D3.4
Report on verification of the atmospheric flow features in sub-seasonal forecasts

Lead Beneficiary: ECMWF
Work Package: WP3
Dissemination level: PUBLIC
Type: Report
Due date: August 2021
Submission date: 01.07.2021
Authors: Nikolaos Mastrantonas, Linus Magnusson, Jörg Matschullat

Disclaimer: This report compiles material from very recent publications of the authors, partly without new wording.

1. Introduction

Extreme events like floods, droughts, and heatwaves, pose a great threat to our societies, economies, and environment. Being able to predict such events long in advance can support mitigating their negative impacts.

In this work, we investigated the applicability of large-scale atmospheric flow patterns to provide useful information about upcoming Extreme Precipitation Events (EPEs) in the Mediterranean for days and weeks in advance.

Extreme precipitation is one of the most frequent threats over the domain, resulting in high economic losses, injuries and casualties (Llasat et al. 2013). A threat that is becoming more crucial in the light of the non-stationary climate (Hannaford et al. 2021) and ongoing climate change, leading to intensifying frequency and magnitude of such extremes in many locations (Kostopoulou and Jones 2005; Papalexiou and Montanari 2019).

Previous works already identified strong connections between precipitation and large-scale atmospheric flow patterns over different parts of the world, like the UK, the Mediterranean, central Europe and India (Grazzini et al. 2020; Hoy et al. 2014; Neal et al. 2020; Richardson et al. 2020; Toreti et al. 2010). Motivated by such works, we investigated how domain-specific large-scale patterns can be used to indirectly predict localized extreme precipitation over the Mediterranean. We made use of data from the ERA5 dataset, and ECMWF extended-range forecasts. Both datasets are freely available (Copernicus Climate Data Store for the former, and S2S database for the latter).

This document provides a summary of the methodology applied and key findings. Section 2 presents the large-scale patterns and their connections to localized extreme precipitation, while section 3 shows the key findings about the predictability of patterns and related extreme precipitation. Finally, section 4 summarises the main conclusions and provides possible pathways for future research.

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

A detailed description of all the information above is available in a peer-reviewed publication (section 2: Mastrantonas et al. 2021a) and a preprint, which was recently submitted for review to the Quarterly Journal of the Royal Meteorological Society (section 3: Mastrantonas et al. 2021b). A recorded talk with many of the findings (https://www.youtube.com/watch?v=W8k5Dy5rlnE), as well as a summary video with the main steps of this work (https://www.youtube.com/watch?v=idWnrXLg-Ao) , are available in the CAFE YouTube channel.

2. Large-scale patterns and extreme precipitation

The large-scale patterns were derived by Mastrantonas et al. (2021a), using ERA5 data between 1979–2019. The analysis was based on Empirical Orthogonal Function (EOF) and subsequent K-means clustering of the daily anomalies of Sea Level Pressure (SLP) and geopotential height at 500 hPa (Z500). The necessary number of modes (principal components) from EOF analysis, explaining at least 90% of the total variance, was kept, meaning 7 and 6 for SLP and Z500, respectively. Each day was allocated to one of nine clusters, the composites of which represent the nine Mediterranean patterns. Domain and variables used for generating the patterns, as well as the number of clusters, were selected so that the derived patterns have a strong association with localized EPEs over the Mediterranean and exhibit distinct synoptic-scale features over the domain.

As presented in figure 1, the patterns indicate negative anomalies over the western Mediterranean (Atlantic/Biscay/Iberian Low), and the eastern Mediterranean (Balkan/Black Sea Low), positive anomalies over the whole domain (Mediterranean High), or, finally, non-anomalous conditions over the entire domain (Minor High, Minor Low).

