Technical Memo

ECMWF
European Centre for Medium-Range Weather Forecasts

851

ECMWF Severe event catalogue for evaluation of multi-scale prediction of extreme weather

Linus Magnusson
(Forecast Department)

October 2019

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our website under:
http://www.ecmwf.int/en/publications

Contact: library@ecmwf.int

Abstract

Since 2013 ECMWF has been collecting material about severe weather events into a web catalogue, with the main aim to keep evaluation material for later use in order to better understand the performance. In this report we describe the outline of the cases in the catalogue and we discuss the lessons learnt so far from the evaluation in terms of multi-scale predictability of severe weather. By studying the cases, we can identify processes that are key for the prediction on different time-scales. By evaluating these processes we can learn about current bottlenecks for severe-weather prediction.

1 Introduction

Predicting extreme events on different time-scales is clearly a probabilistic forecasting problem. Starting from a point-in-time long before the event occurred, the probability distribution (PDF) from the forecast is likely to be very similar to the climatological distribution, with per se a low probability for extreme events. As we approach the event (week(s) before), some slight shifts of the forecasts PDF from the climatology might appear, either because a few members pick up an extreme scenario or most of the members are slightly shifted towards an anomaly due to some large-scale forcing. However, the PDF can still be very broad, and sometimes even broader than the climatological one due to the existence of extreme scenarios. Closer in time to the event (medium-range, 3 to 10 days ahead) the ensemble members centre around an extreme solution and in short-range forecasts (0-3 days) the ensemble (hopefully) sharpens around the (the later) observed value. However, if the magnitude of event is not within the envelope of what the model can simulate, the severity of the event will be missed also by the shortest forecasts. In this report we will discuss processes that influence the PDF on different time-scales and how to define short-comings in the prediction system.

How far ahead is a severe weather event predictable to a level where the forecast information is valuable enough to take action? This is a very obvious question from forecast users, but a very complex question to answer from a verification perspective. First of all, different users have different requirements on the forecast quality needed to take actions and these requirements might be different for different lead times, as forecasting extreme events is important on several time-scales ranging from seasonal to very short time-scales. Different types of extreme weather have different predictability and for similar type of events this can vary with season and location and other factors. Hence it is very difficult to give a generic answer on the question above from a user perspective. Another aspect of forecast verification is to find weaknesses in the prediction system to guide future developments. Here one needs to understand the underlying predictability mechanisms for extreme events to identify key processes to be improved, and such mechanisms varies with the lead time in mind.

Severe events are by definition rare, especially if we narrow the statistics to a short period (e.g a season) and/or a region (e.g Europe). One has to be aware of the trade-off between extending the sample to other seasons/continents and the risk of verifying extremes due to other type of meteorological circumstances. It is therefore difficult to reach a large enough sample to obtain reliable statistics in verification of extremes. Another complication for statistical verification is that severe events often are unique in their composition. For example, river flooding is often compounded by a series of precipitation events together with preconditioning hydrological conditions and geological properties, the life-threatening aspect of a heatwave is a combination of the magnitude and length of the period of extreme heat and coastal flooding can be a combination of storm surge, waves and precipitation. Taking such compositions into account further reduces the verification sample. It is therefore attractive to study cases individually. However, strictly speaking a probabilistic forecast for single extreme event cannot be verified. This is because there is no such quantity as 'true probability distribution' but only the outcome of the event. In

terms of Brier score it means that the reliability component cannot be evaluated for a single event, and one therefore need to aggregate statistics from many events to verify that aspect.

To be able to evaluate severe weather one instead needs to combine several approaches. In order to get reliable statistics for the verification one needs to lower the threshold for the event and focus on simple (non-compounded) events for which observations exists and are distributed, such as 24-hour precipitation, 10-metre mean wind and 2-metre temperature. Such verification of extreme weather for ECMWF forecasts was undertaken in Magnusson *et al.* (2014) and Bouallegue *et al.* (2019). The second approach is to evaluate the model climatology for extreme events to verify if the model is able to produce the extremes with the same frequency as in reality (Magnusson *et al.*, 2014). The third approach is to study individual cases, and from such work one can identify important aspects to verify further.

Since 2013 ECMWF has been collecting material about severe weather events into a web catalogue (Severe Event Catalogue, https://confluence.ecmwf.int/display/FCST/Severe+Event+Catalogue). The main aim of the catalogue is to gather a rich sample of events that can be used for diagnostic studies and forecast performance assessment. For this reason, the catalog is not necessarily a comprehensive archive of severe weather events around the world. The cases are subjectively selected based on media attention, interests from users and prospects for forecasting system evaluation. Since the start and up to autumn 2019 more than 170 cases have been collected, out of which around 100 are from Europe. Once every quarter, one case (or a few similar cases) is selected and forms the basis for an ECMWF newsletter article. Table 1 lists the newsletter cases.

Issue	Title
139	Windstorms in northwest Europe in late 2013
140	Forecasting the severe flooding in the Balkans
141	Recent cases of severe convective storms in Europe
142	Forecasts for a fatal blizzard in Nepal in October 2014
143	Forecasts for US east coast snow storm in January 2015
144	ECMWF forecasts for tropical cyclone Pam
145	Predicting this year's European heat wave
146	Forecasting flash floods in Italy
147	Wind and wave forecasts during Storm Gertrude/Tor
148	Forecasts showed Paris flood risk well in advance
149	Predicting heavy rainfall in China
150	Flash floods over Greece in early September 2016
151	The cold spell in eastern Europe in January 2017
152	ECMWF supports flood disaster response in Peru
153	Predictions of tropical cyclones Harvey and Irma
154	Two storm forecasts with very different skill
155	Predicting extreme snow in the Alps in January 2018
156	Forecasting convective rain events in late May
157	Forecasting the 2018 European heatwave
158	Predicting multiple weather hazards over Italy
159	Forecast freezing rain in Romania
160	ECMWF works with universities to support response to tropical cyclone Idai

Table 1: ECMWF Newsletter articles about severe weather events

The aim of this report is to summarize experience so far from the cases in the Severe Event Catalogue,

and to demonstrate different aspects that were important for the forecast performance in these type of cases on different time-scales. In Section 2 we present the outline of a typical event in the Severe Event Catalogue. In Section 3 we summarize processes important for different types of events on different time-scales. Finally the conclusions are discussed in Section 4. In Appendix A we use the European heatwave from 2015 and the Storm Desmond from the same year to illustrate the evolution of the forecast on different time-scales and important processes for the predictability.

2 Characterization of the severe events in the catalogue - Example from Italy October 2018

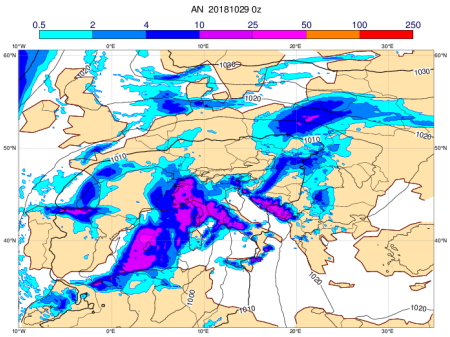
This section describes the outline of the cases presented in the Severe Event Catalogue. The amount and sort of material differs a lot from case to case in the catalogue, and here we will use some plots from a case of extreme wind and rainfall in Italy in October 2018 to serve as an example. A discussion around the case is given in ECMWF Newsletter 158, and a deeper summary in Cavaleri *et al.* (2019). The operational ECMWF products included in this section are further explained in the ECMWF User Guide (https://confluence.ecmwf.int/display/FUG/Forecast+User+Guide).

The first section for each case is **Impact** of the event. The information is usually collected from national hydrometeorological services, news articles on the web and/or Wikipedia. The aim is to give a short overview of the event and the damage caused by it. During the case illustrated here Venice was hit by a storm surge and with extreme waves offshore, while the southern Alps were severely affected of winds that uprooted millions of trees. The second section is **Description of the event** and presents the meteorological evolution of the event. It usually starts with ECMWF analysis maps of mean-sea-level-pressure and precipitation and/or 500hPa geopotential height and 850 hPa temperature. Examples of such plots are present in Figure 1 to illustrate the evolution of the cyclone and associated precipitation (left) and the propagation of the upper level trough that passed Italy in 29 October (right). Regularly satellite images are also included to visualise the evolution. The section can also include observations from national hydrometeorological services related to the extreme event.

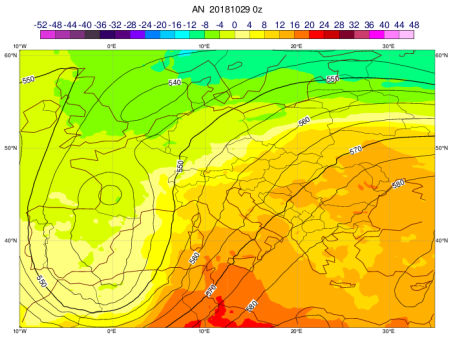
The **Predictability** section includes the subsections **Data assimilation, HRES, ENS, Monthly forecasts, Comparison with other centres** and in the case of flooding **CEMS/EFAS forecasts**. If the plots are presented as series with the same valid time, it starts with the plot from the time closest to the event and the end with the plot with the longest lead time.

The **Data assimilation** part is populated for cases with any interesting aspects regarding observation usage and/or large increment in connection with the extreme. This is often the case for tropical cyclones where we have targeted observations for the feature, but also if the short-range forecasts experienced large synoptic-scale errors.

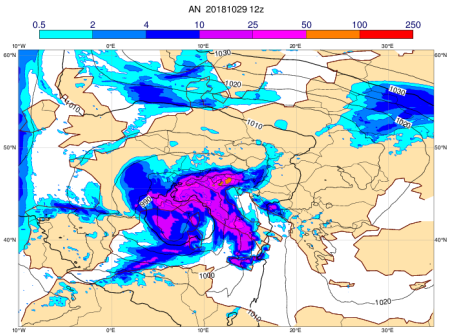
The **HRES** part includes maps of observations and from HRES forecasts with different initial times. Here the main focus is on short lead times to answer the question if the model with the highest operational resolution was able to capture the magnitude of the extreme, when the large-scale setting was captured. Sometimes time-series of the model and observations are included to evaluate the temporal evolution, especially for temperature-related events where the diurnal cycle is an important aspect. For this verification high-resolution (both spatial and temporal) observations are essential. Figure 2 shows observations of 24-hour precipitation and forecasts valid at the same time. The observation data set includes observations collected from national hydro-meteorological centres in the ECMWF HDOBS project (Haiden and Duffy, 2016). For Italy, we receive an extensive number of observations as can be seen in Figure 2(a).



