ENVIRONMENTAL RESEARCH
LETTERS

ACCEPTED MANUSCRIPT • OPEN ACCESS

How well can a seasonal forecast system represent 3 hourly compound wind and precipitation extremes over Europe?

To cite this article before publication: Laura E. Owen *et al* 2021 *Environ. Res. Lett.* in press https://doi.org/10.1088/1748-9326/ac092e

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is **© 2021 The Author(s). Published by IOP Publishing Ltd**.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 77.103.126.25 on 15/06/2021 at 17:14

How well can a seasonal forecast system represent 3 hourly compound wind and precipitation extremes over Europe?

Laura E. Owen, College of Engineering, Mathematics and Physical Sciences, The University of Exeter
Jennifer L. Catto, College of Engineering, Mathematics and Physical Sciences, The University of Exeter
Nick J. Dunstone, Met Office, Exeter, UK
David B. Stephenson, College of Engineering, Mathematics and Physical Sciences, The University of Exeter

Abstract

Extreme precipitation and winds can have a severe impact on society, particularly when they occur at the same place and time. In this study the Met Office's Global Seasonal forecast system version 5 (GloSea5) model ensembles are evaluated against the reanalysis dataset ERA5, to find out how well they represent 3 hourly extreme precipitation, extreme wind and extreme co-occurring events over Europe. Although substantial differences in magnitude are found between precipitation and wind extremes between the datasets, the conditional probability of exceedance above the 99th percentile, which measures the co-occurrence between the two extremes, compares well spatially over Europe. However, significant differences in frequency are found around and over some areas of high topography. Generally GloSea5 underestimates this co-occurrence over sea. The model's co-occurring events at individual locations investigated occur with very similar synoptic patterns to ERA5, indicating that the compound extremes are produced for the correct reasons.

Keywords: compound extremes, precipitation, wind speed, model evaluation.

1 Introduction

Extreme precipitation and winds can have a severe impact on society and the co-occurrence between the two extremes is important when assessing risk, since together they can cause even greater damage than separately [Martius *et al.*, 2016; Raveh-Rubin and Wernli, 2015]. The co-occurrence of precipitation and wind extremes has been studied over Europe at varying temporal scales using observational data [Martius *et al.*, 2016; Ridder *et al.*, 2020; Vignotto *et al.*, 2021]. The spatial pattern of co-occurrence frequency is complex but is consistent between studies, with high co-occurrence found over Europe's western coasts, the north eastern coast of the Mediterranean and south of the Alps. Co-occurring extreme precipitation and wind events are linked to locations with a high frequency of atmospheric rivers and extratropical cyclones [Catto and Dowdy, 2021; De Luca *et al.*, 2017, 2020; Hénin *et al.*, 2021; Owen *et al.*, 2020; Pfahl, 2014; Raveh-Rubin and Wernli, 2015, 2016], and the probability of getting such events is much higher when cyclones and fronts are present [Catto and Dowdy, 2021]. These results indicate that these weather systems are the common drivers of the two extremes.

Given the high impact of co-occurring extreme events, it is vital that climate models, including seasonal forecasting models, can represent extreme precipitation and wind. Such models are an important tool in understanding future changes in high impact weather and climate events. Studies have evaluated extreme wind and precipitation separately in models [Kumar *et al.*, 2015; Wehner *et al.*, 2021], and have also evaluated model ability to represent extratropical cyclones and fronts [Catto *et al.*, 2010; Priestley *et al.*, 2020; Zappa *et al.*, 2013]. These results typically find that higher resolution models represent the structure and intensity of extratropical cyclones better.

Few studies have evaluated compound events in climate models. Recently Ridder *et al.* [2021] found that some of the sixth phase of the Coupled Model Intercomparison Project (CMIP6) models capture the return periods of both co-occurring daily precipitation and wind extremes as well as heat waves and meteorological drought well over North America, Europe and Eurasia but perform less well over Australia. The models considered had horizontal resolution of typically coarser than ∼100 km. Considering much higher resolution simulations, Zscheischler *et al.* [2021] found that simulations from the Weather Research and Forecasting (WRF) model, run at 2 km resolution, captured daily precipitation and wind extremes well over a region in central Europe around the Alps.

Another use for models that can represent rare but high impact events, is to better understand the present day risk of such events. Due to the rarity of co-occurring extremes, observational estimates of the frequency of very extreme events can be uncertain. Additionally, it is very difficult for estimates based on observations alone to help us understand the risk of future rare or unprecedented events. However, using the UNSEEN method (UNprecedented Simulated Extremes using ENsembles) this risk can be estimated from large ensembles of climate simulations. The Met Office's Hadley Centre Global Environment Model (HadGEM3-GC2) has been evaluated and used to investigate unprecedented events [Kent *et al.*, 2017, 2019; Thompson *et al.*, 2017, 2019]. Precipitation and temperature in HadGEM3-GC2 were found to be statistically indistinguishable from observational datasets in Thompson *et al.* [2019], Kent *et al.* [2019] and Thompson *et al.* [2017]. The results from these demonstrate the potential of the UNSEEN methodology to quantify the chance of unprecedented events and understand their causes. However, to date the model has not been evaluated for its ability to represent co-occurring extremes and the dynamical drivers of these events, such as extratropical cyclones.

The aim of this study is to find out how well the Met Office's Global Seasonal forecast system version 5 (GloSea5) model ensembles, which are based on HadGEM3-GC2, represent 3 hourly co-occurring extreme precipitation and wind events and the synoptic situations leading to them by comparing the model ensembles with the reanalysis dataset, ERA5. This is a very stringent test of the model as the correct representation depends on how well it represents synoptic scale dynamical features that drive the extremes, as well as the subgrid-scale features such as extreme precipitation. In contrast to previous studies we investigate a larger region over Europe with a higher horizontal (25 km) and temporal (3 hourly) resolution [Ridder *et al.*, 2021; Zscheischler *et al.*, 2021]. Additionally GloSea5 has not been evaluated in such a way before and is needed before the model can be used to calculate unprecedented events using the UNSEEN method. The following questions are addressed:

1. How well does GloSea5 represent the spatial pattern and magnitude of 3 hourly extreme wind and precipitation?
2. How well does GloSea5 represent the spatial pattern and frequency of 3 hourly co-occurring extreme wind and precipitation?
3. Can GloSea5 represent the correct synoptic situations leading to 3 hourly co-occurring extreme events?