These 9 patterns have a clear seasonality in their frequencies, as shown in Figure 2. The Minor Low and Minor High patterns occur mainly in summer-half years, as during this period the anomalies are generally of substantially smaller magnitude compared to winter-half years. This explains also why the Biscay Low, Iberian Low, Balkan Low, Black Sea Low, and Mediterranean High, patterns with high-magnitude anomalies, occur mainly in winter-half

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

years. The climatological frequencies of the Atlantic Low and Sicilian Low do not deviate depending on the season. For the Atlantic Low, this can be explained by the centre of the negative anomaly, which is located over the Atlantic and in relatively high latitudes. In such a location, phenomena that substantially modulate the pressure fields over large domain take place all year round (e.g., Rossby Waves, cyclonic features). As for the Sicilian Low, it has a low-magnitude negative anomaly centred over central Mediterranean that is confined by strong positive anomalies over the northern parts of the domain. This combination of strong positive anomalies with low negative anomalies results in the stable frequency of this pattern all year-round.

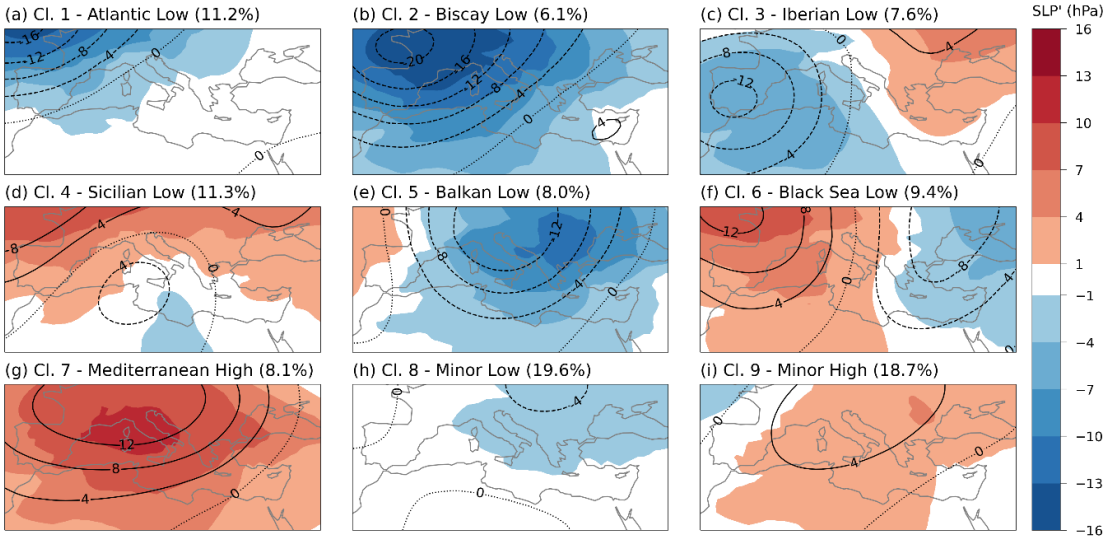

Figure 1. The 9 Mediterranean patterns (Mastrantonas et al. 2021). The figure presents the composites of clusters derived with K-means clustering on the principal components' projections of SLP and Z500 anomalies. Colour shading refers to SLP anomalies (hPa), and contours to Z500 anomalies (dam). Percentages indicate the climatological frequencies of each cluster.

The connection between patterns and EPEs was quantified with the conditional probability of observing EPEs at each grid cell given each of the nine patterns. We analysed the conditional probabilities of 95th percentile extremes (P95) considering full-year statistics for the period 1979–2020. Due to the high seasonal variation of the pattern frequencies, we also applied half-year statistics, with summer-halves referring to April 16 to October 15 (including both dates), while winter-halves represent the remaining dates,

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

respecting the patterns' climatology (Fig. 2). Note that EPEs are always derived based on full-year statistics; the only difference is that the conditional probabilities are refined to winter- and summer-halves in addition to the pattern conditioning. Finer subsetting (e.g., seasonal) was not implemented, as the sample size for some patterns became too small, and conditional probabilities had high fluctuations (a conclusion derived after implementing bootstrapping).