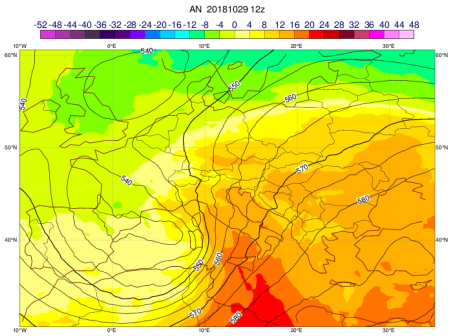
(a) 29 October 00UTC - MSLP/Precip



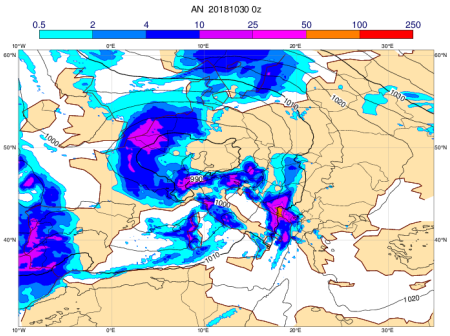
(b) 29 October 00UTC - z500/t850



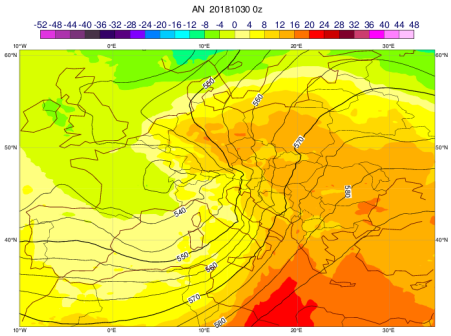
(c) 29 October 12UTC - MSLP/Precip



(d) 29 October 12UTC - z500/t850

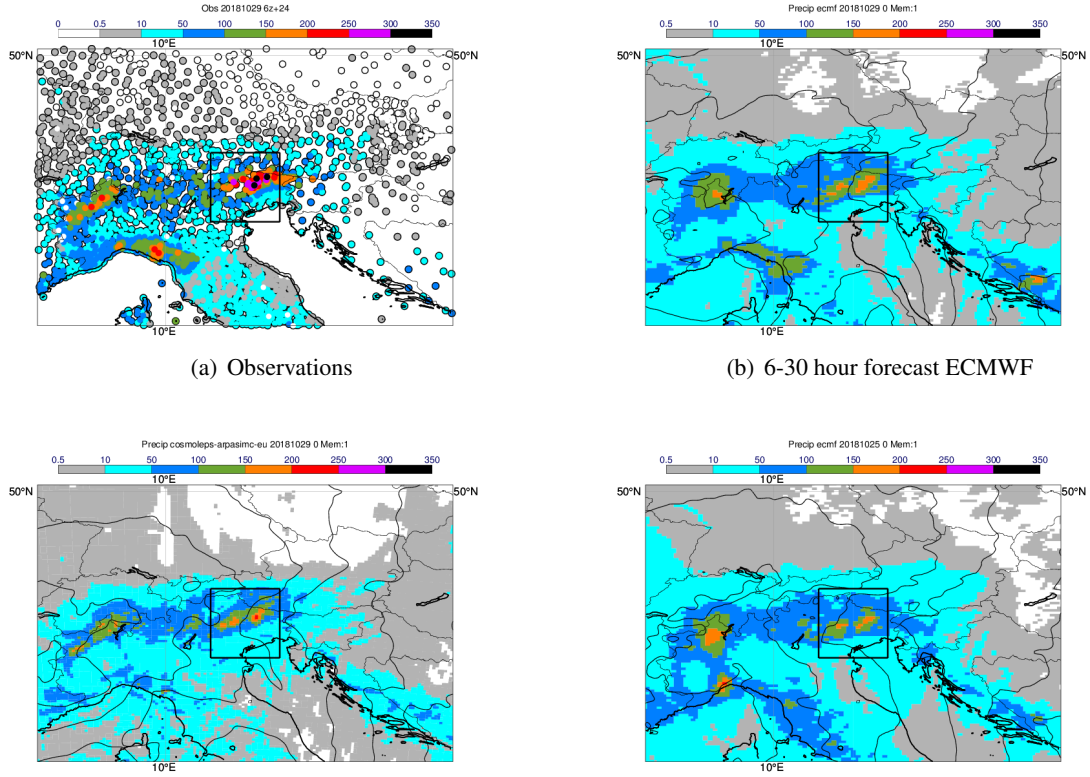


(e) 30 October 00UTC - MSLP/Precip



(f) 30 October 00UTC - z500/t850

Figure 1: Analyses of MSLP, z500 and t850, and 6-hour forecast of precipitation.



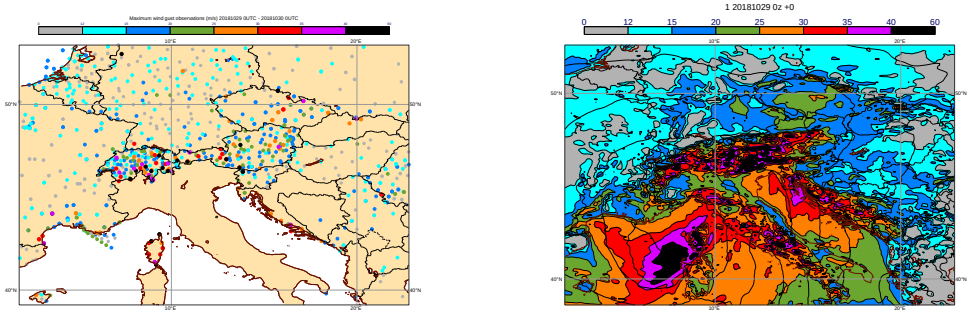
(a) Observations

(b) 6-30 hour forecast ECMWF

(c) 6-30 hour forecast COSMO-LEPS

(d) 102-126 hour forecast ECMWF

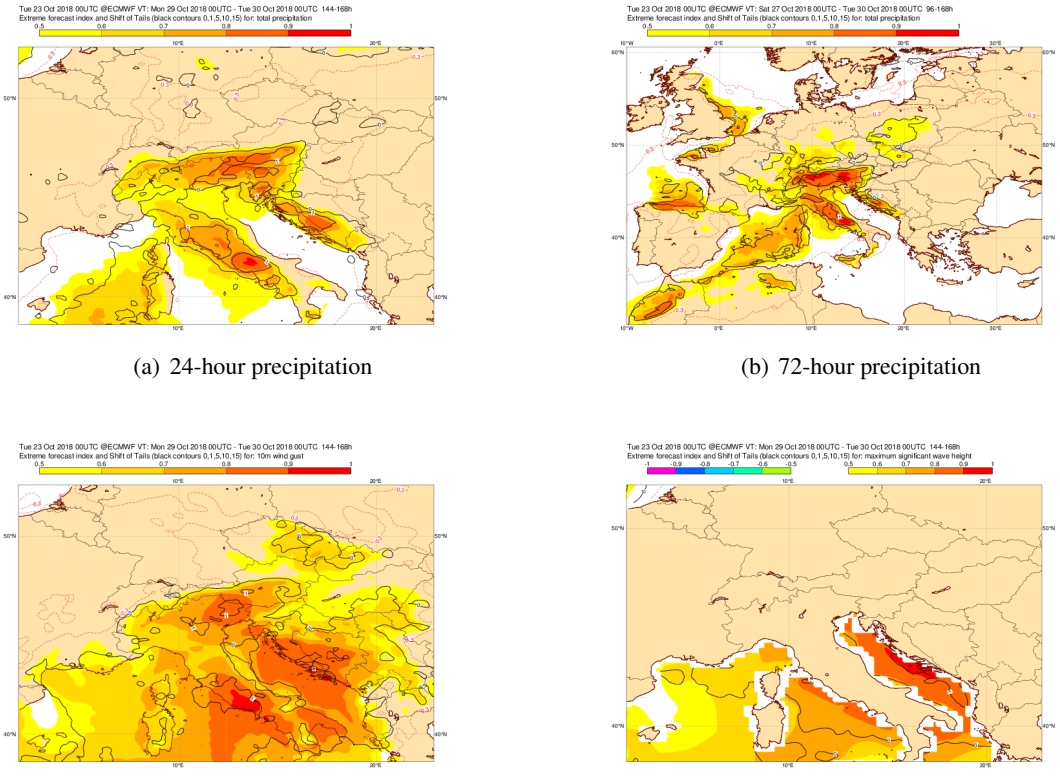
Figure 2: Observations and forecasts for 24-hour precipitation valid 29 October 06UTC to 30 October 06UTC.



(a) Observations

(b) 0-24 hour forecast ECMWF

Figure 3: Observations and forecasts for 24-hour maximum wind gusts valid 29 October 00UTC to 30 October 00UTC.



(a) 24-hour precipitation

(b) 72-hour precipitation

(c) 24-hour maximum wind gusts

(d) 24-hour maximum wave height

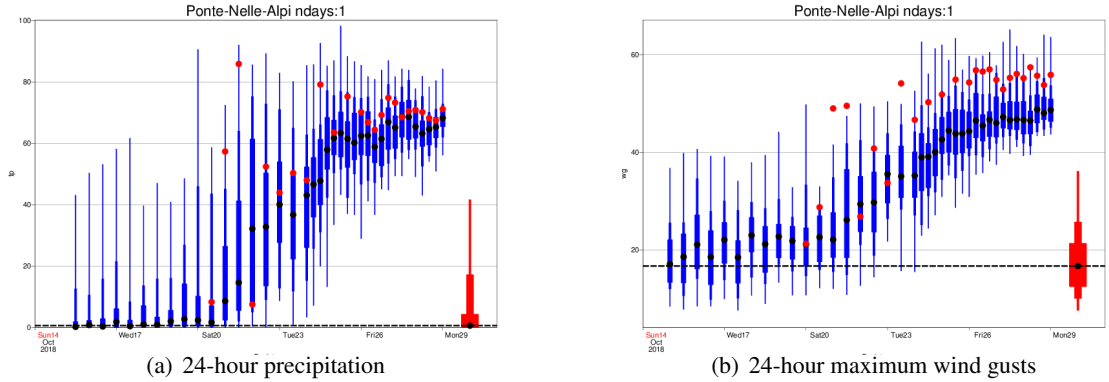
Figure 4: EFI and SOT in forecasts from 23 October 00UTC for precipiation, wind gusts and wave height valid 29 October (24-hour) for 27-29 October (72-hour).

For the short-range HRES forecasts, the pattern of the observed precipitation is captured with the highest numbers over the southern slope of the Alps, but the magnitude is underestimated.

Figure 3 shows observations and a short forecast for 24-hour maximum wind gusts. 24-hour maximum wind gusts is not a reported quantity in the SYNOP messages and here we have aggregated the quantity from the available observations of maximum wind gusts in the SYNOP observations. This introduces inconsistencies between stations and countries, for example due to different period length to find the maximum for each observation. As apparent in Figure 3, we are currently missing observations of wind gusts from Italy. The model wind gusts is diagnostic rather than a prognostic model variable and is calculated from a parameterisation dependent on the 10-metre mean wind, surface momentum flux, and in the case of convection the vertical wind shear (ECMWF, 2017).

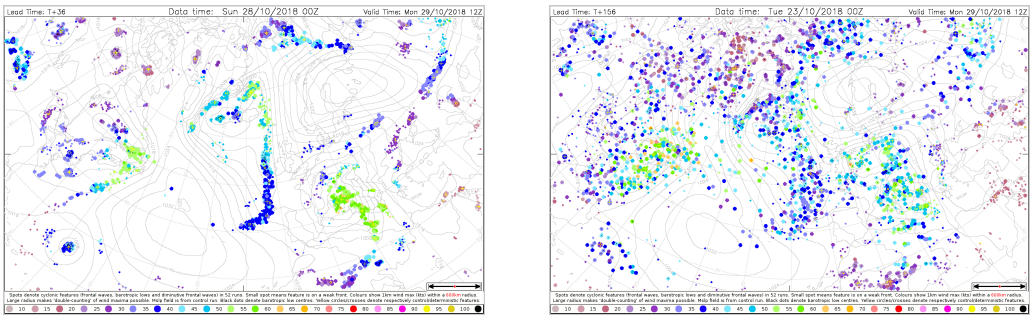
The **ENS** part includes plots of Extreme Forecast Index (EFI, Lalaurette (2003)) and Shift of Tails (SOT) from different initial times valid for time of the event. Depending on the length of the event, either 1-day, 3-day or 5-day averages of the quantities are used. Figure 4 shows examples of EFI and SOT for different parameters and accumulation lengths. The section sometimes include probabilities for some threshold relevant to the event. For the case in the figures, the EFI signals for extreme precipitation, wind gusts and waves were present in the shown forecast from 23 October.

A new diagnostic, in this section, is the diagrams showing the evolution of the forecast ensemble distribution, as in Figure 5. All available forecasts, valid for a specific time or time-window, location and parameter are taken into consideration. The ensemble forecast distribution is illustrated by a box-and-



(a) 24-hour precipitation (b) 24-hour maximum wind gusts

Figure 5: Forecast evolution for 24-hour mean accumulated precipiation and maximum wind gusts in the box outlined in Figure 2 for ensemble (blue box and whisker) and HRES (red dot) valid 29 October. The model climate is included as red box-and-whisker. The bars indicate the 1st, 10th, 25th, 75th, 90th and 99th percentile, and the black dot the 50th percentile.



(a) 28 October 00UTC (b) 23 October 00UTC

Figure 6: Cyclone features (dots) from the 50 ensemble members plus control (yellow circle outline) and HRES (yellow cross) with the colour indicating the maximum wind at 1 km height inside a radius of 600 km valid 29 October 12UTC. Small dots indicates features on weak fronts. The plots also include MSLP from then control forecasts.

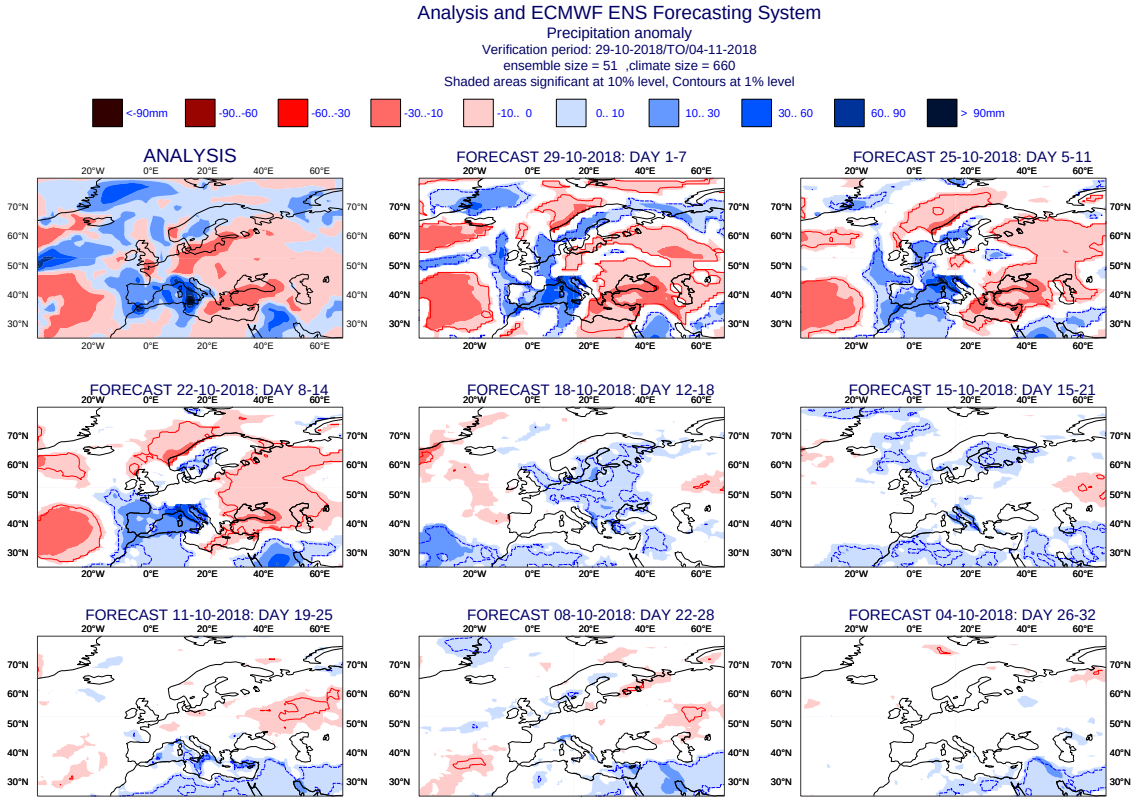


Figure 7: Verification of 7-day precipitation anomaly valid for 29 October to 4 November.

whisker symbol where the mid-line gives the median, the wide box spans the 25-75th percentile, the narrow box 10-90th percentile and the whiskers 1-99th percentile (or ensemble minimum to maximum if less than 100 members). The HRES forecast is represented by red dots. To put the forecasts into a climatological perspective, the model climate distribution, based on the reforecast data set (Vitart, 2014), is included as a box-and-whisker symbol located on the verification date. The climate distribution represents the model climatology at medium-ranges. It is important to note that the model climate at longer lead times, due to biases, can be significantly different from the model climate at medium range. The diagrams in Figure 5 indicate that the ensemble started to be shifted toward the anomalous values from around 22 October. This type of diagnostic visualizes the evolution of forecast signal, including the ensemble spread, as the forecast lead time decreases. The diagram helps identifying the longest lead time at which the forecast produced a signal in terms of a significant shift from the climate distribution. As a consequence it can be used to estimate the predictability level of severe events within the limitation of the current prediction system.