2 Data and Methodology

2.1 Data

The Global Seasonal forecast system version 5 (GloSea5) from the UK Met Office is an 'experimental' high resolution version of the climate model HadGEM3, with a resolution of 25 km. For details of the model setup we refer the reader to MacLachlan *et al.* [2015]. Note that here we use the high horizontal resolution version of GloSea5 for details of this we refer the reader to Scaife *et al.* [2019]. We use the 3 hourly total precipitation and 10 m mean wind speed for the extreme events, and 3 hourly mean sea level pressure (MSLP) for investigating the synoptic patterns. GloSea5 is an ensemble forecast system with 3 start dates per season (25th October, and 1st and 11th November) and 8 ensemble members per start date. Each ensemble member runs for 210 days for 24 seasons. We generate 24 timeseries for 1993 to 2016 by drawing from each of the ensemble members. We also generate 100 timeseries from a random selection of ensemble members using a bootstrapping method which are used in figure 4.

The observationally constrained data against which we evaluate the model is 3 hourly total precipitation accumulation, 10 m mean wind speed and MSLP from the European Centre for Medium-range Weather Forecasts ERA5 reanalysis dataset [Hersbach *et al.*, 2020]. ERA5 is a global dataset with spatial resolution of 31 km and is taken as a good representation of 'real life'. Owen *et al.* [2020] checked the robustness of ERA5 against observational datasets and found it compared well spatially and represented the timings of daily extreme co-occurring events well. Although ERA5 did overestimate the frequency of co-occurrence over high topography .

We focus on the winter season, December to February (DJF), of 1993 to 2016 and the region of 20° W to 40° E and 30° S to 75° N to investigate Europe. GloSea5 is regridded to a spatial resolution of 31 km to match ERA5, using first order conservative mapping.

We use 3 hourly data because this is the highest temporal resolution from GloSea5 and hourly extremes are of most interest when investigating the impact of wind speed. Given that we are considering such a special case of 3 hourly extremes with no relaxation, this is a very stringent test of the model. Furthermore Owen *et al.* [2020] found that a 24 hour lag and lead causes no/very little change in co-occurring precipitation and wind extremes for most of Europe. Additionally, we investigated the longer timescale of 24 hourly extremes where a compound event is found if the daily mean wind speed mean and the daily precipitation accumulation are both extreme, similar to Martius *et al.* [2016]; Ridder *et al.* [2020, 2021]. The results of these are mentioned but not shown.

2.2 Definition of extremes

Precipitation and wind speed above the 99th percentile for each gridpoint, for DJF, are taken as extreme. All events have been included in our dataset including zero precipitation events. The 99th percentile has often been used to define extreme precipitation [Catto and Pfahl, 2013; Pfahl and Wernli, 2012] and wind [Pfahl, 2014], and is used throughout this study. Other thresholds are also investigated for locations of interest (section 3.1.1 and 3.2.1).

The GloSea5 mean 99th percentiles of wind speed and precipitation are created by taking the mean of the 99th percentiles from each of the 24 GloSea5 model timeseries (section 3.1). The mean is taken from the 24 timeseries rather than pooling all the data together due to efficiency. Although analysis using the pooled method was done for multiple gridpoints and results were indistinguishable.

2.3 Definition of extreme co-occurrence

A co-occurrence is recorded if the precipitation and the maximum wind speed occur at the same 3 hourly timestep, at the same gridpoint and are each above the 99th percentile. To quantify the co-occurrence we use the conditional probability measure, χ, which was first introduced by Coles *et al.* [1999], and has also been used in Owen *et al.* [2020]. It is the probability of one variable being extreme given that the other is extreme. At each gridbox

$$\chi(p) = Pr(Y(t) > y_p \mid X(t) > x_p) \tag{1}$$

where Y is precipitation, X is wind speed and y_p and x_p are the pth quantiles of Y and X with threshold probability $p \in [0, 1]$, $\chi(p)$ becomes the measure of extremal dependence in the limit as p tends to 1.

In this paper χ is the probability of a precipitation (or wind) extreme occurring given a wind (or precipitation) extreme and is calculated by

$$\hat{\chi} = \frac{n_a}{(1-p)n} \tag{2}$$

where n is the total number of timesteps, n_a are the number of co-occurring events and for the majority of this study $p = 0.99$. At every gridpoint $\hat{\chi}$ is calculated for each of the 24 GloSea5 timeseries, the mean at each gridpoint is then calculated from all 24 to create the GloSea5 mean $\hat{\chi}$ (section 3.2).

3 Results

3.1 Wind and Precipitation extremes

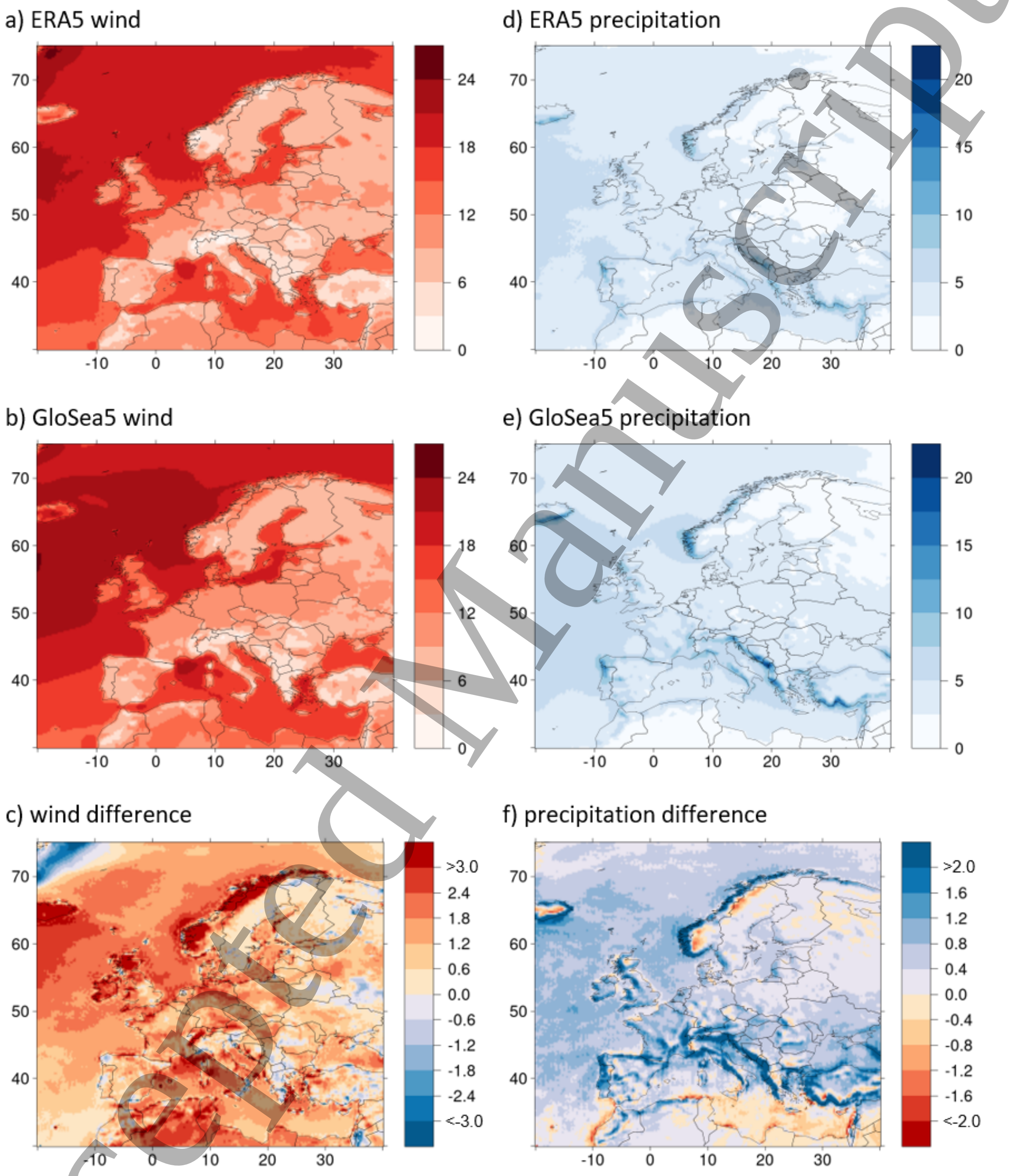
The mean 99th percentiles of wind speed from GloSea5 compare well with ERA5 spatially over Europe (figure 1a-c). However, the model mean overestimates wind speed by up to 9 ms^{-1} over areas of high topography, particularly over the Scandinavian mountain ranges and the Alps. Over most of Europe the model overestimates wind speed by less than 3 ms^{-1}, although there are small scattered regions with underestimations of up to 5 ms^{-1}.

