see Table 3.5). We focus our analysis on the EUR-44 simulations because their resolution roughly corresponds to the resolutions used by PRIMAVERA GCMs, which allows a clean comparison between the two ensembles. However, we evaluate the roles of resolution, regridding, and ensemble size in daily precipitation distribution with equivalent pairs from EUR-11.

Model name	<i>HadGEM3-GC31-HM</i>	<i>EC-Earth3P-HR</i>	<i>CNRM-CM6-1-HR</i>	<i>MPI-ESM1-2-XR</i>	<i>CMCC-CM2-VHR4</i>	<i>ECMWF-IFS-HR</i>
Institute	Met Office	KNMI, SMHI, BSC, CNR	CERFACS	MPI-M	CMCC	ECMWF
Reference	Roberts et al., 2019	Haarsma et al., 2019	Voldoire et al. 2019	Gutjahr et al., 2019	Cherchi et al., 2019	Roberts et al., 2018
Atmosphere horizontal resolution (at 50°N)	N512 (25km)	T1511 (36km)	TI359 (50km)	T255 (34km)	0.25° (18km)	Tco399 (25km, output at 50km)
Ocean resolution (km)	25km	25km	25km	40km	25km	25km
Simulation	hist-1950	hist-1950	hist-1950	hist-1950	hist-1950	hist-1950
Ensemble member	r1i1p1f1	r1i1p2f1	r1i1p1f2	r1i1p1f1	r1i1p1f1	r1i1p1f1

Table 3.4. Information about the PRIMAVERA high-resolution GCMs used in this study, including their spatial resolution (for full details, refer to <https://www.primavera-h2020.eu/modelling/our-models/>). Those listed in italics are of the same family than the CMIP5 GCMs downscaled by CORDEX.

3.4.1.2.3 CMIP5 GCMs

To investigate the added value of CORDEX RCM simulations to CMIP5 GCMs, we constrain our study to the subset of CMIP5 GCMs used to force CORDEX simulations (Table 3.5, second column), available on the ESGF servers. However, we examine the robustness of our findings by also analysing the entire ensemble of CMIP5 simulations. Taking the full set changes the ensemble spread but the main conclusions of the study regarding CMIP5 remain the same (not shown).

We perform our analysis on the full CORDEX and PRIMAVERA ensembles or on a reduced ensemble, which corresponds to PRIMAVERA GCMs and CORDEX RCMs that downscale CMIP5 GCMs that are based on the same GCM family. For example, the PRIMAVERA MPI-

ESM1-2-XR GCM and the EUR-44 RCA4, CCLM4, CCLM5 and REMO2009 that downscaled MPI-ESM-LR (blue in Table 3.5).

HighResMIP equiv.	CMIP5 GCMs		EUR-44 RCMs			
CNRM-CM6-1-HR	CNRM-CM5	r1	RCA4	ALADIN53	ALARO	CCLM5
EC-Earth3P-HR	EC-EARTH	r1	RACMO22E	WRF341E		
	EC-EARTH	r3	HIRHAM5			
	EC-EARTH	r12	RCA4	CCLM5		
MPI-ESM1-2-XR	MPI-ESM-LR	r1	RCA4	CCLM4	REMO09	CCLM5
	MPI-ESM-LR	r2				
HadGEM3-GC31-HM	HadGEM2-ES	r1	RCA4	RACMO22E	CCLM5	RegCM4
	IPSL-CM5A-MR	r1	RCA4	WRF331F		
	NorESM1-M	r1	RCA4			
	GFDL-ESM2M	r1	RCA4			
	MIROC5	r1	RCA4	CCLM5		

Table 3.5. Summary of historical EURO-CORDEX simulations used in this study. The first column indicates HighResMIP models of the same family as the CMIP5 GCM (second column) driving the RCMs. Matching colours show comparable HighResMIP GCMs and EURO-CORDEX RCMs. Within EURO-CORDEX RCMs, dark shaded models are available at both 0.11° (EUR-11) and 0.44 (EUR-44) horizontal resolutions. HIRHAM5* indicates several versions of this model were used.

3.4.1.2.4 Observations

We make use of the best available observational datasets. These are mostly national datasets, such as SAFRAN-V2 (France; Vidal et al., 2010), UKCPobs (British Isles; Perry et al., 2009), ALPS-EURO4M (Alps; Isotta et al., 2014), CARPACLIM (Carpathian region; Szalai et al.). To cover the Iberian Peninsula, we combine Spain02 v2 (Herrera et al., 2012) and PT02 v2 (Belo-Pereira et al., 2011). For other regions, we considered E-OBS v17 (Cornes et al., 2018).

3.4.1.3 Methodology

We look at the daily precipitation distribution in each sub-region. We use a similar method as Berthou et al. (2019) based on the ASOP1 diagnostics tool developed by Klingaman et al. (2017). We calculate the daily precipitation distribution in terms of the actual contribution from 100 different intensity bins to mean precipitation. In order to account for the high frequency of low intensity precipitation events and the low frequency of high intensity events, we use an exponential bin distribution, as described by Berthou et al., 2019. To calculate the contribution to mean precipitation, each bin frequency is multiplied by its average rate. Thus, mean precipitation is split in different contributions of different rates. We consider a logarithmic scale on the x-axis, so the area under the curve is directly proportional to the mean.

For the ‘pie plot’ (Figure 3.43), all data are regridded on a common EUR-44 rotated pole grid. The ensemble mean is calculated for each bin and a bootstrap resampling is used 1000 times on each model ensemble to establish a confidence interval around the ensemble

mean. For the observations, the bootstrap resampling is done on single years, therefore reflecting inter-annual variability. We group the bins as 3 intensity precipitation intervals (low: 1-10 mm/day; mid: 10-60 mm/day; high: >60 mm/day). We evaluate for each interval the percentage of bins over which the ensembles differ (refer to Demory et al. for details).

3.4.1.4 Results overview

3.4.1.4.1 Added value of EUR-11 and EUR-44 compared with CMIP5

Figure 3.41 shows the precipitation distribution for EUR-44, EUR-11, and a selection of CMIP5 GCM models (Table 3.5). There is a clear shift in the precipitation distribution going from CMIP5 to CORDEX (EUR-44 and EUR-11) over all regions. EUR-44 and EUR-11 simulate an overall decrease in low intensity precipitation and an increase in high intensity precipitation. CMIP5 simulate very little high intensity precipitation, while their mid-rate precipitation is much larger than CORDEX. This finding may be attributed to the finer grid box (meaning the rain rates are those of a smaller area), the better representation of orography and coastlines that may enhance the triggering of summer convective precipitation, the use of convective schemes which are more appropriate at the resolution of the RCMs, or the tuning of parameterization schemes. The differences between EUR-44 and EUR-11 are small (and smaller than in DJF), which suppose that such resolution jump does not influence summer precipitation largely when convection parameterization is used. This is also seen in other regions (Figure S3 of Demory et al., 2020), and is in line with previous studies showing no systematic improvement between EUR-44 and EUR-11 for mean precipitation (Kotlarski et al., 2014; Casanueva et al., 2016). The effect of resolution remains, however, large for summer precipitation over orography (e.g. AL), which confirms the findings of Torma et al. (2015) and Prein et al. (2016).

3.4.1.4.2 Precipitation distribution in EUR-44 and PRIMAVERA

Our results show that CMIP5-driven EUR-44 and PRIMAVERA atmosphere-ocean coupled simulation ensembles give equivalent regional climate information in terms of daily precipitation distribution and its contribution to precipitation intervals (Figure 3.42). The differences in their precipitation distribution are generally small, and much smaller than differences between CORDEX and CMIP5 (Figure 3.41 and 3.42).

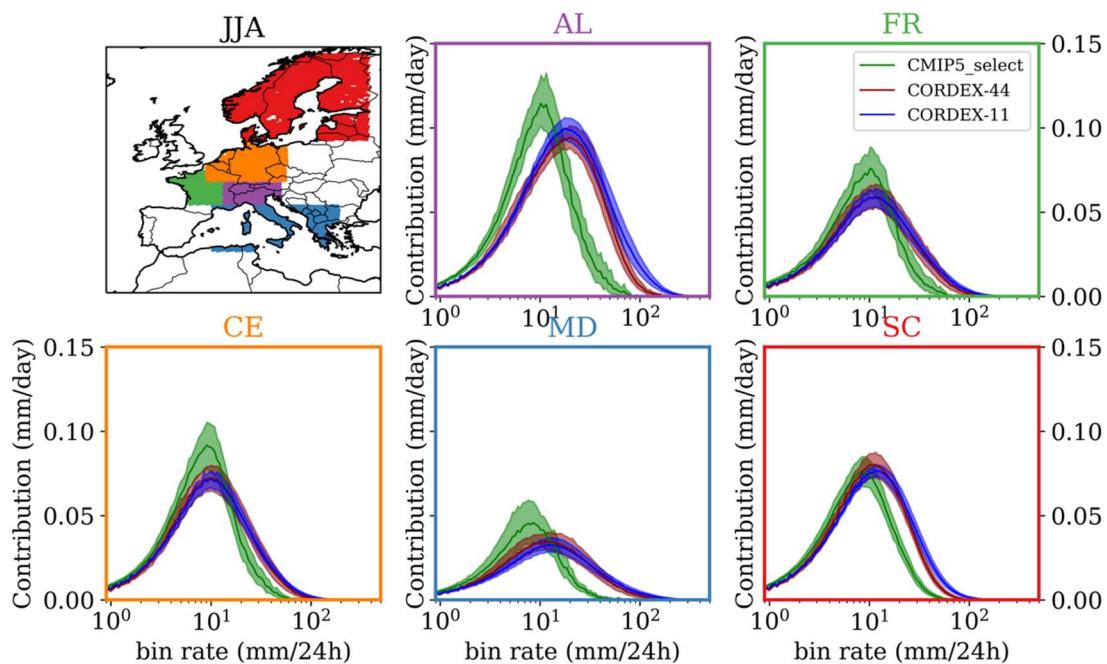


Figure 3.41: Precipitation contribution (frequency x bin rate) per rain rate in JJA over the Alps (AL), France (FR), Central Europe (CE), Mediterranean (MD), Scandinavian (SC) for a selection of CMIP5 GCMs (green), EUR-44 (red), EUR-11 (blue). All data are plotted on the models native grid.

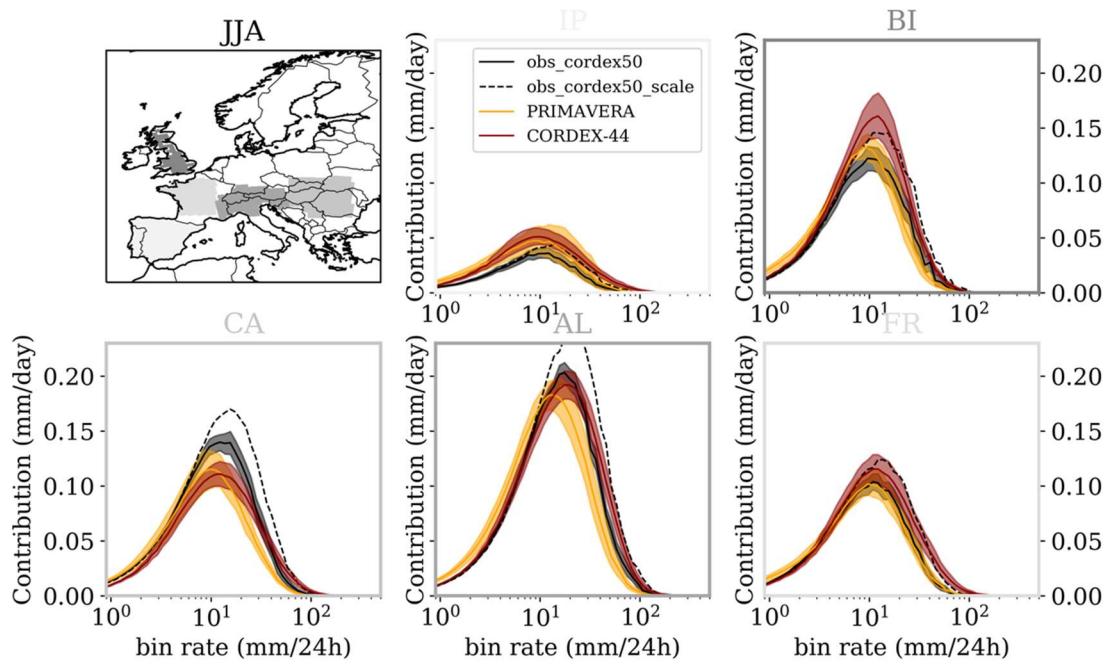


Figure 3.42: Precipitation contribution (frequency x bin rate) per rain rate in JJA over the Iberian Peninsula (IP), British Isles (BI), Carpathians (CA), Alps (AL), France (FR) for EUR-44 (red), PRIMAVERA (orange), observations regredded on EUR-44 (black) and a synthetic observational dataset taking into account an additional 20% under-catch error (dashed line).

Figure 3.43 shows the full EUR-44 ensemble and a reduced ensemble, which includes only the analyses for the EUR-44 RCMs forced by the same GCMs that comprise PRIMAVERA. This means we have 4 RCMs forced by 4 CMIP5 GCMs compared with 4 PRIMAVERA GCMs.

PRIMAVERA and CORDEX ensembles are of good quality in summer and autumn (except in the Carpathians region), but tend to overestimate precipitation in winter and spring. However, there are some precipitation intervals, seasons and regions for which the two ensembles significantly differ. A large difference between the two ensembles is found for heavy precipitation (in all regions in summer, and in some regions in other seasons). PRIMAVERA simulate have less heavy rainfall than EUR-44, and tend to agree better with raw observations, while EUR-44 are closer to synthetic observational datasets when a 20% under-catch error is considered. PRIMAVERA GCMs tend to have more light precipitation than EUR-44, and too much compared with the observations, although this result is not as robust as the former one. It is possible that expert tuning of the convective scheme and land-surface scheme in RCMs has a positive effect towards reducing this “drizzling” problem.

Another conclusion is that when considering only shared GCM families between the two ensembles, differences in the bulk of the distribution (medium rain rates) is mostly found in the central and eastern parts of the European domain, in autumn, winter and spring (Figure 3.43b). PRIMAVERA tend to reduce precipitation overestimation in these regions and seasons compared with EUR-44. This could be linked with better simulation of blocking frequency in PRIMAVERA GCMs (Schiemann et al., 2018), which is not achieved by CORDEX (Jury et al., 2019).

Finally, some results are specific to a few regions: over the Alps and the British Isles, PRIMAVERA underestimate heavy precipitation in summer while EUR-44 overestimate it, although EUR-44 is in good agreement with an approximate correction for precipitation under-catch (rain rates increased by 20%). Over the British Isles, precipitation over 30 mm/day is underestimated in autumn, winter and spring by both ensembles. This could mean that those models are still too coarse to correctly represent the interactions between low pressure systems and local coastal and orographic effects over this region. In the Carpathians, summer precipitation is underestimated by both ensembles and winter precipitation is overestimated, although PRIMAVERA shows improvements in this case.

The performance of PRIMAVERA was not logically expected because these GCMs were developed at a coarser resolution, and only their resolution was increased. The tuning was performed on their low-resolution counterparts, so little additional tuning was performed at these high resolutions (see Roberts et al., 2019, for changes in models when increasing resolution), as opposed to RCMs which are developed at a higher resolution and potentially tuned at each resolution.

The fact that PRIMAVERA results exhibit moderate improvements over CMIP5-driven CORDEX simulations for precipitation over Europe is also an important result of this study, which is consistent with the results of Iles et al. (2019) who used a very different method to compare GCMs and RCMs at different resolutions. It indicates that the potential improvement of large scales dynamics in GCMs due to higher resolution does not have a strong influence on precipitation improvement, which is largely driven by downscaling. The added value of RCMs to CMIP5 GCMs is also an important result, and it emphasizes the

importance of a well-designed, well evaluated model chain when using dynamical downscaling as a method to obtain higher resolution climate data. We show here that considering climate information from various sources is crucial.

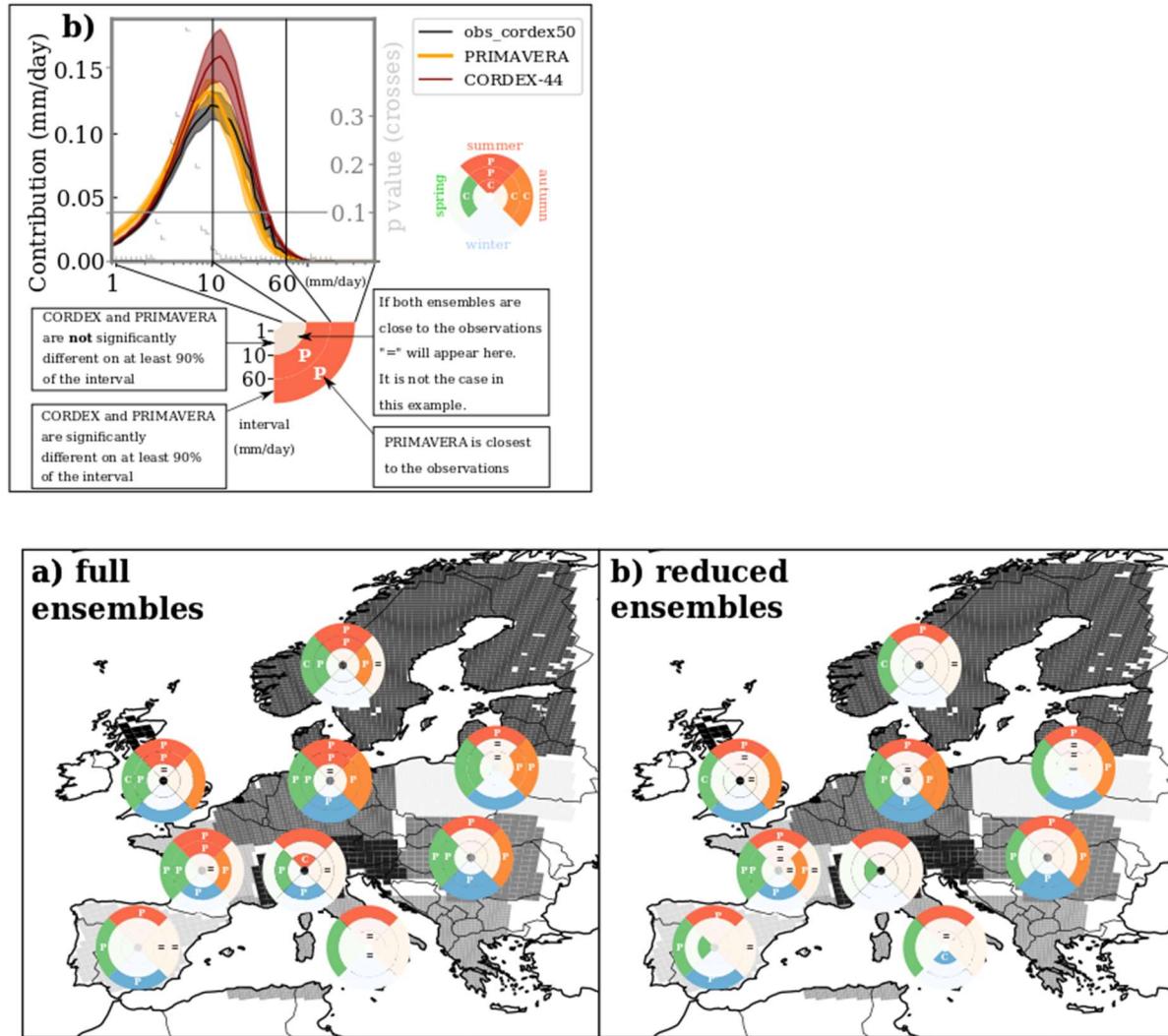


Figure 3.43. Comparison of precipitation simulated by PRIMAVERA and CORDEX models with observations over European sub-domains. A colour indicates where CORDEX and PRIMAVERA are different by at least 90% of the precipitation interval. Where the models are not significantly different, no colour is shown. Letters ('P' for PRIMAVERA, 'C' for CORDEX) indicates which ensemble is closer to observations by at least 10%. An equal sign ('=') indicates ensembles are equally close to observations (at least on 30% of the interval). No letter or sign indicates both ensembles are far from observations, whether they are different (colour) or not (no colour). The top panel illustrates the figure construction and the results are shown in the bottom panels for (a) the full EUR-44 ensemble and (b) the reduced ensemble (i.e., reduced PRIMAVERA and EUR-44 ensembles using GCMs of the same family—coloured in Tables 3.4 and 3.5).

