

Assessing the predictability of Medicanes in ECMWF ensemble forecasts using an object-based approach

Enrico Di Muzio^{1,2}  | Michael Riemer² | Andreas H. Fink¹  | Michael Maier-Gerber¹ 

¹Department of Troposphere Research, Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany

²Institute of Atmospheric Physics, Johannes Gutenberg University Mainz, Mainz, Germany

Correspondence

Enrico Di Muzio, Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe 76131, Germany.
Email: enrico.dimuzio@kit.edu

The predictability of eight southern European tropical-like cyclones – seven Medicanes and the first-ever documented case of such a storm in the Bay of Biscay – is studied evaluating European Centre for Medium-Range Weather Forecasts (ECMWF) operational ensemble forecasts against operational analysis data. Forecast cyclone trajectories are compared with the cyclone trajectory in the analysis by means of a dynamic time warping technique, which allows one to find a match in terms of their overall spatio-temporal similarity. Each storm is treated as an object and its forecasts are analysed using parameters that describe intensity, symmetry, compactness and upper-level thermal structure. This object-based approach allows one to focus on specific storm features, while tolerating their shifts in time and space to some extent.

The high compactness and symmetry of the storms are generally poorly predicted, especially at long lead times. However, forecast accuracy tends to improve strongly at short lead times, indicating that the ECMWF ensemble forecast model can adequately reproduce Medicanes, albeit only a few days in advance. In particular, late forecasts which have been initialised when the cyclone has already developed are distinctly more accurate than earlier forecasts in predicting its kinematic and thermal structure, confirming previous findings of high sensitivity of Medicane simulations to initial conditions.

Findings reveal a markedly non-gradual evolution of ensemble forecasts with lead time, which is often far from a progressive convergence towards the analysis value. Specifically, a rapid increase in the probability of cyclone occurrence (a “forecast jump”) is seen in most cases, generally with lead times between 5 and 7 days. Jumps are also found for the forecast distribution of storm thermal structure. This behaviour is consistent with the existence of predictability barriers. On the other hand, storm position forecasts often exhibit a consistent spatial distribution of storm position uncertainty and bias between consecutive forecasts.

KEYWORDS

dynamic time warping, ensemble forecasts, forecast jumps, Medicanes, object-based approach, predictability

1 | INTRODUCTION

The Mediterranean region has long been known as a hotspot for cyclogenesis (Petterssen, 1956) due to its geography

(Buzzi and Tibaldi, 1978). Despite its relatively high latitude, a small but significant fraction of Mediterranean cyclones display some similarity to tropical cyclones, both in their appearance in satellite images and in their kinematic and ther-

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.

mal structures. Such Mediterranean tropical-like cyclones, also known as *Medicanes* (i.e., Mediterranean hurricanes), have been documented since the beginning of the satellite era (Ernst and Matson, 1983; Mayengon, 1984; Rasmussen and Zick, 1987). *Medicanes* constitute a major threat due to intense winds, torrential rainfall and associated flooding. These storms are usually shorter-lived than North Atlantic hurricanes but may exhibit several tropical-like traits in the mature phase of their life cycle, such as high axial symmetry, a warm core, a strong tendency to weaken after making landfall and a cloud-free, weak-wind region at their centre resembling the eye of a hurricane (Emanuel, 2005; Cavicchia *et al.*, 2014a).

Medicanes are distinguished among Mediterranean cyclones by the complex pathway leading to their formation and maintenance. While hurricanes develop in regions of near-zero baroclinicity and draw their energy from very warm tropical oceans, *Medicanes* arise from pressure lows that are born under moderate to strong baroclinicity. The interaction between the warm sea and cold air associated with a deep upper-level trough provides the necessary thermodynamic disequilibrium for these storms to develop a warm core (Emanuel, 2005; Cavicchia *et al.*, 2014a). It can thus be maintained that *Medicanes* are the result of a synergy between synoptic-scale processes, which provide the necessary environment for their development, and mesoscale processes such as deep convection and latent heat fluxes from the sea, which are crucial for their maintenance (Homar *et al.*, 2003; Emanuel, 2005; Tous *et al.*, 2013). For this reason and because of their small size (Miglietta *et al.*, 2013; Picornell *et al.*, 2014), scarce data availability and the complex orography of the Mediterranean region, predicting *Medicanes* poses a considerable challenge for numerical weather forecasting.

Despite increased research interest in the last two decades, *Medicanes* are by their nature elusive, due to their low frequency of occurrence (less than two per year, according to Cavicchia *et al.*, 2014a) and the fact that they normally occur over the sea, where observations are sparse. For this reason, many studies to date have focused on modelling aspects (Homar *et al.*, 2003; Fita *et al.*, 2007; Davolio *et al.*, 2009; Miglietta *et al.* 2011; 2013; Chaboureaud *et al.*, 2012a; Cioni *et al.* 2016; 2018, Mazza *et al.*, 2017; Pytharoulis *et al.*, 2017) and fewer on observational aspects (Pytharoulis *et al.*, 2000; Reale and Atlas, 2001; Moscatello *et al.*, 2008; Chaboureaud *et al.*, 2012b; Miglietta *et al.*, 2013). Further studies examined *Medicanes* in relation to climate change (Romero and Emanuel, 2013; Cavicchia *et al.*, 2014b; Walsh *et al.*, 2014; Romero and Emanuel, 2017), a critical consideration given the vulnerability of the Mediterranean region to future climate change (Giorgi and Lionello, 2008). Research efforts so far have focused on deterministic simulations of *Medicanes* using high-resolution, convection-permitting models (Fita *et al.*, 2007; Davolio *et al.*, 2009; Cioni *et al.* 2016; 2018; Mazza *et al.*, 2017; Pytharoulis *et al.*, 2017), as they are deemed to best reproduce the small-scale processes that play

a crucial role in storm maintenance during the tropical-like phase. Few studies analysed *Medicanes* using ensemble forecasts (Cavicchia and von Storch, 2012; Chaboureaud *et al.*, 2012a; Pantillon *et al.*, 2013; Mazza *et al.*, 2017), of which only Pantillon *et al.* (2013) used operational ensemble forecasts.