The mean 99th percentiles of precipitation from GloSea5 also have a similar spatial pattern to ERA5 (figure 1d-f). However, GloSea5 underestimates extreme precipitation in the south of the Mediterranean, over Norway and Sweden and in a few small regions over Europe by up to 3 mm (although most of these regions are only underestimated by up to 2 mm). GloSea5 overestimates extreme precipitation around the north eastern coasts of the Mediterranean and western tip of Norway by up to 12 mm. Over most of Europe 99th percentile precipitation is overestimated by less than 2 mm.

Very similar results were found for daily 99th wind speed and precipitation percentiles (not shown).

3.1.1 Evaluation of the distributions

Quantile-quantile plots and histograms are made for three locations of interest: London, England; Madrid, Spain; and Stockholm, Sweden, to compare the distributions between GloSea5 and ERA5 (figure 2). These locations have been picked due to their large population sizes and therefore high



a) ERA5 wind

d) ERA5 precipitation

b) GloSea5 wind

e) GloSea5 precipitation

c) wind difference

f) precipitation difference

Figure 1: 99th percentile 3-hourly 10-m wind speed values (ms^{-1}) for a) ERA5 and b) GloSea5 mean and c) the absolute difference (GloSea – ERA5). d)-f) are the same but for precipitation accumulation values (mm). Scale limits have been calculated using 2×standard deviation of the differences.

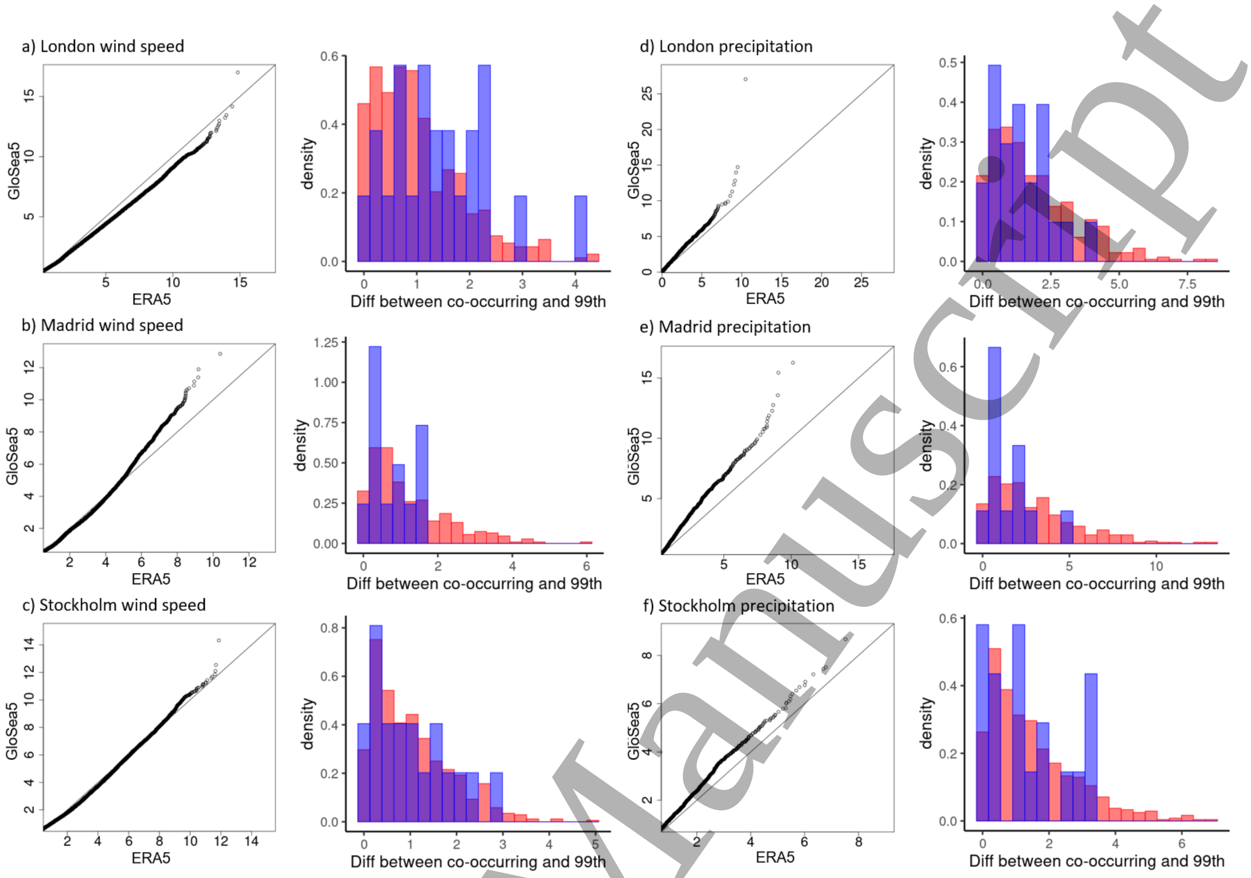


Figure 2: Quantile-quantile plots for wind speed and histograms of the difference between co-occurring extreme wind events and the 99th percentile for London (a), Madrid (b) and Stockholm (c). ERA5 are the blue bars, GloSea5 orange and purple where the datasets overlap. d)-f) are the same but for precipitation. Note that due to the smaller observational sample, a ERA5 single co-occurring extreme bar appears taller than a single model extreme.

