Climate Dynamics (2024) 62:9269–9285
https://doi.org/10.1007/s00382-024-07390-0

ORIGINAL ARTICLE

Pattern-based forecasting enhances the prediction skill of European heatwaves into the sub-seasonal range

Emmanuel Rouges[1,2,3] · **Laura Ferranti**[1] · **Holger Kantz**[2] · **Florian Pappenberger**[1]

Received: 23 October 2023 / Accepted: 4 August 2024 / Published online: 9 August 2024

Abstract
The prediction of European heatwaves at the subseasonal range is of key importance to mitigate their impact. This study builds on previous work which identifies five main European heatwave types based on their atmospheric circulation patterns (CPs). These CPs are potential predictors of heatwaves, as these patterns are connected with a high probability of 2-meter temperature exceeding the 90th percentile. Therefore, the aim of this study is to use these patterns to construct a pattern-based forecast method. The skill of this method to forecast extreme warm temperatures is then assessed and compared with the direct grid-point based forecast (using the direct 2-meter temperature forecast of the model). The extended (or subseasonal) range reforecast data from the European Centre for Medium-Range Weather Forecasts (ECMWF) is used for the skill evaluation. Firstly, the skill of the extended range model is assessed in predicting CPs. The pattern-based methodology is then compared with the direct prediction of extreme warm temperatures. The results show that the pattern-based methodology has a low skill at the short to medium range compared to the direct method, however it maintains skill for longer lead times, extending the forecast skill horizon significantly by up to six days over key heatwave regions. This improvement is localized over regions with the highest conditional probability of extreme warm temperatures. Furthermore, the prediction skill of 4-day periods of high temperatures using CPs lasting at least five days is also assessed. A similar improvement in forecast skill horizon is observed but the improvement is more modest and even more localized. This methodology provides skilful forecast at longer lead times to the end of the medium range and into the subseasonal range, which would be beneficial for early warnings of European heatwaves and therefore support the timely implementation of mitigation plans.

Keywords Heatwaves · subseasonal · ECMWF · Forecast · Europe

1 Introduction

In the recent decades, Europe has experienced some of the most devastating heatwaves (HWs) in its history, with the 2003 HW over western Europe, the 2010 HW over western Russia and the 2018 HW over western and northern Europe. These events can pose serious threats to the health of the population, the infrastructure and agriculture (Brimicombe et al. 2021). Accumulated heat stress was responsible for up to 70,000 and 55,000 casualties during the 2003 and 2010 HWs respectively (Robine et al. 2008; Barriopedro et al. 2011; di Napoli et al. 2018). The damage of HWs can be further exacerbated when they correlate with droughts and/or forest fires (Sutanto et al. 2020; Miralles et al. 2019; Vitolo et al. 2019; Vautard et al. 2007). More recently, the 2021 HW over the Pacific Northwest severely impacted western Canada in particular due to devastating wildfires (Emerton et al. 2022; Burston and Cecco, 2021; Lin et al. 2022). The expected increase in frequency and intensity of HWs (Schär et al. 2004), highlights the necessity to deepen the understanding. The improvement of the prediction of such extremes, allows for better mitigation of their impact.

Lately, forecast at the subseasonal range (S2S; 10 days to 3 months) gained a lot of scientific attention (Domeinsen

✉ Emmanuel Rouges
e.m.rouges@reading.ac.uk

[1] European Centre for Medium-Range Weather Forecasts, Reading, UK

[2] Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

[3] Department of Meteorology, University of Reading, Reading, UK

et al., 2022) as its potential to provide early warnings was highlighted (Vitart & Roberston, 2018). The European Centre for Medium-Range Weather Forecasts (ECMWF) recognised the importance of forecasts at the subseasonal range with the introduction of the extended range forecast model (ECMWF's subseasonal range model) together with operational forecast products (e.g. Ferranti et al. 2015). Further into the text, subseasonal and extended range will be used exchangeably.

At this range, forecasting of surface parameters and extremes is a challenging task. Efforts have been made in identifying the sources of predictability at this time range, especially since the S2S Prediction Project was created by the World Weather Research Program in 2013. At the subseasonal time scale, slower varying processes become more relevant to numerical prediction, such as the Madden-Julian Oscillation, the stratospheric polar vortex and associated teleconnections (Vitart and Robertson 2018; Domeisen et al. 2022). Studies have shown that the predictability of surface extremes at the subseasonal range stems from long-lived large scale flow patterns (Magnusson 2019; Ferranti et al. 2018).

Large scale flow patterns have been identified using in most cases a combination of principal component analysis and clustering algorithms (Michelangeli et al. 1995; Strauss et al., 2007; Ferranti et al. 2015). They are referred to as weather regimes and represent most of the low-frequency variability in the mid-latitudes. Over the Euro-Atlantic region, four weather regimes are classically identified (Michelangeli et al. 1995; Cassou 2008; Ferranti et al. 2015). These regimes are actively used for forecasting purposes at ECMWF for example (Ferranti et al. 2015; Grams et al. 2020). More recently, seven weather regimes were proposed to account more for variability across the year (Grams et al. 2017).

The usefulness of these weather regimes at the subseasonal range arises from the capacity of models to forecast them past two weeks (Ferranti et al. 2015; Matsueda and Palmer 2018; Robertson et al. 2020). This can be explained by their large spatial scale and persistence (Molteni et al. 1990; Vautard, 1990; Michelangeli et al. 1995; Hannachi et al. 2017). Additionally, they are associated with high impact weather such as cold spells and HWs (Cassou et al., 2005; Matsueda 2011). Weather regimes are related to large anomalies in temperature, wind and precipitation (Ferranti et al. 2018; Richardson et al. 2020; Cassou et al. 2004). They are used as predictors to infer surface values and even the production of renewable energy and energy demand (which is partly dependent on temperature; Grams et al. 2017; Bloomfield et al. 2021; van der Wiel et al. 2019b). Large flow patterns are more likely to be predicted at the S2S range than surface parameters alone (Ferranti et al. 2018; Mastrantonas et al. 2022) and are therefore useful for the prediction of extremes at the subseasonal range (Vitart and Robertson 2018).

While the weather regimes are associated with the atmospheric low frequency variability, it is possible to identify circulation patterns (CPs) based on their specific impact on the surface conditions (Mastrantonas et al. 2021; Rouges et al. 2023). These CPs are sometimes referred to as targeted circulation types (Bloomfield et al. 2020). Instead of deriving the main circulation types explaining most of the atmospheric variability over a region, the CPs are classified according to their impact on the surface variables such as surface temperature, surface wind and precipitation. Pfahl (2014) identifies key features responsible for different types of extremes, while others use clustering methods to identify specific or targeted CPs (Mastrantonas et al. 2021; Bloomfield et al. 2021; Rouges et al. 2023). These patterns can then be used to create a pattern-based forecast which provides skilful forecast at a longer range than the conventional forecast (Mastrantonas et al. 2022; Bloomfield et al. 2021; Gonzalez et al. 2022). However, it is important to note that their prediction skill at shorter lead-times is reduced compared to conventional forecast (Mastrantonas et al. 2022). Similarly, weather analogues are used to provide a forecast for precipitation and temperatures at the subseasonal time scale (Yiou and Déandréis 2019; Krouma et al. 2022). Instead of using a set ensemble of weather regimes or patterns, analogues of circulation (other days with similar atmospheric circulation) for example, are computed at each forecast initialisation to provide a forecast (Yiou et al., 2014).

This study aims at quantifying the advantages of CP based forecasting compared to grid-point based forecasting (direct/conventional forecast) for HWs. Across this article, CPs associated with HWs identified in previous research (Rouges et al. 2023) will be used to create a pattern-based (or conditional) forecasting system. This forecasting system will be evaluated in its capacity to forecast extreme warm temperatures. The ECMWF extended range model will be used to assess the capacity of the model to first represent accurately the CPs and then establish if the CPs can produce skilful forecasting of extreme warm temperatures. This will allow to demonstrate if using CPs can provide improvements to the forecast compared to the conventional, grid point forecast.

Firstly, the data and methods used to perform the research will be explained. Then, the results of the forecast evaluation will be described and discussed before the key findings will be summarized.

2 Data and methods

To evaluate the forecast, the ECMWF extended range model is used (Owens and Hewson 2018), providing forecasts with lead time up to 46 days. A set of reforecast data covering the period from 1999 to 2018 is used from the model in operation between April and September 2019. The ECMWF model is updated in yearly cycles (on average), meaning that the real-time forecast benefits from yearly changes to improve its performance. However, it also means that the skill, errors, and biases of the ECMWF model change with each cycle. The changes to the model are updated and published as scorecards for the output variables, but also for different regions and levels (ECMWF, n.d.). It is essential for forecasters to understand the impact of model changes on different variables and regions. While yearly changes are useful to keep the model and forecast at the state of the art, it becomes difficult to account for all model changes when investigating the forecast skill over multiple years. The advantage of using the reforecast data compared to the real-time forecast is that the reforecast data covers a long period provided by the same forecast system, with the same model bias across this period. The model is run twice per week (Monday and Thursday) with one control member and 10 perturbed members, which is significantly lower compared to the 101 ensemble members (for model cycle number 48r1) of the real time forecasts. The perturbed members have small perturbations in the initial conditions and in the physics parameterisation to account for uncertainty in the initial conditions and in the model formulation. Over the forecast period, 880 initialisation dates are covered.

To verify the results, the ECMWF ERA5 reanalysis (Hersbach et al. 2020) dataset is used as truth. Both for the reforecast and ERA5, the daily mean 2-meter temperature (2mT) and geopotential height at 500 hPa (Z500) are used. The data covers the European region (30°*N* – 80°N, 25°W – 60°E) with a horizontal resolution of 1°*1°.