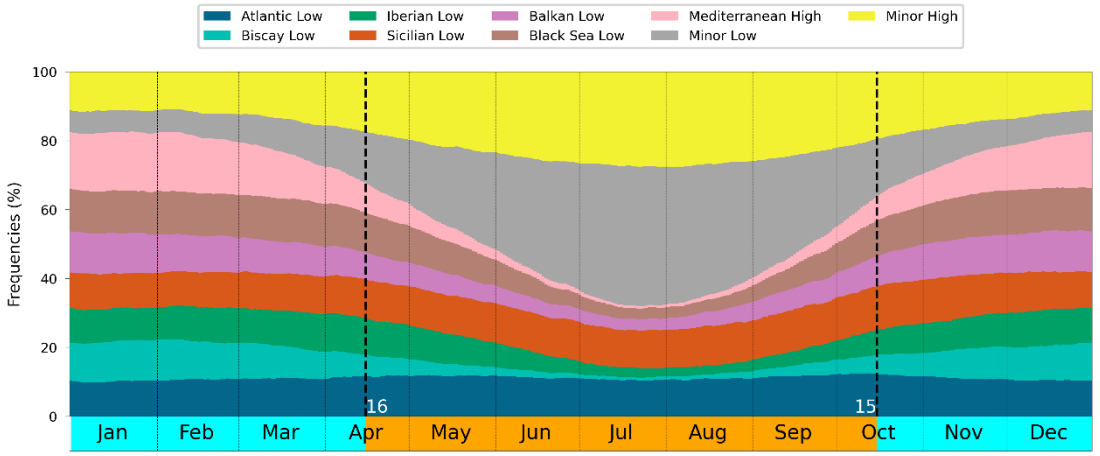

Figure 2. 91-day running mean calendar day climatological relative frequency of the 9 Mediterranean patterns. The dates of the summer-half (April 16–October 15, including both dates) and the winter-half (remaining days) periods were selected, so that the patterns' climatology is respected

Conditional probabilities for the P95 EPEs are presented in Figure 3. Results refer to the full-year statistics (1st column), and winter- and summer-half subsets (2nd and 3rd column respectively). Each pattern is preferentially associated with EPEs at different subdomains. For example, the Biscay Low is the main EPE pattern in parts of Morocco, Iberian Peninsula, France, Italy and western Balkans, while the Black Sea Low mainly affects locations in Turkey. These results can be explained by the air- and subsequent moisture-flow associated with each of the patterns' composites (Fig. 1). Results between the 3 temporal subsets do not vary much; the main differences are identified in the Middle East. These differences are not that crucial for EPEs, as the associated probabilities are very low for these locations and temporal subsets (Fig. 3, 2nd row), meaning that EPEs are generally not expected in these periods and grid cells (in fact, most EPEs over the domain occur in

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

winter-halves; Mastrantonas et al. 2021). The conditional probabilities of EPEs, given the most preferential pattern at each grid cell, are about 3 times higher compared to the climatological ones (5% for P95 EPEs). Especially for locations of high orography and coastal areas, this ratio is even higher (~ 5 times). The half-year subsets drive an increase of conditional probabilities, as they take temporal EPE occurrence into consideration. For most of the domain, most EPEs occur in winter-half years. Thus, the conditional probabilities are higher for this period. Summer-half years is the preferential period for EPEs in north Balkans and Alps; thus, the conditional probabilities are higher for that period for these two regions.