risk (meaning they are of interest to the (re)insurance industry), as well as their spread over Europe. The Quantile-quantile plots show how well the quantiles of the events match between datasets. The histograms show the distributions of the co-occurring wind and precipitation events larger than the 99th percentiles.

GloSea5 underestimates wind speeds at London, except at the very extremes where wind speed is overestimated (figure 2a). GloSea5 co-occurring wind events larger than the 99th percentile are smaller than the observations, suggesting GloSea5 may not be producing extreme enough co-occurring wind events. For Madrid, GloSea5 overestimates wind speed at the upper tail (figure 2b). The distributions of the co-occurring wind events larger than the 99th percentile are similar between the model and observations. For Stockholm, GloSea5 estimates wind speed well, except at the very extremes where wind speed is overestimated (figure 2c).The distributions of the co-occurring wind events larger than the 99th percentile are very similar between the model and observations.

GloSea5 overestimates all quantiles of precipitation for all locations, with overestimation increasing largely in the upper tails (figure 2d-f). Co-occurring precipitation events above the 99th percentile are more extreme in the model, although this may be due to sampling, where there are no ERA5 events at the highest tails because our sample size is so small.

GloSea5 co-occurring events above the 99th percentile have a larger range with more events at the upper tail of the distributions than ERA5 for all of the locations, suggesting that the model can give us insight to unprecedented extreme compound events. It is worth noting that the scarcity of the extreme co-occurring events in the observations results in noisy distributions whereas the model provides a smoother distribution.

3.2 Co-occurring wind and precipitation extremes

Although substantial differences in magnitude are found between the model and ERA5 with the 99th wind and precipitation percentiles, the spatial pattern and frequency of the co-occurrence of precipitation and wind compare well between the GloSea5 mean and ERA5 (figure 3a-c). Although differences in frequency are found around and over some areas of high topography. Over the Scandinavian mountain ranges GloSea5 both under and over estimates $\hat{\chi}$ in distinct bands with significant differences of up to - 0.33 and + 0.27, indicating a shift in the location of the extreme co-occurring events from Sweden in ERA5 to Norway in the model (figure 3c and 3f). A similar pattern of differences between ERA5 and the model mean is also seen in Ridder *et al.* [2021]. Over the Norwegian Sea GloSea5 significantly underestimates $\hat{\chi}$ by up to - 0.1. Along the north eastern coasts in the Mediterranean GloSea5 overestimates $\hat{\chi}$ with differences of up to + 0.27. Generally most places experience differences of less than ± 0.08, which in many regions are not significantly different. Details on how significance is assessed can be found in the Supplementary.

These results are generally quite similar to previous studies [Ridder *et al.*, 2021]. Similar patterns of model mean overestimation over Scotland and underestimation over England are found, along with similar patterns of differences over the Iberian Peninsula. Although the GloSea5 model mean mainly underestimates compound precipitation and wind over central Europe whereas the model in Ridder *et al.* [2021] overestimates.

The values of $\hat{\chi}$ between the model mean and ERA5 have a correlation coefficient of 0.665 over land and 0.649 over sea (figure 3d-e). Over land the scatter is large with GloSea5 both under and overestimating $\hat{\chi}$. Over sea GloSea5 tends to underestimate $\hat{\chi}$.

Very similar results are found for the spatial pattern of co-occurrence between daily extremes (not shown). Although the difference between the datasets has a slightly larger range (most of Europe ± 25% difference) and the correlation coefficients were lower (0.522 over land and 0.485 over sea).

3.2.1 Co-occurrence and threshold

We investigate how $\hat{\chi}$ varies with extremal threshold (from the 90th to 99.9th percentiles) between the datasets at the three locations of interest (figure 4). See Supplementary for how the confidence intervals in figure 4 are calculated.

For London and Stockholm the model represents the frequency of all extreme co-occurring events well. The GloSea5 $\hat{\chi}$ follows the ERA5 $\hat{\chi}$ closely for almost all thresholds, where almost all GloSea5 $\hat{\chi}$ values are within the ERA5 95% confidence intervals (figure 4a and 4c). For Madrid, the Glosea5 $\hat{\chi}$ is larger than the ERA5 $\hat{\chi}$ and outside the 95% confidence intervals between 0.9 and 0.975 (figure 4b). However beyond 0.975 the model and observations match well, meaning that the model is representing the frequency of the most extreme co-occurring events well for Madrid.

All three locations show tail independence ($\hat{\chi}$ goes to 0 as the threshold tends to 1), meaning the largest extremes of wind and precipitation do not co-occur. The rate of convergence is different