3.4.1.5 Conclusions

1. CORDEX and PRIMAVERA differ most for the most intense precipitation in summer (outside part of pies) over all regions (Figure 3.42). EUR-44 overestimate intense precipitation, PRIMAVERA are generally in better agreement with the observations (Figure 3.43a).
2. When using reduced CORDEX and PRIMAVERA ensembles (Figure 3.43b), the ensembles are much more similar (no colour in most parts of the pies), which means that the "bulk" of the distribution in most regions depends on the nature of the GCM rather than on the downscaling method. The main difference is for the intense bins in most seasons and regions, where the conclusions are similar to Figure 3.43.
3. PRIMAVERA seem a better ensemble in general than the full EUR-44 ensemble compared with raw observations, but EUR-44 are closer to the synthetic observational dataset in which we account for an averaged 20% precipitation undercatch (Figure 3.42). PRIMAVERA and CORDEX, being EUR-11 or EUR-44, should therefore be considered equally credible, depending on the user's needs.
4. The added value of RCMs to CMIP5 GCMs emphasizes the importance of a well-designed, well evaluated model chain when using dynamical downscaling as a method to obtain higher resolution climate data. We show here that considering climate information from various sources is crucial.

3.4.1.6 Sensitivity testing

Sensitivity testing has been performed to establish how results change depending on the significance thresholds and bins used. The main conclusions are robust to these methodological changes, but some regions are more sensitive to the thresholds than others (e.g., for the Alps, where observations lie in between CORDEX and PRIMAVERA, one ensemble can be considered better than the other, depending on chosen thresholds). We also performed tests focussed on the impact of resolution (EUR-44 versus EUR-11) and regridding (EUR-11 data on their native grid versus EUR-11 on the EUR-44 grid). The main conclusions do not change (with the EUR-11 ensemble we have used, although more simulations are forthcoming that will require these tests to be repeated). Generally, EUR-11 (regridded on EUR-44) simulates even more intense precipitation than EUR-44 (not shown). Mean precipitation in EUR-11 is generally lower than in EUR-44, which is an improvement, but EUR-11 is generally not as close to observations as PRIMAVERA. These findings are similar on the native grid of EUR-11 compared with observations regridded to EUR-11 (not shown). These results are described in Demory et al. (submitted to GMDD).

3.4.2 The importance of model resolution of simulated precipitation in Europe

Gustav Strandberg (SMHI)

3.4.2.1 Introduction

Extreme precipitation affects many parts of society. Extreme precipitation is also something that we know will change in a different way than the mean precipitation (Collins et al., 2013) and something that we know is highly dependent on model resolution (e.g. Lind et al. 2016). To be able to adapt to future changes in extreme precipitation we need models that can capture these small-scale features. The goal of this study is to examine how precipitation depends on model resolution.

The added value of global high resolution is not entirely clear. It is clear that high horizontal resolution adds value; in that sense there is a value of high global resolution. If the global climate model (GCM) data is used to provide boundary conditions to a regional climate model (RCM) of high resolution the added value is less clear. It is clear however that the large-scale precipitation in regional models to a large degree is governed by the general circulation of the driving global model (e.g., Kjellström et al., 2011). If the global model with higher resolution has an improved circulation it will be beneficial also for the regional simulation. This effect is more important to large scale precipitation but could potentially also affect local precipitation.

3.4.2.2 Method

In this section we look especially at the effect of resolution on the intensity of precipitation. This is done using the ASoP (Analyzing Scales of Precipitation) software (Klingaman et al., 2017; Berthou et al., 2018), which measures the spectrum of precipitation intensities. ASoP gives a distribution of the contributions of each precipitation intensity bin to the mean precipitation rate. The distributions are calculated for each model grid point, and then averaged over desired regions. In the first step, the method defines the precipitation intensity bins such that all bins have a similar number of events, except for the largest bins due to small number of events there. The frequency of events in each bin is then multiplied by the mean precipitation rate of the bin to obtain the actual contribution of the bin to the mean precipitation rate. Note that the sum of all actual contributions gives the mean precipitation rate. Furthermore, dividing the actual contributions by the mean precipitation rate gives the fractional contributions to the mean precipitation. The sum of all fractional contributions equals one, so the information provided by fractional contributions is predominantly about the shape of the distribution.

The models used are a selection of CMIP5 global models (~100-300 km resolution); the high (~40-80 km) and low (~800-160 km) resolution versions of the PRIMAVERA Stream1 global models; and a selection of CORDEX regional models (at 50 and 12.5 km resolution). The CMIP5 and CORDEX ensembles are not complete, but rather “ensembles of opportunity” for which daily precipitation were easily available. Table 3.4 lists the ensembles used.

Ensemble	Model type	No. models	Resolution
CMIP5	GCM	18	100 – 300 km
PRIMAVERA low	GCM	5	80 – 160 km
PRIMAVERA high	GCM	4	40 – 80 km
CORDEX low	RCM	17*	50 km
CORDEX high	RCM	28**	12.5 km

Table 3.4. A description of the model ensembles used in this study.

* The CORDEX low ensemble consists of 7 RCMs combined with 10 GCMs.

** The CORDEX high ensemble consists of 8 RCMs combined with 11 GCMs.

Here, the different model ensembles are analysed and compared. The analysis shows for which intensities precipitation changes depend on resolution. This is done for the Prudence regions in Europe (Figure 3.44). For each region we get a precipitation distribution. AsoP is run on daily precipitation data over land for each model for all regions for annual precipitation.

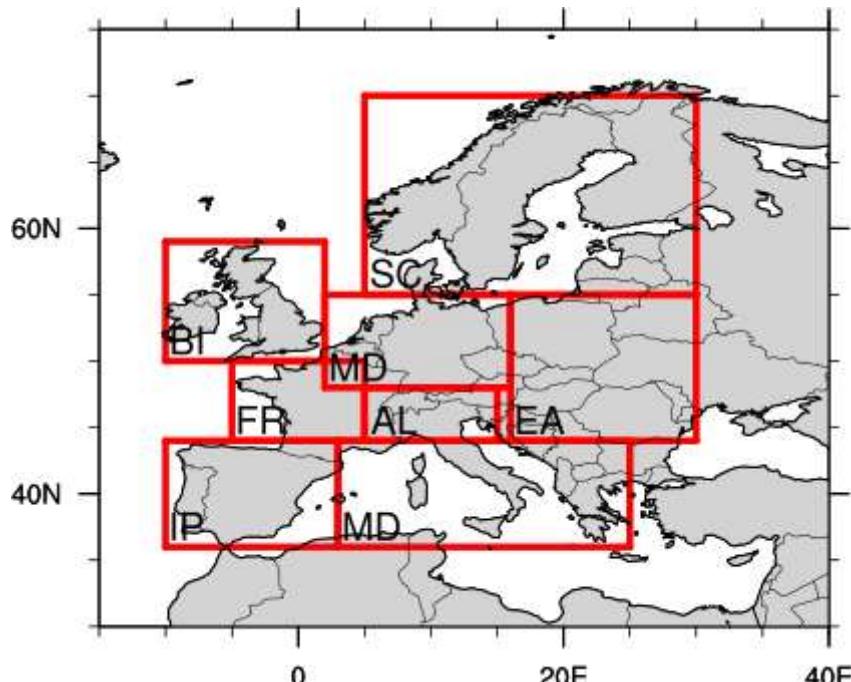


Figure 3.44. The regions for which precipitation data is analysed: Scandinavia (SC), British Isles (BI), Mid-Europe (MD), France (FR), The Alps (AL), Eastern Europe (EA), Iberian Peninsula (IP) and the Mediterranean (MD).

The result of the AsoP analysis is a distribution for each model showing the probability of different precipitation intensities. The distributions of all analysed models are compared with

see how changing the resolution affects different parts of the distribution; i.e., if low and high precipitation intensities change in different ways.

In addition to that a number of precipitation indices are calculated for the same regions. All daily precipitations for all grid points in a region are used to calculate the indices so that each grid point gets one value of the index representing the time period. These values are then pooled to calculate percentiles representing the region for each model. These percentiles are used in the box plots (Figures 3.47-3.50).

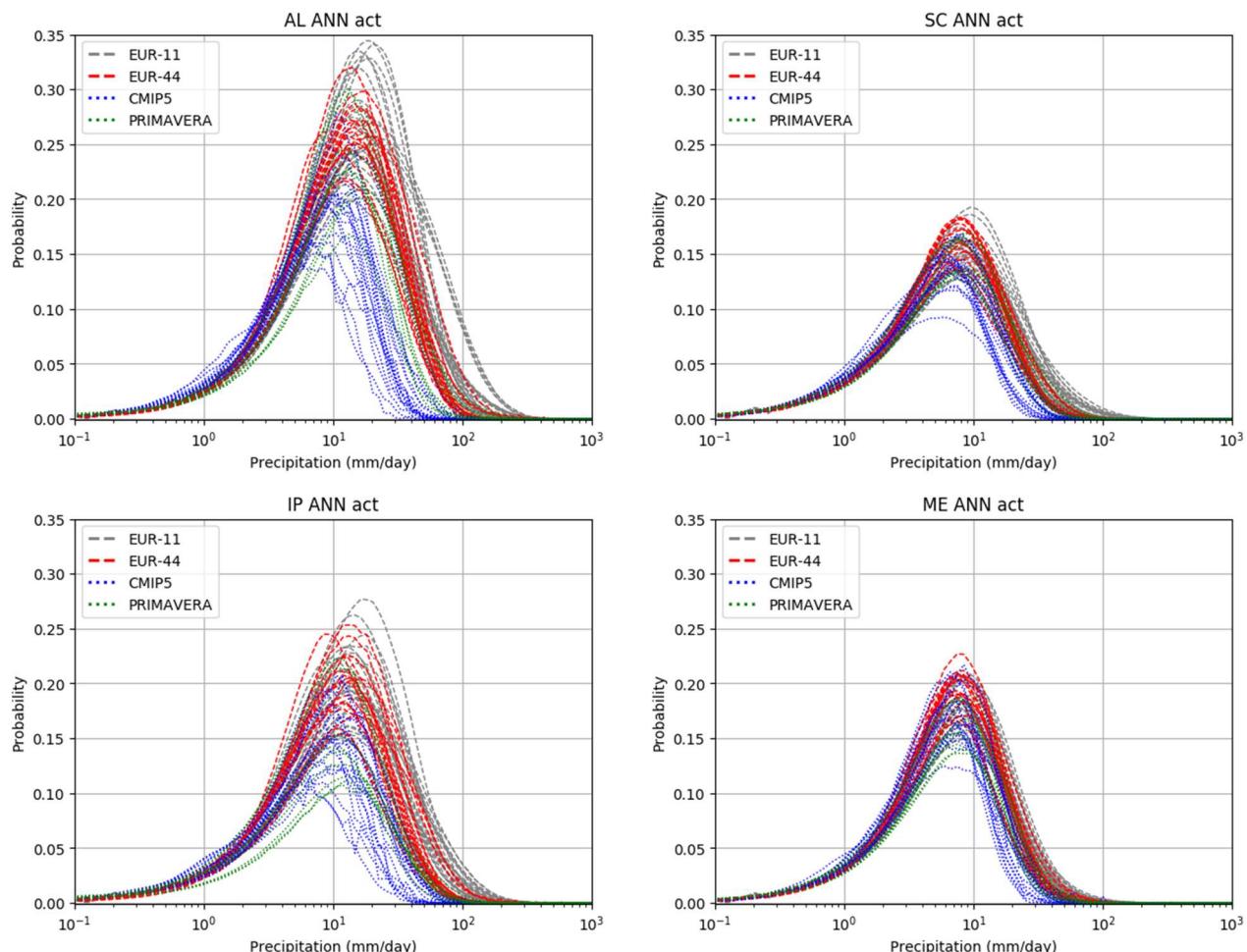


Figure 3.45. Probability of different precipitation intensities in the CMIP5 (blue dotted lines), PRIMAVERA (green dotted lines), CORDEX low resolution (red dashed lines) and CORDEX high resolution (grey dashed lines) ensembles for the Alps (AL, top left), Scandinavia (SC, top right), the Iberian Peninsula (IP, bottom left) and mid-Europe (ME, bottom right).

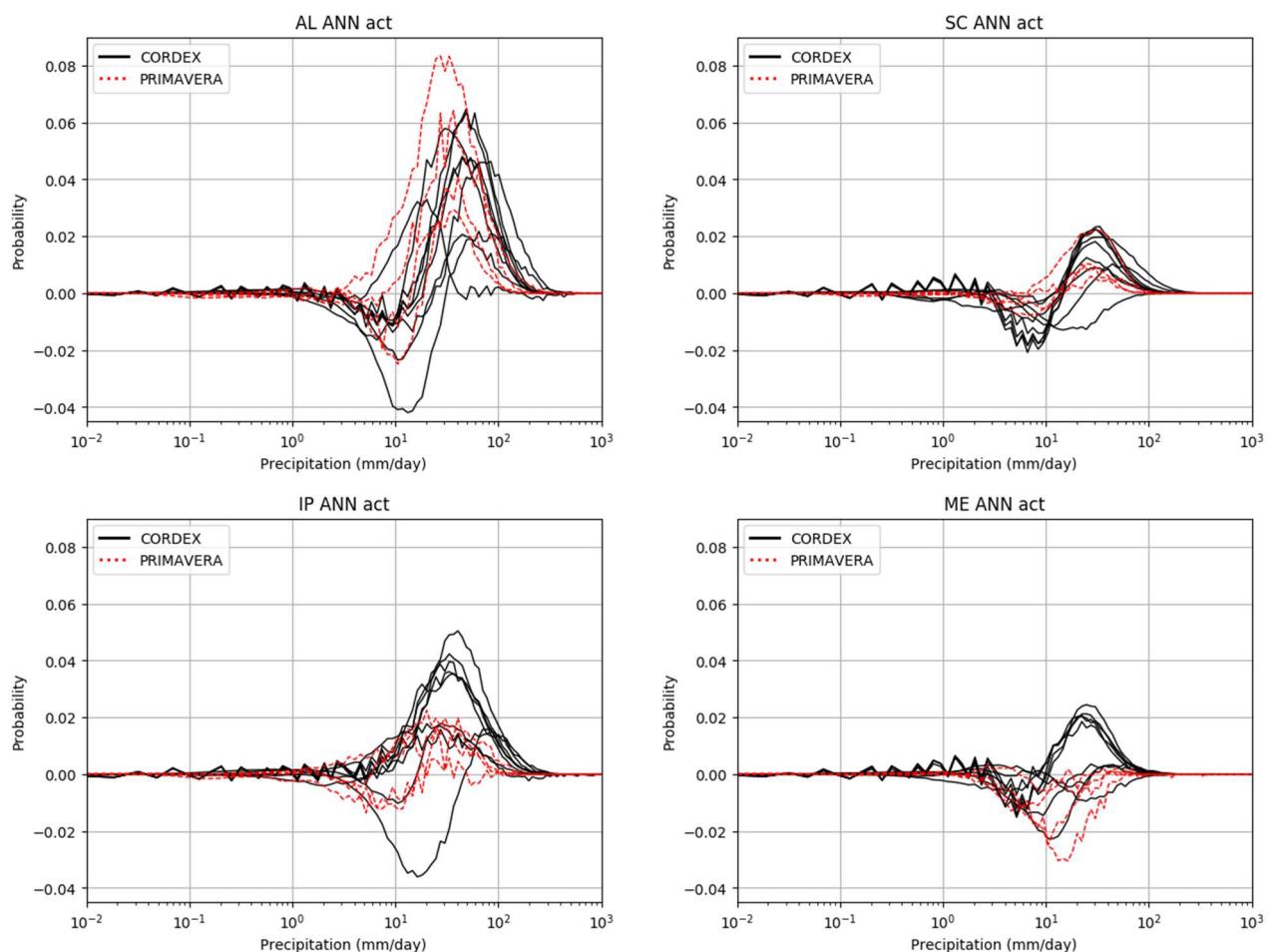


Figure 3.46. Differences in probability of different precipitation intensities within the PRIMAVERA ensemble (red dashed lines) and for selected CORDEX models (black full lines) for the Alps (AL, top left), Scandinavia (SC, top right), the Iberian Peninsula (IP, bottom left) and mid-Europe (ME, bottom right).

3.4.2.3 Results

In general, the difference between the model ensembles increases with higher precipitation intensity, and with the resolution of the model ensemble. For low precipitation intensities the models are rather similar (Figure 3.45). For the high intensity part of the distribution (10-100 mm/day) the CMIP5 models (100-300 km resolution) give weaker precipitation than the PRIMAVERA low resolution (LR) models (80-160 km), which in turn give weaker precipitation than the PRIMAVERA high resolution (HR) models (40-80 km), the CORDEX LR models (50 km) and the CORDEX HR models (12.5 km). The lower resolution models have peaks at lower intensities and shorter distributions, the higher resolution models peak at higher intensities and have wider distributions with more intense extreme precipitation. In the Alps, for example, the upper end of the distribution is around 30 mm/day in some CMIP5 models and 300 mm/day in some CORDEX HR models.

The largest difference between ensembles is seen from CMIP5 to PRIMAVERA; CMIP5 also shows the largest spread within a model ensemble. PRIMAVERA HR and CORDEX LR give comparable distributions as they are of similar resolution. This goes for all regions. There is some tendency for larger differences in areas with complex topography and/or large precipitation amounts (the Alps, the Mediterranean region (not shown), the Iberian Peninsula, Scandinavia, Figure 3.45).

On the model ensemble scale, it is only possible to do a qualitative analysis since resolution is not the only thing that is different between two models; eventual differences could also be explained by differences in the model code, such as differences in parameterizations or process definitions. This means that we can't be sure that any difference comes from higher resolution or from other differences in the model code. For the PRIMAVERA models, however, it's possible to directly compare the low- and high-resolution model versions. For some CORDEX models this is also possible when the high and low version of the RCMs were forced by the same GCM, this is possible for 8 RCM GCM combinations (6 different RCMs driven by 4 different GCMs).