Ensemble forecasts have been shown to be a valuable tool for predicting extreme weather events several days in advance (e.g., Buizza and Hollingsworth, 2002; Palmer, 2002; Lalaurette, 2003; Buizza, 2008; Magnusson *et al.*, 2015) and for analysing tropical cyclones (Torn and Cook, 2013; Rios-Berrios *et al.*, 2016) and their predictability (Munsell *et al.*, 2013; Zhang and Tao, 2013). Among operational ensemble forecast systems, the European Centre for Medium-Range Weather Forecasts (ECMWF) model has shown high predictive skill for extreme weather events (Lalaurette, 2003) including tropical cyclones (Yamaguchi and Majumdar, 2010; Hamill *et al.*, 2011) and it has been successfully used to study their predictability (Magnusson *et al.*, 2014; González-Alemán *et al.*, 2018). Specifically, Pantillon *et al.* (2013) used ECMWF operational ensemble forecasts to study the predictability of a *Medicane* in 2006 and found that they were able to more consistently capture early signals of its occurrence with respect to ECMWF deterministic forecasts.

The present study fills a gap in the existing literature, in that it systematically investigates the predictability of eight recent (2011–2017) southern European tropical-like cyclones – seven *Medicanes* and the first-ever documented case of such a storm in the Bay of Biscay – by evaluating ECMWF operational ensemble forecasts against operational analysis from a fixed-event perspective (Pappenberger *et al.*, 2011). Our goal is to assess whether and how far in advance these forecasts can adequately reproduce *Medicanes*. We also analyse the temporal evolution of the predictability of these storms by identifying rapid changes of the ensemble statistics with lead time that stand out compared with the expected gradual convergence towards the analysis value (Buizza, 2008). We name such changes “forecast jumps,” in line with the terminology adopted in the previous literature (see, e.g., Zsoter *et al.*, 2009) but with the fundamental difference that our analysis focuses on a well-defined event – the occurrence of a tropical-like storm – rather than on the full forecast field. We finally investigate whether there is any consistent bias in the ensemble forecasts, as may be expected given the model’s relatively low horizontal resolution and parameterised convection. This is not a straightforward task, as *Medicanes* are by their very nature extreme events and as such are found near the tail of the forecast distribution, as observed by Majumdar and Torn (2014).

In this study, we evaluate ensemble forecasts against analysis data using an object-based approach. Object-based methods have gained popularity in recent decades for their verification of precipitation forecasts (Ebert and McBride, 2000; Wernli *et al.*, 2008) and have more recently been applied to the analysis of other atmospheric features, such as the jet

stream (Limbach *et al.*, 2012) and Rossby waves (Wiegand and Knippertz, 2014). These methods allow the avoidance of the “double penalty problem” (Ebert and McBride, 2000; Wernli *et al.*, 2008) that arises in the case of the mere displacement of an otherwise well-predicted atmospheric feature if an Eulerian error metric is used. An object-based method can, in contrast, identify that the ensemble spread is due to a displacement of the feature and that only this aspect exhibits reduced predictability. We employ here a *dynamic time warping* (DTW) method (Berndt and Clifford, 1994) to better match the cyclones in the forecast with those in the analysis, allowing (small) temporal shifts of the forecast features in addition to spatial displacements. This method was applied for the first time in meteorology only recently (Maier-Gerber *et al.*, 2019) and is beneficial for studying low-probability weather events, especially at longer lead times.

This article is structured as follows. In section 2, the data and methods used are described; in particular, a dynamic time warping technique that allows one to consider temporal shifts in the forecasts is illustrated in detail. A brief overview of the eight storms analysed is given in section 3, highlighting their salient features. Results are presented in section 4, with a focus on the evolution of ensemble forecasts with lead time; this section is organised into subsections, addressing in sequence forecasts of storm occurrence, position, thermal and kinematic structure, and intensity. The results are finally discussed in section 5, alongside concluding remarks.

2 | DATA AND METHODS

In this section, we provide an illustration of the methods and techniques used to analyse the ECMWF operational analysis and ensemble forecast data in order to apply the object-based approach introduced in section 1. Mean sea level pressure (MSLP) lows are first identified and tracked in both the analysis and the ensemble data. Forecast cyclones are then matched to the reference cyclone in the analysis, using a DTW technique to maximize the similarity of the trajectories. Forecasts

are evaluated over a time window as opposed to a fixed forecast time; the choice of the time window depends on the cyclone’s intensity as well as its dynamical and thermal structure as measured by suitable parameters. A short description is finally provided of the graphics used for evaluating the ensemble forecast statistics.

2.1 | Data

ECMWF operational analysis data are used as reference data to verify ensemble forecasts, which are initialised twice daily (at 0000 and 1200 UTC) and consist of 50 perturbed forecasts or members and a control forecast. The time resolution is 6 hr for both the analysis and the ensemble data.

Both the high-resolution deterministic model (HRES), which is used to generate analysis data, and the ensemble prediction system (ENS) have undergone some changes during the time period considered in this study (2011–2017). Between 2011 and early 2016, horizontal grid spacing is 16 km for HRES and 32 km for ENS; afterwards, grid spacing decreases to 9 km for HRES and 18 km for ENS. Five and three events, respectively, occurred during each of these two time periods (see also Table 1). Vertical resolution also changed for both HRES and ENS during the analysed time period. The reader is referred to the ECMWF website for detailed information on changes and updates to the model (<https://www.ecmwf.int/en/forecasts/documentation-and-support/changes-ecmwf-model>).

2.2 | Cyclone detection and tracking

Many available cyclone detection methods (Neu *et al.*, 2013) are not suitable for Medicanes, which have a much smaller radius compared to most types of cyclones (see, e.g., Miglietta *et al.*, 2011; Picornell *et al.*, 2014). This issue is especially apparent when the input data has a relatively low horizontal resolution (Walsh *et al.*, 2014), which is the case for the ECMWF ensemble forecast data. For this reason, we have developed a new detection method to identify pressure lows in both the analysis and the forecast data. This method has

TABLE 1 Period and region of occurrence, duration, CP (hPa), symmetry, compactness (hPa/100 km), 10 m wind (m/s) and upper-level thermal wind $-V_T^U$ for the eight storms, as inferred from operational analysis data. Values are the lowest (for CP) or the highest (for other quantities) reached in the life cycle of the cyclone. 10 m wind is computed in a 300 km radius around the centre of the storm

Storm	Period	Region	Duration	CP	Symmetry	Compactness	10 m wind	$-V_T^U$
Rolf	November 2011	WM	96 hr	997	0.95	6.6	22	26
Ruven	November 2013	WM, TS, AS	48 hr	990	0.85	3.8	22	-31
Ilona	January 2014	WM, TS, AS	60 hr	991	0.83	3.2	23	8
Qendresa	November 2014	SM	60 hr	986	0.92	10.8	27	-14
Xandra	November/December 2014	WM, TS	84 hr	989	0.95	3.6	19	22
Stephanie	September 2016	BB	54 hr	998	0.96	6.0	22	11
Trixie	October 2016	SM, EM	96 hr	1,005	0.96	4.9	24	18
Numa	November 2017	TS, SM, IS	120 hr	1,002	0.98	5.1	19	20

Notes: WM = Western Mediterranean; SM = Southern Mediterranean; EM = Eastern Mediterranean; TS = Tyrrhenian Sea; AS = Adriatic Sea; IS = Ionian Sea; BB = Bay of Biscay.

been proved to be effective in detecting very small cyclones, while still being capable of detecting larger cyclones as well as filtering out spurious ones produced by noise or orographic effects.