2.1 Connection between extreme high temperatures and circulation patterns

In Rouges et al. (2023), HWs have been identified and clustered based on their atmospheric circulation. The identification of heatwaves follows the same principle as Stefanon et al. (2012) by selecting HWs as period of at least 4 days when the daily mean 2mT exceeds the 90th percentile over a region of 1000 km. The HW days identified are then clustered using a k-means clustering algorithm on the daily Z500 anomaly. The daily Z500 anomaly is filtered using a Principal Component Analysis beforehand to allow for a more effective clustering. This resulted in five HW types characterised by specific CPs for which a composite of Z500 anomaly and 2mT anomaly is found in the supplementary material (S1). Figure 1 shows the probability of 2mT

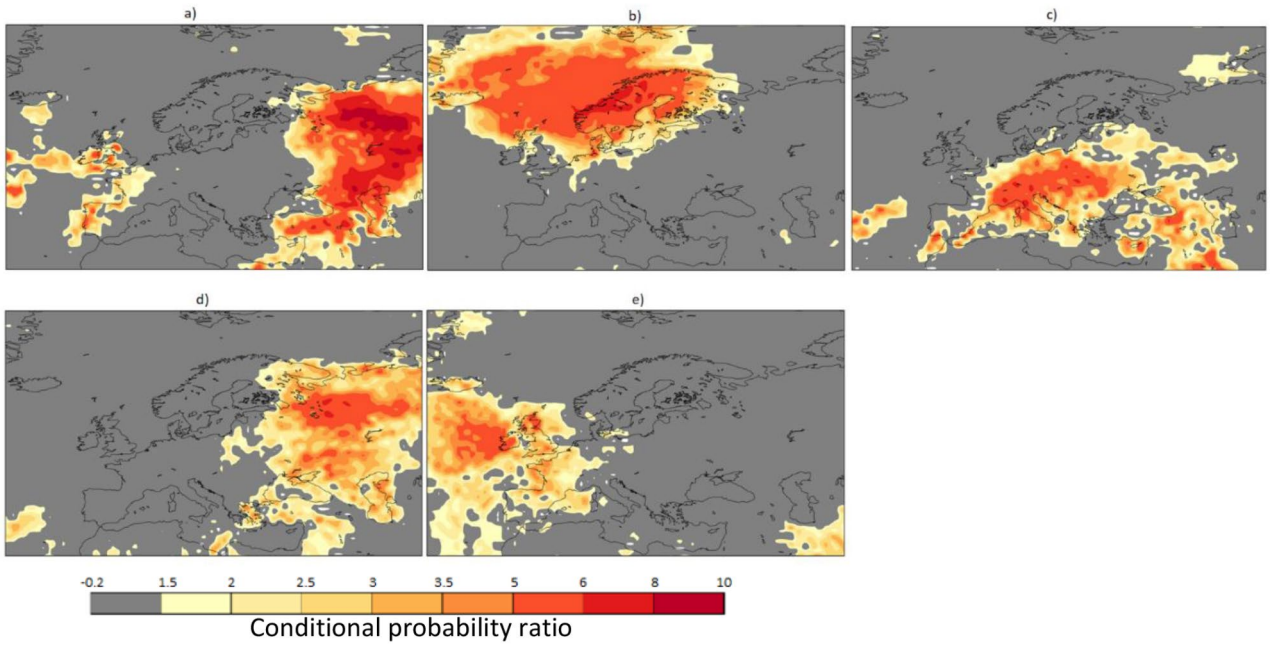

Fig. 1 Increased probability compared to climatology of 2mT exceeding the 90th percentile during the different CPs: Tripole (549 days) **a**), Scandinavia (SC: 794 days) **b**), Southern Europe (SE: 607 days) **c**), Russia (RU: 609 days) **d**) and Western Europe (WE: 721 days) **e**). The increased probability is represented as a ratio of probability of exceeding the 90th percentile during CPs compared to the climatological probability (1979–2020). Grey areas show regions where the conditional probability is less than 1.5 times the climatology. Other colours represent areas with a conditional probability higher than 1.5 times the climatology

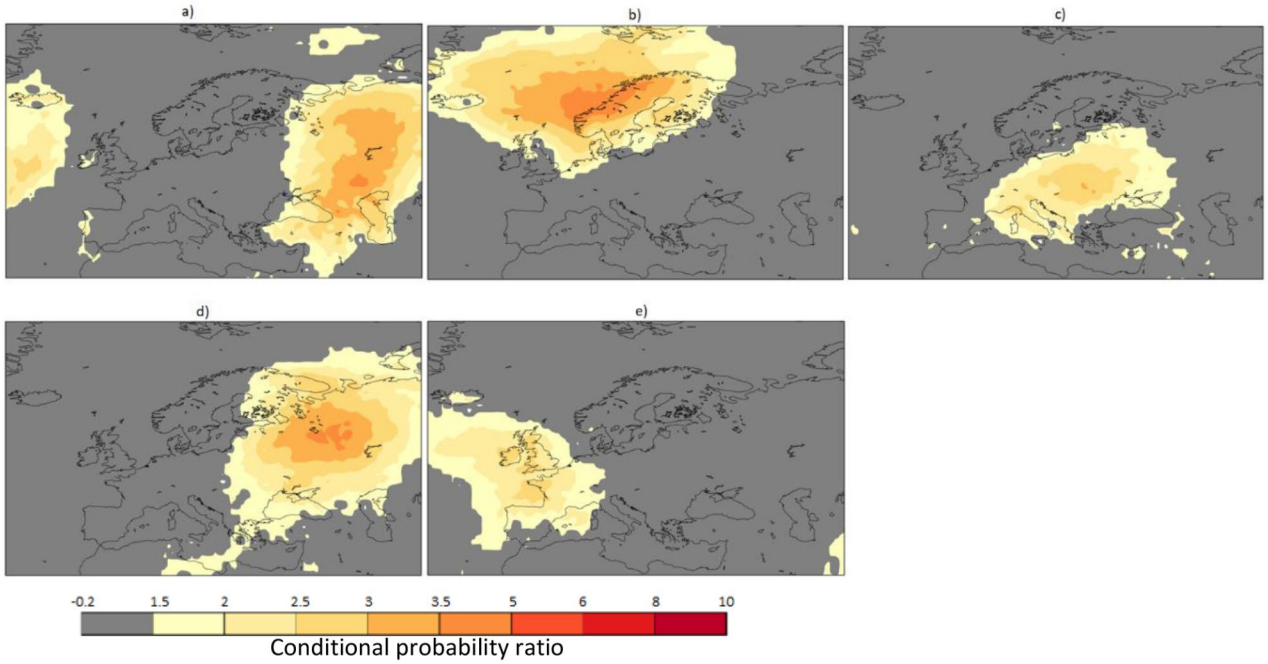

Fig. 2 Increased probability compared to climatology of 2mT exceeding the 90th percentile for 4 days during the different persistent CPs: Tripole (401 days) **a**), Scandinavia (SC: 694 days) **b**), Southern Europe (SE: 459 days) **c**), Russia (RU: 549 days) **d**) and Western Europe (WE: 609 days) **e**). The increased probability is represented as a ratio of probability of exceeding the 90th percentile during CPs compared to the climatological probability (1979–2020). Grey areas show regions where the conditional probability is less than 1.5 times the climatology. Other colours represent areas with a conditional probability higher than 1.5 times the climatology

exceeding the 90th percentile during each of the patterns. This probability is much higher than climatology, highlighting the link between the CPs and HWs.

The probability for 2mT to exceed the threshold is three times higher compared to climatology over the specific HW areas. For some CPs, such as the Russian (RU), Scandinavian (SC) and Tripole CPs, the probability is up to four times higher. One of the findings of Rouges et al. (2023) was that HWs coincide more with persisting CPs (five days or longer). The CPs used have been computed with heatwaves lasting 4 days or longer and have been shown to last on average 5 days or longer. The threshold of 4 days was chosen to follow the methodology by Stefanon et al. (2012) and accounts for the fact that persisting heatwaves are more impactful and more predictable at longer time ranges (Zschenderlein et al. 2019). The conditional probability of 2mT exceeding the 90th percentile for four days during CPs that last for at least five days has been computed. Compared to climatology, the increased likelihood is six to nine times higher, for SC and RU regions (Fig. 2). The signal is however much noisier compared to Fig. 1, showing the impact of the limited sample size.

Importantly, the link between CPs and extreme temperature continues to hold in the reforecast. The increased probability of 2mT exceeding the 90th percentile is also visible at all lead times over the specific regions corresponding to each CPs with, however at lower amplitudes for longer lead times.

This link between CPs and HWs allows us to suggest that CPs could be used as predictors for HWs. Therefore, to determine their capacity as predictors, a pattern-based forecasting method is devised and tested.

2.2 Evaluation of the skill of ECMWF forecasts

The goal of this study is to determine the skill of a pattern-based prediction at the S2S range and compare it with a direct forecast. The direct forecast is defined for this study as the conventional model output. In this instance, the probability of 2mT to exceed the 90th percentile at each gridpoint over the European region is forecasted. The direct forecast consists of the ensemble forecast for 2mT, and the probability will depend on how many ensemble members predict extreme warm temperatures.

Forecast errors can be split into a systematic and a random component (Ferranti et al. 2002). The systematic error is defined as the average forecast error over a certain period. Removing the systematic error or bias can be done by computing the model anomaly, while the random component can not be removed this easily. This approach is often used for longer forecast ranges such as the S2S and seasonal forecast. The forecast information

provided is therefore the anomaly. The model anomaly is computed using a lead time and ensemble member dependent climatology, as the bias depends on lead time and on the ensemble members considered. The climatology is estimated from a set of three initial dates window to replicate as close as possible the construction of the climatology from ERA5 (e.g., the climatology of each individual ensemble member for the forecast initialised on the 3rd of June 2012 is constructed with all 20 forecasts initialised on the 3rd of June and also from the previous and next forecast initialisations, in this case the 30th of May and 6th of June). This process is applied both to the 2mT and the Z500 fields.

While subseasonal forecast products are generally shown as weekly anomalies, in this work the evaluation is performed on daily timescale. This allows to better showcase increases in the forecast skill horizon which are generally at a smaller timescale than weekly.

To forecast the different CPs, the forecasted daily Z500 anomaly field is projected onto the composite Z500 field anomaly from ERA5 (see Rouges et al. 2023) of each pattern for each forecast. This allows for the creation of indices which are a measure of the similarity of the atmospheric conditions compared to the different patterns. Using the same attribution procedure as in Rouges et al. (2023), each ensemble member of each forecast is attributed to one of the CPs or to the 'neutral pattern' (corresponding to days where the atmospheric circulation is too different to any of the CPs).

The conditional or pattern-based forecast is based on the forecast of the CPs. Each pattern is associated with a conditional probability of 2mT exceeding the 90th percentile at each grid point as seen in Fig. 1. The conditional probability is then weighted by the forecast probability of each pattern (including the 'neutral pattern'). The pattern-based forecast can be represented using the following equation:

$$F_c = \sum_{0}^{N_{cp}} P_{cp} * P_x$$

With $P_{cp.}$ being the forecasted probability to have any of the patterns and P_x the probability of extreme temperatures at each grid point (as displayed in Fig. 1) for the corresponding pattern. CP stands for the different patterns (WE, RU, Tripole, SC, SE and neutral) and $N_{cp.}$ is the number of patterns (in this case six including the neutral pattern).

To estimate the benefits of using CPs to predict heatwaves, this study first determines the current prediction skill of the ECMWF extended-range model (direct forecast). The Briers Score (BS; Wilks 2011) is used to determine the capacity of the model in predicting temperatures exceeding the 90th percentile and in identifying the correct CPs. The Brier Skill Score (BSS) allows to compare the BS of the extended range model with a reference score. In this paper, the reference forecast is the climatological probability of extreme warm temperatures or of the different CPs. As the reforecast is constituted of a reduced ensemble size, compared to the 101 ensembles of the ECMWF extended operational forecast, investigating the skill of the reforecast data underestimates the skill of the operational forecast. This is due to probabilistic scores being sensitive to the ensemble size. To account for this, the fair BS introduced by Ferro (2008) and Ferro et al. (2014) is used. The fair BS adjusts the original BS to represent the BS if the ensemble size was infinite.