For the PRIMAVERA models the high-resolution version always simulates more precipitation somewhere in the range 10-100 mm/day, the only exception being mid-Europe. For some models this increase in high precipitation intensities is "compensated" by a decrease in lower intensities (up to 10 mm/day), sometimes the increase covers the whole range (Figure 3.46). The differences are generally larger in regions with larger precipitation amounts.

For the CORDEX models there is a clearer "wavy" pattern with a decrease in lower intensities (roughly below 10 mm/day) and an increase in higher intensities (roughly above 10 mm/day) in the high-resolution simulations compared with the low-resolution simulations. At the same time the spread between models in the ensemble is larger in the CORDEX ensemble; some models even give a decrease in all intensities while the PRIMAVERA ensemble is more similar internally. Since the number of compared CORDEX models is larger than the number of compared PRIMAVERA models (8 compared with 4) the spread between CORDEX models is perhaps expected to be larger. It is worth to note that the difference between different RCM simulations, and how they respond to differences in resolution, may very well be explained by the driving GCM and the state of the atmospheric general circulation in them. Higher resolution is expected to give a better described and more detailed climate, with for example deeper cyclones and more intense local showers; in a sense with more pronounced weather events. If two models are in different states, for example when it comes to where storm tracks cross Europe. If these states are pronounced, that may lead to even larger model differences; instead of a weak storm track in the south and a weak storm track in the north in the low resolution model we may now instead have strong storm tracks which means that the difference between the models increases. To fully answer that would require an analysis of the circulation patterns in the different models. This is not done here, but should be a topic for further studies.

When do these intense precipitation events occur in the high-resolution models? Figure 3.47 shows the number of precipitation days. The number of precipitation days does not change much between the model ensembles. There are differences between models, but it is difficult to establish any significant differences between ensembles. This is true both for regions with more precipitation days (e.g. Scandinavia) and regions with less precipitation days (e.g. the Iberian Peninsula). All models show about the same number of precipitation events over the

whole year, which may suggest that the large-scale weather patterns are not influenced that much by higher resolution. Note, however, that the large-scale circulation in the RCMs to a large extent is governed by the driving GCM which have typical resolutions of around 200 km.

The number of days with large precipitation amounts, above 10 mm/day and 20 mm/day, do, on the other hand, get more frequent with higher model resolution. Figure 3.48 shows the number of days with precipitation over 20 mm. The number of days increases from just a few in CMIP5 to 5-10, or even more, in CORDEX HR. The 10th percentile increases about as much as the median, but the 90th percentile increases more. The spread is larger for models with high resolution. This could partly be explained by higher number of data points in the high-resolution models, but a high-resolution model is more likely to catch the variations in precipitation within a region.

The same number of wet days with increased amounts of (extreme) precipitation means that the amount of precipitation on the wet days will increase. This is shown in the simple precipitation intensity index (SDII, the average precipitation amount on wet days, Figure 3.49). SDII does vary with resolution. The wet days are wetter in the high-resolution simulations. High resolution models give higher SDII and larger intra model spread than low resolution models. The higher SDII is a clear feature in high resolution models both in regions with large intra model spread and large differences between high and low resolution as well as in regions with small spread and smaller differences between high and low resolutions.

That extreme precipitation increases is also seen in the maximum one day (Figure 3.50) and maximum five day precipitation (not shown). These extreme events get more intense in the high-resolution simulations. There is a clear increase in both spread and intensities in the high-resolution models.

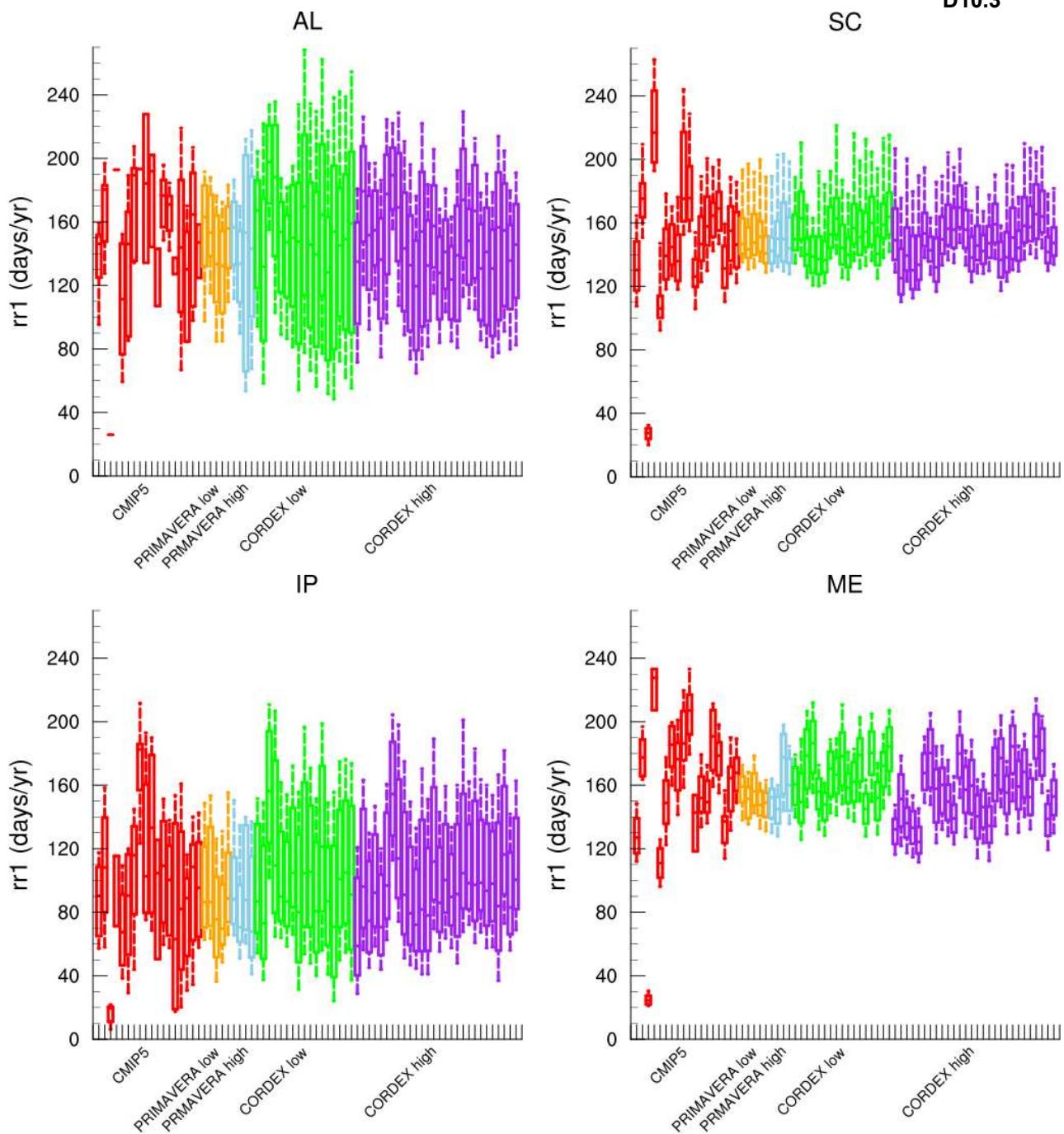


Figure 3.47. Number of precipitation days (rr1 [days/year]) in the Alps (AL, top left), Scandinavia (SC, top right), the Iberian Peninsula (IP, bottom left) and mid-Europe (ME, bottom right) for individual models in the CMIP5 (red), PRIMAVERA LR (orange), PRIMAVERA HR (light blue), CORDEX LR (green) and CORDEX HR (purple) ensembles. Boxes mark the 25th and 75th percentile, with the median inside; whiskers go from the 10th to the 90th percentile.

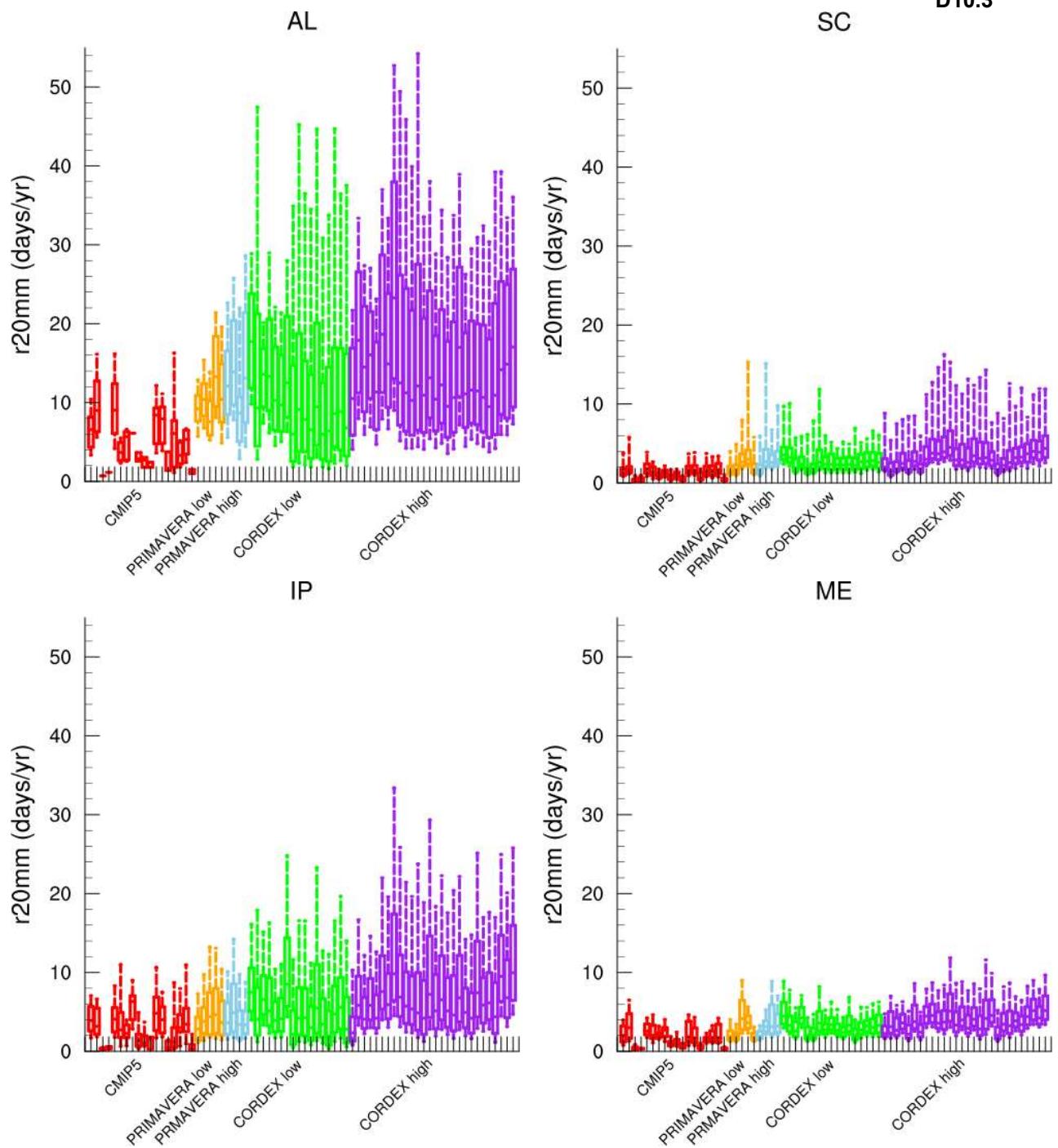


Figure 3.48. Same as Figure 3.47 but for the number of days with precipitation amount over 20 mm (r20mm [days/year]).

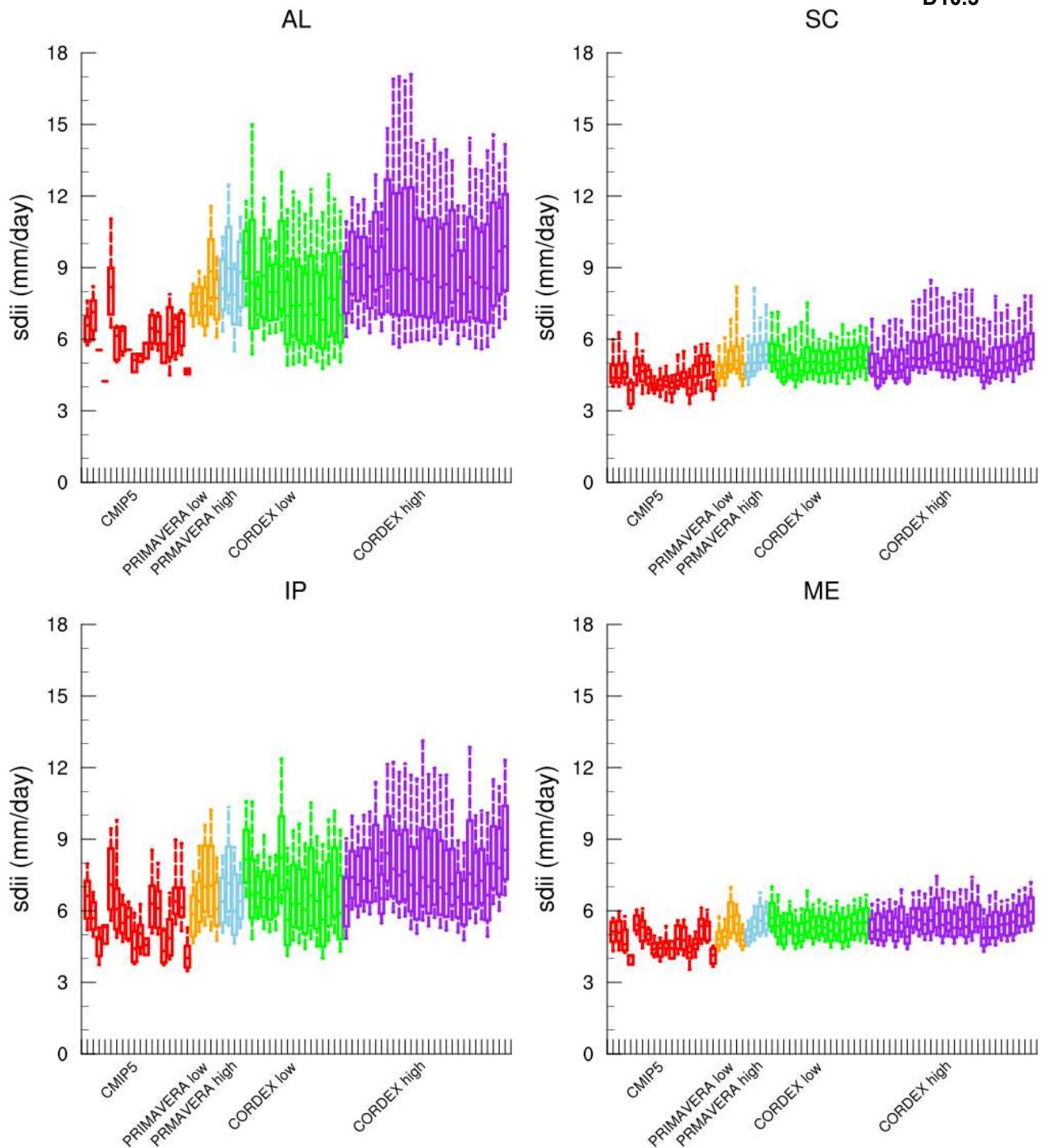


Figure 3.49. Same as Figure 3.47 but for simple precipitation intensity index (SDII [mm/day]).

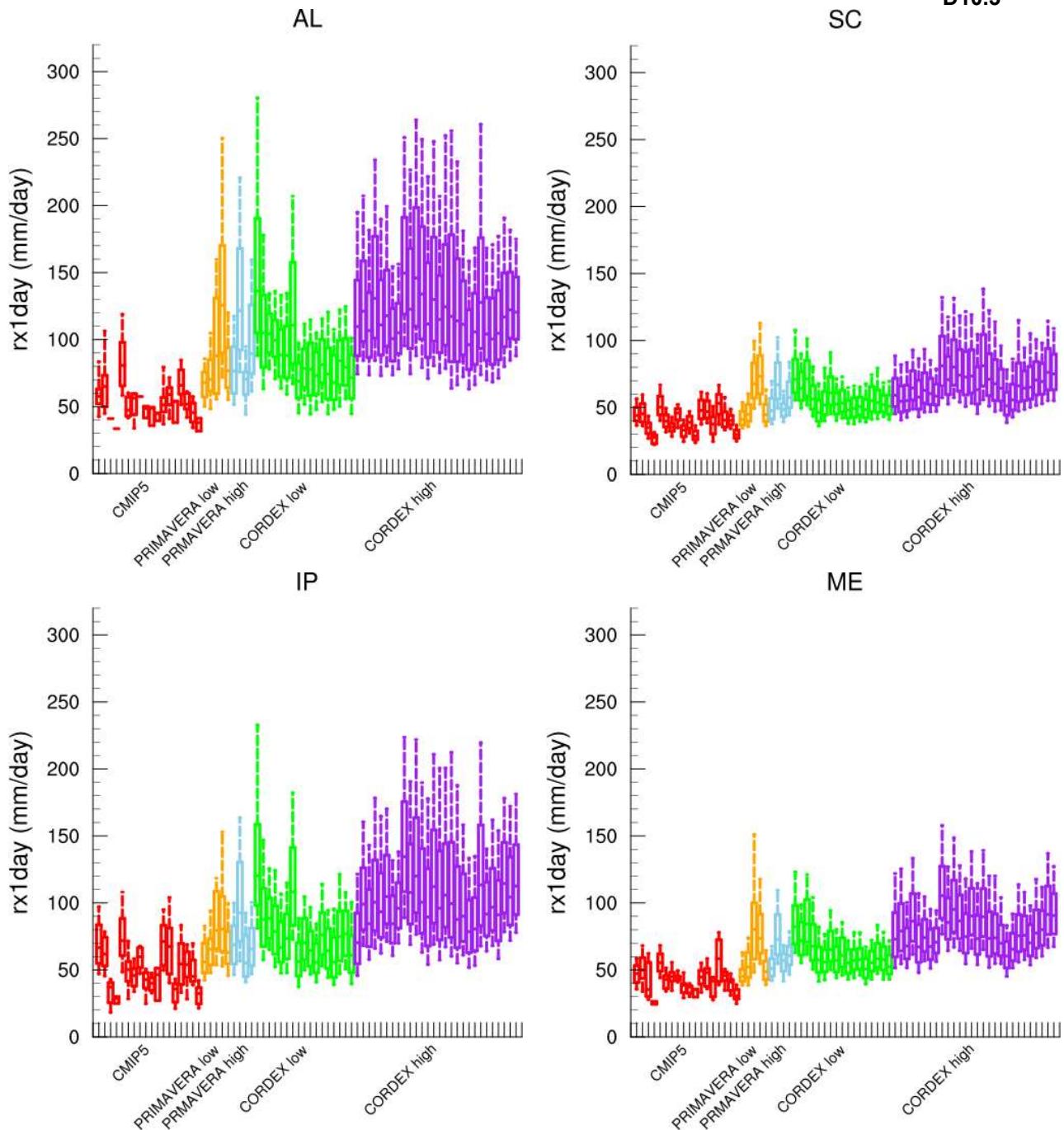


Figure 3.50. Same as Figure 3.47 but for maximum one day precipitation (rx1day [mm/day]).