In cases of extremes connected to extra-tropical cyclones, feature plots from Cyclone Data Base (CDB) are included (Hewson and Titley, 2010). From these plots one can see the spread in the position and in the intensity in the ensemble. Figure 6 shows cyclone feature plots for the maximum wind inside a 600 km radius for two initial dates valid at the same time. In the forecast from 23 October, a lot of the ensemble members have a feature present in the central Mediterranean, but with a considerable spread, while the features are much more confined in the short-range forecast from 28 October.

The part on **Monthly forecast** includes operational plots of weekly anomalies. Such plots are not optimal if an event happened close to a change in calendar week (as the case of Italy when 29 October was on

	Extended range (2-4 weeks)	Medium range (3-10 days)	Short range (0-3 days)	Example of ECMWF evaluations
European heatwave	Soil moisture, SST anomalies	Rossby wave packets	Local heating and evaporation	Newsletter (NL) 145, NL 157
European cold spells	Teleconnections from MJO, SSW to negative NAO	Transition to negative NAO	Surface inversion, precipitation type	NL 151, NL 155, NL 159
NW. European windstorms	Predict positive NAO	Jet-stream propagation	Timing of development, wind gust parameterisation	NL 139 and NL 147
NW. European extreme rainfall	Predict positive NAO	Presence of atmospheric rivers	Exact position of system, strength of orographic precipitation	See Case 2 in Appendix A
S. European extreme precipitation	Negative NAO	Development of cut-off low over Mediterranean	Convection, orographic precipitation	NL 146, NL 150, NL 155 and NL 158
Severe convection	-	Upper-level trough, position of fronts, CAPE and wind shear	Convective triggering, organisation, lifetime, wind gust parameterisation	NL 141, NL 146 and NL 150

Figure 8: Important aspects for predicting extremes on different time-scales. For ECMWF Newsletters, see https://www.ecmwf.int/en/publications/newsletters .

a Monday), or the event was too short to impact the weekly anomaly. This part can also include MJO forecasts if the MJO was considered as a factor influencing the predictability for the event. Figure 7 shows the analysis (top-left) and extended-range forecasts of precipitation anomaly valid for 29 October to 4 November. The picture from the weekly averages agrees with the results in Figure 5, with a good signal from 22 October but not much earlier in the extended-range.

The part **Comparison with other centres** are usually based on forecasts from the TIGGE/S2S archives (Bougeault *et al.*, 2010; Vitart and Robertson, 2018) or forecasts obtained from member states. In Figure 2, an example of ensemble member from the COSMO-LEPS limited-area ensemble has been included to see if the magnitude is better captured with the higher (7-km) resolution.

If the extreme event includes flooding (flash-flood or river), **CEMS/EFAS** products are included in their own sub-section. The plots are usually captured from the EFAS (www.efas.eu) web site.

At the end of the case page we try to add links to similar cases in **Experience from general performance/other cases**, and to list the **Good and bad aspects of the forecasts for the event** as a summary. If we received reports from external sources we add them under **Additional material**.

3 Summary of key factors for predicting extremes on different time-scales

Based on the experience with the Severe Event Catalogue, Figure 8 lists key factors for predicting a selection of extreme event types on extended-range (2-4 weeks; sometimes referred to sub-seasonal), medium-range (3-10 days) and short-range (0-3 days) time-scales for several types of severe weather. Two examples are given in Appendix A where we show key factors for predicting the extreme events on the different time-scales. In the sub-sections below we will review these factors. Many of the forecast deficiencies discussed below are further explained on the ECMWF page for "Known IFS forecasting

issues" - https://confluence.ecmwf.int/display/FCST/Known+IFS+forecasting+issues .

3.1 European heatwaves

On the extended-range time-scale, the temperature anomalies are sensitive both to local soil moisture (Ferranti and Viterbo, 2006; Fischer *et al.*, 2007; Vitart and Robertson, 2018) and to the large-scale flow regimes over Europe. The former directly affects apportioning between sensible and latent heat-fluxes; and in turn cloud cover can also be affected, which itself will provide a negative feedback (in summer) by reflecting back insolation. It is therefore important to correctly initialise the soil conditions in extended-range forecasts (Dirmeyer, 2011). For the large-scale flow, persistent blocks over Europe favour heatwaves. To set up such atmospheric flow conditions in the summer, among other factors, the sea-surface temperature (SST) in north-eastern Atlantic plays a role (Wulff *et al.*, 2017).

To predict onsets of the large-scale flow patterns, medium-range predictability of Rossby Wave Packets (RWP, (Wirth *et al.*, 2018)) are important (Fragkoulidis *et al.*, 2018). On average the presence of long-lived RWP increases the predictability (Grazzini and Vitart, 2015). However, the prediction of the packet propagation is sometimes affected by uncertain elements such as organised convection and/or rapid cyclogenesis, resulting in bad forecasts of onesets of blocking patterns (Rodwell *et al.*, 2013; Magnusson, 2017). There is also a difference in the processes (advection, subsidence, diabatic processes) leading to heatwaves in different parts of Europe as discussed in (Zschenderlein *et al.*, 2019), and the influence of these processes need to be captured in the medium-range forecasts.

Although the relatively high predictability for the large-scale flow, short-range forecasts often result in large 2-metre temperature errors. During heat waves, the ECMWF model has problems to simulate the amplitude of the diurnal cycle. One reason could be that the model does not currently include urban tiles and hence misses the extra heating due to tarmac and concrete; these factors are discussed in e.g Hogan *et al.* (2017).

3.2 European cold spells

Cold spells over Europe are often caused by large-scale flow patterns bringing cold air from north and east, such as Scandinavian blocking and negative North-Atlantic Oscillation (negative NAO, sometimes referred to as Greenland blocking) (Ferranti *et al.*, 2018). These two regimes disrupt the westerly flow towards Europe and replace it with strong meridional flow and often anticyclonic conditions. Under anticyclones, strong surface inversions can form in calm and clear conditions leading to extremely cold wintertime temperatures.

On the extended-range time-scale, these large-scale flow patterns have teleconnections from Madden-Julian Oscillation (MJO) (Cassou, 2008) and/or sudden-stratospheric warmings (SSW) (Baldwin and Dunkerton, 2001). If these precursors are predictable (Vitart, 2014), we should expect some predictability for the flow regimes if the teleconnections are sufficiently captured by the model. The predictive skill of the wintertime regimes and the conditional dependence of skill on MJO was recently presented in (Ferranti *et al.*, 2018).

The transition into a blocked flow-regimes is often related to Rossby wave breaking (Woollings *et al.*, 2008). The ability of extended-range forecasts to, in a climatological way, capture the link between Rossby wave breaking and formation of blockings has recently been evaluated in Quinting and Vitart (2019). The formation and maintenance of these regimes are also suggested to be linked to diabatic

processes related to warm-conveyor belts (Wernli and Davies, 1997; Grams *et al.*, 2011). These onsets are sensitive to small errors in the upstream flow and such a case is discussed in Grams *et al.* (2018).

As for the summer heat-waves, also short-range forecasts during cold spells often experience very large 2-metre temperature errors, and in most cases related to strong surface inversions. The inversions are difficult to simulate in models due to insufficient vertical resolution in the boundary layer. The shallow nature of these inversions lead to large temperature errors for relative small errors in energy fluxes. It could also be the case that physical parameterisations such as vertical diffusion are not perfectly suited for these extreme conditions.

In connection with cold spells, severe weather in the form of heavy snowfall/blizzards can occur. However, the meteorological features causing the extremes can be very different. As an example, proximity to seas or large lakes can give rise to intense, local snow showers on the coasts. With parameterised convection, it is difficult to capture the advection of such showers onto land, and often also the short-range forecasts miss the large amount of snow on the coasts.

Another uncertainty related to precipitation during cold conditions is to predict the precipitation type (Gascon *et al.*, 2018). Here fine boundaries between rain, snow, sleet and even freezing rain make a huge difference in the severity of the event, even if the precipitation rate is similar. An additional complexity here is heavily populated coastal areas with high vulnerability, which magnifies the uncertainty of the impact due the precipitation type.

3.3 North-western European windstorms

While predicting the exact track and timing of Northern European windstorms on the extended-range time-scale is impossible, forecasting the increased likelihood of the features is the target on this time-scale, and predicting the NAO is key (Donat *et al.*, 2010) as its positive phase favours cyclone tracks towards north-western Europe. The positive phase of NAO has a statistical teleconnection from enhanced convection in the Indian ocean due to MJO (Cassou, 2008).

As in the case to make medium-range predictions of many other extreme weather types in the mid-latitudes, capturing RWP is also important for windstorms (Wirth and Eichhorn, 2014), in order to predict the risk of downstream developments that can form extreme cyclones. Explosive developments are often associated with upper-level divergence by the jet stream ('left-exit'), and here the key to capture the phasing of the jet-stream and the lower level cyclones.

For shorter time-scales, another difficulty for predicting extreme winds is to capture structures that cause wind maximum gusts, such as sting jets (Hewson and Neu, 2015) and embedded convection caused by dry intrusions (Raveh-Rubin and Wernli, 2015b). As global models still rely on parameterisation of wind gusts, this naturally causes uncertainties. But the problem of capturing wind gusts also appears in convection-permitting models (Pantillon *et al.*, 2018).

3.4 Precipitation extremes due to North-Atlantic cyclones

Connected to extra-tropical cyclones are so called atmospheric rivers: bands of high mean transport moisture that can bring extreme rainfall when ascending over orography (Ralph and Dettinger, 2011). Lavers *et al.* (2017) showed that using water-vapour flux to trace atmospheric rivers is a good predictor of high precipitation events during positive NAO conditions during European winters. Therefore the extended-raneg predictions of these rely on the same mechanism as the windstorms discussed above.

To capture the magnitude of the precipitation over orography, sufficient model resolution is needed together with accurate model microphysics to capture the time-scale of the rain-formation (Forbes *et al.*, 2015).

3.5 Precipitation extremes in southern Europe

Precipitation extremes in the northern Mediterranean are often connected to large-scale upper level troughs Nuissier *et al.* (2011); Raveh-Rubin and Wernli (2015a) together with interaction with local orography. Statistically the precipitation over the Mediterranean is negatively correlated with the NAO (e.g Trigo *et al.* (2004); Vergni *et al.* (2016); Tsanis and Tapoglou (2019)).

The cases in the database show that a reasonable signal of extreme precipitation compared to the model climatology appears well into the medium-range, due to the prediction of the large-scale troughs. These large-scale precursors are often part of a Rossby wave packet (Martius *et al.*, 2008), and have been shown to have good predictability (Grazzini, 2007). The extreme precipitation often appears on the eastern side of the trough, connected to strong moisture flux (atmospheric rivers). The extreme precipitation often occurs due to orographic enhancement when the moist air is forced to ascend (Khodayar *et al.*, 2018). The presence of atmospheric rivers is largest in the autumn (Lavers and Villarini, 2013). The predictability of the extremes in the medium-range is dependent on the convective influence in the precipitation extremes, with lower predictability in summer-time when the convective part is stronger (Grazzini *et al.*, 2019). An open question is whether the troughs over the Mediterranean have a regime-like behaviour and are predictable on the extended-range time scale.