The cyclone detection method used in this study is based on MSLP contours, spaced at 1 hPa intervals, and low-level vorticity, defined as relative vorticity averaged over the 1,000, 925 and 850 hPa levels. Given our focus on the mature phase of cyclones, only closed contours are considered, thereby neglecting open systems (e.g., diminutive waves; see Hewson, 2009). Pressure lows are identified as objects falling into at least one of two categories: (a) a set of four or more concentric contours; and (b) a set of two or more concentric contours with a radial MSLP gradient of 5 hPa/400 km or larger, calculated within a 400 km distance from the centre of the innermost contour, over at least four consecutive 30°-spaced azimuthal directions. The second category is needed to also include earlier stages of a cyclone, in which its closed circulation is still developing but a small pressure low is already present at the boundary of a larger region of low pressure, with a large MSLP gradient in its vicinity.

Detected lows are discarded when at least one of the three following conditions is met: (a) the area of all MSLP contours exceeds 500,000 km² (the low is too large); (b) the area of the second innermost contour is 50 times or more larger than the area of the innermost contour and the MSLP gradient is smaller than 5 hPa/400 km in all directions (the low is considered noise); and (c) contours are too thin and irregularly shaped (the low is considered noise – this typically occurs in the vicinity of high orography). The centre of each pressure low is finally placed where low-level vorticity reaches a local maximum within a distance of 100 km from the MSLP minimum. The values of thresholds and parameters have been chosen conservatively so as to minimise the number of discarded lows. The outcome of cyclone detection shows little sensitivity to small variations of these thresholds and parameters.

Having been detected, pressure lows are tracked in time using a method adapted from Hewson and Tittley (2010), which uses 1,000–500 hPa geopotential height difference (thickness) and 500 hPa wind speed. While a short description is given here, the reader is referred to the aforementioned article for a detailed explanation. In this tracking scheme, a likelihood score (expressed in km) is computed for each possible pairing of a pressure low at the previous output time and one at the current output time (hereafter referred to as the “past low” and the “present low,” respectively). The score estimates the likelihood of the pairing being correct, that is, how likely it is that the present low is a result of the past low advancing to a new position.

In the present study, the likelihood score is built on two parameters: half-time separation and thickness change. Half-time separation is the distance between the past and the present low, after they have moved forward and backward in time, respectively, for 60% of the time interval, considering

500 hPa wind as the steering flow. Thickness change is the 1,000–500 hPa difference in thickness between the positions of the past and the present low. A third parameter, which was originally used in the likelihood score formula, namely the feature-type transition (Hewson and Tittley, 2010), is kept fixed at 60% (Hewson, 2009, table 2) when calculating the likelihood score, as the only type of feature considered in the present study is the closed low.

The smaller the likelihood score, the more likely it is that the pairing is correct – a low score results from a small half-time separation and a small thickness change. Pairings are discarded if the past and the present low are more than 600 km apart or if their likelihood score is higher than 700 km. After computing the likelihood score for all possible pairings, they are ranked from the lowest (most likely) to the highest (least likely). Finally the ranking is read from top to bottom and each pairing is either accepted, if neither low was already previously paired, or rejected. When a pairing is accepted, the present low becomes the last element of the track that contains the past low. At the end, the remaining present lows form new tracks.

2.3 | Evaluation parameters

In order to evaluate ensemble forecasts, four parameters are used that are deemed to provide an adequate picture of each cyclone’s intensity, kinematics and thermal structure. These are central pressure (CP), symmetry, compactness and the upper-level thermal wind. The statistics of ensemble forecasts of these parameters will be examined in section 4, together with those of storm position forecasts (see also the explanation in section 2.6).

Storm intensity is represented by the cyclone’s central pressure (i.e., its lowest MSLP). The intensity of Medicanes may be slightly underestimated by ECMWF operational analysis data due to insufficient horizontal resolution, an effect that is estimated to be around 2 hPa (see, e.g., Cioni *et al.*, 2016; Pytharoulis, 2018). An even larger underestimation can be expected for ensemble forecasts given their resolution (Picornell *et al.*, 2014; Walsh *et al.*, 2014), which is half that of the analysis data.

In order to quantify the symmetry of the cyclone’s low-level circulation, a symmetry parameter S is defined for any MSLP contour as follows: $S = 1 + \arctan(\pi(4\pi A/P^2 - 1))$, where A is the area and P is the perimeter of the contour. This seemingly complex formula is based on a straightforward expression of symmetry (A/P^2); this function is then scaled ($4\pi A/P^2$) so that maximum symmetry – a perfectly round contour – equals 1, and is finally stretched by applying the arctangent, so that values of high symmetry are more widely spaced (otherwise they would tend to bunch towards 1). The last step allows one to better identify the highly symmetric, mature phase of the cyclone and interpret the ensemble statistics more clearly. The S parameter attains values of approximately 0.5, 0.2 and 0 for ellipse-shaped contours having a minor axis of length 1

and major axes of length 2, 3 and 5, respectively. The symmetry parameter of any pressure low (hereafter referred to as just “symmetry” for the sake of brevity) is obtained by averaging S over the four innermost MSLP contours, spaced at 1 hPa intervals. During their mature, tropical-like phase, Medicanes attain high symmetry, with S values exceeding 0.8, as opposed to their early stages (and the majority of extratropical cyclones) which have much lower S values.

Medicanes also tend to be much smaller than extratropical cyclones, as already pointed out, with strong pressure gradients in the vicinity of their centres. In order to give a measure of the gradient, a “compactness” parameter (hereafter referred to as just “compactness”) is defined as the azimuthally averaged radial MSLP gradient within a 150 km radius around the cyclone centre, expressed in hPa/100 km.