The fair BS is outlined as follows:

$$fair\ BS = \frac{1}{N}\sum_{k=0}^{N}(p_k - o_k)^2 - \frac{p_k(1-p_k)}{m-1}$$

It is defined as the conventional BS (first part) with p_k being the probability of the forecast k (between 0 and 1) and o_k the corresponding observation (1 or 0, if the event occurs or not respectively), subtracted by an adjustment term which accounts for the difference in ensemble size (m being the ensemble size). N corresponds to the number of forecasts considered. F_c ,in the equation describing the pattern-based methodology, is replacing p_k for the pattern-based forecast in the BS equation.

Using the pattern based approach the forecast probability, p_k, at each grid point of exceeding the threshold is dependent on the conditional probability and therefore can never reach 0 or 1. However, the adjustment term in the fair BS works in such a way that it is equal to 0 when p_k is equal to 0 or 1. The conditional forecast would therefore benefit from this adjustment for each forecast while the grid point forecast would not systematically benefit from it. Based on this, the estimation of the skill for the conditional forecast will be shown using the original BS. Comparisons with the direct forecast are performed using the original BS for both the conditional and direct forecast. As the aim is to quantify the benefits of conditional forecasting over direct forecasting, not using the fair BS does not harm the comparison.

Together with the skill scores, an estimation of the frequency bias of both the CPs and the 2mT exceeding the 90th percentile is performed.

The significance of the results is evaluated using bootstrapping of 1000 resamples with replacement, each having the same sample size as the original sample (880 initial dates). A 90th percentile two-tailed confidence interval for the frequencies and a one-tailed confidence interval for the BSS are used.

3 Results

3.1 Direct forecast

The results of the evaluation of the ECMWF extended-range model in forecasting 2mT exceeding the 90th percentile and in identifying the CPs are presented below.

The first step consists in determining if the model has any significant bias in the frequency of these events compared to ERA5. For the 2mT, the frequency is estimated at each lead time for each grid point and then integrated over the region of the HW. For each region, a rectangle is drawn around the region of highest conditional probability (see Fig. 1) and all grid points (including greyed out once in Figs. 1 and 2) are included in the analysis of Fig. 3 and after. As the Tripole CP is defined by two regions of increased temperature, two different regions are integrated over (Iberian Peninsula and southern Russia). To simplify, the frequency shown is the average frequency across the 11 ensemble members. The results presented in Fig. 3.a, show no significant bias in frequency in 2mT exceeding the 90th percentile. The highest frequency bias is seen for the Russian Tripole region with a difference of 0.15% at very short range. The 2mT exceeding the 90th percentile for four days does not show a significant frequency bias either (not shown). The difference in frequency for the CPs is slightly more important (Fig. 3.b) with the RU CP being 1% more frequent and the WE pattern being 1% less frequent at a lead time passed 10 days. However, all other CPs do not have a significant frequency bias. The frequency bias is very similar when looking at long CPs (five days or longer), with the RU and WE CPs showing the strongest difference in frequency, namely

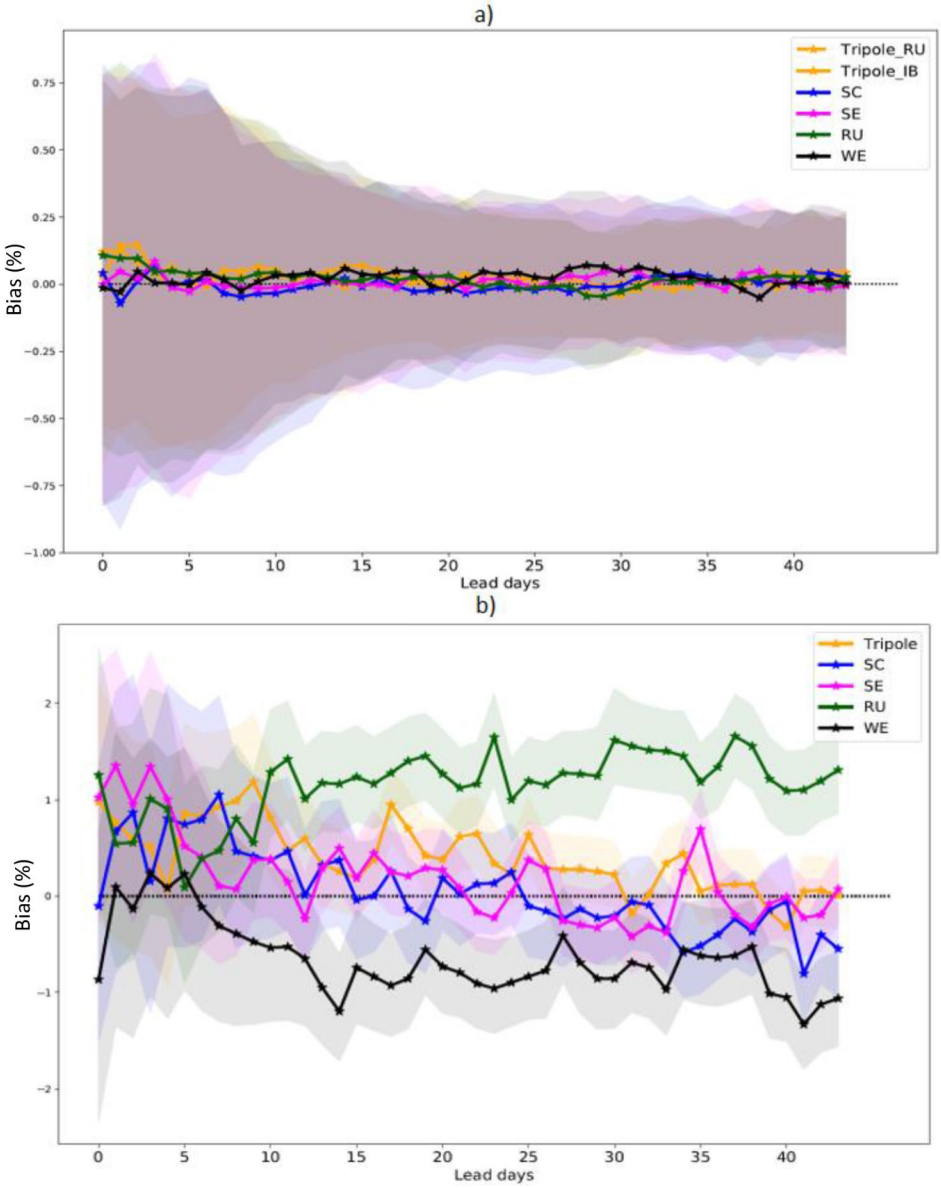

Fig. 3 Frequency difference between the ECMWF model (mean frequency of the 11 members) and the ERA5 climatology (1979–2020), (**a**) for the 2mT above the 90th percentile integrated over the different heatwave regions (by definition the climatological probability is 10%) and (**b**) for the CPs. The shading represents the 90th percentile confidence interval (the climatological frequency CPs varies from 14 to 16%)

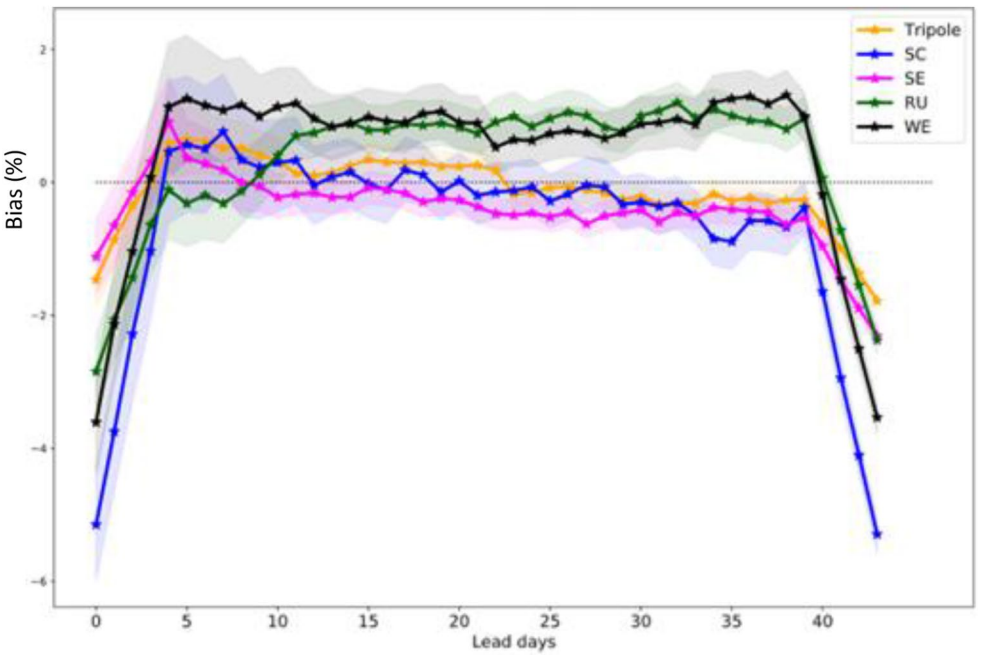

Fig. 4 Frequency difference between the ECMWF model (mean frequency of the 11 members) and the ERA5 climatology (1979–2020), for persistent CPs (5 days or longer). The climatological frequency of persistence CPs varies between 4% for the Tripole CP to ~10% for SC and WE CPs and 5–6% for SE and RU CPs

1.5% higher (Fig. 4). The SE CP has a 1% negative bias from a lead time of 20 days. However, by construction the frequency of long CPs and persistent extreme temperature is strongly negatively biased at short (four days) and very long (35 days+) lead times. If an event starts before the initial dates or finishes after 45 days (maximum lead time of the extended range forecast), the forecast will truncate this event and not identify it as a long event. As this study focuses on the extended range (after 10 days) and very little skill is to be expected passed 30 days, this problem does not impact this exploration.

The small bias in frequency signifies a good representation of the temperature extremes and the CPs, even for persisting CPs. The latter is essential when using these patterns to infer temperature extremes.