In the short-range forecasts the details in the moisture flux, convective initialisation and interactions with orography is important factors to predict the extremes (e.g Gascon *et al.* (2016). These factors are still difficult for global models where the convection is parameterised and the orography is not sufficiently resolved.

3.6 Severe convection

Severe summer-time convection over Europe often results in intense rainfall, hailstorms, severe wind-gusts and on rare occasions tornadoes. However, on the extended-range time-scale it is more difficult to find any key features that would give an early indication of these features. Instead in the medium-range the key feature is to find unstable air that often ahead of cold fronts. To identify such features, the Extreme Index Forecast Index for CAPE has been found to be a useful approach. As convective cells are favoured by a vertical wind shear, another EFI index that is a combination of CAPE and the lower tropospheric vertical wind shear has been developed and tested (Tsonevsky, 2015). The atmospheric models cannot explicitly predict lightning, but promising results with a parameterization is discussed in Lopez (2016).

Above it was stated that the magnitude of wind gusts in connection to synoptic cyclones are reasonably well captured by the model. This is not the case for wind gusts in connection to organised convective systems, that usually appear in summer-times over Europe. Capturing the true magnitude of such systems is challenging for the global models as they do not explicitly resolve convection and the associated downdrafts. Instead convective indices calculated from model quantities are still a useful tool such as the EFI products mentioned in the previous paragraph. Another challenge is the variability inside a grid-cell that can be large for precipitation in convective systems. One way to account for this effect is adopted and outlined in Pillosu and Hewson (2017). This method broadens the PDF to account for the variability

not resolved by the model, and is dependent on the orography and weather situation (more sub-grid scale variability due to mesoscale events compared to synoptic events).

The global models with parameterised convection also have problems to capture the timing of the convection and often miss the extension into the evenings.

Another risk during severe convection is flash-floods (pluvial floods). Here the interaction between large-scale forcings, the moisture flux (atmospheric rivers), orography and convective triggering creates favourable conditions, but also the soil wetness before the event plays a role. Dependent on these factors the predictability of the extreme rainfall can be very different. It is also important here to have a surface models that correctly model the surface runoff and local storage of water.

4 Summary

In this report we have demonstrated the usefulness of case studies to determine factors that influence predictability on different time-scales (short, medium and extended-range). As it is often problematic to obtain a sample large enough for statistical verification of extreme events, identifying key aspects that underpin the predictability of the extremes enables verification of these aspects on a regular basis. Case studies also help to identify how to stratify the statistical verification depending on the meteorological driver of the extreme. One example is wind gusts, which have relatively high predictability if they are associated with mature cyclones but are very unpredictable if they are caused by mesoscale convective systems. However, one needs to be aware of the risk of ignoring evaluation of false alarms by only evaluating cases that became extreme in reality.

The extended-range predictability of extremes comes from long-lived flow patterns in the mid-latitudes together with teleconnections from tropics/stratosphere and boundary conditions such as SST and soil moisture. For medium-range forecasts it is important to capture the large-scale evolution of the atmospheric flow, which requires a good global data assimilation. For short-range forecasts the local data assimilation is important as well as physical parameterisations in the model that are suitable for extreme conditions. However, one also needs to keep in mind that uncertainties even in the shortest forecasts cannot be eliminated due insufficient knowledge about the current state and due to variability inside the grid-box.

In this report we have discussed the evolution of an ensemble PDF as the event of interest approaches. Predicting extreme weather several days in advance is clearly a probabilistic problem, and we have in the appendix of this report shown examples at PDFs on different lead times before an event. The overarching target in ensemble forecasting is to issue as narrow (sharp) PDF as possible whilst maximising the reliability and keep a desirable consistency in the PDF. These two properties (resolution and reliability) can for example be evaluated with the decomposition of the Brier score (Wilks, 1997), but need a large sample as discussed in the introduction.

The fundamental question is how to improve the prediction of extreme events on all time-scales and where to put the resources in terms of research and computer power. To obtain a good reliability by minimising frequency biases, one need to simulate the event with the right climatological frequency, i.e the PDF of the model climate needs to be close to the PDF of the true climatology, including the tails. It is common that the magnitude of the simulated extremes are limited by the model resolution, and increased resolution has in the past improved such biases. Here limited-area models play a role to better resolve the extreme; as well as post-processing techniques to adjust the forecast PDF. However, the frequency bias can also be associated with deficiencies in the model physics connected to the extremes (e.g wind gust

parameterisation and orographic precipitation), and here improved model physics can help to improve the simulation of the extremes.

To increase the sharpness of the PDF without losing reliability, one needs to decrease the forecast error to allow a more confident ensemble. To do this a key is to reduce the analysis error by improving the components involved in the data assimilation, such as observation usage, modelling of background error statistics, minimisation algorithms and also the model used for the first guess forecast. To obtain a reliable ensemble, an accurate simulation of the initial uncertainties is needed as well as simulating the model uncertainties. As predicting extremes in the medium-range is often dependent on resolving the extreme tail of the PDF, a larger ensemble is needed compared to only focusing on the ensemble mean (Leutbecher, 2018).

Finally, to capture signals from boundary conditions (soil, sea-temperature etc.) on the extended-range time-scale we need to include all relevant earth-system modelling components. We also need to make sure that the model is capable to simulate the teleconnections from the sources of predictability.

All these points are associated with resources in terms of research and operational constraints (i.e ensemble configurations and computer power), and the operational forecasting centres need to find a good balance to progress. By evaluating a range of extreme weather events, the current bottlenecks for improving the forecasts can be identified.

The evaluation of the cases here only covers the physical aspects of the forecasts and not the warnings based on the forecasts and the anticipation of the information. This aspects will be evaluated within the WMO/WWRP Hiweather project (Zhang *et al.*, 2019). The future plan is for ECMWF to collaborate with other partners in the project in order to cover the evaluation of the full forecasting chain of hazardous weather events. This would increase our knowledge of the value of the forecast information and shortcomings in the system.

The ultimate question is whether the forecast contains enough information for the user to take preventive actions to reduce the risk of losses due to extreme weather. One simplistic way to answer this question is to evaluate the Potential Economical Values (Richardson, 2000). Albeit building on a simple model of the cost for actions and potential losses, this type of verification can indicate the type of actions for which forecasts are useful. In Magnusson *et al.* (2014) forecasts for moderate extremes (98th percentile) were verified using this metric, and it was found that only action with a relatively low cost compared to the prevented loss is worth taking based on medium-range forecasts. However, in reality, the preventive actions associated with expectations of extreme weather a week ahead is about preparations and redistribution of resources, and these actions are relatively cheap. It is rather in the day(s) just before the event, the relatively expensive actions need to be taken. With a lead time dependency of the cost/loss ratio, it could well be that the forecasts actually are as (or even more) useful in the medium-range than in the short-range and make multi-scale prediction of extreme weather important.

Acknowledgements

We would like to thank Laura Ferranti, Tim Hewson, Florian Pappenberger and David Richardson for their help with this report.

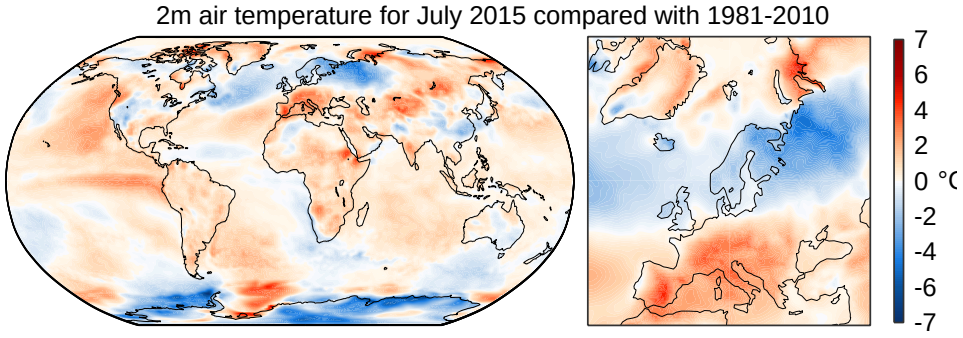


Figure 9: 2-metre temperature anomalies from July 2015 from ERA-Interim.

A Appendix A - Examples of diagnostics for extreme events

In this section we will use two cases to illustrate different diagnostics to understand extremes and give examples of additional experiments that can be run after the events.

A.1 Example 1 - Heat-wave in Western Europe 2015

The summer of 2015 was dominated by very warm weather in southern and western Europe. The warm anomalies clearly stand out on the map of July 2-metre temperature anomalies from ERA-Interim (Figure 9). For this case we will use Paris as the example to illustrate the predictability of the 2-metre temperature. Figure 10 shows time-series of the temperature at 12UTC in Paris (both daily values and 7-day running mean), compared to the climatology. Although most of July was dominated by warmer than normal weather, it was not one continuous heat period, but rather several shorter periods of heat with colder periods in between. In this example we will zoom in on the forecasts for 1 July, which turned out to be the most extreme day in terms of temperature for the summer. The observed temperature at 12 UTC was between 35.5°C and 36.8°C among SYNOP stations in Paris, and later that day one station reached 39.7°C: the second warmest temperature on record for the city. The extreme heat was primarily caused by a ridge that developed over south-western Europe at the end of June (Figure 11) that advected very warm air northward. This event was included as the severe weather event in ECMWF Newsletter 145.

Figure 12 shows the probability density functions (PDFs, a) and cumulative distribution functions (CDFs, b) for 2-metre temperature forecasts valid 1 July 12UTC in Paris, initialised at 00 UTC on 1 June, 18 June, 22 June, 27 June and 1 July. The panels also include the model climatology derived from the ensemble reforecast data set (Vitart, 2014). Starting from the forecast from 1 June, one month before the

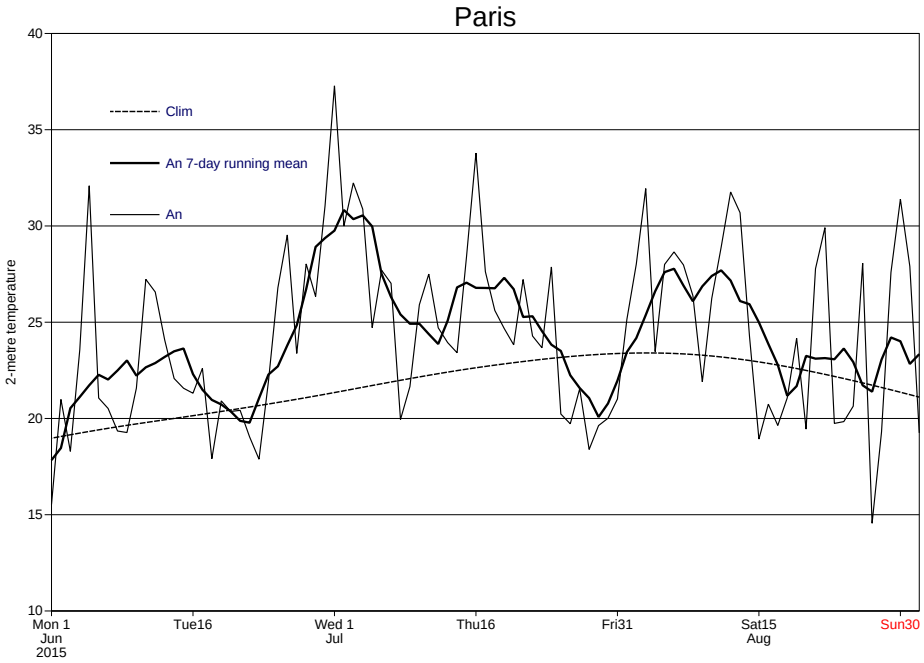
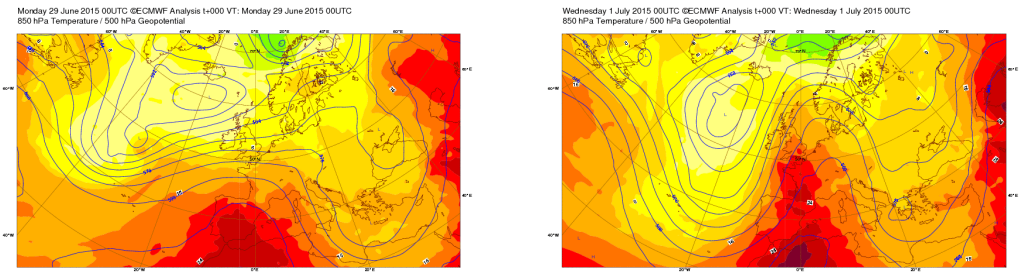


Figure 10: 2-metre temperature at 12UTC in Paris from analyses during the summer 2015. Daily values (thin), 7-day running mean (thick) and climatology (dashed).



(a) 29 June (b) 1 July

Figure 11: Analyses of z500 (contour) and t850 (shade) from 29 June (top) and 1 July (bottom).