Finally, to quantify the cyclone’s upper-level thermal structure (i.e., its cold or warm core) we make use of the three-dimensional cyclone phase space (CPS) introduced by Hart (2003), which is defined by storm-relative thickness asymmetry B and lower- and upper-level thermal wind ($-V_T^L$ and $-V_T^U$, respectively). A positive (negative) sign of the $-V_T^U$ parameter indicates an upper-level warm (cold) core, while the absolute value is proportional to its magnitude. We do not use lower-level thermal wind in our analysis, as positive values of $-V_T^L$ characterise not only Medicanes but also extratropical cyclones with a warm seclusion (Hart, 2003). Given the lower height of the tropopause in the midlatitudes with respect to the tropics (Picornell *et al.*, 2014) and the smaller size of Medicanes compared to tropical cyclones (Miglietta *et al.*, 2013), $-V_T^U$ is calculated in a slightly different way from Hart (2003), using a smaller radius of 100 km and lower levels of 925, 700 and 400 hPa similarly to (Picornell *et al.*, 2014). A 12-hr running mean is used in the present study to smooth the CPS trajectories, differently from Hart (2003) who uses a 24-hr mean: this choice is motivated by the short life of most Medicanes and of their tropical-like phase in particular.

2.4 | Choice of the evaluation time window

Ensemble forecasts are evaluated against analysis data over a time window rather than at a single forecast time. This approach has the benefit of enhancing signals, in that the desired features (e.g., a storm intensity maximum) can be spotted over a larger set of forecast times (see also section 2.6), thereby overlooking small timing errors (e.g., the maximum occurring a few hours earlier or later than forecast). The rationale for our approach is to focus on specific storm features and consider a forecast to be sufficiently accurate if the features are successfully predicted, albeit at a slightly incorrect time. This strategy is especially valuable in extracting information from early forecasts, for which only a few members may have a cyclone and even fewer may have it at the right place and time. In this case, tolerating small timing errors allows more information to be extracted from

the forecast, such as whether a cyclone is predicted at all or whether it exhibits tropical-like traits.

The evaluation time window (hereafter referred to as the ETW) is 24 hr long and corresponds to five points (i.e., five output times) of the cyclone track extracted from analysis data (hereafter the “reference track”). Slightly shorter and longer ETWs were tested before settling on 24 hr, showing little sensitivity. The ETW is subjectively selected to best represent the mature, tropical-like phase of the cyclone, on the basis of the symmetry, compactness and the $-V_T^U$ parameters introduced in section 2.3 (CP is not used, as the mature phase of a Medicanne often does not correspond to the most intense phase).

An example of ETW selection is shown in Figure 1. As a first step, the five consecutive reference track points that have the highest average $-V_T^U$ value are selected, given that $-V_T^U$ is considered the most relevant parameter in distinguishing tropical-like cyclones from fully baroclinic cyclones (see, e.g., Mazza *et al.*, 2017). As a second step, the initial five-point selection is shifted by at most two points, corresponding to a maximum of 12 hr earlier or later. This adjustment is only applied when necessary in order to select output times with as high symmetry and compactness as possible. For storm *Qendresa* (Figure 1), for instance, the initial selection is shifted one point to the left (6 hr earlier), thereby increasing the average values of symmetry and compactness.

2.5 | Track matching

The tracking procedure outlined in section 2.2 is used for each storm to retrieve the reference track as well as tracks of MSLP lows in individual forecasts. The next step is to compare all tracks from a single member of the ensemble with the reference track in order to find the best match, that is, the closest and most similar track, which is to be considered as the given member’s cyclone. In order to avoid penalizing (small) discrepancies in the timings of storm motion and to take into account the overall spatio-temporal similarity between tracks, we use a DTW technique (Berndt and Clifford, 1994) which has been successfully applied to a recent case study of North Atlantic tropical transition (Maier-Gerber *et al.*, 2019). Originally developed for speech recognition (Sakoe and Chiba, 1978), DTW is able to match two time series nonlinearly, thereby taking into account differences in signal speed and providing a more intuitive matching (Keogh and Ratanamahatana, 2005). Using the DTW technique to match cyclone tracks allows us to focus on the spatial accuracy of the forecasts, ignoring small (local) timing errors as long as the forecast track bears a high spatial similarity to the reference one. The average time difference between DTW-paired track points may later be used to assess whether there is an early or late bias. In the following, we briefly describe the structure of the DTW technique to illustrate how it is applied to matching cyclone tracks. The reader is referred to Berndt and

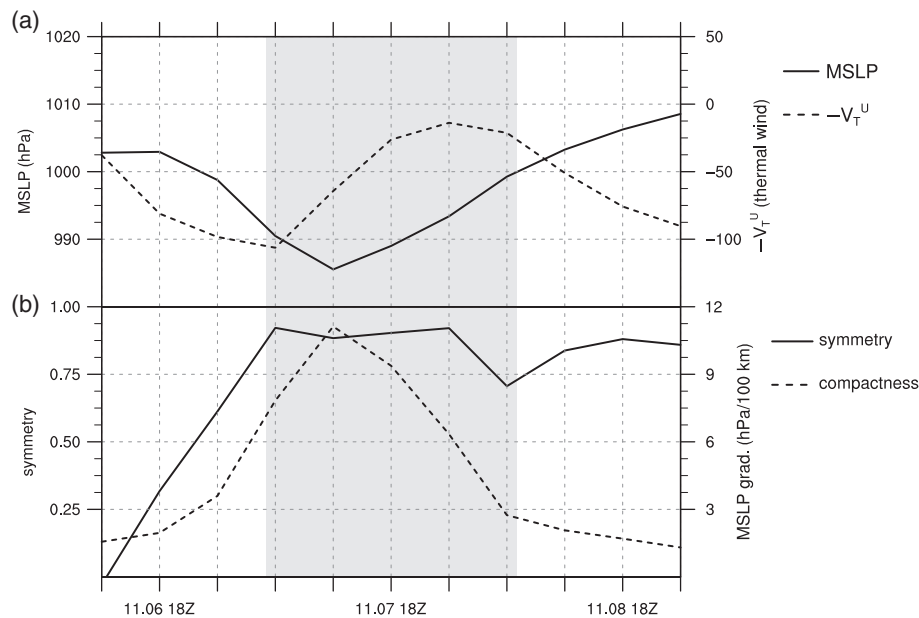


FIGURE 1 Features of storm *Qendresa* (November 2014) as retrieved from the analysis data. (a) CP (hPa) and upper-level thermal wind $-V_T^U$. (b) Symmetry and compactness. The ETW is highlighted in grey; in this case the selected period is from 0600 UTC on November 7, 2014 to 0600 UTC on November 8, 2014