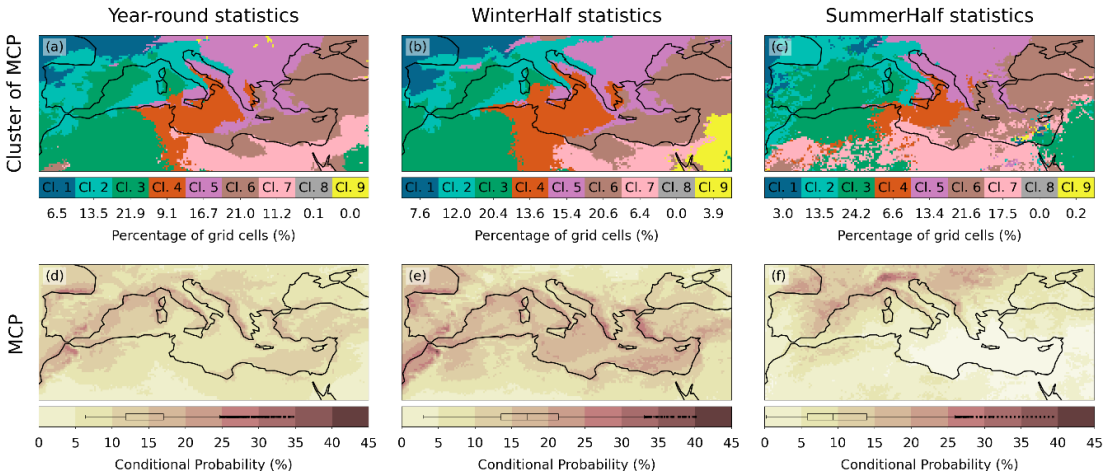

Figure 3. Connection of P95 EPEs and Mediterranean patterns. The 1st row presents the pattern of Maximum Conditional Probability (MCP) at each grid cell; the 2nd row presents associated probabilities. The 1st column is based on full-year conditional probabilities; the 2nd and 3rd are based on the half-year periods (winter- and summer-half years). Note that the EPEs are always derived based on annual analysis. The boxplots at subplots (d)–(f) present the distribution of the conditional probabilities for all grid cells, indicating the median value and extending from the lower to upper quartile. The whiskers extend to the further available value up to 1.5*IQR from the lower and upper quartile; all other values outside this range are presented as outliers.

3. Predictability of patterns and extreme precipitation

We assessed the skill of the ECMWF extended-range forecast in predicting these 9 patterns, as well as predicting EPEs. We used all available data produced with cycle 46r1, with the initiation dates extending between June 11, 2019 and June 30, 2020 (110 dates in total), to provide consistency in model physics and parametrization schemes. As we were interested in long-term statistical analysis, we made use of the reforecasts of these dates. Each

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

reforecast provides ensemble data of 1 control and 10 perturbed members, for the same day-month as the actual forecasts but from 1 up to 20 years in the past. Thus, the total dataset used, consisted of 2,200 initiation dates per lead-time (20 years of reforecasts X 110 dates), with each date having 11 ensemble members.

Each forecasting field was allocated to one of the 9 patterns based on the minimum Euclidian distance. Then, the Brier Score (BS; Wilks 2011) was used to assess the performance of the model in predicting the patterns. The results were compared with the minimum of two reference scores; the BS based on 91-days moving window climatology (centred at the date of interest), and the BS based on half-year transition probabilities (Markov chain). To assess the results' significance, we used bootstrapping of 1,000 resamples with replacement.