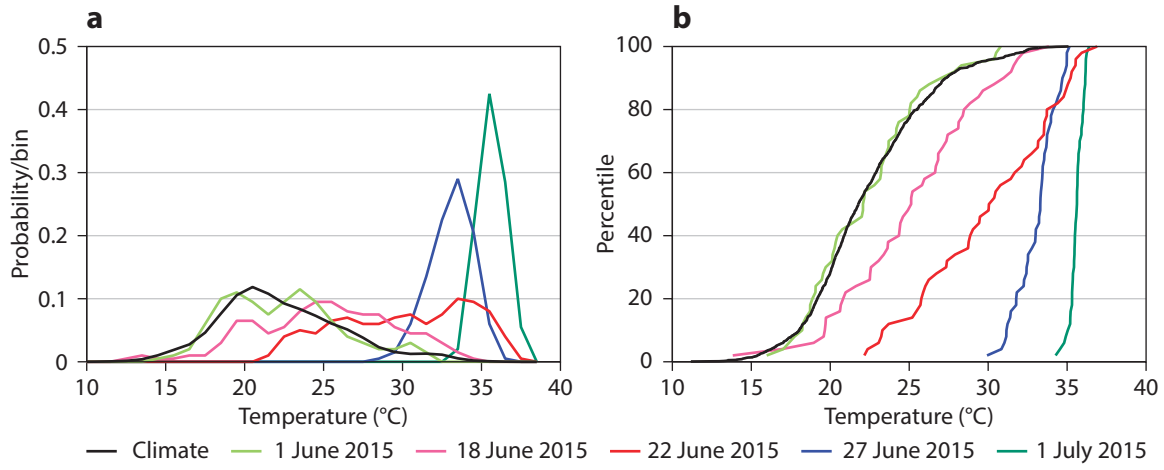


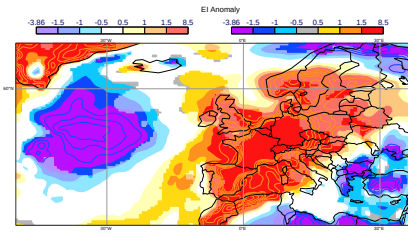
Figure 12: Ensemble forecasts valid 12 UTC on 1 July in Paris visualised by (a) probability density functions (PDFs) for forecasts initialised at 00 UTC on 1 June, 18 June, 22 June, 27 June and 1 July and (b) cumulative distribution functions (CDFs) for forecasts initialised at the same times.

event, one should not expect the forecast to show a strong signal, and indeed the forecast is very similar to the model climatology. Going forward in time and inspecting the PDF from the 18 June, 2 weeks before the event, the PDF has shifted towards warmer temperatures. Looking at the weekly temperature anomaly over the week 1-7 July (Figure 13, see below), a broad-scale warm anomaly was predicted over south-western Europe. This predicted anomaly resembles well the anomaly that was dominating the summer.

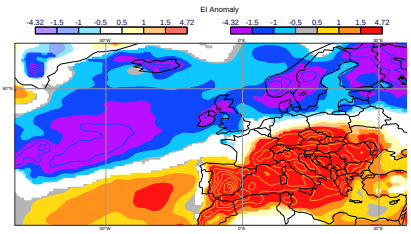
To understand the origin of this anomaly, we have run sensitivity experiments where we (1) replaced the observed SST with climatological values and (2) replaced the initial soil conditions with the conditions of the year before. The results in terms of 2-metre temperature anomalies are presented in Figure 13. The left column in the figure shows the results from the forecast initialised 18 June and valid for the week of 29 June - 5 July, and the right column a composite of 8 different initial dates in June and July 2015 averaged 3-4 weeks into the forecast.

Starting with the forecast from 18 June, the verifying analysis valid 2 weeks later had a strong warm anomaly for more or less the whole of Europe. The southern and eastern part of the anomaly was captured in the control experiment, while the signal was missing over the British Isles and Scandinavia. For this case, the S2S forecasts show similar levels of predictive skill to the one of ECMWF system. In the experiment with the SST anomalies removed, the southern part of the 2-metre temperature anomaly including the warm signal over the Mediterranean sea closely related with the local SST, disappears. In the experiment initialized with the soil conditions of 2014, the eastern part of the anomaly vanishes. This indicates that both the SST and soil anomalies played a role for the anomalies in this forecast.

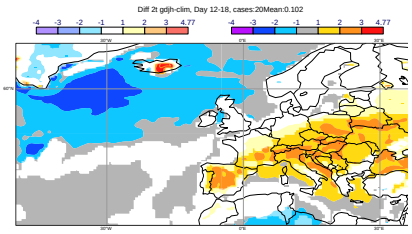
Figure 13(right) shows the ensemble mean 2-metre temperature anomalies of week 3-4 for 8 forecasts initialised between 15 June and 7 July. The control forecasts that used both observed SST and the soil conditions from 2015, captured well the observed warm anomaly over southern Europe. This was also the case in the experiment that used climatological SST, indicating that the SST anomalies did not play a significant role for the anomaly over southern Europe (but to a significant degree for other anomalies on northern hemisphere). In the experiment with the soil conditions replaced with the one from 2014, the 2-metre temperature anomaly over southern Europe is much reduced and mainly persists over sea. This results suggests that the warm anomaly in the 3-4 week forecast over the full summer at least was partly caused by the dry soil conditions starting from the beginning of the summer. Such anomaly would make



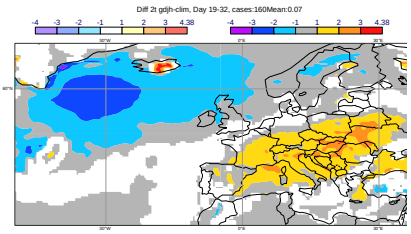
(a) Analysis week 2, 1 date



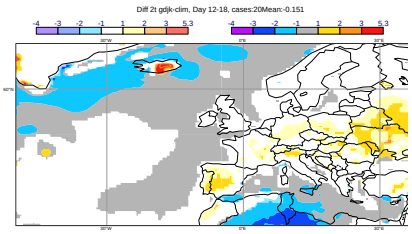
(b) Analysis week 3-4, 8 dates



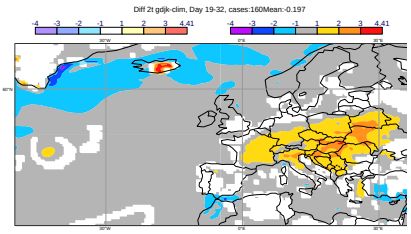
(c) Control week 2, 1 date



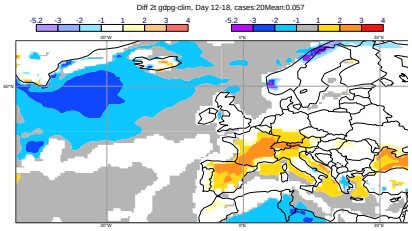
(d) Control week 3-4, 8 dates



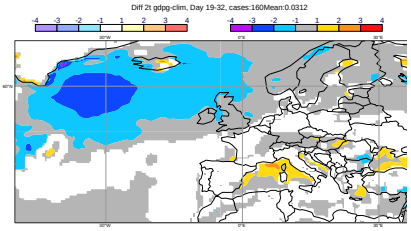
(e) SST clim week 2, 1 date



(f) SST Clim week 3-4, 8 dates



(g) Soil 2014 week 2, 1 date



(h) Soil 2014 week 3-4, 8 dates

Figure 13: 2-metre temperature forecast for 29 June - 5 July from 18 June (left column) and forecasts from 8 different initial dates in June and July 2015 averaged 3-4 weeks into the forecast (right column). Anomalies from analysis (1st row) Control experiment with persisted SST and initialised soil moisture (2nd row), experiment with SST climatology (3rd row) and experiment with soil moisture from 2014 (4th row).

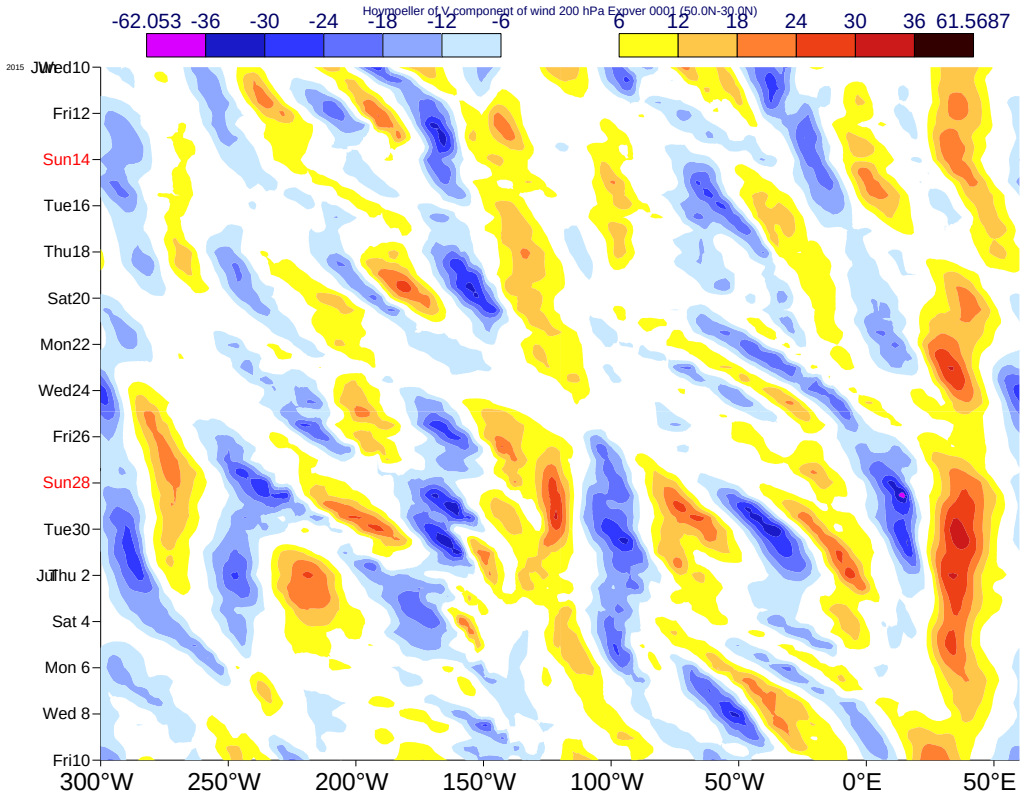


Figure 14: Howmoller diagram of 200 hPa meridional wind averaged between 30°N-50°N from analyses spanning 10 June to 10 July.

a heat-wave more likely but will not tells the exact period it is going to happen. (Ferranti and Viterbo (2006); Fischer *et al.* (2007))

The next ensemble PDF in Figure 12 is from 22 June (9 days before the event). Here the PDF is clearly shifted towards warmer temperatures and is also skewed towards the extreme tail of the distribution. The initialisation of this forecast coincides with the time of the first detection of a Rossby wave packet over Western Pacific. The propagation of the Rossby wave packets are visualised in Figure 14 by plotting a Howmoller diagram of the 200 hPa merional wind averaged between 30°N-50°N. The packet first appeared over the Western Pacific around 22 June and propagated eastward. The packet reached eastern Atlantic in the last days of July with a positive node (winds from south) over western Europe. By capturing the wave packet propagation, the probability for the ridge over western Europe increased and with that the probabilities for the heat-wave. The effect of the presence of Rossby wave packets on the mid-latitude predictability is documented in e.g. Grazzini and Vitart (2015), who showed that the predictability is increased by the presence of long-lived Rossby wave packets. Three days before the event (27 June 00UTC) the ensemble PDF was narrow and centred at the warm end of the climate distribution around 34°C. At this point in time the trough over central Atlantic that later pushed the warm airnorthward had started to develop, giving a strong confidence in the development of the heat-wave.

In the forecast produced on the morning of 1 June, only small uncertainties remained in the forecast. However, when we compare the ensemble forecast with observed temperatures from stations across

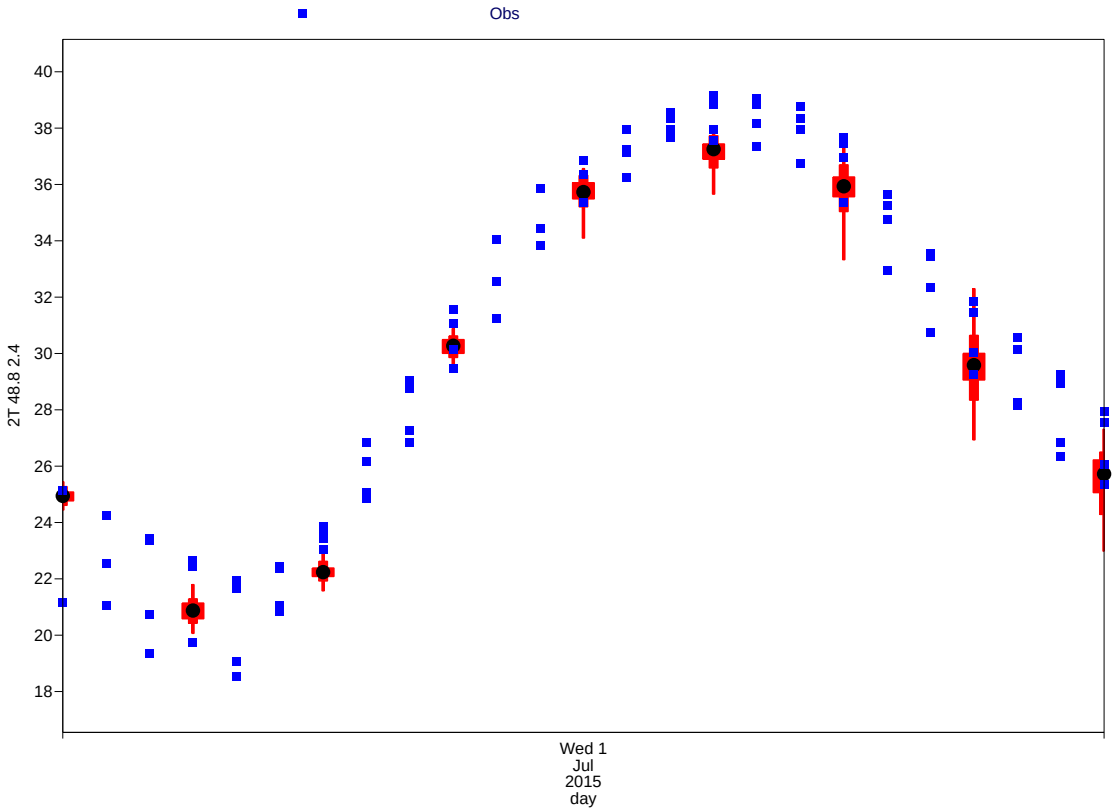


Figure 15: Ensemble forecast from 1 July 00UTC (red box-and-whisker) and observations (blue symbols) for 2-metre temperature in Paris on 1 July.