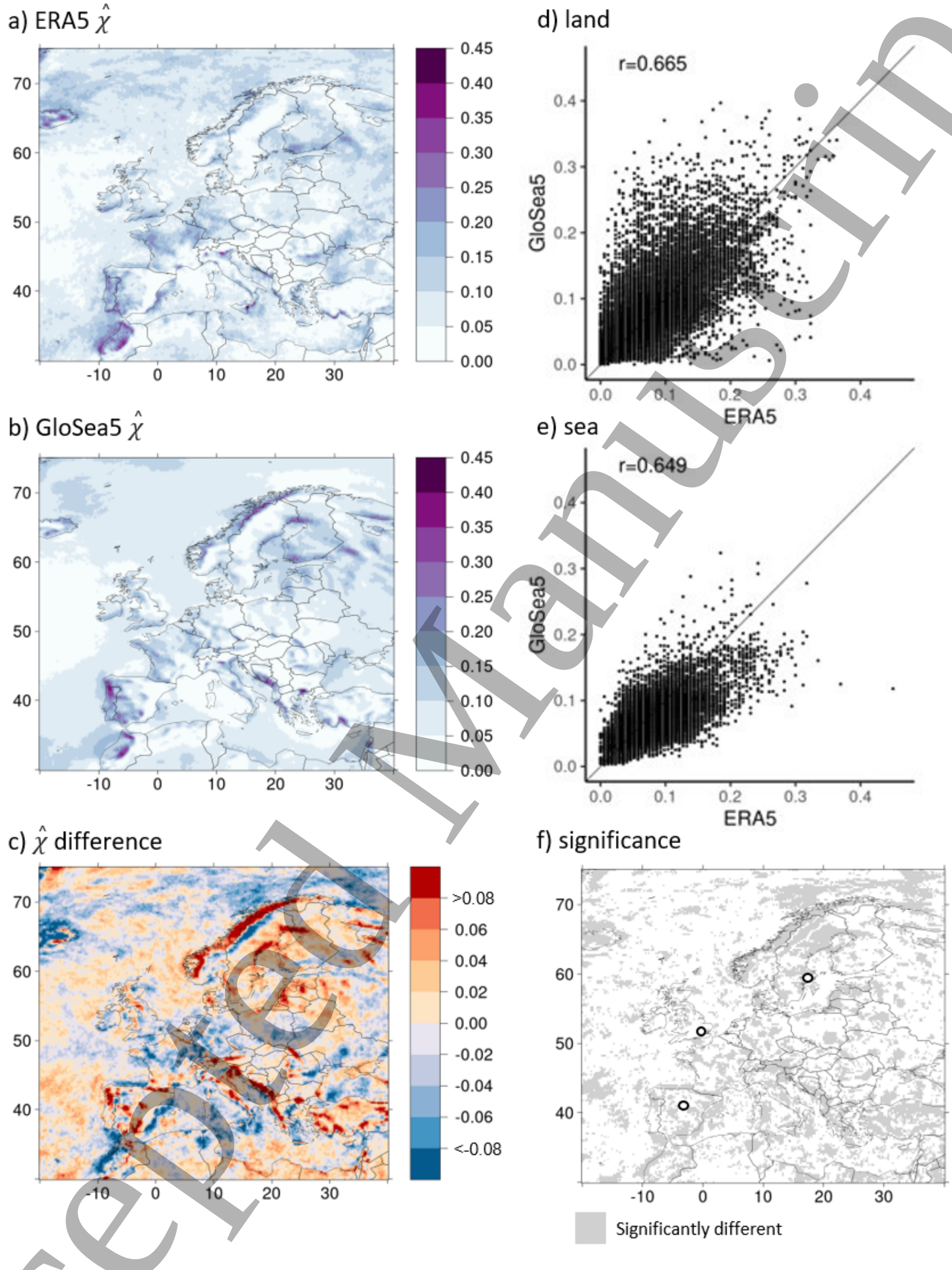
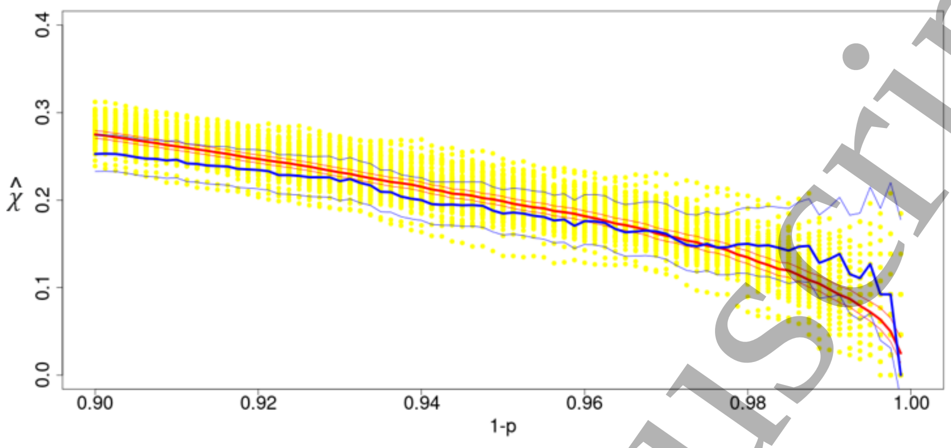
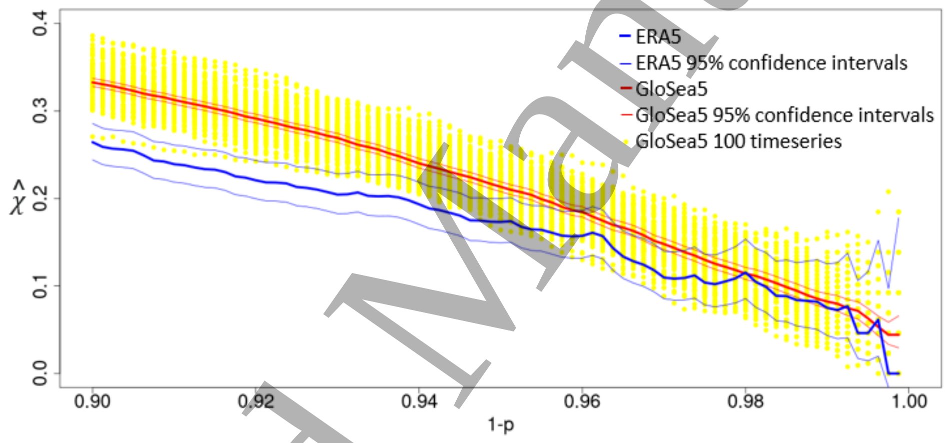


Figure 3: Extremal dependency, $\hat{\chi}$, maps for the 99th percentile 3 hourly extremes from a) ERA5 and b) the GloSea5 mean. Note in the independent case $\hat{\chi} = 0.01$. c) the absolute difference $(\hat{\chi}_G - \hat{\chi}_E)$ between GloSea5 and ERA5 $\hat{\chi}$. Scale limits have been calculated using 2×standard deviation of the difference. $\hat{\chi}$ values for ERA5 and GloSea5 mean for gridpoints d) over land and e) over sea, with corresponding correlation coefficients, r. f) Regions where ERA5 and GloSea5 are significantly different at the 5% level (greyed areas). Circles are London, Madrid and Stockholm.