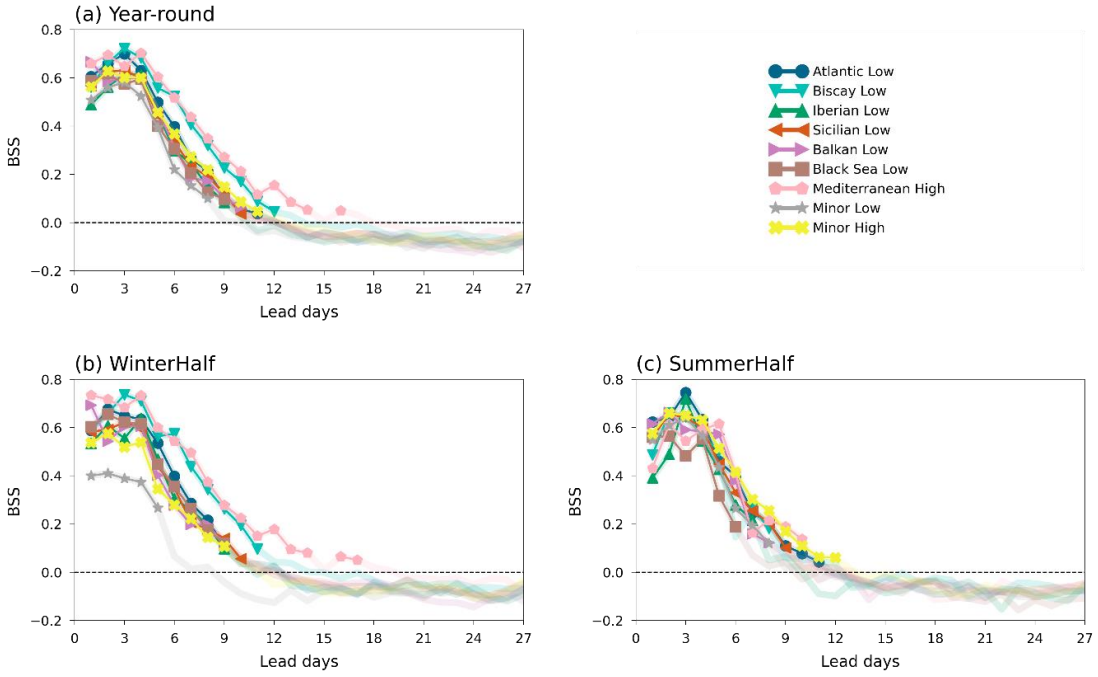

Figure 4. Brier Skill Score (BSS) for the 9 Mediterranean patterns (median value of the bootstraps), considering the annual (a), winter-half (b), and summer-half (c) periods. The points (connected with bold lines) indicate the lead times that the model significantly outperforms the reference score.

Figure 4 presents the performance of the ECMWF model (BSS) in predicting the nine patterns (median value of the bootstraps). The points, connected

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

with bold lines, indicate that the model significantly outperforms the reference scores (90% 1-tailed confidence interval). The model performance is similar for all patterns and both temporal subsets, with a forecasting horizon of about 11 days. The main difference is observed for the winter half and the patterns Mediterranean High and Minor Low. The former has a forecasting horizon of over 2 weeks, while the latter is constrained up to 5 days only. Biscay Low, a pattern highly associated with EPEs (Fig. 2), is also more predictable during winter half.

The skill of the model in predicting EPEs (P95) was assessed also based on BS. We calculated the direct BS for EPEs when considering the model-predicted precipitation, and the indirect BS when using the model-predicted patterns and their climatological connections to EPEs. The latter was derived by substituting each predicted pattern with the relevant conditional probabilities for EPEs at each grid cell. As each reforecast has 11 members, the final conditional probabilities used for the indirect BS were the average of all 11 members at each grid cell. To assess the significance of the results (BSS>0), we used bootstrapping of 1,000 resamples with replacement and considered 90% one-tailed confidence interval. To obtain a larger sample of data at each bootstrap (the actual set had 2,200 at each lead time), and to respect the full-year statistics that EPEs are based on, each resample had 3,000 dates, with climatologically stable number of winter, spring, summer, and autumn days (741, 756, 756, 747 days, respectively). For each resample, we finally calculated the BS for EPEs, given a perfect forecast of the patterns, so we could assess the discriminatory skill of the patterns for the localised EPEs, and its statistical significance.

The forecast performance (median value of the bootstraps) in the predictability of P95 EPEs is presented in Figure 5. The plot shows the performance over selected (large) domains (area-weighted mean of all included grid cells), in terms of BSS for P95 EPEs, based on direct forecasting (using model-predicted precipitation) and indirect one (using the forecasted patterns and their half-year conditional probabilities for EPEs). The direct forecasts have a high skill for short and medium range predictions; the skill drops below the reference score (pale-coloured lines) at about 8 days lead time. The indirect EPE forecasting outperforms the reference score over 10

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

days lead time in general, extending the forecasting horizon. As the discriminatory skill of the patterns nowhere exceeds 20% even for perfect forecasts (Mastrantonas et al. 2021b), the BSS does not exceed 0.2, even at 0-day lead time. From the selected subregions (subplots b-f), northern Morocco has a better skill. This relates to its small domain and homogeneity, with the mountain ranges forming a barrier that direct the flow and force moisture to precipitate.