Figure 5 represents the performance (fair BSS) of the ECMWF model at forecasting 2mT exceeding the 90th percentile and at forecasting the different CPs. As for the frequency, the skill for extreme 2mT is computed for each grid point and then integrated over the different HW regions. The 2mT fair BSS shows a large disparity at short lead times with the Iberian Tripole region having a skill score of only 0.5 compared to most other regions having a score closer to 0.6 and the RU region close to 0.7. The difference can be partly attributed to the selection of the regions. The selection was made based on the regions of highest conditional probability of 2mT exceeding the 90th percentile in Fig. 1. The extent of the area selected for the Iberian Peninsula is smaller than for other regions, potentially impacting the aggregated score. At longer lead times, the scores come closer together and passed ten days, all regions have very similar scores. This signifies that at longer lead times, there is little difference in skill across regions. Only, the RU regions (RU and southern Russia) have a larger fair BSS than other regions at longer lead times. This might indicate a source of predictability (e.g., Boreal Summer Intraseasonal Oscillation: BSISO) that the model is able to represent correctly. The impact of the BSISO has been investigated and showed a reduction in spread at the extended range (Rouges et al. 2023). However, this impact is limited to cases with very high BSISO values and did not result in a significant improvement in the accuracy of the forecast, potentially showing that the model can capture the forcing from the BSISO but not the teleconnection pathway.

The forecast skill horizon is defined as the forecast lead time when the prediction has some level of skill (Buizza and Leutbecher 2015). In this case, the last lead time when the fair BSS is significantly above the climatology (based on a one-tailed test at the 90th percentile) is used as threshold for the forecast skill horizon. This test shows that the skill is not significantly above the climatology passed 15 days for the SC, SE, WE and Tripole IB regions while the RU regions extend the forecast horizon by up to 22 days. Again, this disparity could be explained by the RU regions being sensitive to teleconnections. In Rouges et al. (2023), the BSISO (Lee et al. 2013; Kiladis et al. 2014) was identified as a possible driver for some of the RU heatwave events, potentially impacting the predictability.

In comparison, the CPs all start with a very similar score of 0.85 and a similar forecast skill horizon of about 15 days. Only the RU CP has a lower forecast skill horizon at 11 days and the Tripole CP has better skill than climatology until

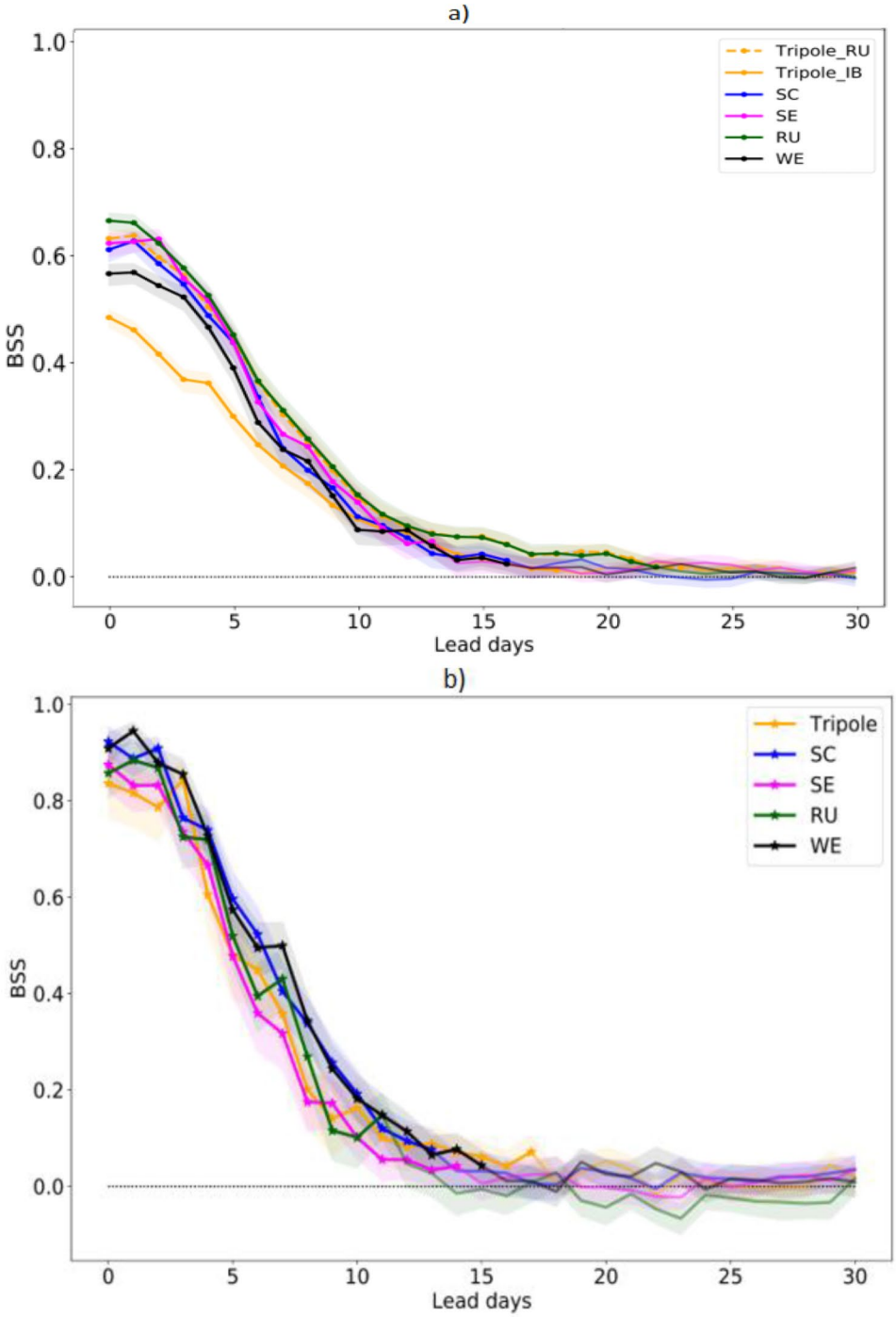

Fig. 5 Fair BSS for 2mT exceeding the 90th percentile for (**a**) the different heatwave regions and (**b**) for the different CPs. Full lines with symbols represent a fair BSS significantly above 0 using a one tailed test at 90th percentile. Shading represents the 10th to 90th percentile range

17 days. The fair BSS for the CPs is limited by a sample of 880 forecasts per lead time, as is the 2mT forecast, however it does not benefit from the aggregation of grid points as for the 2mT. This is visible by the spread in skill across the different patterns and the large overlap between skill scores (Fig. 5). Therefore, the difference in forecast skill horizon between CPs has a limited significance. In comparison, the fair BSS of 2mT shown in integrated over multiple grid points, allowing for a greater sample size. The uncertainty of fair BSS is visibly less important for 2mT, making the difference between regions more significant.

Rouges et al. (2023), identified how HWs are driven predominantly by persisting CPs. Therefore, it is key to assess the capacity of the model in identifying longer persisting CPs. In this study, persisting CPs are defined as lasting five days or longer. Figure 6 compares the BSS of forecasting the CPs with the BSS of forecasting persisting CPs. The difference in forecast skill is low across most CPs. However, due

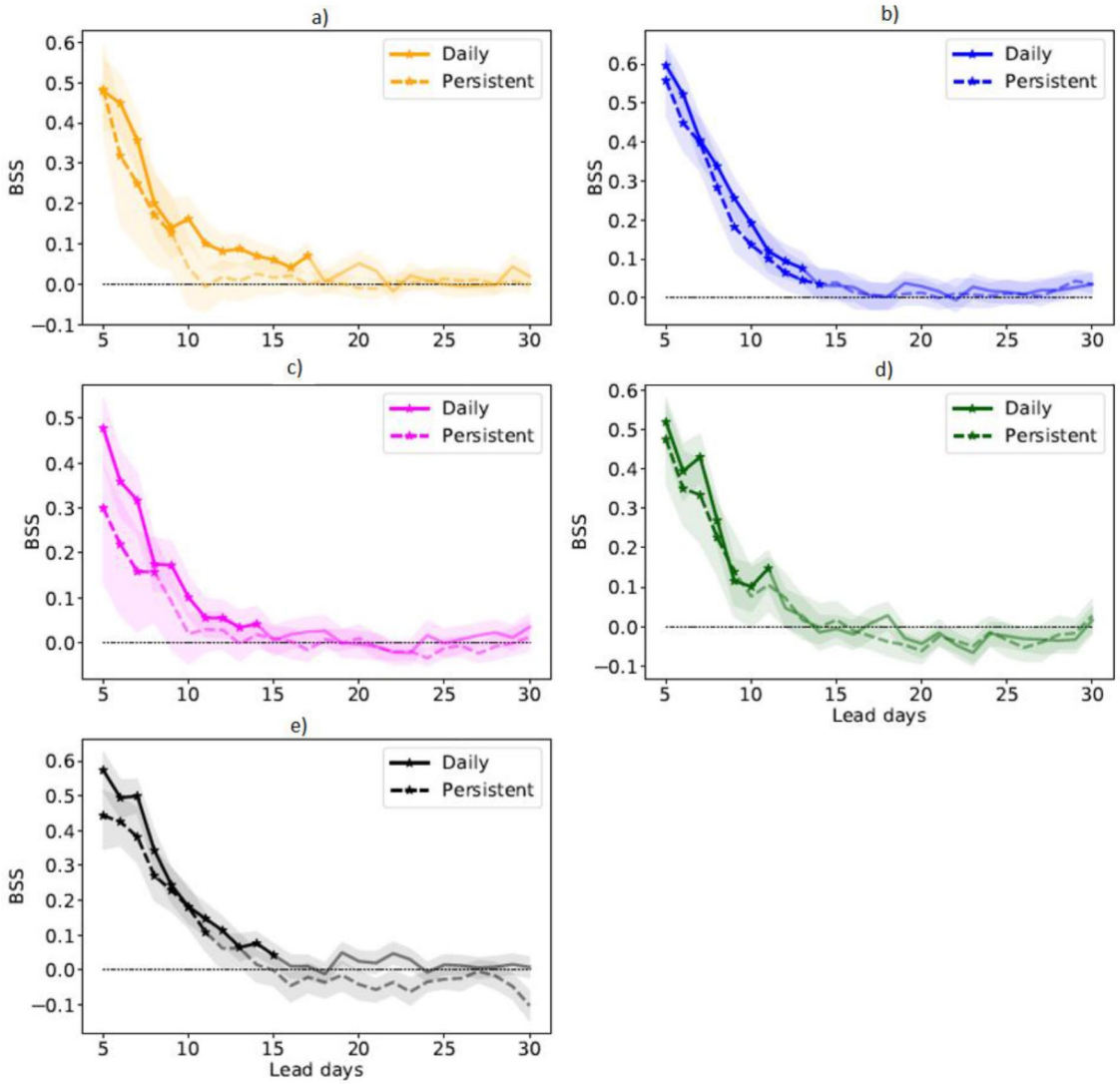

Fig. 6 Fair BSS for CPs. Daily attribution (continuous line) and persistent patterns (dashed line): (**a**) Tripole, (**b**) SC, (**c**) SE, (**d**) RU and (**e**) WE CPs. Full lines with symbols represent a fair BSS significantly above 0 using a one tailed test at 90th percentile. Shading represents the 10th to 90th percentile range

to the reduced sample size in persisting CPs, this leads to an increase in spread and results in a significant decrease in forecast skill horizon. For instance, the BSS for the Tripole CP overlaps over most lead times for daily and persistent CPs, but due to the larger spread, the skill horizon is reduced by eight days. The forecast horizon of the SE and WE CPs is reduced by six and four days respectively, compared to daily CPs (without persistence criterium). The forecast skill horizon of the RU persistent CPs is however reduced by only two days and the SC persistent patterns experience an increase in forecast range of one day. While the significance tests highlight an important decrease in forecast skill horizon when forecasting persistent CPs compared to daily CPs, the uncertainty of the fair BSS visible in the shading of Fig. 6, shows a large overlap between both fair BSSs. This underlines that the reduction in forecast skill horizon might be mainly due to the sample size. However, the lower skill could be explained by a slightly higher occurrence rate of neutral regimes and/or different regime transitions compared to ERA5.