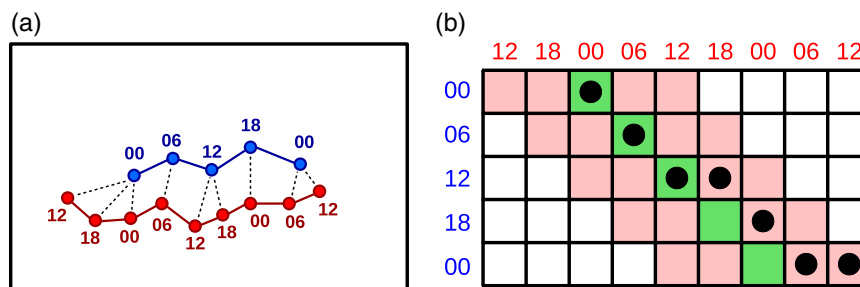


FIGURE 2 Example of DTW matching of the reference track (blue) and a forecast track (red). Numbers denote times in UTC. (a) Spatial match; matched track points are highlighted by a black dashed line. (b) Cumulative distance matrix M and warping path, represented as black filled circles, for the track match in (a). The warping window is highlighted in red, while the equal-time match is highlighted in green [Colour figure can be viewed at wileyonlinelibrary.com]

Clifford (1994) and Keogh and Ratanamahatana (2005) for more detailed explanations of the algorithm.

DTW requires first a suitable metric to express the spatial distance between each pair of track points. We choose the great-circle distance, but note that any distance metric could in principle be used. The aim is then to minimise the overall distance between the two input tracks $R = r_1, r_2, \dots, r_m$ and $S = s_1, s_2, \dots, s_n$ by finding the best possible way of matching them. To do so, an $m \times n$ distance matrix D is first computed: $D(i, j) = d(r_i, s_j)$ for each $i = 1, \dots, m$ and $j = 1, \dots, n$, where $d(r_i, s_j)$ is the spatial distance between the r_i and s_j track points. A cumulative distance matrix M is then defined recursively as follows: $M(i, j) = D(i, j) + \min[D(i-1, j), D(i-1, j-1), D(i, j-1)]$. The optimal match is finally obtained as the *warping path*, defined as the succession of M elements minimising the cumulative distance at every point. Each element of the warping path represents a pair of matched track points, as shown in the fictitious example in Figure 2. We note that the two tracks here have different lengths and that multiple points of one track may be matched to a single point of the other.

A DTW technique is usually applied with some constraints that introduce physically meaningful requirements (Berndt and Clifford, 1994). Monotonicity and continuity constraints are first imposed to ensure that all track points are matched at least once and with increasing time. A warping window (highlighted in red in Figure 2) only allows the warping path to exist in the vicinity of the diagonal of the M matrix (i.e., the succession of equal-time elements), thereby restricting the time difference between any pair of matched track points to a maximum absolute value of 12 hr. Using a warping window ensures a physically meaningful track match, preventing the match of two points that are spatially close but too distant in time. Finally, boundary conditions require the warping path to start from (end at) the forecast track point that is closest to the first (last) analysis track point, to prevent the algorithm from matching too many distant forecast track points to the first or last analysis point, which it would be forced to do if the forecast cyclone moved very quickly (an example is seen in Figure 2, where the first two forecast track points are not matched to the first analysis track point). These conditions ensure that similarity is maximised in the matching process.

DTW is applied to match the reference track's 24-hr ETW to each track in an ensemble member. Only forecast tracks that are at least 24 hr long (five output times) are considered. Furthermore, a 48-hr interval is selected from each track, spanning the ETW plus a further 12 hr (two output times) at both ends, to meet the warping window constraint. If the forecast track only exists for a fraction of these 48 hr, only the existing part is considered (the example in Figure 2 shows the longest possible forecast track, at 48 hr or nine output times). The spatio-temporal distance between (the selected interval of) the forecast and the reference tracks is finally computed in two steps: first, the average distance between a single reference track point and all associated forecast track points is computed; secondly, the final track distance is obtained by averaging the result of the above calculation over all reference track points.

For a given ensemble member, forecast tracks having a 600 km or larger spatio-temporal distance from the reference track are discarded. This threshold has been chosen after testing the sensitivity of the results to its value, similarly to Maier-Gerber *et al.* (2019). If no forecast tracks are left, the member is considered to have no cyclone (such members are hereafter called "no-storm members"). Otherwise, the track with the shortest spatio-temporal distance is considered to be the best match, that is, the most similar to the reference track. Members having a best match are hereafter called "storm members". The DTW technique has ultimately a twofold purpose: it yields the exact point-by-point matching between any forecast track and the reference track, and as a result it assists in finding the best match among all tracks for a given member.

2.6 | Evaluation of ensemble forecasts

The ensemble forecasts of the eight storms are evaluated in section 4 from a fixed-event perspective (Pappenberger *et al.*, 2011), that is, by examining multiple consecutive forecasts while focusing on a fixed period of forecast time (in this case the ETW). Given that forecasts are evaluated over a time window rather than at a fixed forecast time, lead times refer to the central time of the ETW. For each storm, the latest forecast considered is the one initialised either at the beginning of the ETW, if this begins at 0000 or 1200 UTC, or 6 hr earlier if the ETW begins at 0600 or 1800 UTC. For the sake of simplicity, the latest forecast is always labelled as "0.5 day" (see section 4) despite actually being an 18-hr forecast in cases where the ETW begins at 0600/1800 UTC (this small 6-hr difference between cases does not affect the results). A total of 16 forecasts are examined for each case, the earliest being an 8-day forecast.

Ensemble forecast statistics for the eight cases are shown in section 4. Box-percentile plots (Esty and Banfield, 2003) are preferred over standard box-and-whisker plots because they display the whole distribution of the input data. The width of each irregular "box" is proportional to the percentile p of the ordinate if $p \leq 50$, or to $100-p$ if $p > 50$; the maximum width

is thus reached at the median, while outliers are revealed by thin spikes at each tail.

As a further step to enhance the signals and relax the requirement of an exact match between forecasts and analysis, the extreme value of each parameter (the lowest value for CP; the highest value for symmetry, compactness and $-V_T^U$) is considered when evaluating forecasts. The extreme value of a given parameter is computed for each storm member within the DTW-matched interval of its track (the best match found by applying DTW as described in section 2.5) and compared with the extreme value computed within the reference track's ETW.