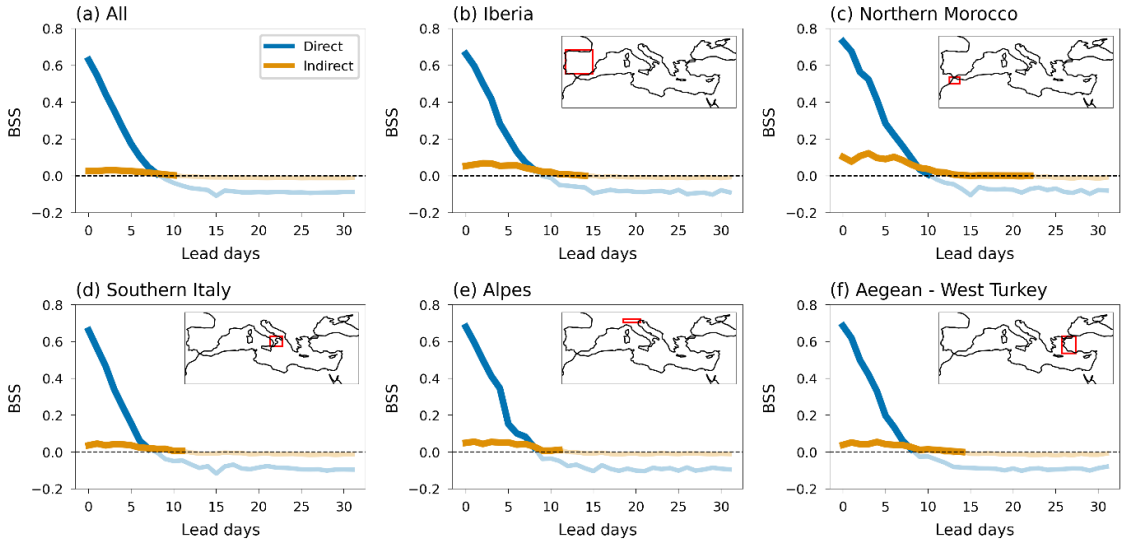

Figure 5. Brier Skill Score for direct (blue colour) and indirect (orange colour) P95 EPEs forecasting for selected subregions.

Figure 6 presents similar results as Figure 5, yet now for each grid cell of the studied domain. Subplots (a) and (b) present the forecasting horizon up to when the model beats the reference score for direct and indirect forecasting (statistically significant results). The results of the indirect forecasting are masked to exclude any grid cell that has no significant discriminatory skill for EPEs when conditioning on the patterns (i.e., assuming a perfect pattern forecast). Significance is derived based on the 1,000 bootstraps per lead day (and considering all lead times used for the analysis) with a 90% one-tailed confidence interval. The indirect forecasting outperforms the reference score even for over 8 days lead time for many regions, especially so for Iberian Peninsula, southern Balkans and western Turkey. In contrast, direct forecasting has skilful predictions only up to 8 days lead time for most locations. It can be noticed from subplot (c) that indirect forecasting

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

extends the forecasting horizon for EPEs by between 3 and 6 days for many of the locations. This is a substantial increase in lead time, which could support informative decision making for various domains, as for example (re)insurance companies and agricultural sector. As the discriminatory skill of the patterns for inferring EPEs is not very strong, indirect forecasting is not beneficial at short lead times. This information outperforms the direct forecasting only after the end of week 1 for most locations as shown in subplot (d). The use of the large-scale patterns for inferring information about EPEs is not helpful for most of the locations over northern Africa and Middle East, as these 2 regions fall within the sinking masses of the Hadley cell, making the connection of localized EPEs and large-scale patterns notably weaker.