3.4.2.4 Conclusions

This section investigates the importance of model resolution on the simulated precipitation in Europe. The aim is to investigate the differences between models and model ensembles rather than evaluating models. It is clear that the type of model has a large effect on precipitation, mostly on the more extreme precipitation. The number of precipitation days does not depend much on resolution as this is mostly depending on large scale weather patterns and not so much on local topography and convection, at least not for annual

precipitation. For extreme precipitation events that are more local and short-lived model resolution is more important. A high-resolution model better resolves such events and distinguishes better between different parts of a region. Thus, extreme precipitation is more extreme and more frequent in high resolution models compared with low resolution models. With the same amount of wet days this means that precipitation intensifies so that the wet days are wetter.

The largest effect with increasing precipitation comes at low resolution; increasing the resolution from low to less low has a greater effect than increasing from relatively high to high resolution. This does not, however, mean that increased resolution gets less and less worthwhile, once the models reach convection permitting resolutions it will have a large effect. This is not shown here as the highest resolution in this study is 12.5 km and not the ~3 km that is required to be convection permitting.

The effect of higher resolution is seen in regions with small amounts of precipitation as well as regions with high amounts of precipitation, and in regions with small and large geographical differences. The higher percentiles change more than the low percentiles for all studied indices. Increasing resolution has about the same effect on both GCMs and RCMs, furthermore GCMs and RCMs of comparable resolution simulate comparable precipitation climates.

The results presented here are expected and in line with previous results, but this is the first time it is done across such relatively large model ensembles of different resolutions, and with a method studying all parts of the precipitation distribution.

3.4.3 The impact of atmospheric resolution on simulated extreme European precipitation

Alexander Baker, Reinhard Schiemann and Pier Luigi Vidale (UREAD)

3.4.3.1 Aim

We evaluated the impact of increased horizontal atmospheric resolution on extreme daily precipitation across the Stream 1 ensemble of atmosphere-land only and fully coupled simulations. Here, we show the added value of increased atmospheric resolution for winter (DJF) precipitation over Europe and the North Atlantic.

3.4.3.2 Methodology

We employed generalised extreme value (GEV) analysis and applied the parametric block maxima method globally. At each model grid point, globally, 1-day precipitation maxima were computed for each canonical season. GEV distributions were fitted to these seasonal precipitation block maxima time series, described by the location (μ), scale (σ) and shape (ξ) parameters, which determine the change in return value as a function of return period. Here,

we focus on two quantities: μ determines the vertical position of the GEV curve and thereby ‘typical’ return values, σ determines the slope of the GEV curve and thereby the year-to-year variability in extremes, and ξ determines the curvature of the distribution and thereby whether or not an upper bound exists. We show results for highresSST-present simulations for winter (DJF), but all seasons were analysed. An example of the application of GEV analysis to global climate model integrations is given in Schiemann et al. (2018).

3.4.3.3 Key results

For winter, PRIMAVERA models underestimate μ over land, the Mediterranean, and northeast North Atlantic (Figure 3.51). Increasing resolution increases μ across the midlatitudes in all models (Figure 3.52a). Increased extremes are simulated over much of the North Atlantic, particularly the storm track region in winter (and the equinoctial seasons – not shown). Simulated μ is closer to observational Global Precipitation Climatology Project data (Huffman et al. 2001) over this region (Figure 3.52b). However, simulated σ is further from observational estimates, indicating that typical return values are better simulated in high-resolution forced simulations than inter-seasonal variability (Figure 3.53). Importantly, increased extreme precipitation is coterminous with reduced error over the north-eastern North Atlantic, Mediterranean and European orographic regions, exhibiting the added value of high-resolution integrations across much of the Euro-Atlantic domain of immediate interest to PRIMAVERA partners and stakeholders.

3.4.3.4 Forthcoming research

- Assess observational uncertainty in GEV parameters over land using national datasets.
- Analysis of hist-1950 coupled simulations plots.
- Link GEV evaluation to analyses of ETC activity and associated precipitation as well as North Atlantic eddy-driven jet variability (CMCC collaboration).

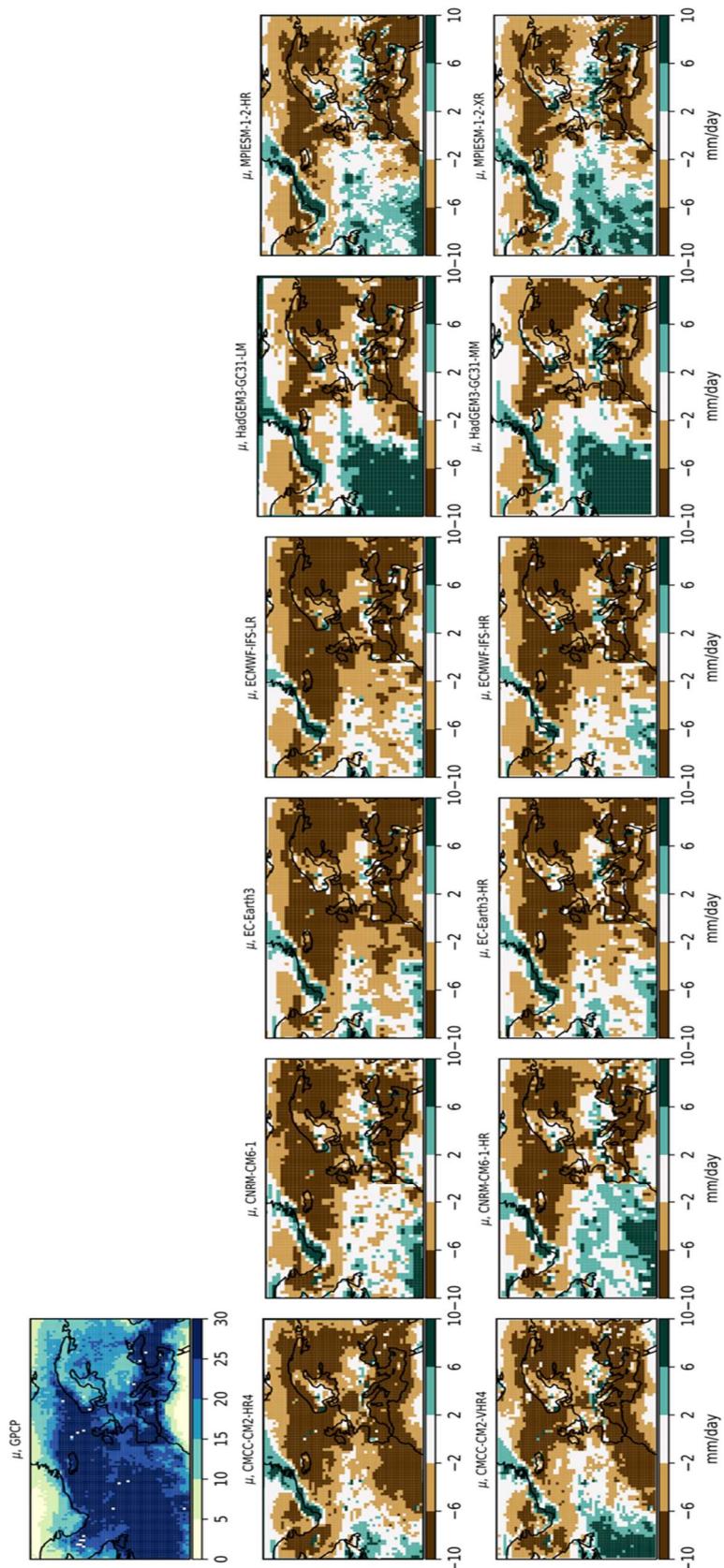


Figure 3.51. Wintertime mean bias (model-observation) in μ for lowest (middle row) and highest (bottom row) available resolution. Biases are computed versus GPCP daily, gridded (1°) precipitation data, available for 1996-2013 (Huffman et al., 2001).

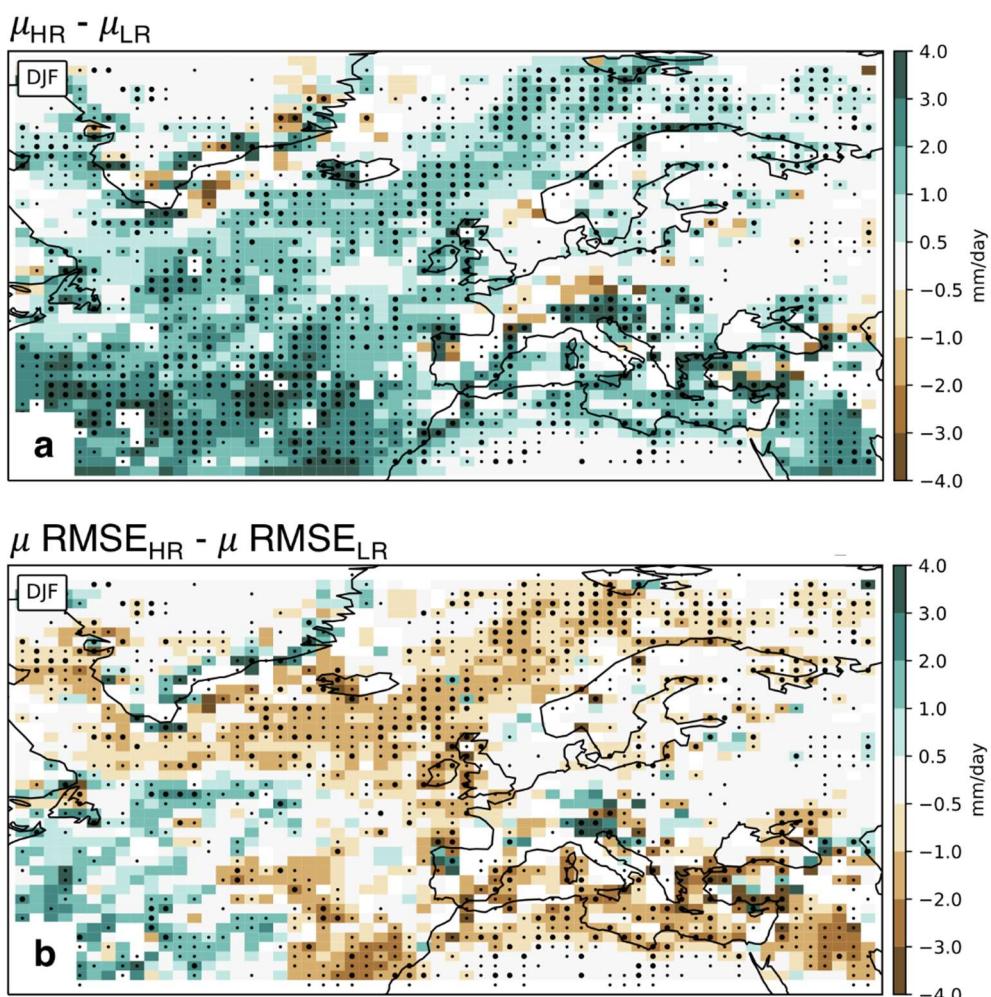


Figure 3.52. Multi-model wintertime mean difference in (a) μ (high-resolution minus low-resolution) and (b) μ root-mean-square error (RMSE). Positive (negative) values indicate increased (decreased) μ or μ RMSE at high-resolution. Large (small) stippling indicates all six (five out of six) models agree on sign of μ change with resolution increase. RMSE is computed versus GPCP data.

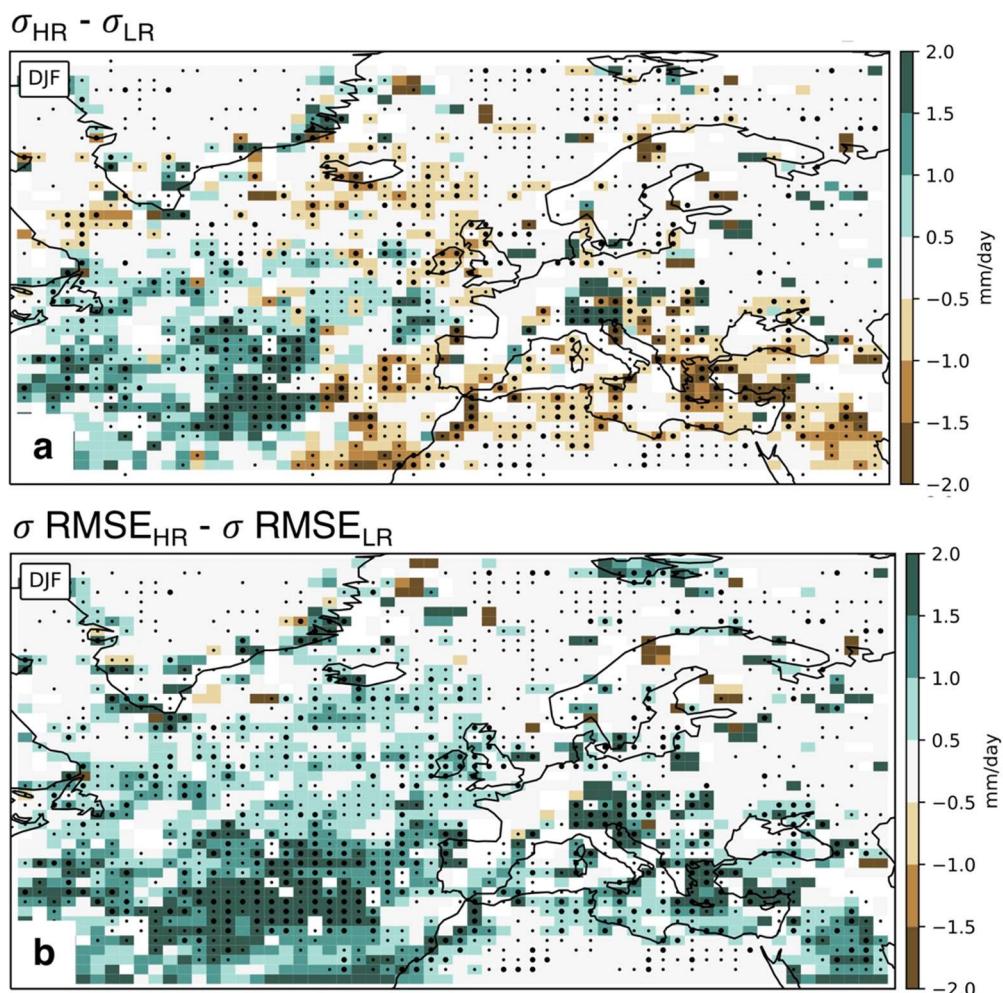


Figure 3.53. As Figure 3.52 for σ .

3.5 Potential emerging hazards for Europe

3.5.1 Post-tropical cyclones impacting Europe

Alex Baker (UREAD) and Rein Haarsma (KNMI)

3.5.1.1 Introduction

Post-tropical cyclones (PTCs) are an important midlatitude natural hazard across the North Atlantic (Dekker et al. 2018; Evans et al. 2017; Jones et al. 2003; Keller et al. 2019). The occurrence of such storms exposes populous midlatitude regions, such as Europe, to hurricane-force wind speeds and extreme precipitation. A recent, high-impact example is ex-hurricane Ophelia (9th–15th October, 2017), which caused loss of life as well as severe damage across Ireland, the UK, and Scandinavia (Stewart 2018). Ophelia possessed tropical-cyclone-like structural characteristics at or near landfall, and significant human and economic impacts occurred during its post-tropical storm phase.

PRIMAVERA supported an observational study into PTCs, based on reanalyses and National Hurricane Centre (NHC) best-track dataset—HURDAT2. This section is currently undergoing peer review at *Geophysical Research Letters*. Two key foci in this section are quantifying historical variability of European PTCs and their structural characteristics at the point of landfall over Europe. A follow-up study into how well represented PTCs are in PRIMAVERA highresSST-present (atmosphere-land-only) simulations is reported here, which also focusses on interannual variability and landfall characteristics. A second manuscript describing this follow-up work is forthcoming.

3.5.1.2 Methodology

3.5.1.2.1 Tropical cyclone tracking

To identify and track the evolution of PTCs in reanalysis and PRIMAVERA data (Table 3.1, column 2) in this study, the objective feature-tracking algorithm—*TRACK*—of Hodges (1995) is used. *TRACK* was used by Hodges et al. (2017) to evaluate the representation of tropical cyclones in reanalyses (except ERA5). The *TRACK* algorithm was applied to six-hourly, spectrally filtered and vertically averaged vorticity at the 850, 700 and 600 hPa isobaric levels, which was computed from the zonal and meridional wind fields at those levels. Spectral filtering (i.e., truncation to T63 resolution) was performed to remove large-scale, planetary motion (total wavenumbers 0-5) and small-scale noise (total wavenumbers >63). Vorticity maxima exceeding $5 \times 10^{-6} \text{ s}^{-1}$ were identified and formed into tracks using a nearest-neighbour approach and subsequently refined by minimising a cost function for track smoothness, subject to adaptive constraints on track displacement and smoothness (Hodges 1995, 1999). Following Hodges et al. (2017), tropical cyclones were considered to be the identified vorticity features (i) whose genesis occurs equatorward of 30°N , (ii) have a vorticity decrease with increasing height between 850 and 250 hPa that exceeds $5 \times 10^{-6} \text{ s}^{-1}$ for at least 1 day over ocean (i.e., a warm-core), and (iii) whose total lifetime exceeds 2 days. This method is consistent with published studies of tropical cyclones in climate models (e.g., Roberts et al., 2019b).

3.5.1.2.2 Post-tropical cyclone filtering

The tropical cyclone identification criteria described above minimise inclusion of spurious short-lived or relatively weak vorticity features, thereby allowing geographical identification of post-tropical cyclones. From objectively identified tropical cyclones, post-tropical cyclone tracks are those systems which pass within Europe ($36\text{-}70^\circ\text{N}$, $10^\circ\text{W}\text{-}30^\circ\text{E}$). We also applied this geographical filtering criterion to HURDAT2 data.

3.5.1.2.3 Cyclone phase-space analysis

Hart (2003) devised a phase-space analysis to describe cyclone core structure and its temporal evolution, which involves three parameters: the thermal symmetry of the cyclone (B) and the lower- (T_L) and upper-tropospheric cyclone thermal wind (T_U). This is used to

classify storms according to whether they possess tropical, extratropical or hybrid structural characteristics.