Storm position forecast statistics (Figure 6) are investigated by means of empirical orthogonal function (EOF) analysis. For each storm member, a two-dimensional storm position error is expressed as the average difference in longitude and latitude between the reference track and the forecast track, computed using the DTW-matched points in a similar manner as for the spatio-temporal distance. EOF analysis (Wilks, 2011) is then performed on all two-dimensional error values for each forecast (one value for each storm member). The eigenvectors of their covariance matrix define a rotated coordinate system where variability is maximized along the x axis. The spread of the storm position errors is proportional to their variance in this coordinate system, and is represented in Figure 6 as an ellipse whose axes are aligned to those of the rotated system and have lengths proportional to the variance along each eigenvector. This compact representation of storm position errors provides an immediate visual overview of their extent and spatial distribution.

3 | OVERVIEW OF THE STORMS

The eight storms analysed in this study are briefly illustrated here. A summary of their main features as retrieved from the analysis data is given in Table 1, where storm duration, period and region of occurrence are provided along with extreme intensity, symmetry, compactness, 10 m wind speed and upper-level thermal wind. Storm names were chosen by the Institute of Meteorology at the Free University of Berlin. The storm trajectories, intensity and upper-level thermal wind values are displayed in Figure 3.

Of the eight storms, four developed or spent a significant part of their lifetime over the Western Mediterranean and three over the Southern Mediterranean and Ionian Sea; these two regions are indeed hotspots for Medicanes (Cavicchia *et al.*, 2014a). In contrast, storm *Stephanie* is technically not a Medicanne, in that it occurred outside of the Mediterranean Sea. *Stephanie* has been included in our study as it exhibited the same tropical-like traits as a Medicanne (Maier-Gerber *et al.*, 2017), formed under similar large-scale circulation patterns and occurred in a region that is geographically and climatologically close to the Mediterranean. Five of the eight storms occurred in November, the most frequent month in our

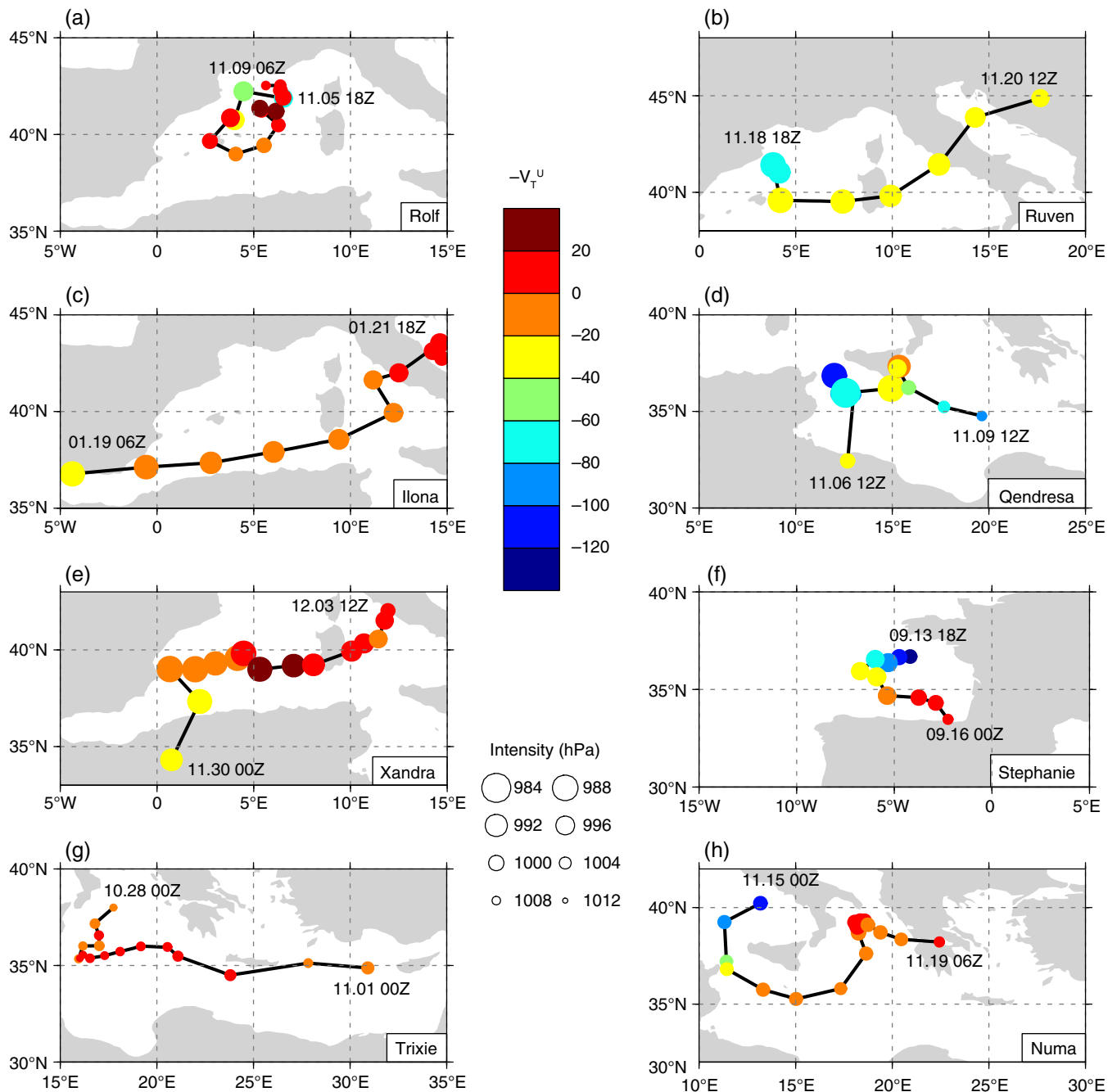


FIGURE 3 Tracks of the eight storms. The colour of each circle represents the $-V_T^U$ value, while its size represents the CP value [Colour figure can be viewed at wileyonlinelibrary.com]

sample; the three remaining storms occurred in September, October and January, respectively. We note in passing that this temporal distribution differs from the one relating to the 1948–2011 climatology produced by Cavicchia *et al.* (2014a), which has a maximum in January.