In summary, Fig. 6 shows that the ECMWF extended range forecast skill level is very similar in predicting persistent CPs compared to daily CPs, with significant skill at a forecast range of 10 to 15 days. This highlights that the ECMWF extended forecast model is able to represent persisting CPs at a similar level to daily patterns, making it appropriate to use for conditional/pattern-based forecast. Overall, the model is able to forecast the CPs at similar skill level to traditional weather regimes (Ferranti et al. 2018; Bueler et al., 2021).

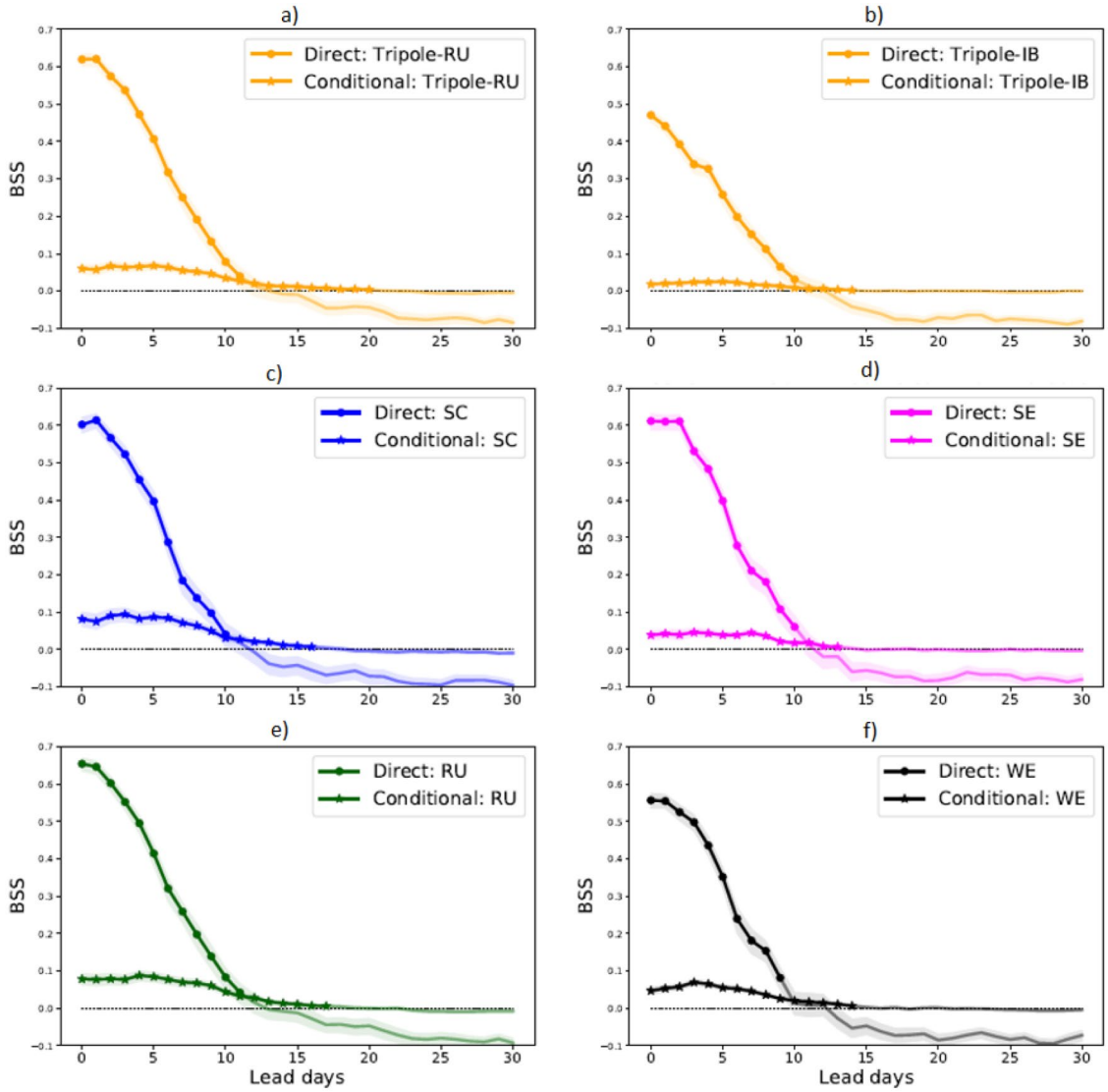

Fig. 7 Direct forecast (circle markers) and conditional (star markers) for 2mT exceeding the 90th percentile. Faded lines without markers represent a BSS not significantly higher than 0. Shading represents the 10th to 90th percentile

3.2 Conditional forecast

Using the approach detailed in the methods section, the skill of the conditional forecast is evaluated and compared with the direct forecast. The first comparison is made for 2mT exceeding the 90th percentile integrated over each key region.

As explained previously, the conditional forecast is evaluated using the original BS and not the fair BS. For the purpose of the comparison, the direct forecast will also be evaluated using the original BS. This significantly reduces the forecast range over the different regions from a lead time of more than 15 days to 10–12 days (Fig. 7). Therefore, the skill shown here for the conditional forecast will be an underestimation of the potential skill of the ECMWF model which uses 101 ensemble members operationally.

Figure 7 shows the BSS for both direct and conditional forecast for the different regions. The first observation is the low score of the conditional forecast at short lead times compared to the direct forecast. This was also seen in previous studies using similar methods (Mastrantonas et al. 2022; Bloomfield et al. 2021). It can be explained by the fact that the conditional skill is dependent on the conditional probability, therefore even a 'perfect' forecast (perfect knowledge of the CP) cannot reach a perfect skill score. This was demonstrated by Bloomfield et al. (2021), who used a pattern-based method for energy forecasting. To showcase the relevance of their approach, the skill for the pattern-based method was shown when having perfect knowledge

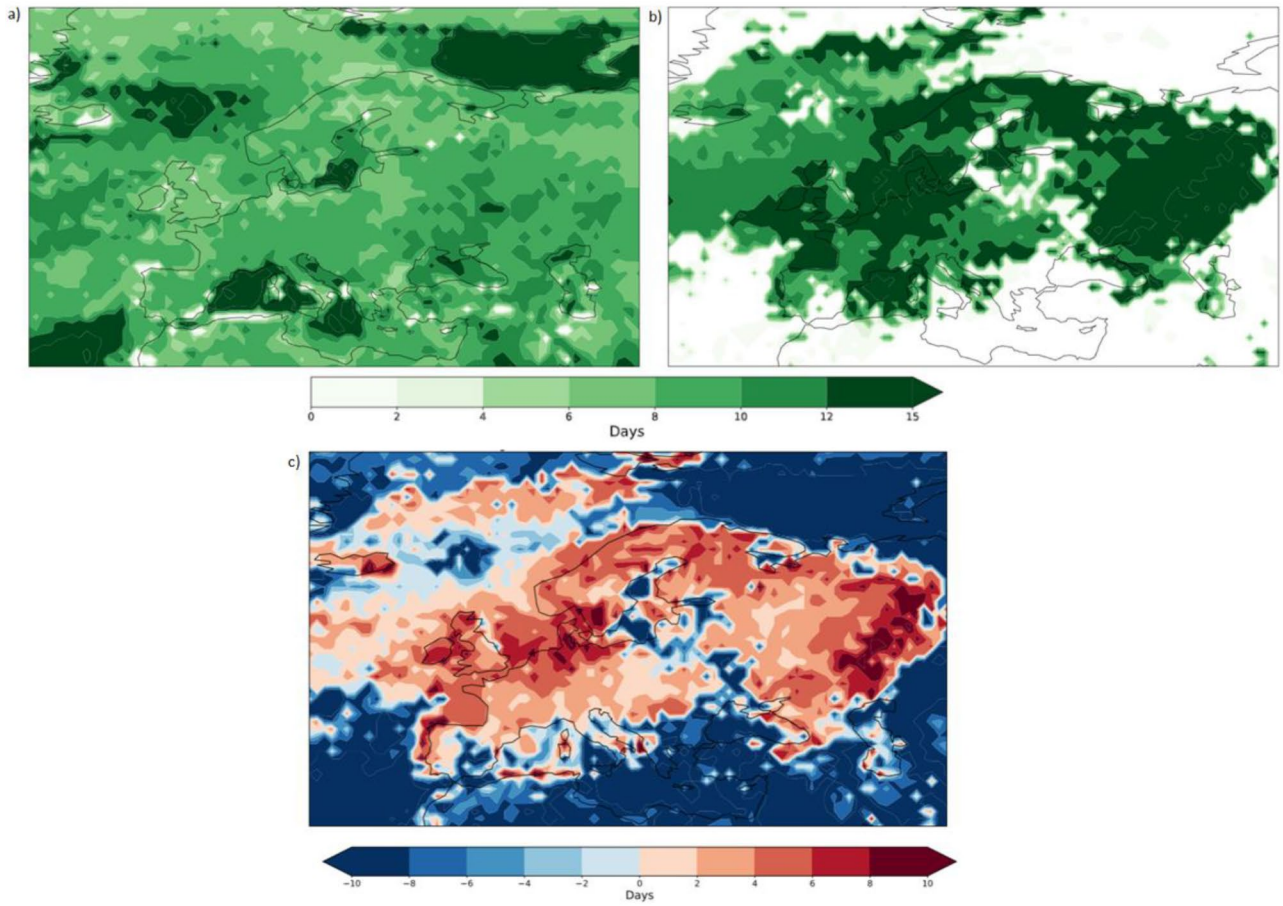

Fig. 8 Forecast skill horizon for direct and conditional forecast of 2mT exceeding the 90th percentile at each grid point based on 1000 bootstraps. (**a**) Direct forecasting; (**b**) Conditional forecasting; (**c**) Conditional minus direct forecasting

of patterns (pattern forecast done with patterns from ERA5 at a given date). Perfect knowledge of the patterns allowed for the skill to remain constant and perform better at an extended range compared to grid-point forecast The skill is dependent on the conditional probability of the variable in question which is therefore the limiting factor of the pattern-based method.