B is defined as:

$$B = h \left(\overline{Z_{600} - Z_{900}}|_R - \overline{Z_{600} - Z_{900}}|_L \right)$$

where $h = 1$ for the Northern Hemisphere, Z_p is geopotential height (m) at isobaric level p (hPa), and the subscripts R and L denote the right- and left-hand semicircles, respectively, relative to the cyclone's propagation direction. Following previous research (Dekker et al. 2018; Hart 2003), we defined thermally symmetrical cyclones (i.e., non-frontal) as those with near-zero B values and asymmetrical (i.e., frontal) systems as those whose $B \geq 10$ m. The lower- (900-600 hPa) and upper-tropospheric (600-300 hPa) thermal wind, T_L and T_U , respectively, are defined as vertical derivatives of the horizontal geopotential height gradient:

$$T_L \equiv -|V_T^L| = \frac{\partial(\Delta Z)}{\partial \ln p} \Big|_{P_l}^{P_u}$$

$$T_U \equiv -|V_T^U| = \frac{\partial(\Delta Z)}{\partial \ln p} \Big|_{P_l}^{P_u}$$

where p is pressure and $\Delta Z = Z_{max} - Z_{min}$, where Z_{max} and Z_{min} are the maximum and minimum geopotential height, respectively, at a given isobaric level within a 5° radius of the cyclone centre. The slope of the regression between ΔZ and $\ln p$ is used as the derivative of ΔZ relative to $\ln p$ to determine the mean ΔZ over the pressure range P_l to P_u – the lower and upper isobaric levels, respectively. The isobaric levels 900, 600 and 300 hPa were used, except for JRA-25, where 925, 600 and 300 hPa were used. Positive (negative) T_L or T_U indicate the presence of a warm- (cold-)core in the upper or lower troposphere. A deep warm- or cold-core structure is defined as occurring at both the lower and upper levels. As such, cyclone-phase-space analysis is a crucial tool for the classification of PTCs and their evolution.

Note that the CMCC-CM2 model is absent from this analysis because the six-hourly geopotential height data output on pressure levels is not available.

3.5.1.2.4 Cyclone lifecycle analysis

Post-tropical cyclone structures were classified using the phase-space parameters determined at the point of landfall—defined as entering the Europe landfall domain. To construct composite lifecycles, storms within each phase-space category were aligned by their landfall point (i.e., $t = 0$) and an average v_{max} timeseries computed. Averaging was performed for every landfall-normalised timestep (i.e., negative (positive) denoting before (after) landfall), where the number of cyclones exceeded 5. We employed this criterion to maximise sampling of pre- and post-landfall intensity evolution. Again, CMCC-CM2 is excluded.

3.5.1.3 Key results

3.5.1.3.1 Interannual variability

Interannual variability in PTC frequency is not well captured across the ensemble of highresSST-present simulations (Figure 3.54), with some models overestimating (e.g., CMCC-CM2, CNRM-CM6) and others underestimating (e.g., MPI-M) annual counts. Three models—EC-Earth3, ECMWF-IFS and HadGEM3-GC3.1—simulate similar average annual counts since 1950 to HURDAT2 observations. However, no systematic dependence of PTC counts on resolution across models is seen (Figure 3.54).

3.5.1.3.2 Post-tropical cyclone structures

The proportion of each phase-space structural type, categorised at landfall, is highly consistent across reanalyses: roughly 60–70 % of landfalling PTCs undergo transition to frontal extratropical cyclones prior to reaching Europe, but—crucially—approximately 10–20 % make landfall with warm-core characteristics, of which approximately half make landfall as an axisymmetric tropical cyclone (Baker et al., 2019b). The retention of tropical structural characteristics at least until landfall is associated with greater impacts than cold-core landfalling PTCs. There is much greater spread in these proportions among PRIMAVERA models (Figure 3.55). Additionally, most models simulate an increased proportion of warm-core landfalls at high resolution, compared with low resolution.

3.5.1.3.3 Composite lifecycles

Reanalyses consistently show warm-core landfalls are more intense than cold-core landfalls, as measured by v_{max} (Baker et al., 2019b). However, this distinction is not reproduced all PRIMAVERA models (Figure 3.56, rows 1 and 3). Only ECMWF-IFS and MPI-M (at both low- and high-resolution) and CNRM-CM6 and EC-Earth3 (at high-resolution) show this behaviour. HadGEM3-GC3.1 simulates similar landfall intensities between warm- and cold-core PTCs at both low- and high-resolution. There appears to be little systematic sensitivity to resolution in these composite PTC lifecycles. However, the number of warm-core systems increases with resolution across models (Figure 3.56, rows 2 and 4).

3.5.1.4 Forthcoming work

- Analysis of coupled simulations, focussing on SST and OHC conditions, as well as large-scale modes of atmospheric circulation variability, favourable to PTC activity.

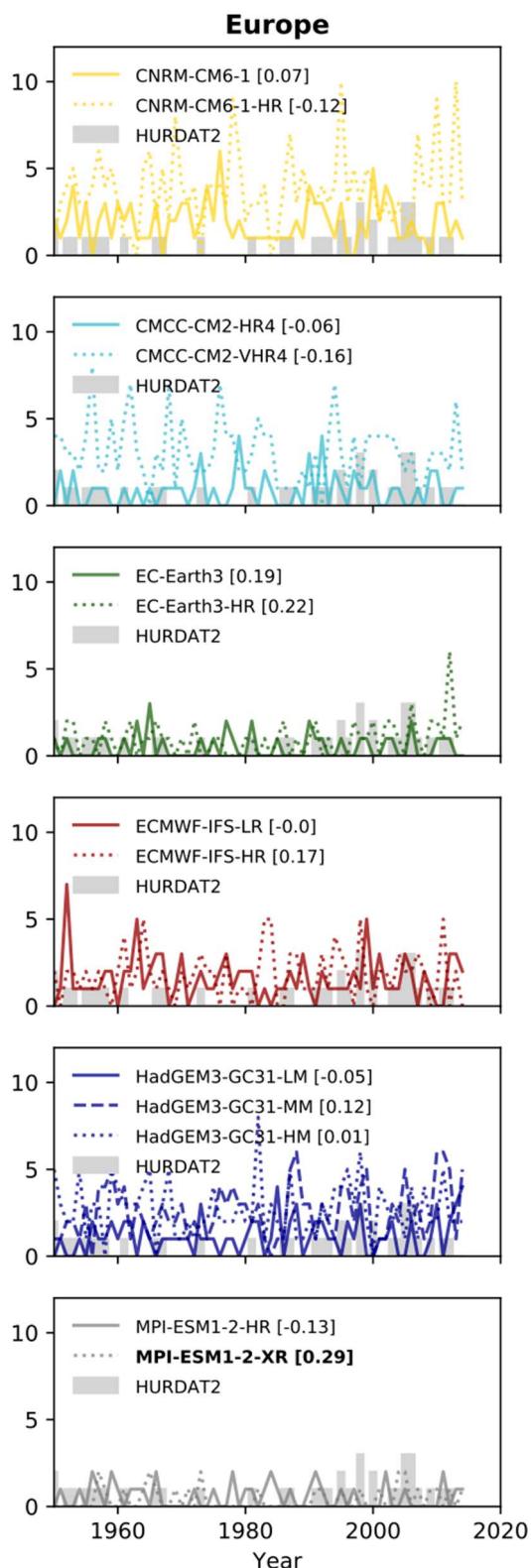


Figure 3.54. Interannual variability in simulated PTC count [cyclones yr⁻¹] for Europe simulated by PRIMAVERA models under highresSST-present forcing compared with HURDAT2 (grey bars). Solid and dashed lines indicate low- and high-resolution simulations, respectively. Spearman's correlation coefficients are given in the legends, with significance (i.e., $p \leq 0.05$) indicated by bold type.

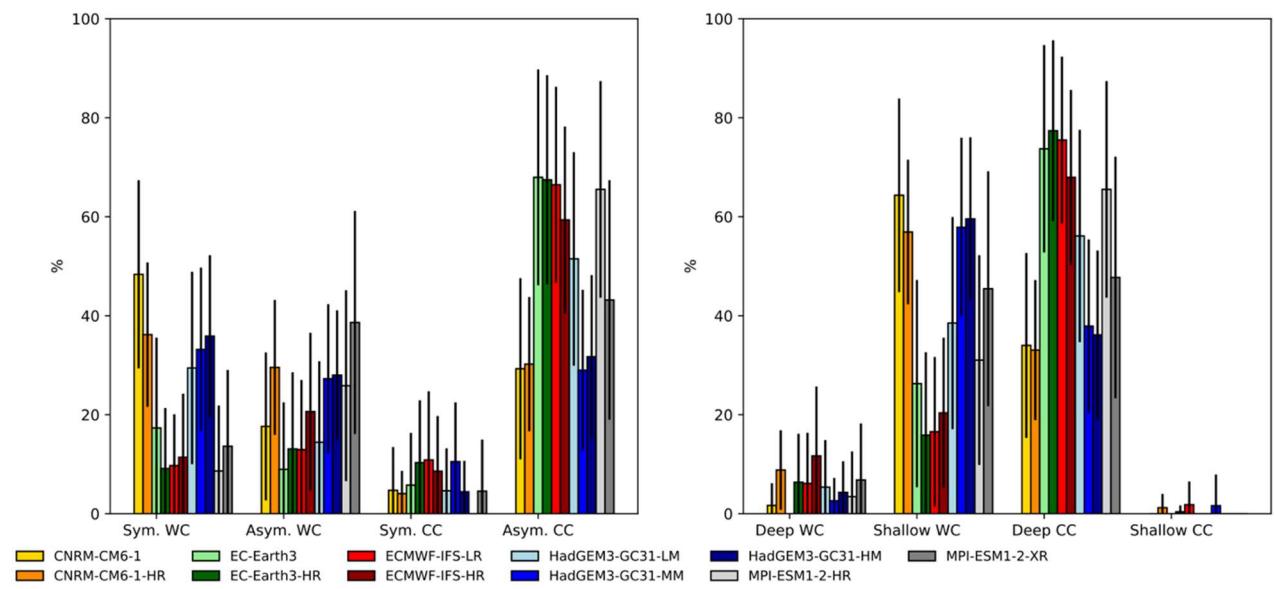


Figure 3.55. Relative proportions of PTC structures at time of landfall over Europe. Symmetrical (Sym.) / asymmetrical (Asym.) warm-core (WC) / cold-core (CC). Lighter colours represent low-resolution simulations; darker colours represent high-resolution simulations. Structure type defined by (left) B and T_L and (right) T_L and T_U .

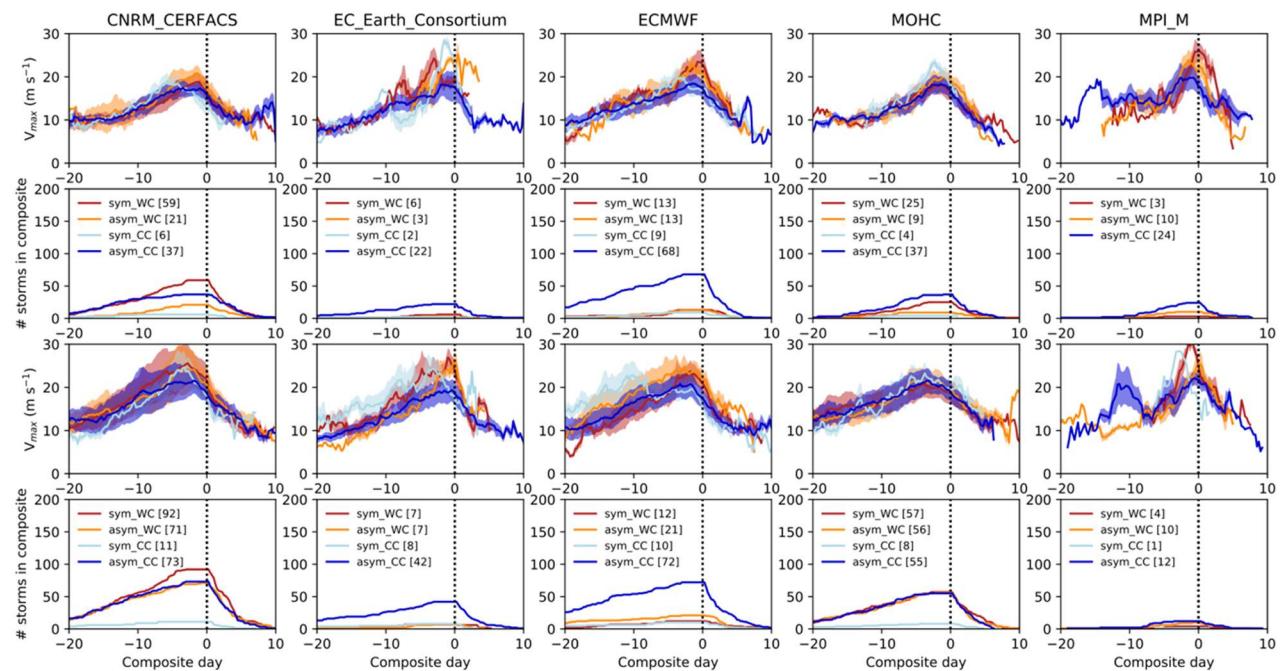


Figure 3.56. Landfall-centred v_{max} composites lifecycles. Row 1 & 2 are low-resolution lifecycles and number of cyclones in composite, respectively. Row 3 & 4 are high-resolution lifecycles and number of cyclones in composite, respectively. Warm (cold) colours indicate cyclones classified as warm- (cold-) core at the point of landfall.

3.6 Temperature extremes

3.6.1 Temperature trends over Europe

Gerard van der Schrier (KNMI)

3.6.1.1 Introduction

A change in the frequency and intensity of heat extremes strongly impacts society and the economy, as these events have serious effects on, in particular, agriculture, energy demand, transportation, and health. For this reason, prediction of future climate that realistically assesses changes in the evolution of extreme events is fundamental to understand the challenges that need to be met. Numerous studies have shown that their nature, scale and frequency is changing and will change further due to climate change (IPCC 2012). In particular, heat waves and high temperatures are shown to increase significantly in frequency and severity, also in Europe.

A reproduction of historic trends in temperature and temperature extremes adds to the confidence in the models that are the basis for climate scenarios. Testing for differences in temperature trends is done by making simulations which have as input verified historical observed boundary conditions (e.g. land use) and observed forcings (greenhouse gases and aerosol concentration, volcanic eruptions and solar radiation).

Earlier work (e.g. Min et al. 2013) studied trends in the hottest day of the year (i.e., the annual maximum of the daily maximum 2 m temperature Tmax), as produced by a large ensemble of RCMs run for the historical period 1961–2000. In that study, a strong and significant trend was found in observations over this period. However, the ensemble median of the 14 RCM ensemble under study strongly underestimates this trend. Despite a large inter-model variability, over extended areas not a single RCM could match the trends in observations. The aim of this section is to assess whether current-generation PRIMAVERA models improve on the poor representation of trends in temperature extremes, as found by Min et al. (2013).

There are important shortcomings issues with the Min et al. (2013) study, which need to be avoided in the current study. These issues relate to the observational dataset and to the metric of extreme temperature. Although the observational dataset used by Min et al. (2013) (E-OBSv6.0) is based on a large number of meteorological stations in Europe, time series from these stations are not homogeneous. Although there is a check on homogeneity in the underlying station dataset (ECA&D), all available series are used in the construction of the E-OBS—homogeneous or not. This version of the E-OBS dataset is therefore not well suited to trend calculations. Additionally, Min et al. (2013) used calculated trends in the hottest day of the year—a metric that is very sensitive to local conditions, meaning that the observed maximum temperature may not be representative for a larger area.

In this study, a comparison between trends in observations and models is assessed using a new version of the E-OBS (E-OBSv19.0HOM) which is based on a dataset of homogenized temperature records of ECA&D. The metric used to assess trends in temperature extremes is the trend in the number of warm daytimes (TX90p). This metric uses the 90th percentile of daily maximum temperatures (over the 1981–2010 period) as a threshold to define a warm daytime. Similarly, trends in cold extremes are assessed using TN10p, the number of cold nights. The benefit of these less-extreme metrics for temperature extremes (in comparison to the hottest day of the year/coldest night in the year) is that the percentile-based metrics are much more robust and much less dependent on local conditions.

3.6.1.2 Observational dataset

The reference used for the evaluation of the models is the E-OBS for TN and TX (Haylock et al. 2008, Cornes et al. 2018). It comes as a 100-member ensemble, whose spread increases in areas with low station density, indicating a larger uncertainty. In this work, only the ensemble mean is considered. E-OBS is based on the station data of the European Climate Assessment & Dataset (ECA&D) (Klein Tank et al. 2002), which collects data of thirteen variables from more than 19000 stations located in all countries of the European and Mediterranean region. Almost 10000 of these stations provide temperature data. These are provided by national meteorological and hydrological Services, universities or private companies, and range from late 18th century to present. However, relocation of stations, instrumentation changes and variations in the surroundings of the meteorological stations affect the quality of ECA&D temperature temporal series related to such stations (and therefore E-OBS), reducing the reliability for temporal analyses.

For this analysis, a modified version of E-OBS is constructed based on recent work on the homogenization of the temperature series of ECA&D (Squintu et al. 2019, 2020). This work has removed a large part of the inhomogeneities and makes it possible to smoothly combine series that belong to neighbouring stations, gathering data into one long-running homogeneous series, called blended series. This considerably improves the input data for E-OBS. More details on the homogenization, assessments of the effects of the homogenization and results on trend analysis based on the homogenized data of ECA&D can be found in Squintu et al. (2019, 2020).

3.6.1.3 Models

Six models have been analysed in both low- and high resolution (LR/HR) version, focusing on the period from 1970–2014 and considering the region enclosed between 22°W and 50°E and 20°N and 76°N. The variables considered in this study are minimum temperature (TN) and maximum temperatures (TX) on a daily resolution. Each model taking part in PRIMAVERA has contributed with several experiments, the one that has been used for this work is named ‘highresSST-present’. This consists of an atmosphere-land-only integration, forced with observed SST, observed sea-ice concentration and external radiative forcings over the period 1950–2014. Each model has a different spatial resolution and a different number of ensemble members. Table 3.2 summarises the characteristics of the used models and the availability of the ensemble members (as at September, 2019).

The ECMWF model has native resolution Tco399 (± 25 km) for HR and Tco199 (± 50 km). In the frame of PRIMAVERA they have been provided in a regressed version, respectively to 0.25- and 0.5-degree constant latitude-longitude regular grids. The Ec-Earth3P model runs at the resolution TL511 for HR and TL255 for LR on a non-regular latitude-longitude grid, which has been regressed to a regular longitude-latitude grid.

3.6.1.4 Results

3.6.1.4.1 Mean bias in the climate models

a) Bias in winter averages

The considered HR models show strong differences in the reproduction of TNavg-DJF (Figure 3.57). The largest mean bias is found for CMCC (+2.96°C), while Ec-Earth3, ECMWF and HadGEM3 underestimate on average the minimum temperatures, with a common exception on Northern Scandinavia, far from the coast. At the same time, an underestimation of winter TN over Italy and Norway is found in 5 models. MPI and CNRM perform best in terms of mean biases and present considerably lower extension of the shaded area. These are present when the simulated TNavg-DJF is significantly different from the observed one (i.e., absence of overlap between the 95% confidence interval of the two terms of the difference).

b) Bias in summer averages

The comparison of summer maximum temperatures has started from the evaluation of the bias of the models. Four of them give a mean bias that has a lower absolute value compared with what is observed for winter TN (Figure 3.58). CMCC presents a large underestimation (-3.83°C), similar, but with opposite sign, to the corresponding result for TN. The remaining 5 models show a common north-south gradient, with a warm bias along the Mediterranean and Black Sea coast and a general underestimation over Northern Europe, with different patterns and small exceptions. A large overestimation common to all models is found in Northern African regions. Nevertheless, in these areas the large biases (above +10°C) can be in part related to the high uncertainty of E-OBS, due to a lower station density.