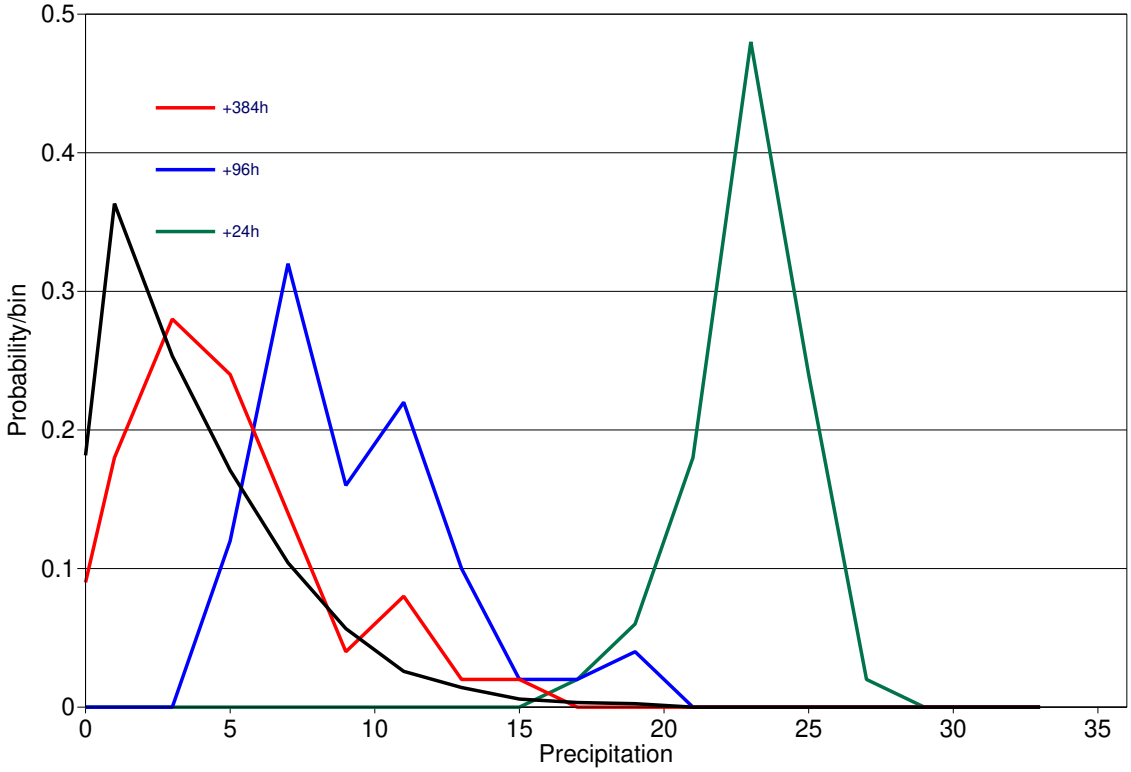


Figure 16: PDF for 24-hour precipitation over the northern Great Britain, defined as 53°N-58°N, 10°W-0°W, on 5 December from the reforecast climatology (black) and the ensemble PDF from the forecast issued 19 November (red), 1 December (blue) and 4 December (green).

Paris, we find that the maximum temperature was underestimated by all ensemble members, as seen in Figure 15. One plausible explanation is that the model does not include urbanisation and can therefore not capture heat islands in cities. But the ECMWF verification of 2-metre temperature also shows a more general underestimation of the diurnal cycle over Europe during summer-time that is currently under investigation.

To summarize this case, we found indications that anomalies in soil-moisture and Atlantic sea-surface temperatures shifted the PDF towards warmer temperatures on the extended-range time-scale. In the medium-range, capturing the Rossby-wave packet seems to have been the key to predicting the meridional flow that brought warm air northward over western Europe. However, short-range forecasts also suffered from an underestimation of the maximum temperatures.

A.2 Example 2 - Storm Desmond

On 5 December 2015, the extra-tropical cyclone Desmond caused severe flooding, travel disruption and a power outage across northern England and parts of Scotland (Ferranti *et al.*, 2017), and locally more than 340 mm of rain over 24 hours was reported (Honister Pass, Cumbria). Flooding and travel disruptions were reported from Ireland as well and the cyclone also led to flooding in southern Norway. In this

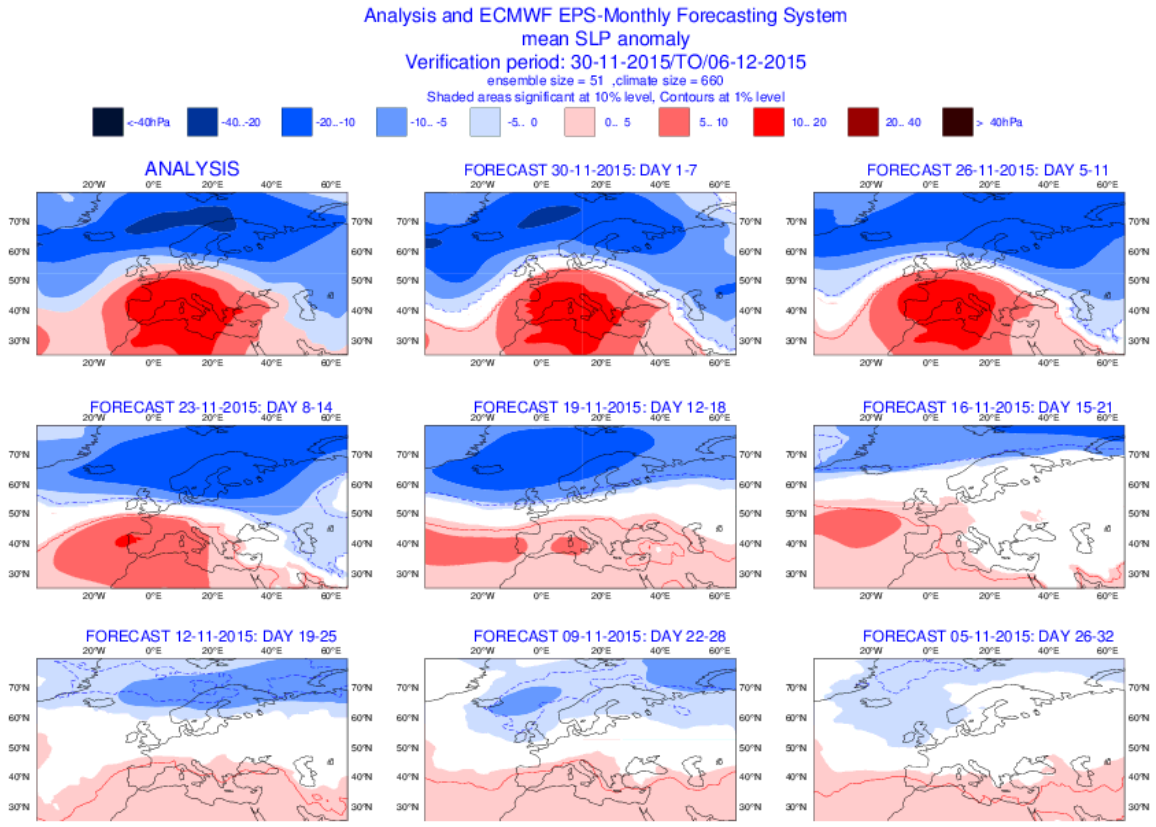


Figure 17: MSLP anomalies for 30 November to 6 December from analysis (first panel) and ensemble mean from monthly forecasts.

section we will focus on the prediction of the rainfall in a region over the northern Great Britain, defined as 53°N-58°N, 10°W-0°W, and Figure 16 shows the PDF for 24-hour precipitation on 5 December for the climatology from the reforecasts and the ensemble PDF from the forecast issued 19 November, 1 December and 4 December.

Already the ensemble PDF two weeks before the event was slightly shifted towards more wet conditions. The weekly ensemble mean anomalies of MSLP from the monthly forecasts for the week 30 November-6 December (Figure 17) shows a positive North-Atlantic Oscillation (NAO) pattern in the analysis and the shortest forecasts. But also at longer lead-times (3-4 week before the event), the ensemble mean indicated a positive phase of NAO, which increase the likelihood for unsettled weather in north-western Europe (Donat *et al.*, 2010). The accurate prediction of the NAO phase for this case raises the question about the origin of the signal. One candidate is the Madden-Julian Oscillation (MJO). Looking into the MJO forecast from 19 November (Figure 18), we see that an MJO was predicted to be in the phases 2-3 in the beginning of the forecast. Climatologically, MJO phase 3 is a precursor to a positive NAO with a lag of 10 days (Cassou, 2008). However, to tap this predictability one need a model that is able to capture this teleconnection. Figure 19 shows the teleconnection pattern from the model climate from ECMWF Seasonal forecasting system 4 and ERA-Interim reanalysis, both based on 30 years of data. The result shows that the model is at least partly able to capture this linkage.

The forecast PDF from 1 December was clearly shifted towards wetter conditions, but not necessarily extreme. Using a product like the Extreme Forecast Index - EFI (Lalaurette, 2003), we see a band of the higher risk for extreme rainfall over the Atlantic and towards the British Isles. This was connected to an

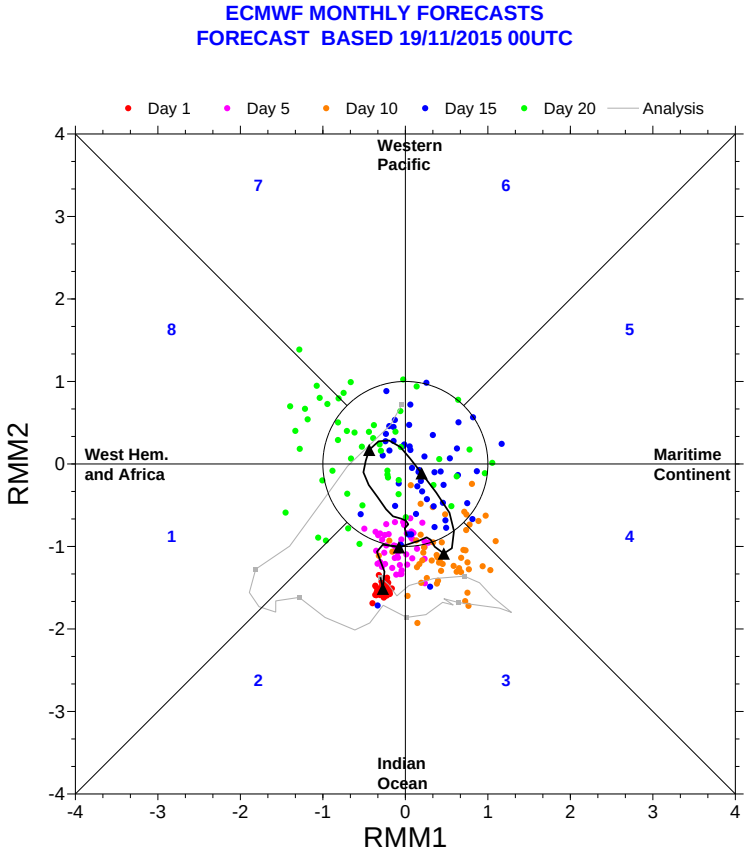
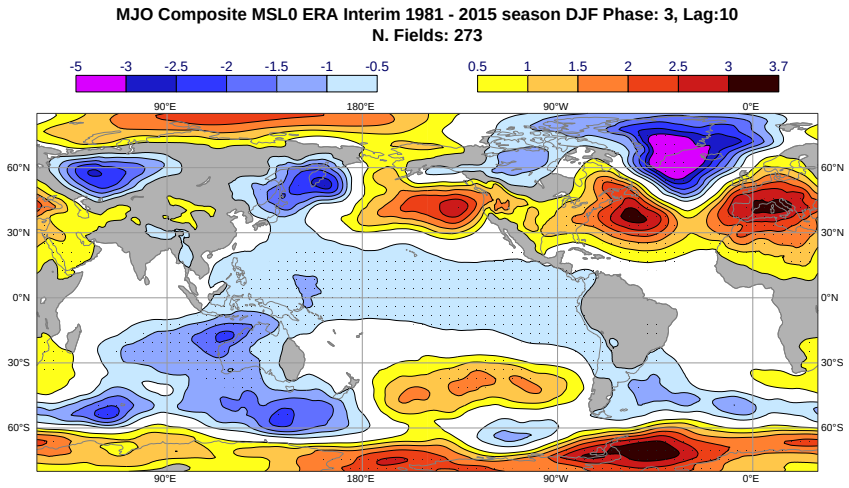
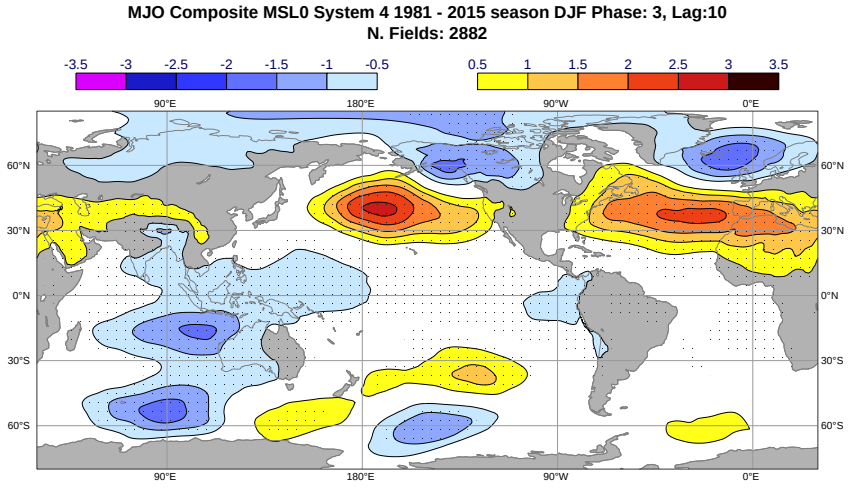