a) London



b) Madrid



c) Stockholm

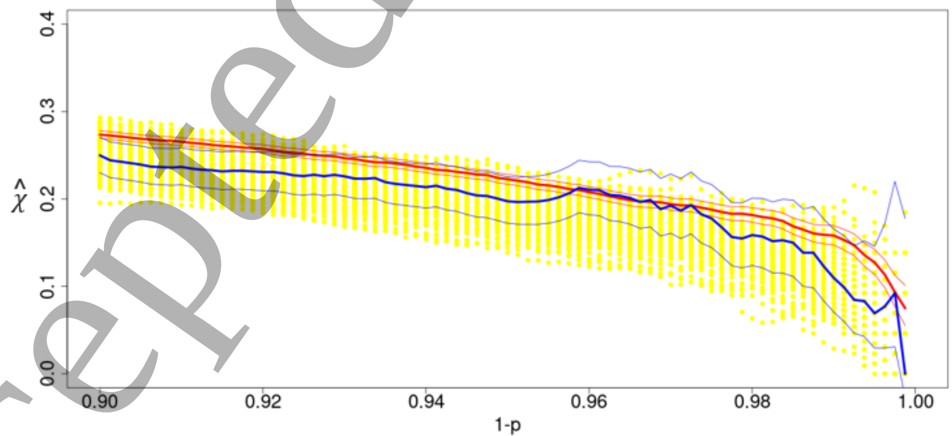


Figure 4: Threshold vs $\hat{\chi}$ for London (a), Madrid (b) and Stockholm (c). The blue lines represent ERA5 $\hat{\chi}$ and corresponding 95% confidence intervals, the red lines represent GloSea5 $\hat{\chi}$ and corresponding 95% confidence intervals and the yellow dots are 100 timeseries from the GloSea5 ensembles.

for each location. For Stockholm $\hat{\chi}$ converges very suddenly whereas at London and Madrid $\hat{\chi}$ converges more slowly.

Even at the largest threshold, some of the GloSea5 100 timeseries give very large values of $\hat{\chi}$ (>0.2) (yellow dots on figure 4). The model is producing $\hat{\chi}$ values that are unprecedented in the observations suggesting that it is possible to have a series of years that give very large $\hat{\chi}$ values at the upper thresholds and hence consist of a large number of very extreme co-occurring events.

The slope of GloSea5 is a smooth slope down to the highest thresholds, whereas ERA5 is variable due to the smaller sample size. Therefore GloSea5 can be used quantify the frequency of extreme co-occurring events in the very high thresholds better than the observations can.

3.3 Synoptic patterns

To evaluate the synoptic patterns between ERA5 and GloSea5, MSLP anomaly composites for 3 hourly co-occurring extreme events are made for the three locations of interest (figure 5). This anomaly is the difference from the DJF 24 year mean. For all three locations GloSea5 MSLP winter anomaly matches well to ERA5, suggesting that the extreme co-occurring events found in the model have been caused for the correct synoptic reasons.

When extreme co-occurring events occur in London, there is a large negative MSLP anomaly of 45 hPa over the North Atlantic north west of the UK (figure 5a). This indicates strong south-westerly flow over the English Channel in a region of the cyclone that is likely associated with frontal precipitation. GloSea5 shows a similar anomaly pattern with the largest magnitude negative anomaly of 36 hPa (figure 5d).

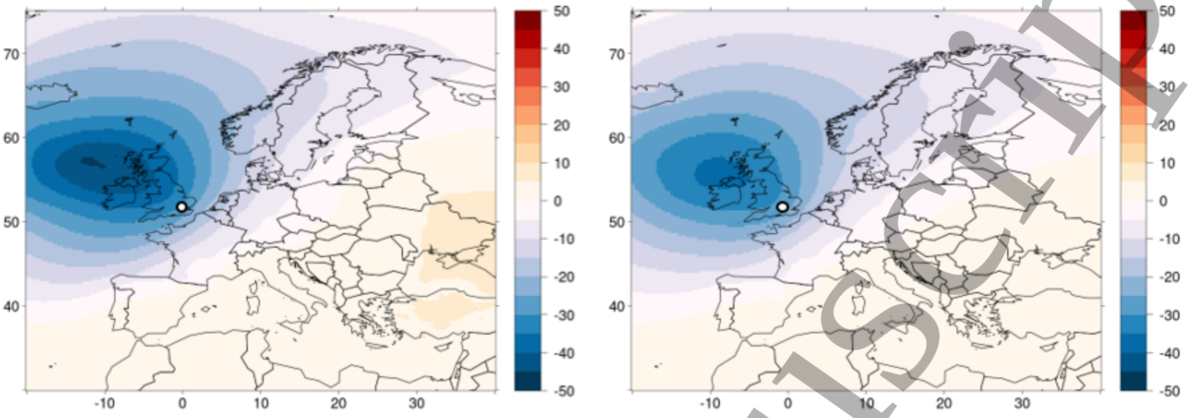
For extreme co-occurring events in Madrid, the negative MSLP anomaly is situated to the north of Spain, indicating westerly flow over the Iberian Peninsula. The anomaly is weaker than for the London case, at 32 hPa (figure 5b). The model shows a very similar pattern with the largest anomaly shifted slightly west with a smoother pattern due to the larger number of events (figure 5e). These patterns are consistent with Catto *et al.* [2010] where composites of strong cyclones showed the strongest winds occur to the south east of the cyclone where the warm conveyor belt is found.

For the Stockholm cases, the lowest MSLP is found directly to the west indicating winds that are more southerly with a MSLP anomaly of 25 hPa (figure 5c). This location is also likely associated with frontal zones and the different position relative to the cyclone centre associated with the time in the cyclone lifecycle and the poleward movement of the cyclones. GloSea5 shows a similar pattern with the largest magnitude negative anomaly of 28 hPa (figure 5f).

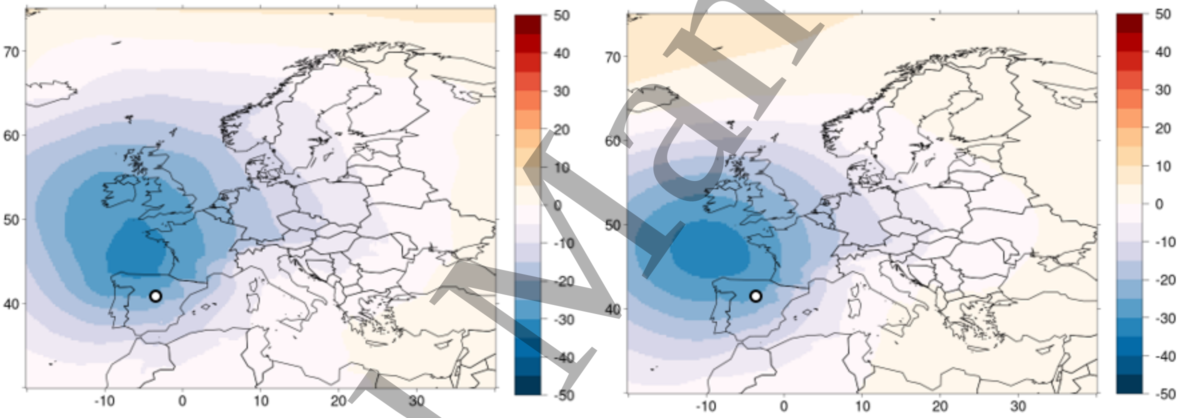
4 Discussion & Conclusions

This study has investigated how well the Met Office's GloSea5 seasonal forecast model ensembles represent extreme precipitation, extreme winds, and their co-occurrence and the synoptic situations leading to them by comparing them with ERA5. This is a difficult test for a model due to the number of factors that play a role. The main conclusions are given below in reference to the questions posed in the introduction, along with the main discussion points.