3.6.1.4.2 Trends in mean maximum and minimum temperature

a) Trends in winter averages

Trends on the TNavg-DJF of the models in the 1970-2014 period are compared against the same indices of E-OBS. All models reproduce very well the trends in winter TN. The mean trend biases (Figure 3.59) range between -0.16°C/decade (CNRM) and +0.02°C/dec (ECMWF). This indicates a tendency in simulating lower trends over the continent but mainly for Eastern Europe. Nevertheless, recurring positive biases are found over the Kola Peninsula (North Western Russia, 6 models out of 6), in the Balkans (5/6) and along the European coast of the Western Mediterranean (4/6).

b) Trends in summer averages

The difference in trends of TXavg-JJA ranges between $-0.17^{\circ}\text{C}/\text{dec}$ (ECMWF) and $+0.03^{\circ}\text{C}/\text{dec}$ (CMCC)—see Figure 3.60). The models tend to slightly underestimate the warming of summer temperatures. This is more evident over the Mediterranean and Eastern Europe, especially for models as EC-Earth and ECMWF. These present large areas (over Italy and Balkans) where the differences are significant, implying an inaccurate reproduction of the changes in the climate of these areas. On the other side almost all the models tend to overestimate the trends over Southern Scandinavia, especially over Norway.

3.6.1.4.3 Trends in temperature extremes

a) Trends on cold extremes

Trends in winter cold extremes as TN10p-DJF are more challenging and Figure 3.61 shows these trends for the HR version of the models. While HadGEM3 (mean trend bias: $-1.25\%/\text{dec}$) and, less strongly, ECMWF (mean trend bias: $-0.65\%/\text{dec}$) simulates a lower trend of number of days below the 10th percentile almost over the whole continent, thus having warmer trends than observed, CMCC presents a large contrast between Western and Eastern Europe. In particular a wide area over Iberia, Southern France and the Alps is found where the differences are significant and exceed $-4\%/\text{dec}$, indicating a poor representation of the trends in these areas. The overestimation of the trends (colder trends than what is observed) over Eastern Europe is common to 4 models, while warmer trends over the Mediterranean area are simulated by all models but MPI. MPI is the only model whose mean bias is not significantly negative and does not present pronounced patterns, with the exception of having a too strong warming in TN10p-DJF over Sweden and Norway, in common with three other models.

The combination of the tendency to warmer simulated trends in TN10p-DJF together with the fair representation of trends in average values, imply that these HR models have simulated a winter climate with similar average characteristics but fewer cold events, indicating a narrowing daily minimum temperature distribution. A similar analysis for the LR models gives that they show similar patterns in the trend biases as HR models (not shown).

Figure 3.62 presents the difference in absolute trends biases of TN10p-DJF between HR and LR. Negative values (green) indicate that HR has lower absolute trend bias than LR for that specific grid-point, thus it is performing better. Only CMCC clearly shows an area with worse absolute trend biases over Central Europe (where very large trends are simulated), which contrasts with the strong performance of the same model over Eastern Europe. Despite of this, the mean absolute trend biases over the whole continent are reduced for almost all the models, indicating a general improvement in the description of the cold extremes between low and high resolution. The best improvement is found for HadGEM3 ($-0.51\%/\text{dec}$), while the only worsening, out of the considered models, is for ECMWF ($+0.17\%/\text{dec}$) whose LR version is found to perform the best among the others. The model with the lowest mean absolute trend bias in high resolution is MPI ($0.61\%/\text{dec}$).

3.6.1.4.4 Trends in warm extremes

Figure 3.63 evaluates the reproduction of TX90p-JJA, which describes warm temperature extremes. The results show a large underestimation of the trends for EC-Earth(-0.73%/dec), ECMWF (-0.59%/dec) and HadGEM3 (-0.56%/dec). In all cases stronger trends, consistent with what found for the trends on the averages, are simulated over Norway and Sweden. The overall bias in the MPI model is very small, but the underestimation of trends in southeastern Europe apparently compensates the overestimation of trends in northwestern Europe. This aspect (as found for the simple seasonal averages as well) is simulated by most of the models, indicating a general tendency to reproduce lower trends of warm extremes on the Mediterranean and Black Sea region and slightly larger ones around Northern Sea. In these areas large significant differences are found in particular for EC-Earth3, ECMWF and HadGEM3.

Figure 3.64, showing the difference in absolute trend biases between the HR and LR model configuration, does not show a common pattern. Best improvement in the passage from HR to LR is for MPI (-0.16%/dec), while HadGEM3 presents the largest worsening (+ 0.25%/dec). No particular geographical structures are found in this case. This result indicates that the reproduction of trends of warm extremes with High Resolution models hasn't considerably improved over Europe for most of the models.

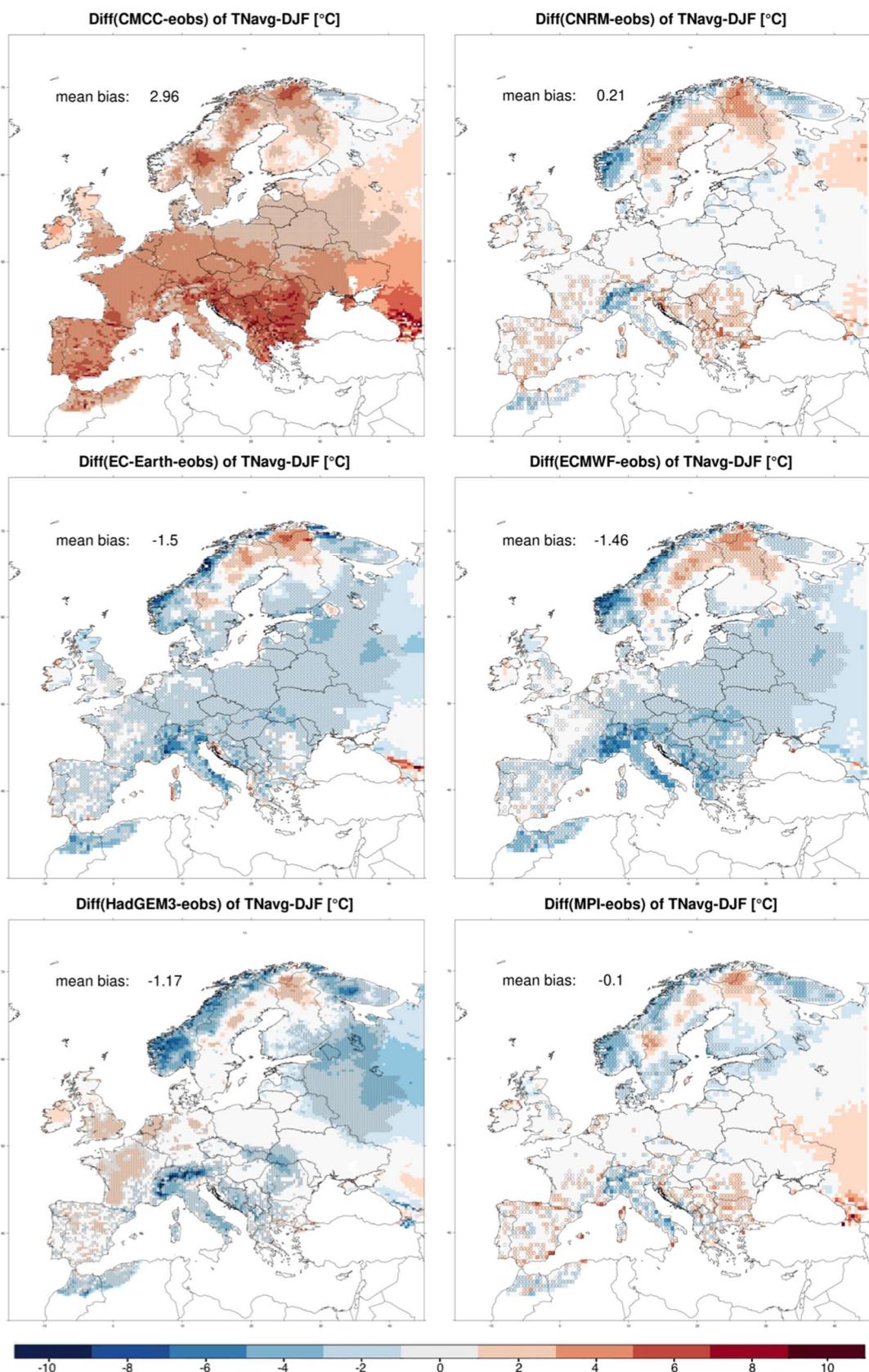


Figure 3.57. Difference in the winter average of TN between the HR models and E-OBS. Red indicates overestimation, blue indicates underestimation. Significant differences are indicated by small thin circles for each grid-point, which result in shaded areas.

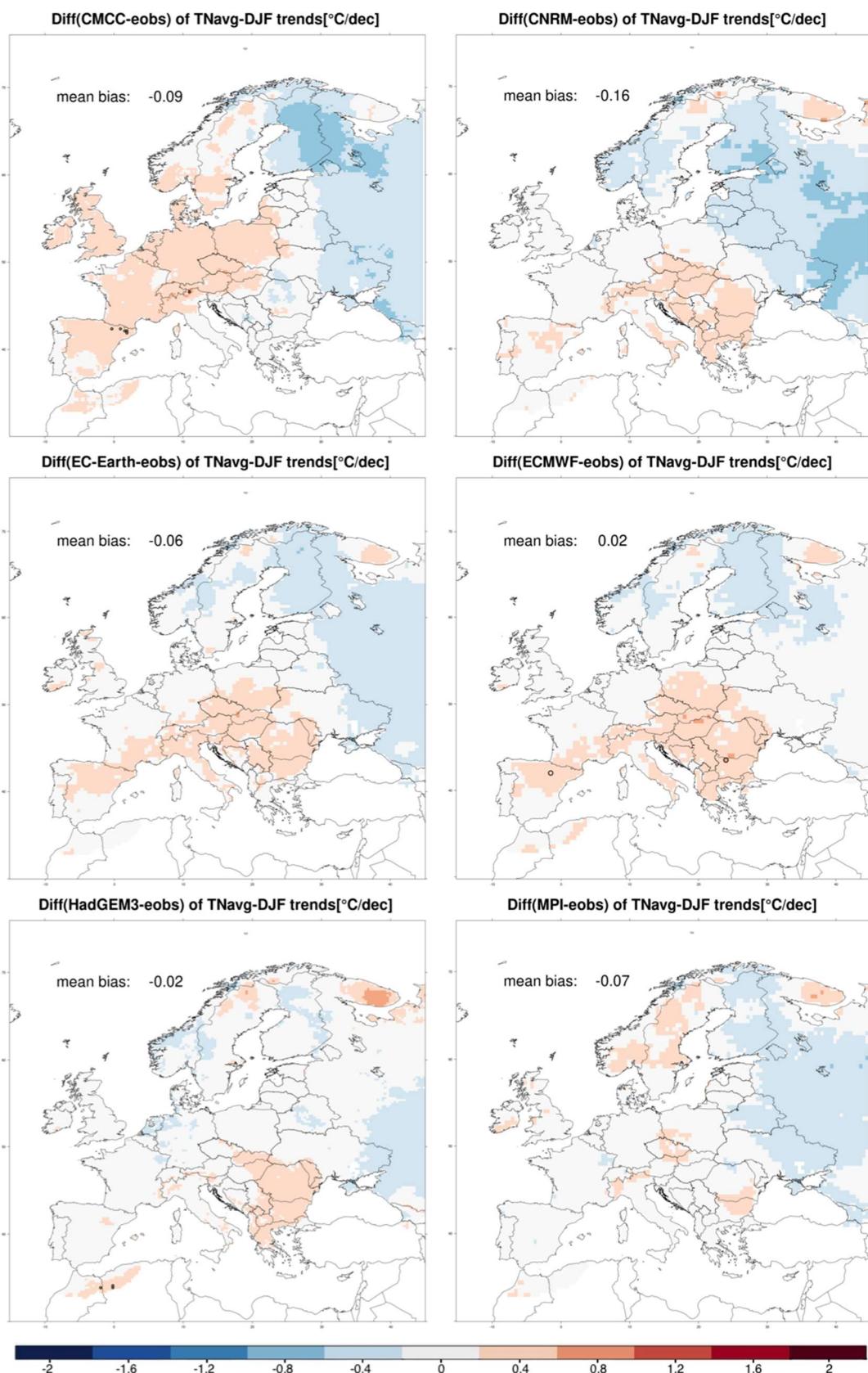


Figure 3.58. Similar to Figure 3.57, but for the bias in daily maximum temperature for the summer (JJA) season.

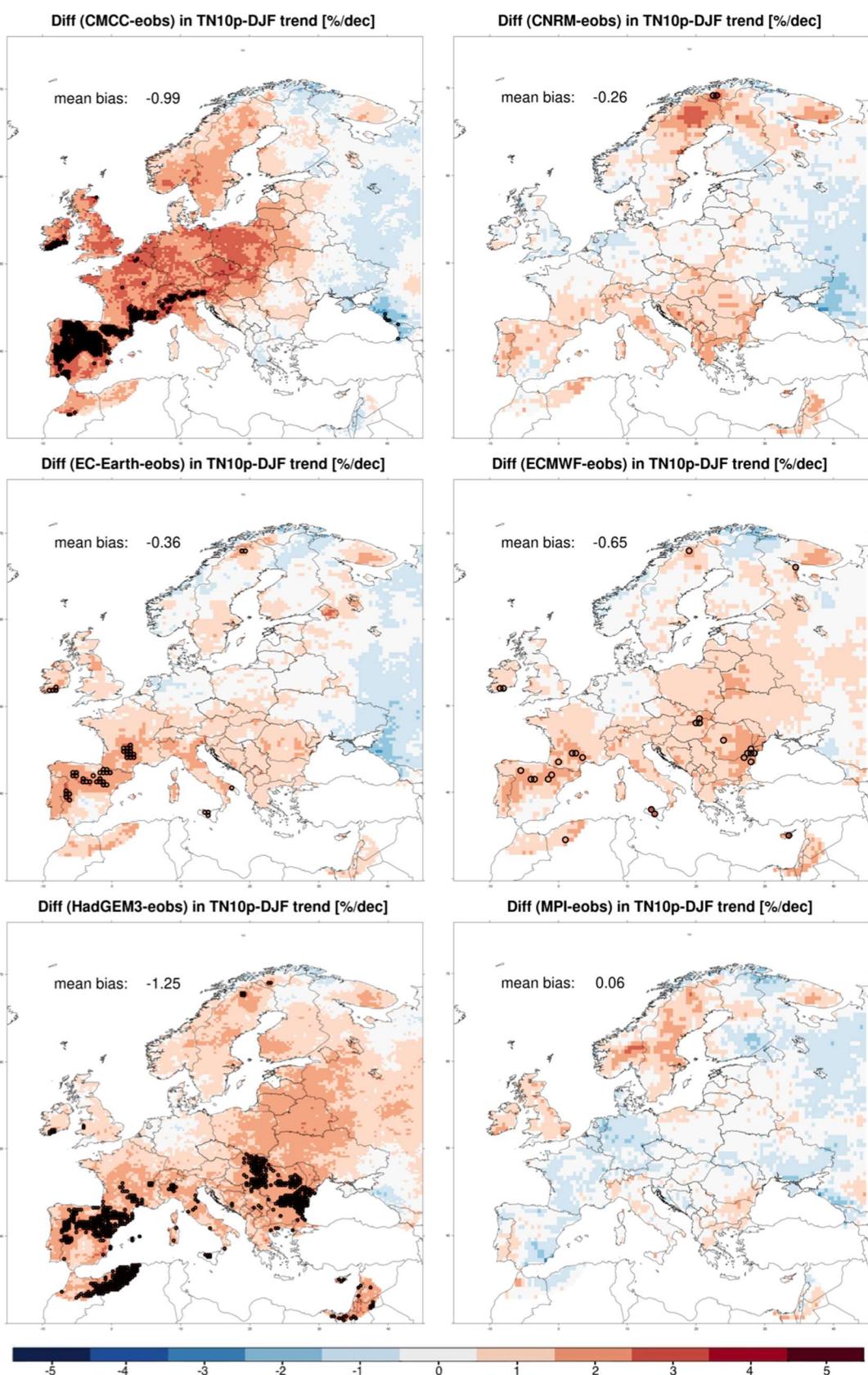


Figure 3.59. Difference in the trends of TNavg-DJF for HR models. Red indicates overestimation (warmer simulated trends), blue indicates underestimation (colder simulated trends). Significant differences are indicated by black circles.

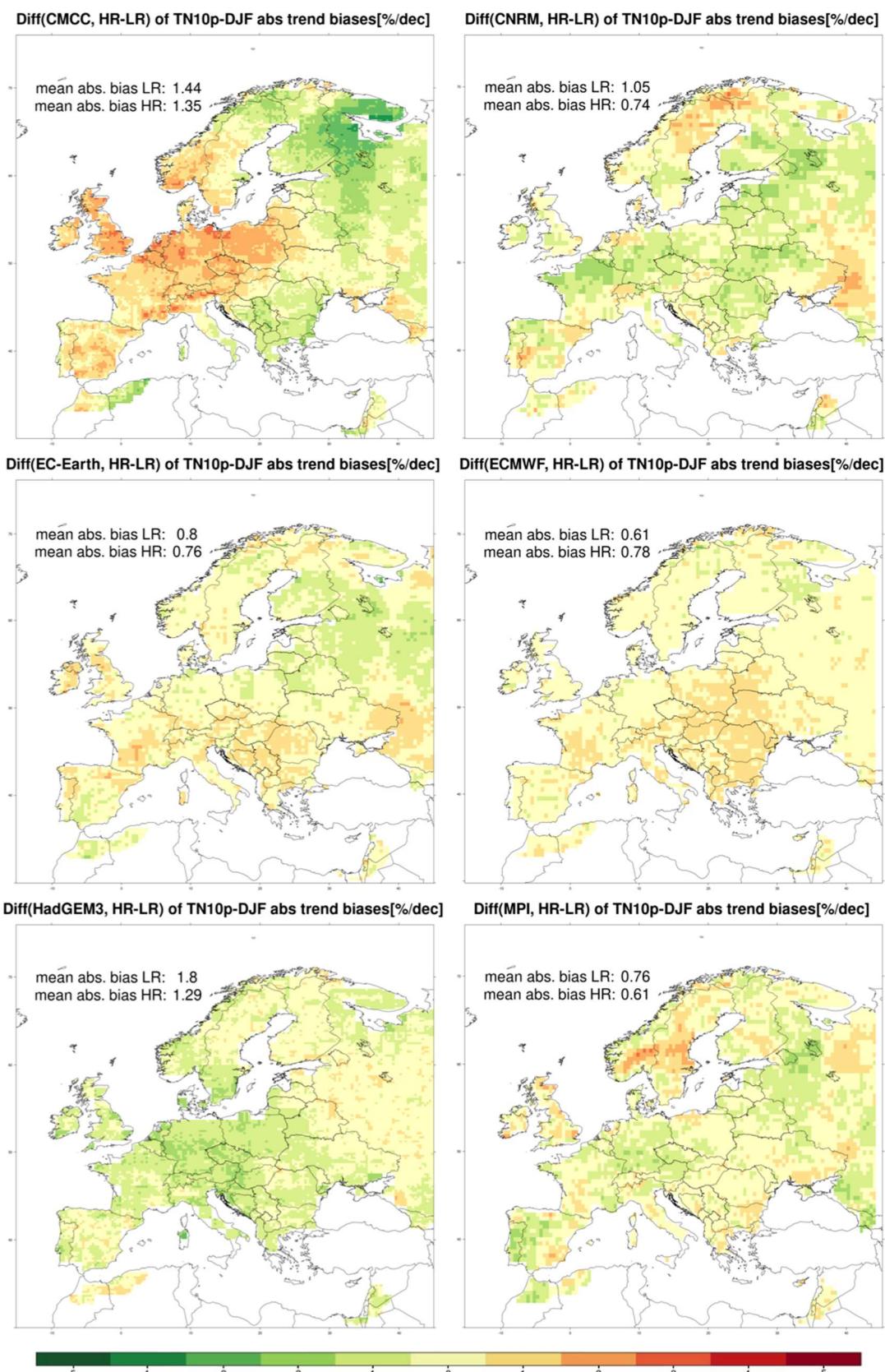


Figure 3.60. Similar to Figure 3.59, but for the summer-averaged trends in daily maximum temperature.