Figure 18: Forecast of MJO from 19 November.



(a) Reanalysis



(b) Model

Figure 19: Teleconnenctions from MJO phase 3 to MSLP for DJF based on composites with 10 days lag.

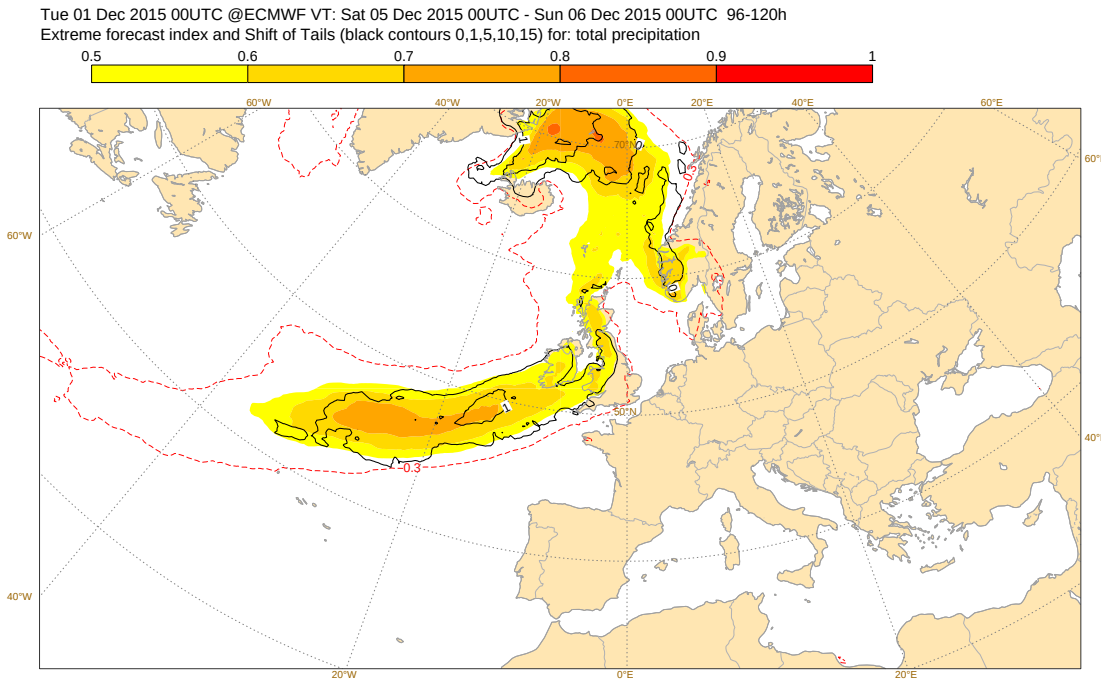


Figure 20: Extreme forecast index (EFI) and Shift of Tails (SOT) for total precipitation on 5 December in the forecast from 1 December.

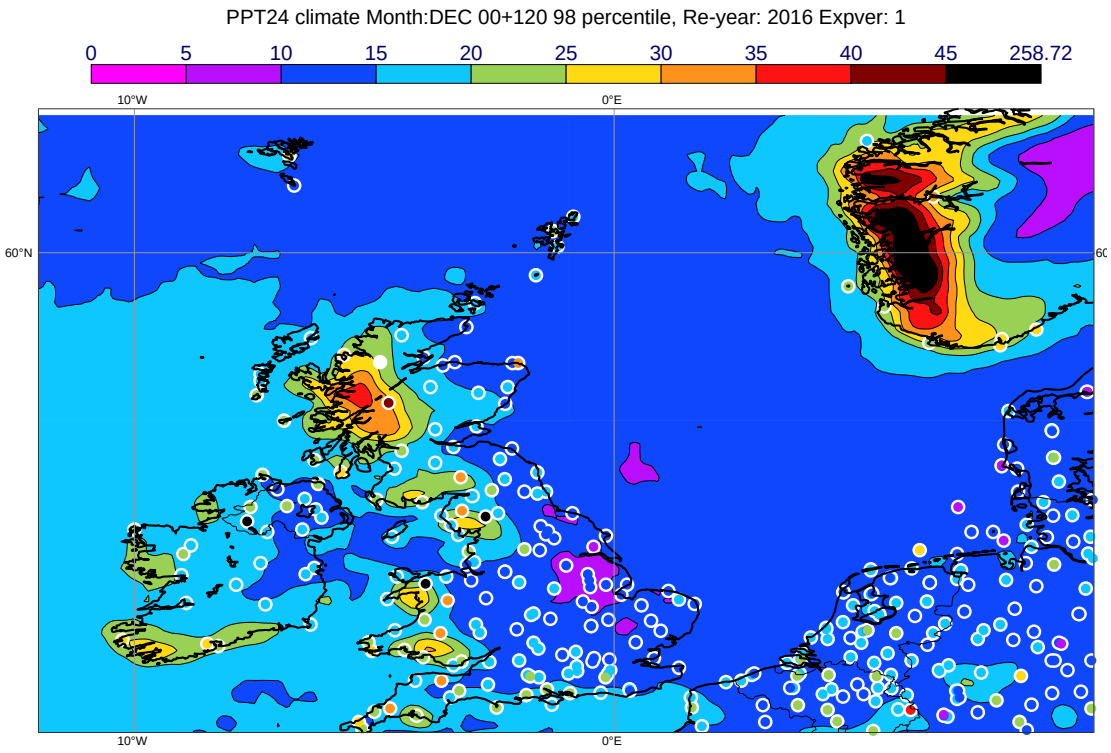


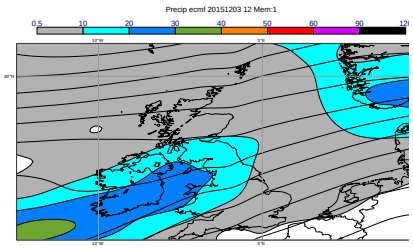
Figure 21: 98th percentile of the model climatology (shade) and observations (circles) for 24-hour precipitation valid for December.

atmospheric river, bringing humid air, which would precipitate out when hitting the orography along the coasts. It was therefore suggested in Lavers *et al.* (2017) the atmospheric rivers can be used a predictor for extreme coastal rainfall.

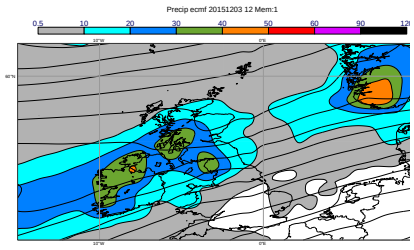
Reaching the day of Storm Desmond, the last forecast before the event showed a clear shift to extreme rainfall over northern British Isles. However, if we compare the last HRES forecast with observations of 24-hour precipitation, we find that the model underestimated the rainfall in the most extreme places. The underestimation of orographic precipitation has been improved with increased resolution and new model physics (Haiden *et al.*, 2014; Forbes *et al.*, 2015)), but is still underestimated. This can be seen when we evaluate the value of the 98th percentile of the model climate with the observation climatology (Figure 21), derived over the period 1980-2009, in a similar way as presented in Magnusson *et al.* (2014).

Figure 22 shows the effect of model resolution in 2-day forecasts for Storm Desmond. The figure includes results from 6 different resolutions of the ECMWF model, ranging from 300 km to 5 km. The figure also includes the 2.5km limited area-model from UK Metoffice (Hagelin *et al.*, 2017). Inspecting the results we see a clear benefit of increased resolution for orographic precipitation. But also the 5 km forecast is not as intense as the 2.5km UKMO forecast along the high coasts of northern England.

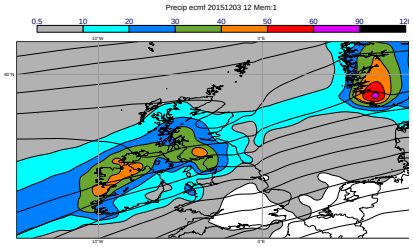
To summarize the multi-scale aspects of this case, the important aspect to capture on the extended-range time-scale was the MJO and the linkage to NAO. In the medium-range the prediction of the atmospheric river seems to have been the key to capture the rainfall over the northern British Isles. In the short-range



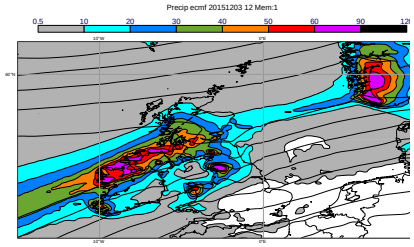
(a) 300 km



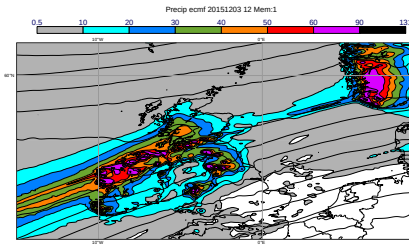
(b) 80 km



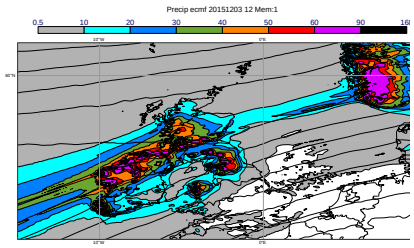
(c) 36 km



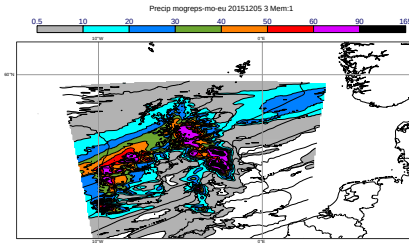
(d) 18 km



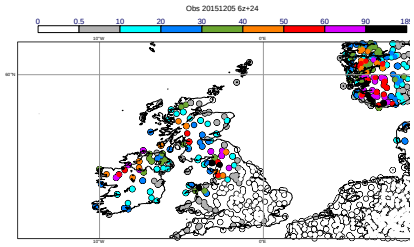
(e) 9 km



(f) 5 km



(g) UKMO 2.5 km



(h) Observations

Figure 22: 24-hour precipitation (5 December 06UTC to 6 December 06UTC) ans MSLP valid 5 December 18UTC from IFS model simulations with different resolutions (a-f) and from UKMO regional EPS member 1 (g) and observations (h).

capturing the magnitude of the rainfall and the orographic enhancement was important to capture and created forecast errors.

References

Baldwin MP, Dunkerton TJ. 2001. Stratospheric Harbingers of Anomalous Weather Regimes. *Science* **294**: 581–584, doi:10.1126/science.1063315.

Bouallegue ZB, Magnusson L, Haiden T, Richardson DS. 2019. Monitoring trends in ensemble forecast performance focusing on surface variables and high-impact events. *Quarterly Journal of the Royal Meteorological Society* **145**(721): 1741–1755, doi:10.1002/qj.3523.

Bougeault P, Toth Z, Bishop C, Brown B, Burridge D, Chen D, Ebert E, Fuentes M, Hamill T, Mylne K, Nicolau J, Paccagnella T, Park YY, Parsons D, Raoult B, Schuster D, Silva Dias P, Swinbank R, Takeuchi Y, Tennant W, Wilson L, Worley S. 2010. THORPEX Interactive Grand Global Ensemble (TIGGE). *Bull. Amer. Meteor. Soc.* **91**: 1059–1072.