The second observation is that the BSS for the conditional forecast crosses the BSS for the direct forecast at around 10 days. Around this time range, the BSS for the direct forecast decreases significantly below 0 while the conditional forecast significantly stays above 0. The BSS of the conditional forecast decreases at a slower rate with lead time, showing that beyond 10 days, the use of CPs can be advantageous compared to the direct forecast. The pattern-based forecast remains skilful at a longer lead time than the direct forecast. The RU regions (RU and Tripole-RU), with the SC and WE regions, experience the greatest benefit with a gain of about five days lead time. The SE and Tripole-Iberian regions both show a gain of three and four days respectively. For most regions the improvement in skill extends the forecast skill horizon to the limit of the medium range, for RU and SC regions this extends into the start of the subseasonal range (in particular Tripole-RU with skilful forecast up to day 20). The higher skill seen for RU regions has been seen in previous studies (Domeisen et al. 2023; Wulf & Domeisen, 2019).

As mentioned in the previous paragraph, the skill is dependent on the conditional probability. The regions with the most skill across lead time and the highest increase in forecast skill horizon correspond to regions with the highest conditional probability. Figure 8 highlights this result even more and compares the forecast skill horizon of the direct forecast with the skill horizon of the pattern-based forecast, at each grid-point. This representation points out how using the CPs can improve the prediction locally. The improvement is specific over regions with the highest conditional probability (Fig. 1). Comparatively to Fig. 7, the forecast range is enhanced by up to ten days at the grid point scale (Fig. 8.c), in particular over land (RU, SC, British Isles and northern Europe). With the direct forecast, over the ocean, the predictability is already quite high therefore

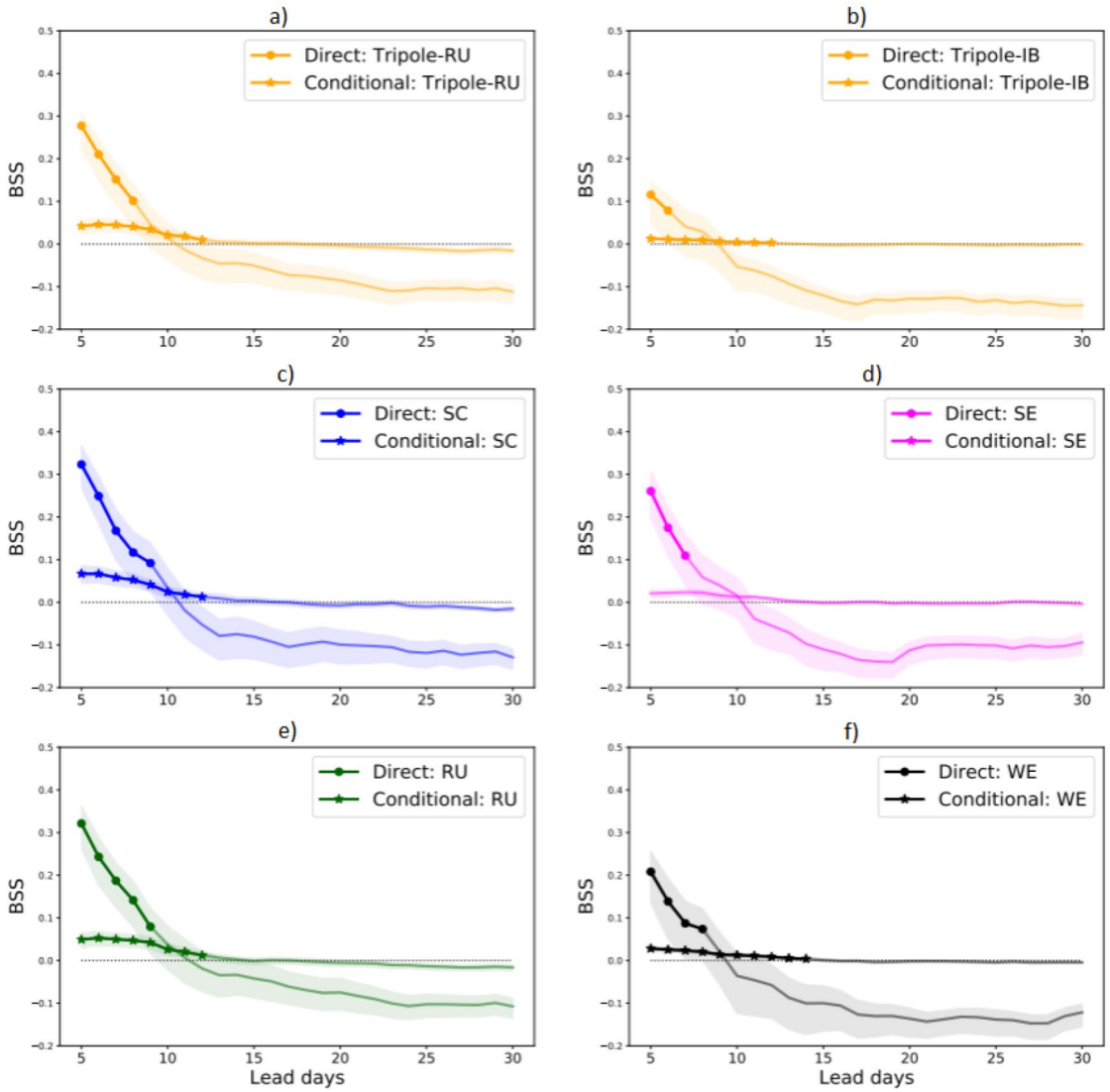

Fig. 9 Direct forecast (circle markers) and conditional (star markers) for 2mT exceeding the 90th percentile for 4 days. Faded lines without markers represent a BSS not significantly higher than 0. Shading represents the 10th to 90th percentile

there is no benefit in using the CPs. This is due to the high memory of the ocean. An important point to highlight, is the relatively small improvement over Greece and the Balkans region, which coincides with a lower conditional probability, explaining the limited improvement in forecast skill horizon.

Using the same methodology, the direct and conditional forecast of persistent extreme warm temperatures are investigated. Here, the 2mT exceeds the 90th percentile for four days is forecasted while the CPs used for the conditional forecast persist for 5 days or more. Similarly, the BSS of the conditional forecast is comparatively low at short to medium range (Fig. 9). Around day 10, the BSSs cross and most regions observe an increase in forecast skill horizon (besides the SE region). This increase is more modest with the RU (including the Tripole Russian region) and SC regions seeing an increase of three to four days while the Tripole Iberian and WE regions experience a larger increase of up to six days. While there is significant improvement, the forecast skill horizon for persistent extreme heat only reaches the edge of the medium range. The difference in skill and improvement has a similar explanation as for the forecast of daily extreme warm temperatures. Regions with the highest skill (Tripole RU, SC, RU) also have the largest conditional probability of persistent extreme warm temperatures. While the Tripole Iberian region sees the largest benefit, the skill is overall low across all lead times compared to other regions. The SE region is the only region with no significant improvement. As for daily 2mT extremes, the conditional probability

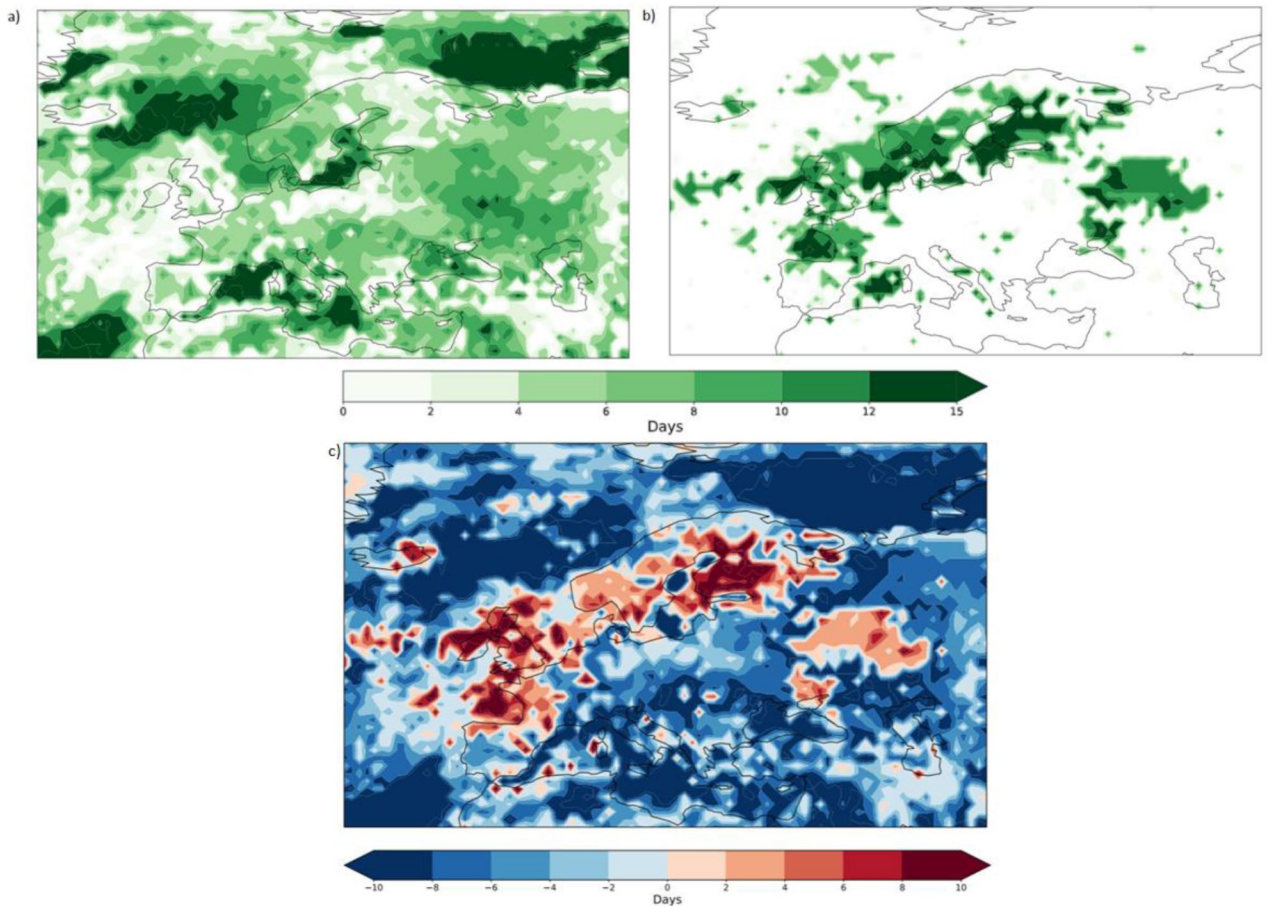

Fig. 10 Forecast skill horizon for direct and conditional forecast of 2mT exceeding the 90th percentile during 4 days at each grid point based on 1000 bootstraps. (**a**) Direct forecasting; (**b**) Conditional forecasting; (**c**) Conditional minus direct forecasting

over this region is among the lowest compared to other regions. Additionally, as seen in Rouges et al. (2023), the SE heatwaves coincide less with persistent CPs compared to other heatwave types. In fact, the average length of CPs coinciding with heatwaves is among the lowest for the SE heatwaves (Rouges et al. 2023). This combination of factors might explain the limited influence of persistent CPs and therefore the reduced gain in skill compared to other regions.