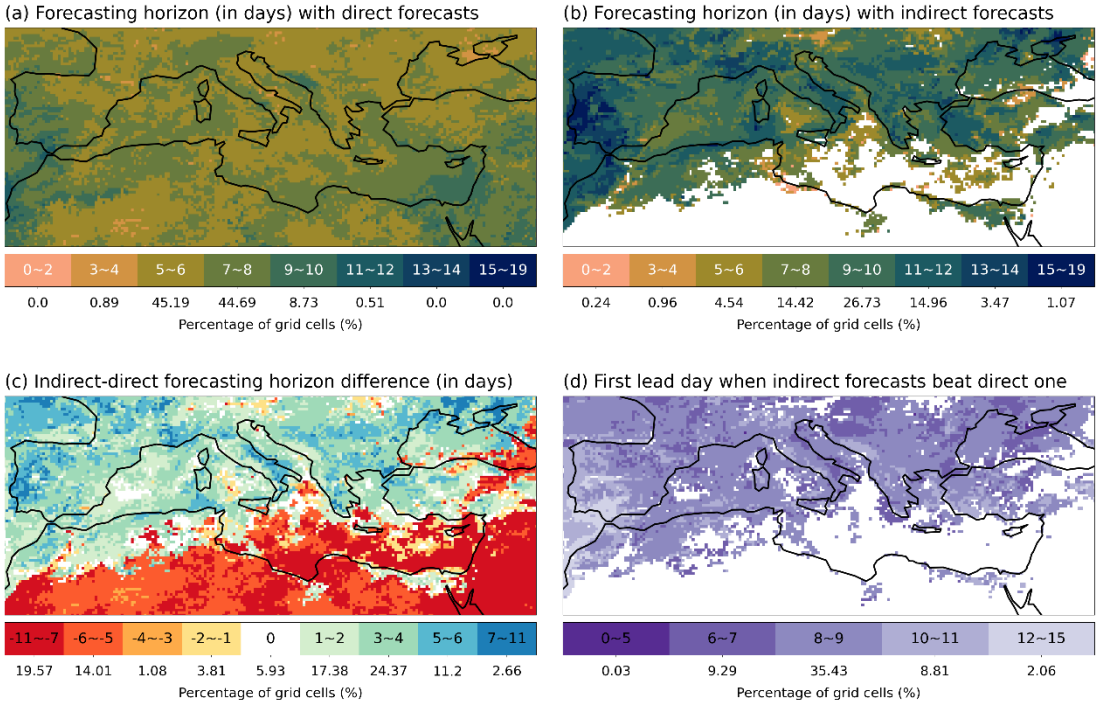

Figure 6. Forecasting horizon (maximum forecast day) up to when the ECMWF model outperforms the reference score (BSS>0) in predicting P95 EPEs, when assessing EPEs based on the forecasted precipitation (a), or the forecasted Mediterranean patterns (b). (c) Difference of the forecasting horizon between subplot (b) and subplot (a). (d) Minimum lead day from when indirect forecasting of EPEs (based on patterns) outperforms direct forecasting (based on precipitation).

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

4. Conclusions

This work analysed if and where large-scale atmospheric flow patterns over the Mediterranean can be used as skilful predictors for localized EPEs at medium and extended range forecasts. The nine large-scale patterns, selected by Mastrantonas et al. (2021a), based on EOF analysis and subsequent K-means clustering of daily anomalies of sea level pressure and geopotential height at 500 hPa, depict atmospheric variability in the lower and middle troposphere over the Mediterranean. The EPEs were derived from the 95th percentile of annual daily precipitation at each grid cell. The ERA5 dataset was used as the reference dataset, while the ECMWF extended range reforecasts (cycle 46r1) were used as forecasting product. Long-term statistics of the product, regarding pattern predictability and indirect EPE predictability, based on the predicted patterns, was assessed with the Brier Skill Score, considering 2,200 reforecasts at each lead time ranging from 0 up to 45 days ahead. Bootstrapping was also implemented to assess results' significance.

The results show that the ECMWF model well represents the Mediterranean patterns and provides skilful predictions of the patterns up to 2 weeks in advance, outperforming results based on climatological frequencies and persistence. Its performance does not show noticeable deviations between the different patterns and the winter-half and summer-half periods. The only differences worth-noticing are observed during winter-half years for Mediterranean High and Minor Low patterns. The former has a forecasting horizon extending up to week 3, while the latter is limited to week 1.

Using the forecasted patterns for indirect EPE predictability provides skilful predictions up to ~10 days for many locations in the Mediterranean. Especially for areas with high orography and coastal locations (e.g., parts of Iberian Peninsula, Morocco, western Italy/Balkans/Turkey), the use of these patterns outperforms climatological EPE allocations by more than 10 days ahead. In fact, using these patterns, rather than the actual forecasted precipitation fields, extends the EPE forecasting horizon by 3–6 days for many locations.