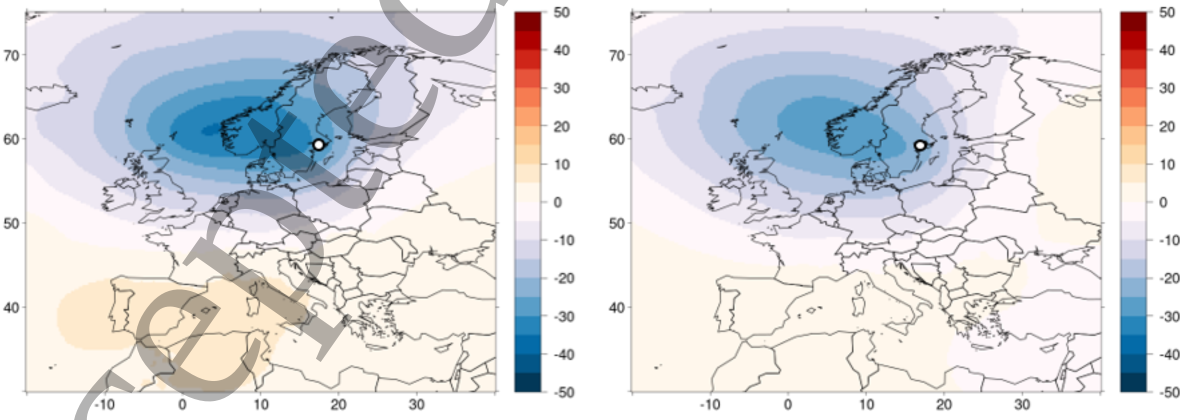
1. How well does GloSea5 represent the spatial pattern and magnitude of 3 hourly extreme wind and precipitation?


a) London 23 events

b) London 410 events M = 17.1, R = 8-25


c) Madrid 12 events

d) Madrid 342 events M = 14.3 R = 9-22


e) Stockholm 17 events

f) Stockholm 660 events M = 27.5 R = 22-35

Figure 5: Mean mslp DJF anomaly (hPa) when 3 hourly extreme co-occurring events occur at London, Madrid and Stockholm from ERA5 (a-c) and GloSea5 (d-f). The total number of extreme co-occurring events from ERA5 (left) and all 24 GloSea5 timeseries and the corresponding mean (M) and range (R) from the 24 timeseries (right) are written above each plot.

The 99th percentiles of 3 hourly 10 m wind speed and precipitation accumulation have been compared between ERA5 and GloSea5. The spatial pattern of these both compare well over Europe. However, for most of Europe the model overestimates extreme wind speed. This is particularly true over areas of high topography. Histograms and quantile-quantile plots for London, Madrid and Stockholm show that the model under and over estimates extreme wind and the number of extreme co-occurring wind events depending on location. At the most extreme thresholds these are always larger in GloSea5. For most of Europe the model overestimates extreme precipitation, which is true for London, Madrid and Stockholm. From quantile-quantile plots it is seen that extreme precipitation is particularly overestimated.

2. How well does GloSea5 represent the spatial pattern and frequency of 3 hourly co-occurring extreme wind and precipitation?

   The GloSea5 conditional probability of exceedance above the 99th percentile compares well spatially against ERA5 over Europe. And for most places the frequency of the events matches very well. Nevertheless, significant differences in frequency are found around and over some areas of high topography, particularly over Scandinavia, and the Norwegian Sea. For London, Madrid and Stockholm the model $\hat{\chi}$ values are larger for almost all thresholds (0.9 to 0.999), although for much of these the model is still within the ERA5 95% confidence intervals.

3. Can GloSea5 represent the correct synoptic situations leading to 3 hourly co-occurring extreme events?

   MSLP winter anomaly composites for extreme co-occurring events have been made for London, Madrid and Stockholm using ERA5 and GloSea5. The GloSea5 model creates the correct synoptic situations leading to extreme co-occurring events at all three of these locations. This is seen in the spatial pattern of the MSLP anomaly as well as the magnitude. The model could be particularly useful to look at weather systems associated with co-occurring extremes since the corresponding synoptic situations were found to be so similar to ERA5.

Although the 99th percentiles of precipitation and wind are generally overestimated in the model, $\hat{\chi}$ is underestimated over regions of western Europe, the Mediterranean and the Norwegian Sea. This means the model has more extreme wind (or precipitation) events that co-occur with non extreme precipitation (or wind) events in these regions. We investigated whether the model has a lag, where precipitation and wind values are out of phase with each other but found this is not the case for most locations (not shown).

Another reason $\hat{\chi}$ could differ from the observations may be due to a spatial offset in GloSea5, particularly within the cyclones causing these co-occurring events. The positioning of extreme precipitation within cyclones in the model may be inaccurate due to parametrizations of diabatic processes [Catto *et al.*, 2010]. Although the synoptic patterns themselves do look very similar between the model and observations which means that the dynamical features are well represented.

It is also worth noting that ERA5 may not be suitable as a benchmark for compound precipitation and wind events over complex terrain such as the Alps [Zscheischler *et al.*, 2021]. Zscheischler *et al.* [2021] showed ERA5 has different behaviour for $\hat{\chi}$ compared to high resolution weather model simulations over such regions. Owen *et al.* [2020] also found that ERA5 overestimated daily co-occurrence over high topography compared to observational data. Hence differences in $\hat{\chi}$ occurring over high topography may be down to inaccuracies in ERA5 rather than GloSea5.

Although precipitation and wind extremes can differ in magnitude largely between the model and observations, the frequency of co-occurring events compare much better, highlighting the potential of GloSea5 to investigate unprecedented and very rare extreme co-occurring events. The potential of GloSea5 is also seen in the larger sample of extreme and co-occurring events, which gives us smoother distributions and more confidence in our results. The results in this paper along with Ridder *et al.* [2021] and Zscheischler *et al.* [2021] give us confidence that climate models of different horizontal resolutions can simulate compound precipitation and wind extremes well. This suggests that such models can be used to investigate future changes in compound events as well as assessing the likelihood of unprecedented and very rare events.