3.6.1.5 Discussion and conclusions

Six models in their High (HR) and Low Resolution (LR) versions have been compared (over the 1970-2014 period) to E-OBS.hom, a version of the gridded dataset E-OBS based on homogenized daily series (each covering at least 1970-2014) of observed temperatures. The analysis has been performed focusing first on the biases and the trend biases of mean values of winter minimum temperatures (TNavg-DJF) and mean values of summer maximum temperatures (TXavg-JJA) and then on two ETCCDI (ETCCDI, 2009) defined indices. These are the number of days with minimum temperatures below the 10th percentile of winter values ('cold nights', TN10p-DJF) and the number of days with maximum temperatures exceeding the 90th percentile of summer values ('warm day-times', TX90p-JJA). The percentile thresholds have been calculated using the 1981-2010 period.

For both winter-mean TN as summer-mean TX strong biases have been found in the simulations, with the strongest ones for CMCC. This model shows mean bias of +2.96°C/dec for TNavg-DJF and -3.83°C/dec for TXavg-JJA, indicating an underestimation of the seasonal cycle all over the continent. The other models present smaller biases (averaged over Europe). Nevertheless, common patterns are found, such as an underestimation of winter minimum temperatures over Italy and Norway and a shared overestimation in the north of Sweden and Finland. This last issue may be related to a lack of snow coverage simulated by the models.

As for maximum temperatures in summer: the models share a common North-South gradient in the bias, with warmer values along the European coasts of the Mediterranean. This may be related to excessive moisture in Northern Europe and a lack of moisture in the Southern sector. Evaluation of results for TXavg-JJA shows a slight overestimation of trends for HR compare to E-OBS on Northern Europe and an underestimation on Southern Europe, especially over Italy and the Balkans.

In Southern Europe, the combination of an excessively large negative bias in summer maximum temperatures with a too-weak increase in the seasonal average and with a much weaker (compared with observations) increase in the extreme indices point to issues in the representation of soil moisture in the models. In a climate which is too warm the soil can be expected to lack more moisture than in cooler conditions due to enhanced evaporation. Once the soil is dry the radiation balance is shifted to a state where sensible heat is dominant over latent heat. Under boundary conditions where the incoming energy flux (due to increase of greenhouse gases) raises, this implies a further increase in sensible heat and surface warming. Nevertheless, in conditions of moist soil, the simulated warming trend in temperatures would be even stronger, due to the shift from latent, thus getting close to the observed conditions.

The most interesting aspects on the trends in winter temperatures (Figure 3.59) is the simulation of colder trends in Eastern Europe (excluding the Kola peninsula) common to all models. Such anomaly might be linked to too small simulation of the reduction trend of snow coverage compared with the observations (van Oldenborgh et al. 2009) and an inspection on the performances of models on this particular variable, which will be subject of future studies.

The too warm simulated trend on the peninsula of Kola is found in the trends on TNavg-DJF and TN10p-DJF (as an underestimation of the number of days below the 10th percentile) and is related to E-OBS station density issues. The only series with observed values in the area (Krasnoshelye) starting before 1970 has missing data between 1972 and 1980. The interpolation of data coming from series in surrounding stations, in the case of TN, brings to higher values in the 1972-1980 compared with the following years, introducing a too cold trend that doesn't take place in the models. This behavior, limited to only one series, motivates the ECA&D group to work on further data collection and in increasing the station density in this and other areas. This will allow to increase the quality of the interpolation and avoid such criticisms.

Trends in extreme values have presented several anomalies, often with common geographical patterns among the models. While the underestimation of trends of TNavg-DJF simulated over Eastern Europe is found also for TN10p-DJF five models indicate an underestimation of the percentage of cold days, thus warmer trends, over Southern Europe. At the same time the underestimation of percentage of warm day is found for the trends of TX90p-JJA, indicating colder trends than the observed ones, consistent with the findings of Min et al. (2013) for CMIP5.

The combination of these two aspects indicates that around the Mediterranean the model trends in the tails of the distribution are closer to each other than what is observed. Therefore, in Southern Europe the distribution of simulated daily temperatures tends to get narrower compared with the distribution of observed daily temperatures, underestimating the intensity of the extremes, especially the warm ones.

As a last step, the analysis of the absolute trend bias evolution in the models from LR to HR does not show a general improvement. Each model presents different patterns and diverse behavior in terms of change of mean absolute trend bias. Nevertheless, this index decreases for TN10p-DJF in all models except ECMWF, indicating a better improvement compared with what is found for TX90p-JJA, where only 3 methods slightly improve, and the others present worsening up to +0.25%/dec.

Finally, it appears that the new high-resolution models, even though they do not significantly increase or decrease their absolute bias on the trends of the extremes, still present some criticisms especially on the area of the Mediterranean. In this region the most serious discrepancy to observations is the large underestimation of the increasing trends of warm extremes. Considering the high economic and societal vulnerability of these areas to very warm events in summer and the importance of the prediction of heatwaves intensity and frequency for the next decades, it is fundamental to improve the simulation of these phenomena and of their projections to future decades.

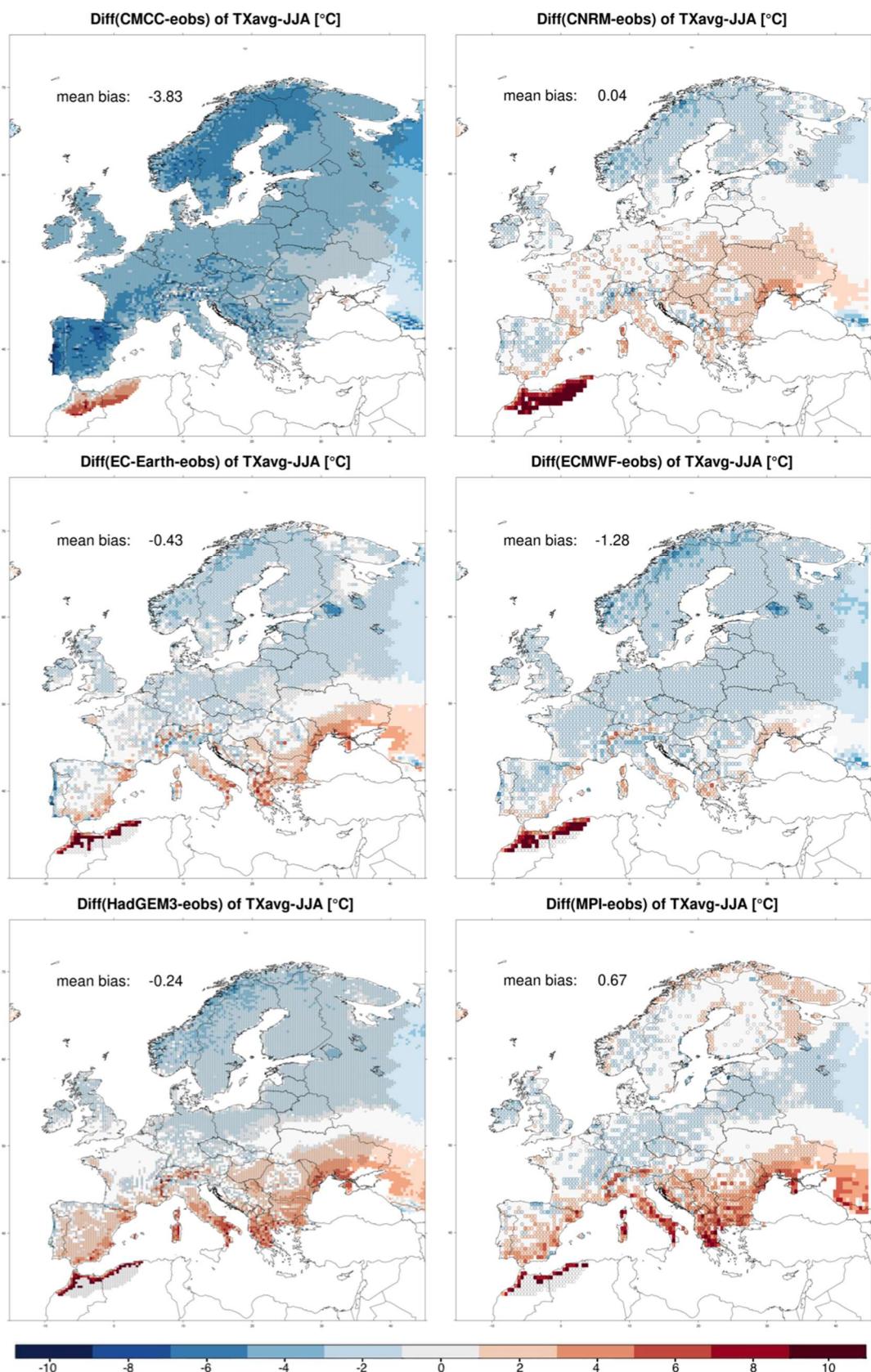


Figure 3.61. Difference in trends of TN10p-DJF between the HR models and E-OBS. Red(blue) indicates an underestimation (overestimation) of the trend, related to a warmer (colder) trend in the model.

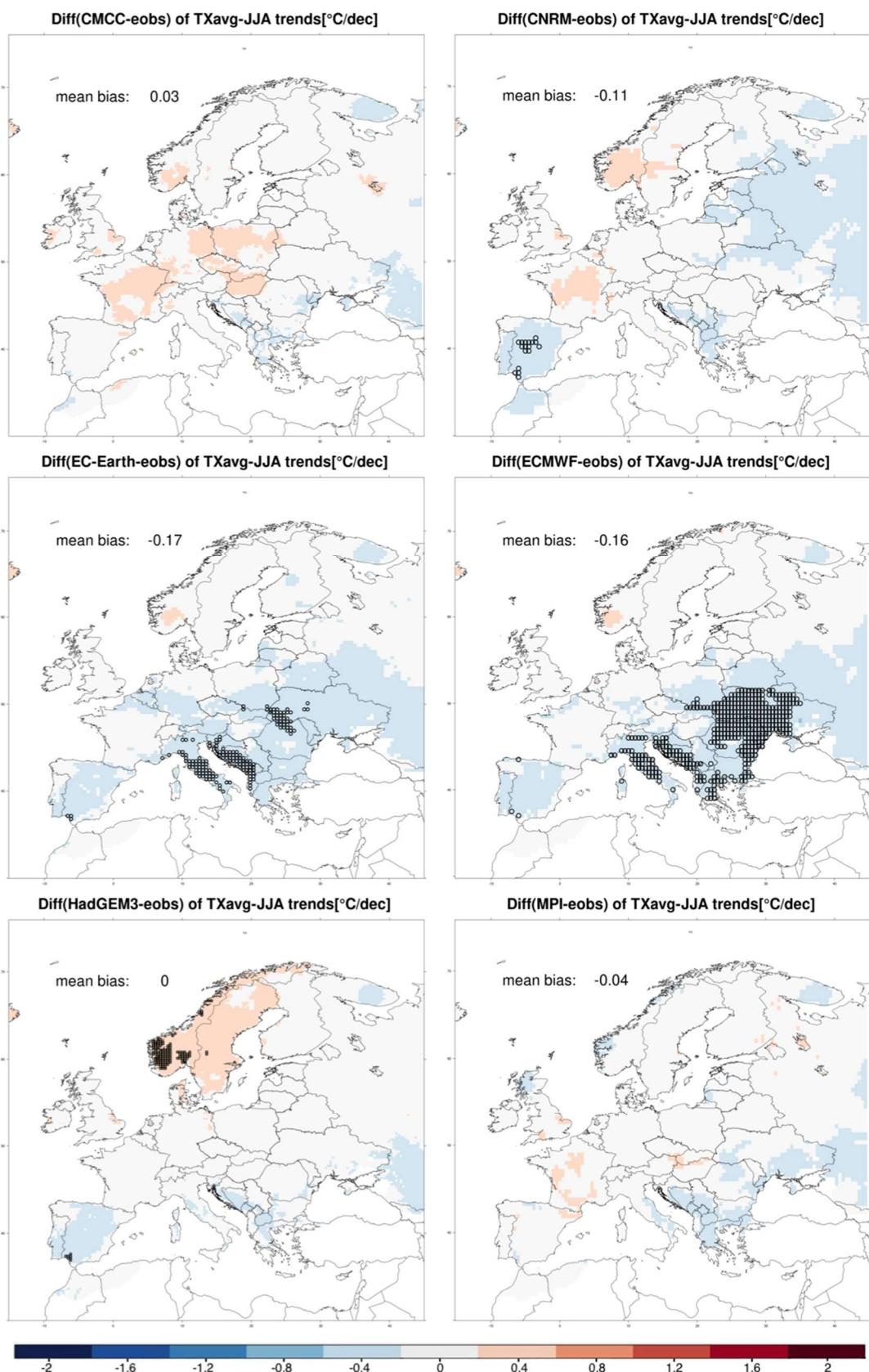


Figure 3.62. Difference in absolute trend bias of TN10p-DJF between HR and LR models. Red (green) pixels indicate an increase (decrease) of the absolute trend bias, thus a better (worse) performance.

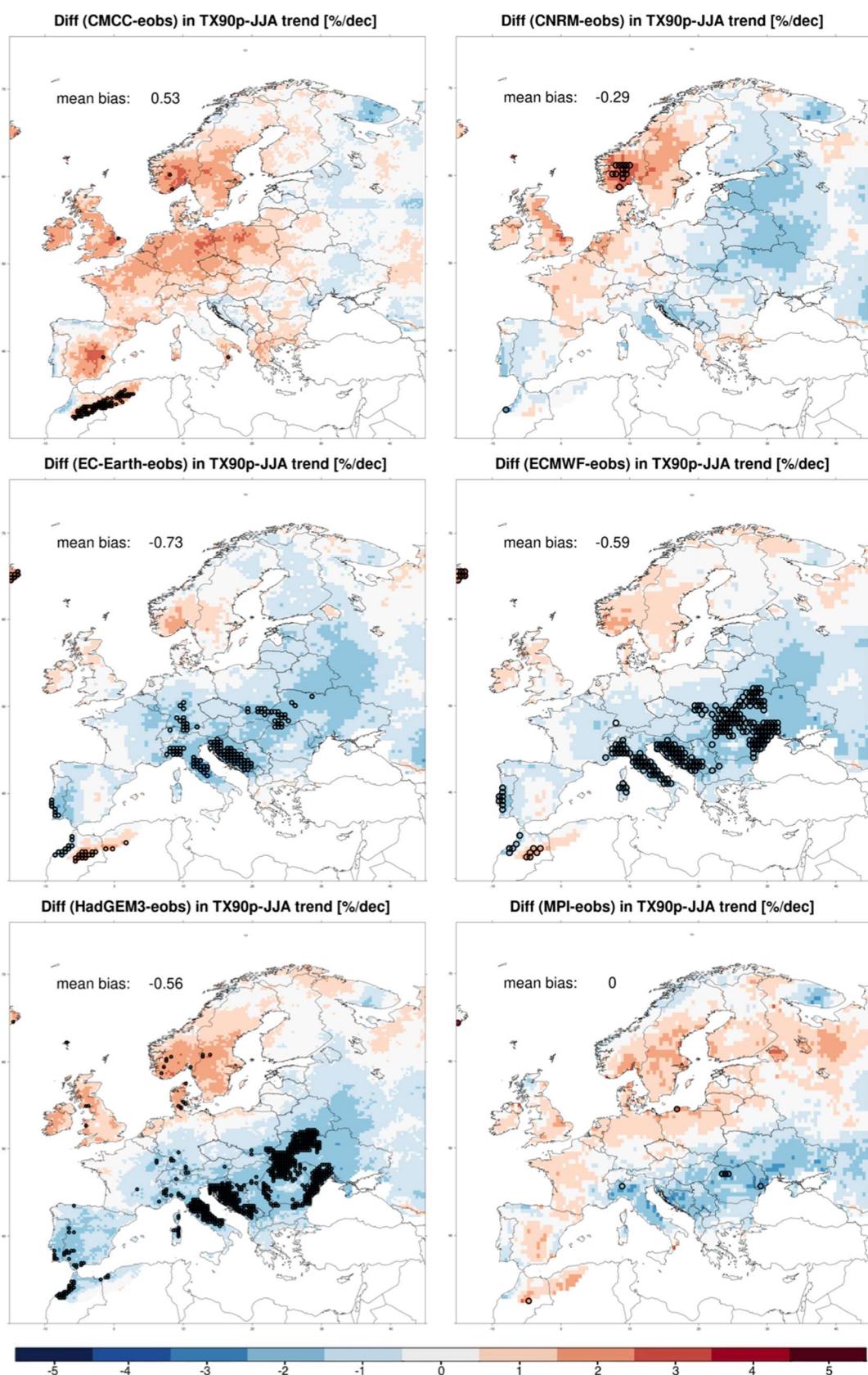


Figure 3.63. Similar as Figure 3.61, but for the trends in TX90p over the summer (JJA) season.