This difference is further highlighted in Fig. 10. The forecast skill horizon is improved very locally with almost no improvement across central to southern Europe. Most of the improvement is focused on western Europe, southern Scandinavia, and RU. The conditional probability for 2mT exceeding the 90th percentile for four days is significantly more important than climatology but remains low in comparison with the conditional probability of extreme temperatures for a given day. It is also generally more localised, therefore further limiting the potential in skill improvement. Additionally, persistent CPs are less frequent, therefore the forecast is done on a very limited sample. This partly explains why the improvement in forecast skill horizon is low and limited to key regions, and why the SE region shows no benefit.

Operationally, forecast products at the subseasonal time scale are shown as weekly average anomalies. While this paper discusses the potential of CPs as predictors in the subseasonal range, and improvements on the daily timescale are more evident (Mastrantonas et al. 2022), BSS of weekly forecast has been computed (not shown). Beyond week 2 (days 8–14; as per Pyrina & Domeisen, 2019) the pattern-based forecast is significantly more skilful than the direct methodology. At week 3 (days 15–21) for most regions the forecast is not significantly skilful, but for the SC, RU and Tripole-IB regions the pattern-based forecast is still skilful into the subseasonal range.

Using CPs has shown that additional forecast information can be extracted, extending the forecast skill horizon for extreme warm temperatures to the edge of the medium range and for some regions into the subseasonal range.

4 Conclusion

Subseasonal forecast has received a lot of support and interest in recent years with the establishment of the Subseasonal to Seasonal Prediction project by the World Weather Research Programme/World Climate Research Programme. At this timescale, the forecast of surface parameters is limited in skill while the atmospheric circulation can be represented skilfully (Ferranti et al. 2015; Matsueda and Palmer 2018; Robertson et al. 2020). The atmospheric circulation can be characterised into different main circulation patterns (CPs) or weather regimes which impact the surface weather (Ferranti et al. 2018; Grams et al. 2017; van der Wiel et al. 2019a; Mastrantonas et al. 2021). The use of these patterns to infer surface conditions has been the focus of multiple recent studies (Grams et al. 2020; Mastrantonas et al. 2022; van der Wiel et al. 2019b; Bloomfield et al. 2021). In this research, CPs associated with extreme warm temperatures over Europe are investigated as potential predictors. To this aim, a forecast method using these CPs is constructed and its skill is compared to the conventional or direct forecast method.

The five CPs identified (Rouges et al. 2023), Tripole, Scandinavia (SC), Southern Europe (SE), Russia (RU) and Western Europe (WE), are all associated with an important increase in probability of 2mT exceeding the 90th percentile, by up to four times compared to climatology, over the heatwave (HW) region.

Based on these results, the CPs are used to provide a forecast for the probability of 2mT exceeding the 90th percentile. The principle of this forecast is to forecast the CPs and use the conditional probability of 2mT exceeding the 90th percentile threshold to provide a prediction for extreme heat. The skill of this pattern-based forecast or conditional forecast has been evaluated with the Brier Skill Score (BSS) using 20 years of reforecast from the ECMWF extended range model.

The results show that the ECMWF extended range model is able to well represent the CPs, with limited bias. Additionally, the BSS showed a similar skill by the model to forecasting the CPs compared to traditional weather regimes (Ferranti et al. 2018; Bueler et al., 2021). Interestingly, persistent CPs generally have a similar skill than daily CPs with a limited reduction in forecast skill horizon. This reduction is partly due to the limited sample size of persistent CPs. This further highlights that the model is able to represent the patterns sufficiently well to be used to construct a pattern-based forecasting method.

The forecast skill horizon of the conditional forecast is compared to the forecast skill horizon of the direct forecast. While at short to medium range, the pattern based forecast has a lower skill, beyond 10 days the conditional forecast exhibits enhanced skill compared to the conventional forecast, indicating its clear advantage. In addition, the conditional BSS remains above zero (better than the climatological reference) for longer, providing a skilful forecast at longer lead times. This allows for an increase of forecast skill horizon by five days on average and by more than six days for some regions (RU, SC, Tripole RU). This increase allows for skilful forecast to extend to the edge of the medium range and into the subseasonal time range for some regions (RU, SC and Tripole RU). By plotting the forecast skill horizon of the conditional forecast for each grid point over Europe, the benefit provided by the CPs specifically over the regions of HWs is more evident. In particular, the regions with the most important improvement are the regions with the highest conditional probability for 2mT to exceed the 90th percentile.

This methodology is also applied for 2mT exceeding the 90th percentile for four days with CPs which persist five days or longer. A more modest increase in skill is observed, in particular for the SE region which does not benefit from this method. This could be attributed to the low conditional probability and generally lower persistence in CPs associated with these HWs.

The impact of HWs and extreme heat is a growing concern that requires accurate forecast in particular at longer lead times. Providing skilful forecast at longer lead times enables early warnings which helps anticipating and mitigating the impact of such disasters. Overall, this research highlights the potential of CPs in predicting extreme warm temperatures at a longer time range than the direct forecast. Integrating this type of forecasting system within a decision-making framework could help the agricultural and energy sector in addition to local authorities in taking key decisions earlier and better prepare for extreme events. The energy sector for instance already uses weather regimes in their operations. Using patterns specifically designed for extreme heat might in some cases be more appropriate, as highlighted by Bloomfield et al. (2021). Further studies to evaluate the usefulness and value of such forecast for the different sectors are however essential. The skill metrics used here only allow to determine that the pattern based forecast is better than climatology until a certain lead time, but does not inform on the usefulness to stakeholders. Determining this requires metrics tailored to the users needs (Bloomfield et al. 2021; Dorrington et al. 2020).

Other predictors should also be tested and taken into consideration when relevant. For instance, Mastrantonas et al. (2021) also concluded of the necessity of testing teleconnections or other parameters such as water vapor flux to predict extreme precipitation. In Rouges et al. (2023), pre-existing drought conditions have been linked to more extreme cases. Teleconnections with the tropics also highlighted the possibility of

'windows of opportunities' i.e. periods of higher predictability and therefore higher confidence in the forecast, however only for specific cases and with limited significance. Long-lived Rossby wave packets which can be triggered by enhanced convection over the tropics have been shown to improve predictability of heatwaves (Domeisen et al. 2023). Over the USA, bad coupling between soil moisture and temperature in the models could be responsible for lower prediction skill of heatwaves (Ford et al. 2018), highlighting the importance of surface-atmosphere interactions. In Bloomfield et al. (2021) it was also highlighted how the use of CPs can enable to identify 'windows of opportunity'. Lastly, it is important to mention the rise of machine learning in numerical prediction with some machine learning models reaching similar performance than the ECMWF ensemble mean using 2mT, Z500 and soil moisture as predictors (Weirich-Benet et al. 2023).

This study highlighted the potential of using CPs to extend the forecast skill horizon for the prediction of extreme warm temperatures into the subseasonal range.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00382-024-07390-0.

Acknowledgements This work is part of the EU International Training Network (ITN) Climate Advanced Forecasting of sub-seasonal Extremes (CAFE). The project receives funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 813844.

Data availability The ERA5 data used are available through the ECMWF archive (https://www.ecmwf.int/en/forecasts/access-forecasts/access-archive-datasets).

Declarations

Conflict of interest The authors declare that there are no conflicts of interest.

References

Barriopedro D, Fischer EM, Luterbacher J, Trigo RM, García-Herrera R (2011) The hot summer of 2010: redrawing the temperature record map of Europe. Science 332(6026):220–224. https://doi.org/10.1126/science.1201224Epub 2011 Mar 17. PMID: 21415316

Bloomfield HC, Brayshaw DJ, Gonzalez PLM, Bloomfield HC, Brayshaw DJ, Gonzalez PLM, Charlton-Perez A (2020) Sub-seasonal forecasts of demand and wind power and solar power generation for 28 European countries. *Earth Syst. Sci. Data*, *13*. https://doi.org/10.5194/essd-13

Bloomfield HC, Brayshaw DJ, Gonzalez PLM, Charlton-Perez A (2021) Pattern-based conditioning enhances sub-seasonal prediction skill of European national energy variables. Meteorol Appl 28(4). https://doi.org/10.1002/met.2018

Brimicombe C, di Napoli C, Cornforth R, Pappenberger F, Petty C, Cloke HL (2021) Borderless heat hazards with bordered impacts. Earth's Future 9(9). https://doi.org/10.1029/2021EF002064

Büeler D, Ferranti L, Magnusson L, Quinting JF, Grams CM (2021) Year-round sub-seasonal forecast skill for Atlantic–European weather regimes. Q J R Meteorol Soc 147(741):4283–4309

Buizza R, Leutbecher M (2015) The forecast skill horizon. Q J R Meteorol Soc 141:3366–3382. https://doi.org/10.1002/qj.2619

Burston C (2021) and Cecco L. 'There's nothing left in Lytton': the Canadian village destroyed by wildfire – picture essay. *The Guardian*. https://www.theguardian.com/world/2021/jul/25/lytton-canada-heat-wildfire-record-temperatures [accessed 17 February 2022]

Cassou C (2008) Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. Nature 455(7212):523–527. https://doi.org/10.1038/nature07286

Cassou C, Hurrell JW, Deser C (2004) North Atlantic Winter Climate Regimes: Spatial Asymmetry, Stationarity with Time, and Oceanic Forcing

di Napoli C, Pappenberger F, Cloke HL (2018) Assessing heat-related health risk in Europe via the Universal Thermal Climate Index (UTCI). Int J Biometeorol 62(7):1155–1165. https://doi.org/10.1007/s00484-018-1518-2

Domeisen DI, White CJ, Afargan-Gerstman H, Muñoz ÁG, Janiga MA, Vitart F, Tian D (2022) Advances in the subseasonal prediction of extreme events: relevant case studies across the globe. Bull Am Meteorol Soc 103(6):E1473–E1501

Domeisen DI, Eltahir EA, Fischer EM, Knutti R, Perkins-Kirkpatrick SE, Schär C, Wernli H (2023) Prediction and projection of heatwaves. Nat Reviews Earth Environ 4(1):36–50