Acknowledgements

L.E.Owen is supported by The University of Exeter's College of Engineering, Mathematics and Physical Sciences studentship.

N.J.Dunstone was supported by the Met Office Hadley Centre Climate Programme funded by BEIS and Defra

References

Agresti, A. and Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, **52**(2), 119–126.

Catto, J., Shaffrey, L., and Hodges, K. (2010). Can climate models capture the structure of extratropical cyclones? *Journal of Climate - J CLIMATE*, **23**, 1621–1635.

Catto, J. L. and Dowdy, A. (2021). Understanding compound hazards from a weather system perspective. *Weather and Climate Extremes*, **32**, 100313.

Catto, J. L. and Pfahl, S. (2013). The importance of fronts for extreme precipitation. *Journal of Geophysical Research: Atmospheres*, **118**(19), 10,791–10,801.

Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, **2**, 339–365.

De Luca, P., Hillier, J., Wilby, R., Quinn, N., and Harrigan, S. (2017). Extreme multi-basin flooding linked with extra-tropical cyclones. *Environmental Research Letters*, **12**.

De Luca, P., Messori, G., Pons, F. M. E., and Faranda, D. (2020). Dynamical systems theory sheds new light on compound climate extremes in europe and eastern north america. *Quarterly Journal of the Royal Meteorological Society*, **146**(729), 1636–1650.

Dorai-Raj, S. (2014). *binom: Binomial Confidence Intervals For Several Parameterizations*. R package version 1.1-1.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*.

Hénin, R., Ramos, A. M., Pinto, J. G., and Liberato, M. L. R. (2021). A ranking of concurrent precipitation and wind events for the iberian peninsula. *International Journal of Climatology*, **41**(2), 1421–1437.

Kent, C., Pope, E., Thompson, V., Lewis, K., Scaife, A. A., and Dunstone, N. (2017). Using climate model simulations to assess the current climate risk to maize production. *Environmental Research Letters*, **12**(5), 054012.

Kent, C., Pope, E., Dunstone, N., Scaife, A. A., Tian, Z., Clark, R., Zhang, L., Davie, J., and Lewis, K. (2019). Maize Drought Hazard in the Northeast Farming Region of China: Unprecedented Events in the Current Climate. *Journal of Applied Meteorology and Climatology*, **58**(10), 2247–2258.

Kumar, D., Mishra, V., and Ganguly, A. R. (2015). Evaluating wind extremes in cmip5 climate models. *Climate Dynamics*, **45**, 441–453.

MacLachlan, C., Arribas, A., Peterson, K. A., Maidens, A., Fereday, D., Scaife, A. A., Gordon, M., Vellinga, M., Williams, A., Comer, R. E., Camp, J., Xavier, P., and Madec, G. (2015).

Global seasonal forecast system version 5 (glosea5): a high-resolution seasonal forecast system. *Quarterly Journal of the Royal Meteorological Society*, **141**(689), 1072–1084.

Martius, O., Pfahl, S., and Chevalier, C. (2016). A global quantification of compound precipitation and wind extremes. *Geophysical Research Letters*, **43**(14), 7709–7717.

Owen, L. E., Catto, J. L., Stephenson, D. B., and Dunstone, N. J. (2020). Submitted, compound precipitation and wind extremes over europe and their relationship to extratropical cyclones. *Weather and Climate Extremes*.

Pfahl, S. (2014). Characterising the relationship between weather extremes in europe and synoptic circulation features. *Natural Hazards and Earth System Sciences*, **14**(6), 1461–1475.

Pfahl, S. and Wernli, H. (2012). Quantifying the relevance of cyclones for precipitation extremes. *Journal of Climate*, **25**(19), 6770–6780.

Priestley, M. D. K., Ackerley, D., Catto, J. L., Hodges, K. I., McDonald, R. E., and Lee, R. W. (2020). An overview of the extratropical storm tracks in cmip6 historical simulations. *Journal of Climate*, **33**(15), 6315 – 6343.

Raveh-Rubin, S. and Wernli, H. (2015). Large-scale wind and precipitation extremes in the mediterranean: a climatological analysis for 1979–2012. *Quarterly Journal of the Royal Meteorological Society*, **141**(691), 2404–2417.

Raveh-Rubin, S. and Wernli, H. (2016). Large-scale wind and precipitation extremes in the mediterranean: dynamical aspects of five selected cyclone events. *Quarterly Journal of the Royal Meteorological Society*, **142**(701), 3097–3114.

Ridder, N. N., Pitman, A. J., Westra, S., Ukkola, A., Hong, X. D., Bador, M., Hirsch, A. L., Evans, J. P., Di Luca, A., and Zscheischler, J. (2020). Global hotspots for the occurrence of compound events. *Nature Communications*, **11**(5956), 2041–1723.

Ridder, N. N., Pitman, A. J., and Ukkola, A. M. (2021). Do cmip6 climate models simulate global or regional compound events skillfully? *Geophysical Research Letters*, **48**(2), e2020GL091152.

Scaife, A. A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., MacLachlan, C., Martin, N., Nie, Y., Ren, H.-L., Roberts, M., Robinson, W., Smith, D., and Vidale, P. L. (2019). Does increased atmospheric resolution improve seasonal climate predictions? *Atmospheric Science Letters*, **20**(8), e922.

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., and Belcher, S. E. (2017). High risk of unprecedented uk rainfall in the current climate. *Nature Communications*, **8**(107).

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Hardiman, S. C., Ren, H.-L., Lu, B., and Belcher, S. E. (2019). Risk and dynamics of unprecedented hot months in south east china. *Climate Dynamics*, **52**(5), 2585–2596.

Vignotto, E., Engelke, S., and Zscheischler, J. (2021). Clustering bivariate dependencies of compound precipitation and wind extremes over great britain and ireland. *Weather and Climate Extremes*, **32**, 100318.

Wehner, M., Lee, J., Risser, M., Ullrich, P., Gleckler, P., and Collins, W. D. (2021). Evaluation of extreme sub-daily precipitation in high-resolution global climate model simulations. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **379**(2195), 20190545.

Zappa, G., Shaffrey, L. C., and Hodges, K. I. (2013). The ability of cmip5 models to simulate north atlantic extratropical cyclones. *Journal of Climate*, **26**(15), 5379 – 5396.

Zscheischler, J., Naveau, P., Martius, O., Engelke, S., and Raible, C. C. (2021). Evaluating the dependence structure of compound precipitation and wind speed extremes. *Earth System Dynamics*, **12**(1), 1–16.