Cassou C. 2008. Intraseasonal interaction between the madden–julian oscillation and the north atlantic oscillation. *Nature* **455**(7212): 523–527.

Cavaleri L, Bajo M, Barbariol F, Bastianini M, Benetazzo A, Bertotti L, Chiggiato J, Davolio S, Ferrarin C, Magnusson L, Papa A, Pezzutto P, Pomaro A, Umgiesser G. 2019. The october 29, 2018 storm in northern italy - an exceptional event and its modeling. *Progress in Oceanography* **178**: 102 178, doi:https://doi.org/10.1016/j.pocean.2019.102178.

Dirmeyer PA. 2011. A history and review of the global soil wetness project (gswp). *Journal of Hydrometeorology* **12**(5): 729–749, doi:10.1175/JHM-D-10-05010.1.

Donat MG, Leckebusch GC, Pinto JG, Ulbrich U. 2010. Examination of wind storms over central europe with respect to circulation weather types and nao phases. *International Journal of Climatology* **30**(9): 1289–1300, doi:10.1002/joc.1982, URL http://dx.doi.org/10.1002/joc.1982.

ECMWF. 2017. IFS Documentation CY43R3. Technical report, ECMWF.

Ferranti E, Chapman L, Whyatt D. 2017. A perfect storm? the collapse of lancaster's critical infrastructure networks following intense rainfall on 4/5 december 2015. *Weather* **72**(1): 3–7, doi:10.1002/wea.2907, URL http://dx.doi.org/10.1002/wea.2907.

Ferranti L, Magnusson L, Vitart F, Richardson DS. 2018. How far in advance can we predict changes in large-scale flow leading to severe cold conditions over europe? *Quarterly Journal of the Royal Meteorological Society* **144**(715): 1788–1802, doi:10.1002/qj.3341.

Ferranti L, Viterbo P. 2006. The european summer of 2003: Sensitivity to soil water initial conditions. *Journal of Climate* **19**(15): 3659–3680, doi:10.1175/JCLI3810.1.

Fischer EM, Seneviratne SI, Vidale PL, Luthi D, Schar C. 2007. Soil moisture-atmosphere interactions during the 2003 european summer heat wave. *Journal of Climate* **20**(20): 5081–5099, doi:10.1175/JCLI4288.1.

Forbes R, Haiden T, Magnusson L. 2015. Improvements in IFS forecasts of heavy precipitation. Newsletter 144, ECMWF.

Fragkoulidis G, Wirth V, Bossmann P, Fink AH. 2018. Linking northern hemisphere temperature extremes to rossby wave packets. *Quarterly Journal of the Royal Meteorological Society* **144**(711): 553–566, doi:10.1002/qj.3228.

Gascon E, Hewson T, Haiden T. 2018. Improving predictions of precipitation type at the surface: Description and verification of two new products from the ecmwf ensemble. *Weather and Forecasting* **33**(1): 89–108, doi:10.1175/WAF-D-17-0114.1.

Gascon E, Laviola S, Merino A, Miglietta M. 2016. Analysis of a localized flash-flood event over the central mediterranean. *Atmospheric Research* **182**: 256 – 268, doi:10.1016/j.atmosres.2016.08.007.

Grams CM, Magnusson L, Madonna E. 2018. An atmospheric dynamics perspective on the amplification and propagation of forecast error in numerical weather prediction models: A case study. *Quarterly Journal of the Royal Meteorological Society* **144**(717): 2577–2591, doi:10.1002/qj.3353.

Grams CM, Wernli H, Bottcher M, Campa J, Corsmeier U, Jones SC, Keller JH, Lenz CJ, Wiegand L. 2011. The key role of diabatic processes in modifying the upper-tropospheric wave guide: a north atlantic case-study. *Quarterly Journal of the Royal Meteorological Society* **137**(661): 2174–2193, doi:10.1002/qj.891.

Grazzini F. 2007. Predictability of a large-scale flow conducive to extreme precipitation over the western alps. *Meteorology and Atmospheric Physics* **95**(3), doi:10.1007/s00703-006-0205-8.

Grazzini F, Craig GC, Keil C, Antolini G, Pavan V. 2019. Extreme precipitation events over northern italy. part (i): a systematic classification with machine learning techniques. *Quarterly Journal of the Royal Meteorological Society* **0**(ja), doi:10.1002/qj.3635.

Grazzini F, Vitart F. 2015. Atmospheric predictability and rossby wave packets. *Quarterly Journal of the Royal Meteorological Society* **141**(692): 2793–2802, doi:10.1002/qj.2564, URL http://dx.doi.org/10.1002/qj.2564.

Hagelin S, Son J, Swinbank R, McCabe A, Roberts N, Tennant W. 2017. The met office convective-scale ensemble, mogreps-uk. *Quarterly Journal of the Royal Meteorological Society* : n/a–n/adoi:10.1002/qj.3135, URL http://dx.doi.org/10.1002/qj.3135. QJ-16-0327.R2.

Haiden T, Duffy S. 2016. Use of high-density observations in precipitation verification. Newsletter 146, ECMWF.

Haiden T, Magnusson L, Tsonevsky I, Wetterhall F, Alfieri L, Pappenberger F, de Rosnay P, Munoz-Sabater J, Balsamo G, Albergel C, Forbes R, Hewson T, Malardel S, Richardson D. 2014. ECMWF forecast performance during the June 2013 flood in Central Europe. Technical Memorandum 723, ECMWF.

Hewson TD, Neu U. 2015. Cyclones, windstorms and the imilast project. *Tellus A: Dynamic Meteorology and Oceanography* **67**(1): 27 128, doi:10.3402/tellusa.v67.27128.

Hewson TD, Titley HA. 2010. Objective identification, typing and tracking of the complete life-cycles of cyclonic features at high spatial resolution. *Meteorological Applications* **17**(3): 355–381, doi:10.1002/met.204.

Hogan R, Ahlgrimm M, Balsamo G, Beljaars A, Berrisford P, Bozzo A, Giuseppe FD, Forbes R, Haiden T, Lang S, Mayer M, Polichtchouk I, Sandu I, Vitart F, Wedi N. 2017. Radiation in numerical weather prediction. Technical Memorandum 816, ECMWF.

Khodayar S, Kalthoff N, Kottmeier C. 2018. Atmospheric conditions associated with heavy precipitation events in comparison to seasonal means in the western mediterranean region. *Climate Dynamics* **51**(3): 951–967, doi:10.1007/s00382-016-3058-y.

Lalaurette F. 2003. Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Quarterly Journal of the Royal Meteorological Society* **129**(594): 3037–3057, doi:10.1256/qj.02.152, URL http://dx.doi.org/10.1256/qj.02.152.

Lavers DA, Villarini G. 2013. The nexus between atmospheric rivers and extreme precipitation across europe. *Geophysical Research Letters* **40**(12): 3259–3264, doi:10.1002/grl.50636.

Lavers DA, Zsoter E, Richardson DS, Pappenberger F. 2017. An assessment of the ecmwf extreme forecast index for water vapor transport during boreal winter. *Weather and Forecasting* **32**(4): 1667–1674, doi:10.1175/WAF-D-17-0073.1.

Leutbecher M. 2018. Ensemble size: How suboptimal is less than infinity? *Submitted to Quarterly Journal of the Royal Meteorological Society* .

Lopez P. 2016. A lightning parameterization for the ecmwf integrated forecasting system. *Monthly Weather Review* **144**(9): 3057–3075, doi:10.1175/MWR-D-16-0026.1.

Magnusson L. 2017. Diagnostic methods for understanding the origin of forecast errors. *Q. J. R. Met. Soc.* **143**(706): 2129–2142, doi:10.1002/qj.3072.

Magnusson L, Richardson D, Haiden T. 2014. Verification of extreme weather events: Discrete predictands. Technical Memorandum 731, ECMWF.

Martius O, Schwierz C, Davies HC. 2008. Far-upstream precursors of heavy precipitation events on the alpine south-side. *Quarterly Journal of the Royal Meteorological Society* **134**(631): 417–428, doi:10.1002/qj.229.

Nuissier O, Joly B, Joly A, Ducrocq V, Arbogast P. 2011. A statistical downscaling to identify the large-scale circulation patterns associated with heavy precipitation events over southern france. *Quarterly Journal of the Royal Meteorological Society* **137**(660): 1812–1827, doi:10.1002/qj.866.

Pantillon F, Lerch S, Knippertz P, Corsmeier U. 2018. Forecasting wind gusts in winter storms using a calibrated convection-permitting ensemble. *Quarterly Journal of the Royal Meteorological Society* **144**(715): 1864–1881, doi:10.1002/qj.3380.

Pillosu F, Hewson T. 2017. New point-rainfall forecasts for flash flood prediction. Newsletter 153, ECMWF.

Quinting JF, Vitart F. 2019. Representation of synoptic-scale rossby wave packets and blocking in the s2s prediction project database. *Geophysical Research Letters* **46**(2): 1070–1078, doi:10.1029/2018GL081381.

Ralph FM, Dettinger MD. 2011. Storms, floods, and the science of atmospheric rivers. *Eos, Transactions American Geophysical Union* **92**(32): 265–266, doi:10.1029/2011EO320001.

Raveh-Rubin S, Wernli H. 2015a. 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, doi:10.1002/qj.2531.

Raveh-Rubin S, Wernli H. 2015b. 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, doi:10.1002/qj.2891.

Richardson DS. 2000. Skill and relative economic value of the ECMWF Ensemble Prediction System. *Q. J. R. Met. Soc.* **126**: 649–668.

Rodwell MJ, Magnusson L, Bauer P, Bechtold P, Bonavita M, Cardinali C, Diamantakis M, Earnshaw P, Garcia-Mendez A, Isaksen L, Källén E, Klocke D, Lopez P, McNally T, Persson A, Prates F, Wedi N. 2013. Characteristics of occasional poor medium-range weather forecasts for Europe. *Bull. Amer. Meteor. Soc.* **94**: 1393–1405.

Trigo RM, Pozo-Vazquez D, Osborn TJ, Castro-Diez Y, Gamiz-Fortis S, Esteban-Parra MJ. 2004. North atlantic oscillation influence on precipitation, river flow and water resources in the iberian peninsula. *International Journal of Climatology* **24**(8): 925–944, doi:10.1002/joc.1048.

Tsanis I, Tapoglou E. 2019. Winter north atlantic oscillation impact on european precipitation and drought under climate change. *Theoretical and Applied Climatology* **135**(1): 323–330, doi:10.1007/s00704-018-2379-7.

Tsonevsky I. 2015. New EFI parameters for forecasting severe convection. Newsletter 144, ECMWF.

Vergni L, Todisco F, Di Lena B, Mannocchi F. 2016. Effect of the north atlantic oscillation on winter daily rainfall and runoff in the abruzzo region (central italy). *Stochastic Environmental Research and Risk Assessment* **30**(7): 1901–191", doi:10.1007/s00477-015-1194-2.

Vitart F. 2014. Evolution of ECMWF sub-seasonal forecast skill scores. *Q. J. R. Met. Soc.* **140**: 1889–1899, doi:10.1002/qj.2256.

Vitart F, Robertson AW. 2018. The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. *npj Climate and Atmospheric Science* **1**: N3, doi:10.1038/s41612-018-0013-0.

Wernli H, Davies HC. 1997. A lagrangian-based analysis of extratropical cyclones. i: The method and some applications. *Quarterly Journal of the Royal Meteorological Society* **123**(538): 467–489, doi:10.1002/qj.49712353811.

Wilks DS. 1997. Resampling Hypothesis Tests for Autocorrelated Fields. *J. Climate* **10**: 65–82.

Wirth V, Eichhorn J. 2014. Long-lived rossby wave trains as precursors to strong winter cyclones over europe. *Quarterly Journal of the Royal Meteorological Society* **140**(680): 729–737, doi:10.1002/qj.2191.

Wirth V, Riemer M, Chang EKM, Martius O. 2018. Rossby wave packets on the midlatitude waveguide - a review. *Monthly Weather Review* **146**(7): 1965–2001, doi:10.1175/MWR-D-16-0483.1.

Woollings T, Hoskins B, Blackburn M, Berrisford P. 2008. A new rossby wave-breaking interpretation of the north atlantic oscillation. *Journal of the Atmospheric Sciences* **65**(2): 609–626, doi:10.1175/2007JAS2347.1.

Wulff CO, Greatbatch RJ, Domeisen DIV, Gollan G, Hansen F. 2017. Tropical forcing of the summer east atlantic pattern. *Geophysical Research Letters* **44**(21): 11,166–11,173, doi:10.1002/2017GL075493.

Zhang Q, Li L, Ebert B, Golding B, Johnston D, Mills B, Panchuk S, Potter S, Riemer M, Sun J, Taylor A, Jones S, Ruti P, Keller J. 2019. Increasing the value of weather-related warnings. *Science Bulletin* **64**(0): 647 – 649, doi:10.1016/j.scib.2019.04.003.

Zschenderlein P, Fink AH, Pfahl S, Wernli H. 2019. Processes determining heat waves across different european climates. *Quarterly Journal of the Royal Meteorological Society* doi:10.1002/qj.3599.