The data in Table 1 and the cyclone tracks in Figure 3 show the high heterogeneity of the eight storms in terms of their duration (*Ruven* developed rapidly and only lasted 48 hr, *Numa* remained almost statically over the Ionian Sea for 36 hr and lasted 120 hr), intensity (almost 20 hPa difference between the most intense, *Qendresa* at 986 hPa, and the least intense, *Trixie* at 1005 hPa), compactness (*Qendresa* reached

a 10.8 hPa/100 km MSLP gradient, while *Ilona* reached only 3.2 hPa/100 km) and thermal structure (most storms developed a moderate upper-level warm core, yet storms *Ruven* and *Qendresa* failed to attain one, with upper-level thermal wind $-V_T^U$ peaking slightly short of zero). We observe here that even though storm *Qendresa* did not attain an upper-level warm core, it is widely recognised as a Medicanne (Pytharoulis *et al.*, 2017; Pytharoulis, 2018; Cioni *et al.*, 2018). In fact, a unique, objective definition of what constitutes a Medicanne has not yet been established in the literature (Fita and Flaounas, 2018). However, all storms analysed in this study share some distinctive traits, in that at some

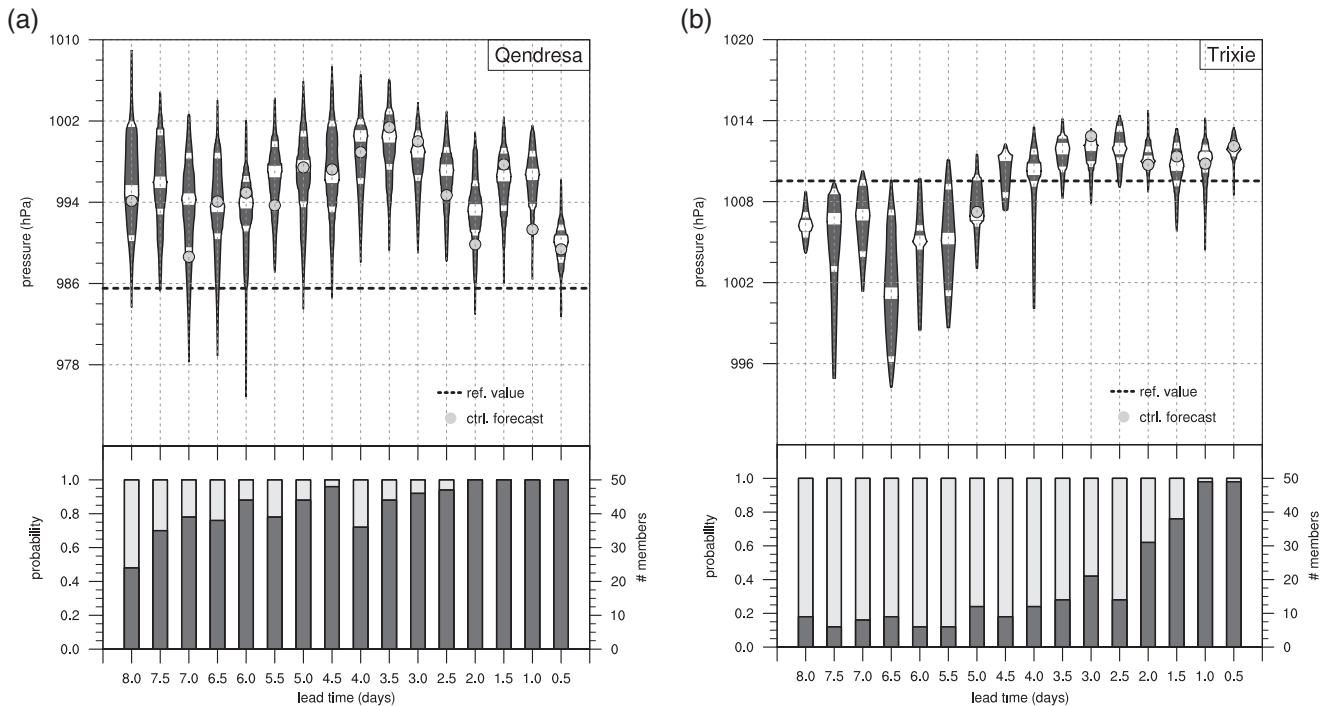


FIGURE 4 Ensemble forecasts of CP for storms *Qendresa* (a) and *Trixie* (b). Upper panels: Box-percentile plots, with white stripes marking the 25th, 50th (median) and 75th percentiles; light grey circles show the control forecast values (provided the storm has a cyclone), while the dashed line is the operational analysis value. Lower panels: The number of members having a cyclone (no cyclone) are represented by dark (light) grey bars

point during their life cycle they shrank notably, acquiring a highly axisymmetric circulation with strong MSLP gradients while quickly moving towards positive values of upper-level thermal wind (i.e., building a warm core). Another feature shared by most storms is their weakening or even fading after making landfall, which is consistent with the fact that Medicanes, similarly to tropical cyclones, are strongly influenced by surface fluxes (Fita *et al.*, 2007; Tous *et al.*, 2013).

4 | RESULTS

In this section, forecasts of several parameters are analysed, with a focus on the evolution of ensemble statistics with lead time. Forecasts of cyclone occurrence are first examined in section 4.1. Cyclone position forecasts are then explored in section 4.2. To analyse the storms' thermal structure, upper-level thermal wind forecasts are examined in section 4.3. The kinematic structure and intensity of the eight storms are finally discussed in section 4.4.

Two examples of the evolution of ensemble forecasts are provided in Figure 4, which shows CP forecasts for storms *Qendresa* and *Trixie*. These two cases illustrate the high variability among both Medicane features (see also section 3) and their forecasts. *Qendresa* is the deepest cyclone of the eight cases, with 986 hPa minimum pressure in the ETW. For this storm, the probability of cyclone occurrence (i.e., the number of storm members) is already high at 7.5 days lead time (days LT) (Figure 4a) and remains high throughout. Conversely, the ensemble median CP is much higher (by up to 14 hPa) than

the analysis value, with the latter consistently lying at the far lower end of the forecast distribution or even well below the lowest member. A small but evident dip is seen around 4 days LT for both occurrence probability and storm intensity. On the other hand, *Trixie* is the weakest cyclone in our list, with over 1,009 hPa, although it is very long-lived with a lifetime of 96 hr (Table 1). For this storm, occurrence probability is much lower than 0.5 at lead times longer than 3 days, with a considerable increase between 2.5 and 1 day LT (Figure 4b). The distribution of CP forecasts also shifts from having a large spread and being mostly or entirely below the analysis value (up to 5 days LT) to being mostly above it (lead times shorter than 3 days) with a much smaller spread.