Dorrington J, Finney I, Palmer T, Weisheimer A (2020) Beyond skill scores: exploring sub-seasonal forecast value through a case-study of French month-ahead energy prediction. Q J R Meteorol Soc 146(733):3623–3637

Emerton R, Brimicombe C, Magnusson L, Roberts C, Di Napoli C, Cloke HL, Pappenberger F (2022) Predicting the unprecedented: forecasting the June 2021 Pacific Northwest heatwave. Weather 77:272–279. https://doi.org/10.1002/wea.4257

European Centre for Medium-range Weather Forecast (ECMWF). (accessed in (2023) 47r3 ENS Scorecard. https://sites.ecmwf.int/ifs/scorecards/scorecards-47r3ENS.html

Ferranti L, Klinker E, Hollingsworth A, Hoskins BJ (2002) Diagnosis of systematic forecast errors dependent on flow pattern. Q J R Meteorol Soc 128:1623–1640. https://doi.org/10.1002/qj.200212858312

Ferranti L, Corti S, Janousek M (2015) Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. Q J R Meteorol Soc 141(688):916–924. https://doi.org/10.1002/qj.2411

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? Q J R Meteorol Soc 144(715):1788–1802. https://doi.org/10.1002/qj.3341

Ferro CAT (2014) Fair scores for ensemble forecasts. Q J R Meteorol Soc 140(683):1917–1923. https://doi.org/10.1002/qj.2270

Ferro CAT, Richardson DS, Weigel AP (2008) On the effect of ensemble size on the discrete and continuous ranked probability scores. Meteorol Appl 15(1):19–24. https://doi.org/10.1002/met.45

Ford TW, Dirmeyer PA, Benson DO (2018) Evaluation of heat wave forecasts seamlessly across subseasonal timescales. NPJ Clim Atmospheric Sci 1(1):20

Gonzalez PLM, Howard E, Ferrett S, Frame THA, Martínez-Alvarado O, Methven J, Woolnough SJ (2022) Weather patterns in Southeast Asia: enhancing high-impact weather subseasonal forecast skill. Q J R Meteorol Soc. https://doi.org/10.1002/qj.4378

Grams CM, Beerli R, Pfenninger S, Staffell I, Wernli H (2017) Balancing Europe's wind-power output through spatial deployment informed by weather regimes. Nat Clim Change 7(8):557–562. https://doi.org/10.1038/NCLIMATE3338

Grams CM, Ferranti L, Magnusson L (2020) How to make use of weather regimes in extended-range predictions for Europe, Newsletter (165), 14–19. https://www.ecmwf.int/en/newsletter/165/meteorology/how-make-use-weather-regimes-extended-range-predictions-europe

Hannachi A, Straus DM, Franzke CLE, Corti S, Woollings T (2017) Low-frequency nonlinearity and regime behavior in the Northern Hemisphere extratropical atmosphere. Reviews of Geophysics, vol 55. Blackwell Publishing Ltd, pp 199–234. 1 https://doi.org/10.1002/2015RG000509

Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, Thépaut JN (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146(730):1999–2049

Kiladis GN, Dias J, Straub KH, Wheeler MC, Tulich SN, Kikuchi K, Weickmann KM, Ventrice MJ (2014) A comparison of OLR and circulation-based indices for tracking the MJO. Mon Weather Rev 142(5):1697–1715. https://doi.org/10.1175/MWR-D-13-00301.1

Krouma M, Yiou P, Déandreis C, Thao S (2022) Assessment of stochastic weather forecast of precipitation near European cities, based on analogs of circulation. Geosci Model Dev 15(12):4941–4958. https://doi.org/10.5194/gmd-15-4941-2022

Lee JY, Wang B, Wheeler MC, Fu X, Waliser DE, Kang IS (2013) Real-time multivariate indices for the boreal summer intraseasonal oscillation over the Asian summer monsoon region. Clim Dyn 40(1–2):493–509. https://doi.org/10.1007/s00382-012-1544-4

Lin H, Mo R, Vitart F (2022) The 2021 western north American heatwave and its subseasonal predictions. Geophys Res Lett 49:e2021GL097036. https://doi.org/10.1029/2021GL097036

Magnusson L (2019) Technical Memo (851). ECMWF Severe event catalogue for evaluation of multi-scale prediction of extreme weather

Mastrantonas N, Herrera-Lormendez P, Magnusson L, Pappenberger F, Matschullat J (2021) Extreme precipitation events in the Mediterranean: spatiotemporal characteristics and connection to large-scale atmospheric flow patterns. Int J Climatol 41(4):2710–2728. https://doi.org/10.1002/joc.6985

Mastrantonas N, Magnusson L, Pappenberger F, Matschullat J (2022) What do large-scale patterns teach us about extreme precipitation over the Mediterranean at medium- and extended-range forecasts? Q J R Meteorol Soc 148(743):875–890. https://doi.org/10.1002/qj.4236

Matsueda M (2011) Predictability of Euro-Russian blocking in summer of 2010. Geophys Res Lett 38(6). https://doi.org/10.1029/2010GL046557

Matsueda M, Palmer TN (2018) Estimates of flow-dependent predictability of wintertime Euro-Atlantic weather regimes in medium-range forecasts. Q J R Meteorol Soc 144(713):1012–1027. https://doi.org/10.1002/qj.3265

Michelangeli P, Vautard R, Legras B (1995) Weather regimes: recurrence and quasi stationarity. J Atmos Sci 52(8):1237–1256. https://doi.org/10.1175/1520-0469(1995)052<1237:WRRAQS>2.0.CO;2

Miralles DG, Gentine P, Seneviratne SI, Teuling AJ (2019) Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges. Ann N Y Acad Sci 1436(1):19–35. https://doi.org/10.1111/nyas.13912

Molteni F, Tibaldi S, Palmer TN (1990) Regimes in the wintertime circulation over northern extratropics. I: observational evidence. Q J R Meteorol Soc 116:31–67. https://doi.org/10.1002/qj.49711649103

Owens RG, Hewson T (2018) ECMWF forecast user guide. *Reading: ECMWF*, *10*(870), m1cs7h

Pfahl S (2014) Characterising the relationship between weather extremes in Europe and synoptic circulation features. Nat Hazards Earth Syst Sci 14(6):1461–1475. https://doi.org/10.5194/nhess-14-1461-2014

Pyrina M, Domeisen DI (2023) Subseasonal predictability of onset, duration, and intensity of European heat extremes. Q J R Meteorol Soc 149(750):84–101

Richardson D, Fowler HJ, Kilsby CG, Neal R, Dankers R (2020) Improving sub-seasonal forecast skill of meteorological drought: a weather pattern approach. Nat Hazards Earth Syst Sci 20(1):107–124. https://doi.org/10.5194/nhess-20-107-2020

Robertson AW, Vigaud N, Yuan J, Tippett MK (2020) Toward identifying subseasonal forecasts of opportunity using north American weather regimes. Mon Weather Rev 148(5):1861–1875. https://doi.org/10.1175/MWR-D-19-0285.1

Robine JM, Cheung SLK, le Roy S, van Oyen H, Griffiths C, Michel JP, Herrmann FR (2008) Death toll exceeded 70,000 in Europe during the summer of 2003. Comptes Rendus - Biologies 331(2):171–178. https://doi.org/10.1016/j.crvi.2007.12.001

Rouges E, Ferranti L, Kantz H, Pappenberger F (2023) European heatwaves: link to large-scale circulation patterns and intraseasonal drivers. Int J Climatol 43(7):3189–3209. https://doi.org/10.1002/joc.8024

Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Liniger MA, Appenzeller C (2004) The role of increasing temperature variability in European summer heatwaves. Nature 427(6972):332–336. https://doi.org/10.1038/nature02300

Stefanon M, Dandrea F, Drobinski P (2012) Heatwave classification over Europe and the Mediterranean region. Environ Res Lett 7(1). https://doi.org/10.1088/1748-9326/7/1/014023

Straus DM, Corti S, Molteni F (2007) Circulation regimes: chaotic variability versus SST-forced predictability. J Clim 20(10):2251–2272. https://doi.org/10.1175/JCLI4070.1

Sutanto SJ, Vitolo C, di Napoli C, D'Andrea M, van Lanen HAJ (2020) Heatwaves, droughts, and fires: exploring compound and cascading dry hazards at the pan-european scale. Environ Int 134. https://doi.org/10.1016/j.envint.2019.105276

van der Wiel K, Bloomfield HC, Lee RW, Stoop LP, Blackport R, Screen JA, Selten FM (2019a) The influence of weather regimes on European renewable energy production and demand. Environ Res Lett 14(9). https://doi.org/10.1088/1748-9326/ab38d3

van der Wiel K, Stoop LP, van Zuijlen BRH, Blackport R, van den Broek MA, Selten FM (2019b) Meteorological conditions leading to extreme low variable renewable energy production and extreme high energy shortfall. Renew Sustain Energy Rev 111:261–275. https://doi.org/10.1016/j.rser.2019.04.065

Vautard R, Yiou P, D'Andrea F, de Noblet N, Viovy N, Cassou C, Polcher J, Ciais P, Kageyama M, Fan Y (2007) Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit. Geophys Res Lett 34(7). https://doi.org/10.1029/2006GL028001

Vitart F, Robertson AW (2018) The sub-seasonal to seasonal prediction project (S2S) and the prediction of extreme events. Npj Clim Atmospheric Sci 1(1). https://doi.org/10.1038/s41612-018-0013-0

Vitolo C, di Giuseppe F, Krzeminski B, San-Miguel-ayanz J (2019) Data descriptor: a 1980–2018 global fire danger re-analysis dataset for the Canadian fire weather indices. Sci Data 6. https://doi.org/10.1038/sdata.2019.32

Weirich-Benet E, Pyrina M, Jiménez-Esteve B, Fraenkel E, Cohen J, Domeisen DIV (2023) Subseasonal Prediction of Central European Summer heatwaves with Linear and Random Forest Machine Learning Models. Artif Intell Earth Syst 2(2):e220038. https://doi.org/10.1175/AIES-D-22-0038.1

Wilks DS (2011) Statistical methods in the atmospheric sciences, vol 100. Academic

Wulff CO, Domeisen DI (2019) Higher subseasonal predictability of extreme hot European summer temperatures as compared with average summers. Geophys Res Lett 46:11520–11529

Yiou P (2014) Anawege: a weather generator based on analogues of atmospheric circulation. Geosci Model Dev 7(2):531–543

Yiou P, Déandréis C (2019) Stochastic ensemble climate forecast with an analogue model. Geosci Model Dev 12(2):723–734. https://doi.org/10.5194/gmd-12-723-2019

Zschenderlein P, Fink AH, Pfahl S, Wernli H (2019) Processes determining heat waves across different European climates. Q J R Meteorol Soc 145(724):2973–2989. https://doi.org/10.1002/qj.3599