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

Our results demonstrate that using large-scale patterns as predictors can provide useful information for localised extreme precipitation already at medium-range forecasts, extending up to sub-seasonal scales. Such information can be promising for various users, for example the agricultural sectors, emergency response units and (re)insurance companies. To further advance this direction, it would be useful to research additional aspects:

- Are there teleconnections influencing the occurrence/predictability of the nine Mediterranean patterns?
- How skilful is indirect EPE forecasting when using other predictors, such as water vapour flux that is highly related to precipitation (e.g. Lavers et al. 2016)

Answering such questions can provide guidelines about which predictor is beneficial for different spatiotemporal resolutions, locations, and forecasting horizons, making better use of already available Numerical Weather Prediction model outputs. This can ultimately support the development of new operational products towards seamless predictions of extreme precipitation that will provide higher confidence to decision makers and users of different sectors. Such steps, that are considered a priority for international research (Majumdar et al., 2021), will be addressed by future studies.

Finally, another interesting pathway would be to assess whether these patterns can be used for the indirect predictability of other surface extremes, as for example droughts or heatwaves.

References

Grazzini F, Craig GC, Keil C, Antolini G, Pavan V, 2020. Extreme precipitation events over northern Italy. Part I: A systematic classification with machine-learning techniques. Q J R Meteorol Soc 146: 69–85. doi 10.1002/qj.3635

Hannaford J, Mastrantonas N, Vesuviano G, Turner S, 2021. An updated national-scale assessment of trends in UK peak river flow data: how robust are observed increases in flooding? Hydrol Res 52: 699–718. doi 10.2166/nh.2021.156

Hoy A, Schucknecht A, Sepp M, Matschullat J, 2014. Large-scale synoptic types and their impact on European precipitation. Theor Appl Climatol 116: 19–35. doi 10.1007/s00704-013-0897-x

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.

Kostopoulou E, Jones PD, 2005. Assessment of climate extremes in the Eastern Mediterranean. Meteorol Atmos Phys 89: 69–85. doi 10.1007/s00703-005-0122-2

Lavers DA, Pappenberger F, Richardson DS, Zsoter E, 2016. ECMWF Extreme Forecast Index for water vapor transport: A forecast tool for atmospheric rivers and extreme precipitation. Geophys Res Lett 43: 11,852-11,858. doi 10.1002/2016GL071320

Llasat MC, Llasat-Botija M, Petrucci O, Pasqua AA, Rosselló J, Vinet F, Boissier L, 2013. Towards a database on societal impact of Mediterranean floods within the framework of the HYMEX project. Nat Hazards Earth Sys Sci 13: 1337–1350. doi 10.5194/nhess-13-1337-2013

Majumdar SJ, Sun J, Golding B, Joe P, Dudhia J, Caumont O, Gouda KC, Steinle P, Vincendon B, Wang J, Yussouf N, 2021. Multiscale forecasting of high-impact weather: Current status and future challenges. Bull Am Meteorol Soc 102: E635–E659. doi 10.1175/BAMS-D-20-0111.1

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: 2710–2728. doi 10.1002/joc.6985

Mastrantonas N, Magnusson L, Pappenberger F, Matschullat J, 2021. What do large-scale patterns teach us about extreme precipitation over the Mediterranean at medium- and extended-range forecasts? https://arxiv.org/abs/2106.16008

Neal R, Robbins J, Dankers R, Mitra A, Jayakumar A, Rajagopal EN, Adamson G, 2020. Deriving optimal weather pattern definitions for the representation of precipitation variability over India. Int J Climatol 40: 342–360. doi 10.1002/joc.6215

Papalexiou SM, Montanari A, 2019. Global and regional increase of precipitation extremes under global warming. Water Resour Res 55: 4901–4914. doi 10.1029/2018WR024067

Richardson D, Neal R, Dankers R, Mylne K, Cowling R, Clements H, Millard J, 2020. Linking weather patterns to regional extreme precipitation for highlighting potential flood events in medium- to long-range forecasts. Meteorol Appl 27: e1931. doi 10.1002/met.1931

Toreti A, Xoplaki E, Maraun D, Kuglitsch FG, Wanner H, Luterbacher J, 2010. Characterisation of extreme winter precipitation in Mediterranean coastal sites and associated anomalous atmospheric circulation patterns. Nat Hazards Earth Sys Sci 10: 1037–1050. doi 10.5194/nhess-10-1037-2010

Wilks DS, 2011. Statistical methods in the atmospheric sciences. Academic Press, San Diego, CA.

CAFE project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 813844.