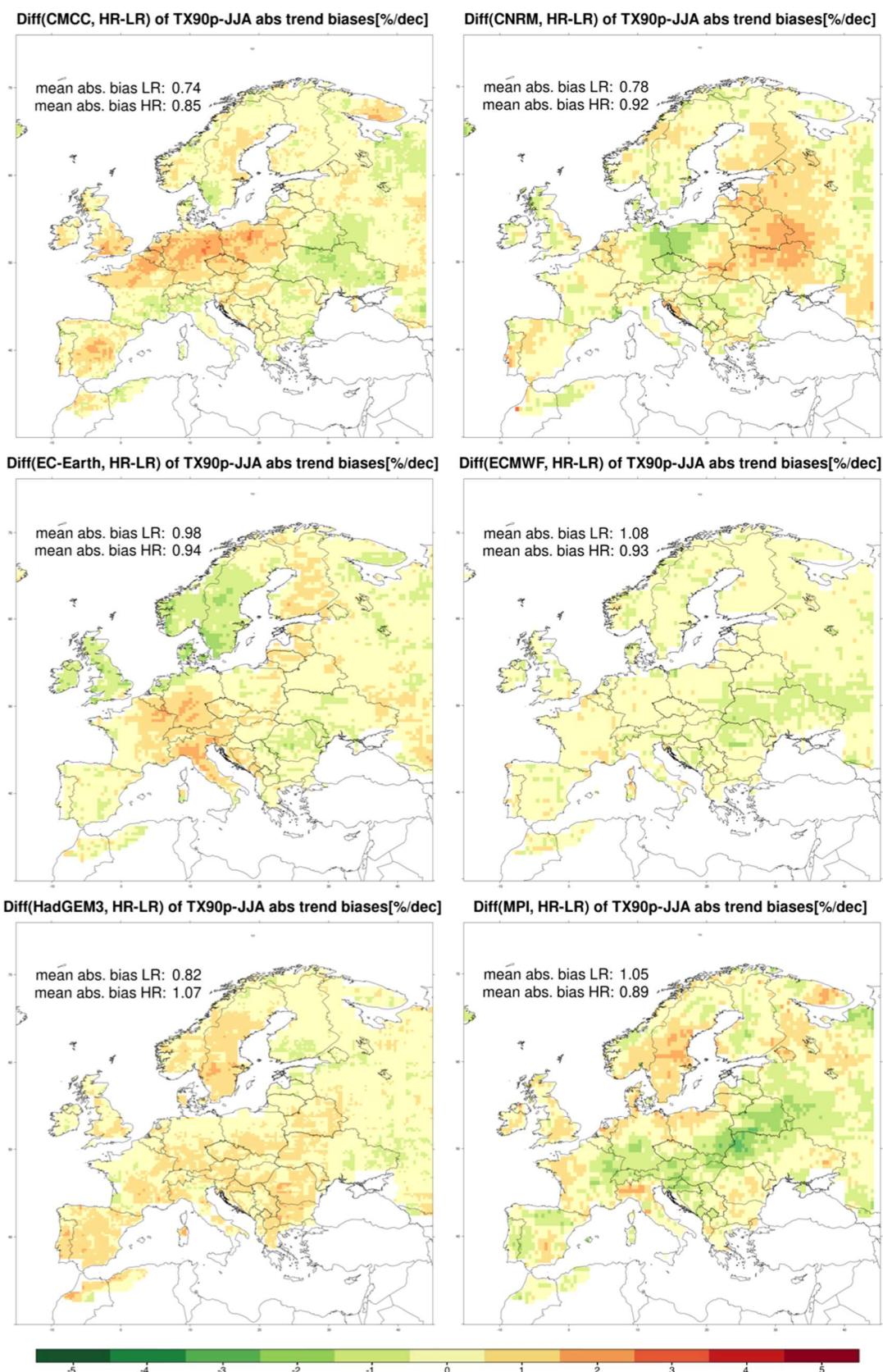


Figure 3.64. Similar to Figure 3.63, but for the trends in TX90p in the summer season.

4. Lessons Learnt

The general improvement resulting from increased resolution in both forced and coupled simulations is often difficult to quantify, not least because improved sampling at higher resolutions should be accounted for and mean-state improvements may arise for incorrect mechanistic reasons. Improvements for specific processes are more easily quantifiable, particularly:

- Aspects of large-scale circulation of the atmosphere are better represented at high resolution, namely Euro-Atlantic blocking frequency and the North Atlantic eddy-driven jet variability. This is key because these large-scale phenomena steer synoptic-scale weather systems and govern the synoptic-scale spatiotemporal distribution of weather regime occurrences, and thereby extreme events.
- Extratropical cyclone frequency and wind speeds extrema are better represented in the PRIMAVERA models compared with CMIP5.
- As may be expected from improved representation of large- and synoptic-scale processes, extreme precipitation is better reproduced by high-resolution models, particularly during winter.
- PRIMAVERA models outperform EURO-CORDEX RCMs for extreme precipitation over larger parts of Europe (this deliverable) and other climate variables (see D10.2). Regions of complex orography are more suited to RCM-based risk assessments.

However, inter-model differences exceed differences between resolutions for other phenomena, particularly summer and autumn post-tropical cyclone occurrence, which is poorly represented at low-resolution across all models, and by some models at both low and high-resolution.

PRIMAVERA models therefore hint at generally improved performance—and greater added value to users due to increased resolution—for wintertime hydroclimate extremes and temperature and their drivers. In particular, model developments to improve the representation of blocking are anticipated to, in turn, improve simulated extremes. This warrants greater mechanistic explanation based on analysis of (i) multiple ensemble members from models where available and (ii) coupled PRIMAVERA simulations across WPs.

5. Links Built

5.1 Outside PRIMAVERA

- Section 3.2.1 work on blocking contributed to IPCC AR6 WG1.
- Section 3.4.1 work on precipitation contributed to IPCC AR6 WG1, chapter 10.
- Work on blocking and extreme precipitation includes application of metrics developed in WP1, which were contributed to the ESMValTool. Priority was given to those

metrics that needed for evaluations related to WP10, as outlined in D1.2 and presented at GA4. This will continue to GA5.

- Sections 3.1.1 and 3.1.2 have been presented to members of the insurance industry, and generated interest in developing a windstorm event set for use in catastrophe modelling (to be described in D10.4). This will be developed in a splinter session during the 2020 European Geosciences Union General Assembly.

5.2 Between PRIMAVERA Work Packages

- WP10 and WP11 are closely connected. Results from this report will be built into user cases and dissemination by WP11. The WP10-11 link is bi-directional: following D10.2, the scientific studies undertaken for this deliverable aim to address at least some of the use cases from D10.1, which were, in turn, compiled based on methods and information from WP11 interviews/survey. These user cases informed the priorities for cross-partner collaborations on topics relevant to extremes, and helped define the scope of this deliverable report.
- Manuscripts are in preparation describing evaluations of the representation of blocking and eddy-driven jet variability. Manuscript preparation is taking place in tandem to link these related phenomena. It is anticipated that, with the publication of two coherent narratives, strong links between this WP1 research and impacts-related work in WP10, particularly storm and low-wind event occurrence, may be drawn, which will directly connect these publications to users' interests.
- Further inter-WP research is anticipated related to (i) decadal variability and (ii) climate change, particularly based on coupled (hist-1950 and highres-future) simulations.

6. References

Baker, A. J. et al., 2019a: Enhanced Climate Change Response of Wintertime North Atlantic Circulation, Cyclonic Activity, and Precipitation in a 25-km-Resolution Global Atmospheric Model. *Journal of Climate* **32**, 7763-7781

Baker, A. J., Hodges, K. I., Schiemann, R., and Vidale, P. L., 2019b: Historical variability of North Atlantic post-tropical cyclones. In review, *Geophysical Research Letters*, 2019.

Berthou S, Kendon E J, Chan SC, Ban N, Leutwyler D, Schär C and FosserG 2018: Pan-european climate at convection permitting scale: a model intercomparison study Clim. Dyn. (<https://doi.org/10.1007/s00382-018-4114-6>)

Cannon, D. J., Brayshaw, D. J., Methven, J., Coker, P. J. & Lenaghan, D. Using reanalysis data to quantify extreme wind power generation statistics: a 33 year case study in Great Britain. *Renew. Energy* **75**, 767–778 (2015).

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., Masina, S., Soccimarro, E., Materia, S., Bellucci, A., and Navarra, A. (2019). Global mean climate and

main patterns of variability in the CMCC-CM2 coupled model. *Journal of Advances in Modeling Earth Systems*, 11, 185–209. <https://doi.org/10.1029/2018MS001369>

Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A.J. Weaver and M. Wehner, 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J., & Jones, P. D. (2018). An ensemble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123(17), 9391-9409.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P. , Poli, P. , Kobayashi, S. , Andrae, U. , Balmaseda, M. A., Balsamo, G. , Bauer, P. , Bechtold, P. , Beljaars, A. C., van de Berg, L. , Bidlot, J. , Bormann, N. , Delsol, C. , Dragani, R. , Fuentes, M. , Geer, A. J., Haimberger, L. , Healy, S. B., Hersbach, H. , Hólm, E. V., Isaksen, L. , Källberg, P. , Köhler, M. , Matricardi, M. , McNally, A. P., Monge-Sanz, B. M., Morcrette, J. , Park, B. , Peubey, C. , de Rosnay, P. , Tavolato, C. , Thépaut, J. and Vitart, F. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*, 137: 553-597. doi:10.1002/qj.828

Dekker, M. M. et al., 2018: Characteristics and development of European cyclones with tropical origin in reanalysis data. *Climate Dynamics* **50**, 445-455

Economou, T., Stephenson, D. B., Pinto, J. G., Shaffrey, L. C. and Zappa, G. (2015), Serial clustering of extratropical cyclones in a multi-model ensemble of historical and future simulations. *Q.J.R. Meteorol. Soc.*, 141: 3076-3087. doi:10.1002/qj.2591

Evans, C. et al., 2017: The Extratropical Transition of Tropical Cyclones. Part I: Cyclone Evolution and Direct Impacts. *Monthly Weather Review* **145**, 4317-4344

Fabiano, F. H.M. Christensen, K. Strommen, P. Athanasiadis, A. Baker, R. K. Schiemann, S. Corti: Euro-Atlantic wintertime Weather Regimes in coupled historical climate simulations: model performance and improvements with increased resolution, in review.

Gelaro, R. and others, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate*, 30, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>

Gonzalez, P. L. M. et al., 2019: The contribution of North Atlantic atmospheric circulation shifts to future wind speed projections for wind power over Europe. *Climate Dynamics* **53**, 4095-4113

Grams, C., Beerli, R., Pfenninger, S. et al. Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Clim Change* 7, 557–562 (2017) doi:10.1038/nclimate3338

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., Haak, H., Stössel, A. (2019). Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP), *Geosci. Model Dev.*, 12, 3241–3281, <https://doi.org/10.5194/gmd-12-3241-2019>.

Haarsma, R, M Acosta, R Bakhshi, P-A Bretonnière, L-P Caron, M Castrillo, S Corti, P Davini, E Exarchou, F Fabiano, U Fladrich, R Fuentes, J García-Serrano, J von Hardenberg, T Koenigk, X Levine, V Meccia, T van Noije, G van den Oord, F Palmeiro, M Rodrigo, Y Ruprich-Robert, P Le Sager, E Tourigny, S Wang, M van Weele, K Wyser, (2019). HighResMIP versions of EC-Earth: EC-Earth3P and EC-Earth3P-HR. Description, model performance, data handling and validation. *Geosci. Model Dev.*, submitted.

Hannachi A. et al. (2013). Behaviour of the winter North Atlantic eddy-driven jet stream in the CMIP3 integrations. *Clim. Dyn.* 41: 995–1007.

Hart, R. E., 2003: A Cyclone Phase Space Derived from Thermal Wind and Thermal Asymmetry. *Monthly Weather Review* **131**, 585-616

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research: Atmospheres*, 113(D20).

Hodges, K. I., 1995: Feature Tracking on the Unit Sphere. *Monthly Weather Review* **123**, 3458-3465

Hodges, K. I., 1999: Adaptive Constraints for Feature Tracking. *Monthly Weather Review* **127**, 1362-1373

Hodges, K. et al., 2017: How Well Are Tropical Cyclones Represented in Reanalysis Datasets? *Journal of Climate* **30**, 5243-5264

Huffman, G. J. et al., 2001: Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *Journal of Hydrometeorology* **2**, 36-50

Iqbal W. et al. (2018). Analysis of the variability of the North Atlantic eddy-driven jet stream in CMIP5. *Clim. Dyn.* 51: 235.

Jones, S. C. et al., 2003: The Extratropical Transition of Tropical Cyclones: Forecast Challenges, Current Understanding, and Future Directions. *Weather and Forecasting* **18**, 1052-1092.

Keller, J. H. et al., 2019: The Extratropical Transition of Tropical Cyclones. Part II: Interaction with the Midlatitude Flow, Downstream Impacts, and Implications for Predictability. *Monthly Weather Review* **147**, 1077-1106

Kjellström, E., Nikulin, G., Hansson, U., Strandberg, G. and Ullerstig, A. 2011: 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations. *Tellus* 63A, 24-40.

Klawa, M. and Ulbrich, U.: A model for the estimation of storm losses and the identification of severe winter storms in Germany, *Nat. Hazards Earth Syst. Sci.*, 3, 725–732,

Klingaman, N. P., Martin, G. M., and Moise, A.: ASoP (v1.0): a set of methods for analyzing scales of precipitation in general circulation models, *Geoscientific Model Development*, 10, 57–83, doi:10.5194/gmd-10-57-2017, 2017.

Klein Tank, A. M. G., Wijngaard, J. B., Können, G. P., Böhm, R., Demarée, G., Gocheva, A., & Heino, R. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 22(12), 1441-1453.

Kwon, YO. et al. (2018). North Atlantic winter eddy-driven jet and atmospheric blocking variability in the Community Earth System Model version 1 Large Ensemble simulations. *Clim. Dyn.* 51: 3275. doi:10.5194/nhess-3-725-2003, 2003.

Lamb, H. H.: *Historic Storms of the North Sea, British Isles and Northwest Europe*, Cambridge University Press, 1991.

Lind, P., Lindstedt, D., Kjellström, E. and Jones, C., 2016: Spatial and Temporal Characteristics of Summer Precipitation over Central Europe in a Suite of High-Resolution Climate Models. *Journal of Climate*, 29, 3501-3518.

Madonna, E., Li, C., Grams, C.M. and Woollings, T. (2017), The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Q.J.R. Meteorol. Soc.*, 143: 2960-2972. doi:10.1002/qj.3155.

Min, E., Hazeleger, W., Van Oldenborgh, G. J., & Sterl, A. (2013). Evaluation of trends in high temperature extremes in north-western Europe in regional climate models. *Environmental Research Letters*, 8(1), 014011.

Pinto, J. G., Fröhlich, E. L., Leckebusch, G. C., and Ulbrich, U.: Changing European storm loss potentials under modified climate conditions according to ensemble simulations of the ECHAM5/MPI-OM1 GCM, *Nat. Hazards Earth Syst. Sci.*, 7, 165–175, <https://doi.org/10.5194/nhess-7-165-2007>, 2007.

Roberts, C. D., Senan, R., Molteni, F., Boussetta, S., Mayer, M., and Keeley, S. P. E. (2018). Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP, *Geosci. Model Dev.*, 11, 3681–3712, <https://doi.org/10.5194/gmd-11-3681-2018>.

Roberts, M. J., Baker, A., Blockley, E. W., Calvert, D., Coward, A., Hewitt, H. T., Jackson, L. C., Kuhlbrodt, T., Mathiot, P., Roberts, C. D., Schiemann, R., Seddon, J., Vannière, B., and Vidale, P. L. (2019). Description of the resolution hierarchy of the global coupled HadGEM3-GC3.1 model as used in CMIP6 HighResMIP experiments, *Geosci. Model Dev.*, 12, 4999–5028, <https://doi.org/10.5194/gmd-12-4999-2019>.

Roberts, M. J., Joanne Camp, Pier Luigi Vidale, Kevin Hodges, Benoit Vanniere, Jenny Mecking, Rein Haarsma, Alessio Bellucci, Enrico Scoccimarro, Louis-Philippe Caron, Fabrice Chauvin, Dian Putrasahan, Christopher Roberts, 2019b. Impact of model resolution on tropical cyclone simulation using the HighResMIP-PRIMAVERA multi-model ensemble. Submitted.

Scaife, A. a., Woollings, T., Knight, J., Martin, G., & Hinton, T. (2010). Atmospheric Blocking and Mean Biases in Climate Models. *Journal of Climate*, 23(23), 6143–6152. <https://doi.org/10.1175/2010JCLI3728.1>

Scherrer, S. C., Croci-Maspoli, M., Schwierz, C. and Appenzeller, C. (2006), Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *Int. J. Climatol.*, 26: 233-249. doi:10.1002/joc.1250

Schiemann, R., Demory, M.-E., Shaffrey, L. C., Strachan, J., Vidale, P. L., Mizielinski, M. S., et al. (2017). The Resolution Sensitivity of Northern Hemisphere Blocking in Four 25-km Atmospheric Global Circulation Models. *Journal of Climate*, 30(1), 337–358. <https://doi.org/10.1175/JCLI-D-16-0100.1>

Schiemann, R.; Athanasiadis, P. 2018: Representation of blocking in a multi-model ensemble of high-resolution coupled global climate models. 20th EGU General Assembly, EGU2018. Geophysical Research Abstracts, Vol. 20, EGU2018-7884, 2018

Schiemann, R. et al., 2018: Mean and extreme precipitation over European river basins better simulated in a 25km AGCM. *Hydrology and Earth System Sciences* 22, 3933-3950.

Schiemann, R., P. Athanasiadis, D. Barriopedro, F. Doblas-Reyes, K. Lohmann, M. J. Roberts, D. Sein, C. D. Roberts, L. Terray and P. L. Vidale. (2020): The representation of Northern Hemisphere blocking in current global climate models. *Wea. and Clim. Dyn.* <https://doi.org/10.5194/wcd201919>.

Scherrer, S. C., Croci-Maspoli, M., Schwierz, C., & Appenzeller, C. (2006). Two-dimensional indices of atmospheric blocking and their statistical relationship with winter climate patterns in the Euro-Atlantic region. *International Journal of Climatology*, 26(2), 233–249. <https://doi.org/10.1002/joc.1250>

Silverman BW (1986). Density Estimation for Statistics and Data Analysis, Vol. 26, Monographs on Statistics and Applied Probability, Chapman and Hall, London, 1986.

Squintu, A. A., van der Schrier, G., Brugnara, Y., & Klein Tank, A. (2019). Homogenization of daily temperature series in the European Climate Assessment & Dataset. *International journal of climatology*, 39(3), 1243-1261.

Squintu, A. A., van der Schrier, G., Štěpánek, P., Zahradníček, P., & Klein Tank, A. M. G. (2020) Comparison of homogenization methods for daily temperature series against an observation-based benchmark dataset. *Theoretical and Applied Climatology*, 1-17.

Stewart, S. R., 2018: Hurricane Ophelia (AL172017). *National Hurricane Center Tropical Cyclone Report*, www.nhc.noaa.gov/data/tcr/AL172017_Ophelia.pdf

van Oldenborgh, G. J., Drijfhout, S. S., Van Ulden, A., Haarsma, R. J., Sterl, A., Severijns, C., ... & Dijkstra, H. A. (2009). Western Europe is warming much faster than expected. *Climate of the Past*, 5(1), 1-12.

Volodire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., et al. (2019). Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, 11, 2177–2213. <https://doi.org/10.1029/2019MS001683>

Walz, M.A., Kruschke, T., Rust, H.W., Ulbrich, U. and Leckebusch, G.C. (2017), Quantifying the extremity of windstorms for regions featuring infrequent events. *Atmos. Sci. Lett.*, 18: 315-322. doi:10.1002/asl.758

Woollings T. et al. (2010). Variability of the North Atlantic eddy-driven jet stream. *Q.J.R. Meteorol. Soc.*, 136: 856-868. doi: 10.1002/qj.625.

Woollings, T. (2010). Dynamical influences on European climate: an uncertain future. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, 368, 3733–3756. <https://doi.org/10.1098/rsta.2010.0040>

Woollings, T., E. Barnes, B. Hoskins, Y. Kwon, R.W. Lee, C. Li, E. Madonna, M. McGraw, T. Parker, R. Rodrigues, C. Spensberger, and K. Williams, 2018: Daily to Decadal Modulation of Jet Variability. *J. Climate*, 31, 1297–1314, <https://doi.org/10.1175/JCLI-D-17-0286.1>