In both these cases, the evolution of ensemble forecasts with lead time is far from gradual, with storm intensity forecasts showing little convergence towards the analysis value for *Qendresa*, while an early convergence is followed by a plateau for *Trixie*; the probability of cyclone occurrence is consistently high for *Qendresa*, whereas for *Trixie* it is very low for early forecasts, but grows rapidly for late forecasts: we refer to such rapid changes in the ensemble statistics as “forecast jumps,” as already noted in section 1. We stress here that the distributions shown in the upper halves of Figures 4, 7 and 8, as well as in the ellipses in Figure 6, are conditional distributions, that is, they only comprise storm members and omit no-storm members. Storm occurrence probabilities (appearing in the lower halves of Figures 4, 7 and 8 and shown as colour saturation in Figure 6) then need to be taken into account to get the full picture and correctly interpret the forecast evolution.

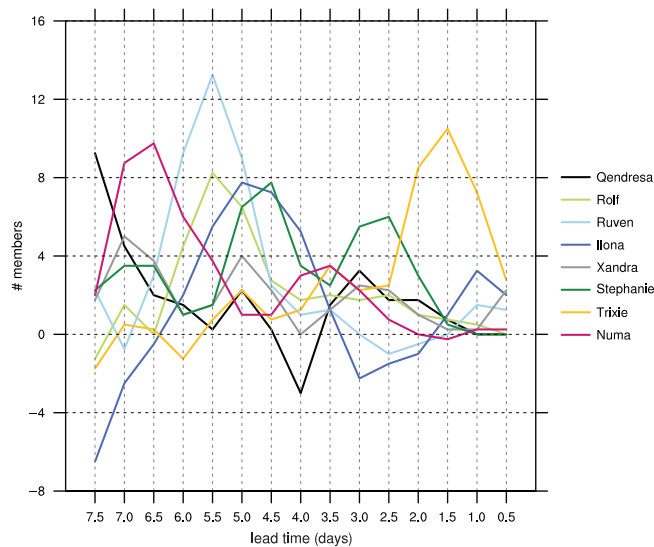


FIGURE 5 Difference in the number of storm members (proportional to cyclone occurrence probability) between the current and previous forecast, for each lead time and for all cases. Values are smoothed with a 1–2–1 running mean to reduce noise [Colour figure can be viewed at wileyonlinelibrary.com]

4.1 | Cyclone occurrence forecasts

Medicanes develop because of a combination of factors spanning multiple spatial and temporal scales and are therefore low-frequency events (Cavicchia *et al.*, 2014a). Early signals of the occurrence of a cyclone, as seen in ensemble forecasts 5–8 days in advance, are thus to be considered a valuable first piece of prognostic information. For this reason, cyclone occurrence forecasts are examined as a first insight into the predictability of the eight storms.

A benefit of using ensemble forecasts instead of deterministic forecasts is better consistency between consecutive forecasts (Buizza, 2008; Zsoter *et al.*, 2009). One could therefore expect a somewhat gradual increase of the probability of cyclone occurrence with decreasing lead time. This is not the case for most storms analysed in the present study, as forecasts often exhibit a distinctly rapid increase in occurrence probability at some lead times (known as a forecast jump). In order to extract such signals, the difference in the number of storm members between consecutive forecasts is computed and is shown in Figure 5. Seven of the eight cases exhibit distinct positive peaks, signifying a rapid increase in occurrence probability over a short interval of lead time: *Qendresa* (7.5 days LT), *Numa* (around 7 days), *Ruven* (6 to 5 days), *Rolf* (around 5.5 days), *Ilona* (around 5 days), *Stephanie* (double peak around 5 and 3 days) and *Trixie* (2 to 1 day). These peaks stand out above the bulk of the values, which are contained between -1 and 5 . Two cases also exhibit distinct negative peaks: *Qendresa* (4 days LT) and *Ilona* (3 to 2 days). Only one case, *Xandra*, shows a gradual increase of occurrence probability throughout all forecasts.

We note here that six of the seven occurrence forecast jumps are found at lead times longer than 4 days. A notable exception

is *Trixie*, for which occurrence probability does not increase above 50% until 2 days LT. These results are consistent with the low likelihood of Mediane occurrence and the hypothesis that occurrence probability increases significantly only when the forecast model's initial data contain sufficient information on all processes impacting Mediane formation. Such predictability barriers constitute a source of significant unpredictability, as discussed, for example, by Riemer and Jones (2014) in relation to bifurcation points in the context of extratropical transition, and by Pantillon *et al.* (2016) with regard to the interaction between a hurricane and an upper-level cut-off.

4.2 | Cyclone position forecasts

The impacts of Medicanes can be considerable (Cavicchia *et al.*, 2014a), but are spatially limited to small regions due to their small size. For this reason, an accurate prediction of their trajectory is key in preventing and mitigating the damage they cause locally. The next step in evaluating the ensemble forecasts of the eight storms is thus to examine their predicted position during their mature, tropical-like phase.

Cyclone position forecast statistics are shown in Figure 6, where the median of the position errors is represented as an arrow and the forecast spread as an ellipse whose axes (and hence area) are proportional to the variance of the position errors (see section 2.6). These forecasts appear to converge more gradually compared with cyclone occurrence forecasts, as demonstrated by the overall slow variation of the size and tilt of both the arrows and the ellipses over lead time (the convergence is towards a median and spread value which is very close to zero, but not exactly zero as forecasts are evaluated in a time window). However, rapid changes of one or more of these quantities (referred to as “position forecast jumps”) are visible at some lead times for some cases. For instance, a sudden decrease in spread, with the ellipse decreasing in size by 30% or more between consecutive forecasts, is seen for *Ilona* (3 days LT), *Numa* (5.5 and 4.5 days), *Qendresa* (2 days), *Rolf* (5.5 days), *Stephanie* (3 days) and *Xandra* (6 days). Rapid changes in the magnitude of the median error are also apparent for *Ilona* (3 days LT), *Qendresa* (5.5 days), *Ruven* (5.5 and 4 days), *Stephanie* (3 days), *Trixie* (1 day) and *Xandra* (3.5 days). Position forecast jumps occur in most cases at slightly shorter lead times than occurrence forecast jumps (the difference is 2 days or less for *Qendresa*, *Ilona*, *Numa*, *Ruven* and *Trixie*, while the jumps occur at the same lead time for *Rolf* and *Stephanie*). This suggests the existence of a causal link between increased occurrence probability and higher accuracy of position forecasts.

It is worth noting that the spatial distribution of position errors tends to evolve slowly with lead time. For instance, forecasts exhibit a consistent northwestern bias for *Ilona* (i.e., the storm is predicted to occur too far to the northwest), a southern to southeastern bias for *Numa*, a southwestern bias for *Rolf*, a northeastern bias for *Stephanie* (at least until 3 days LT) and a large western to southwestern bias for *Trixie*

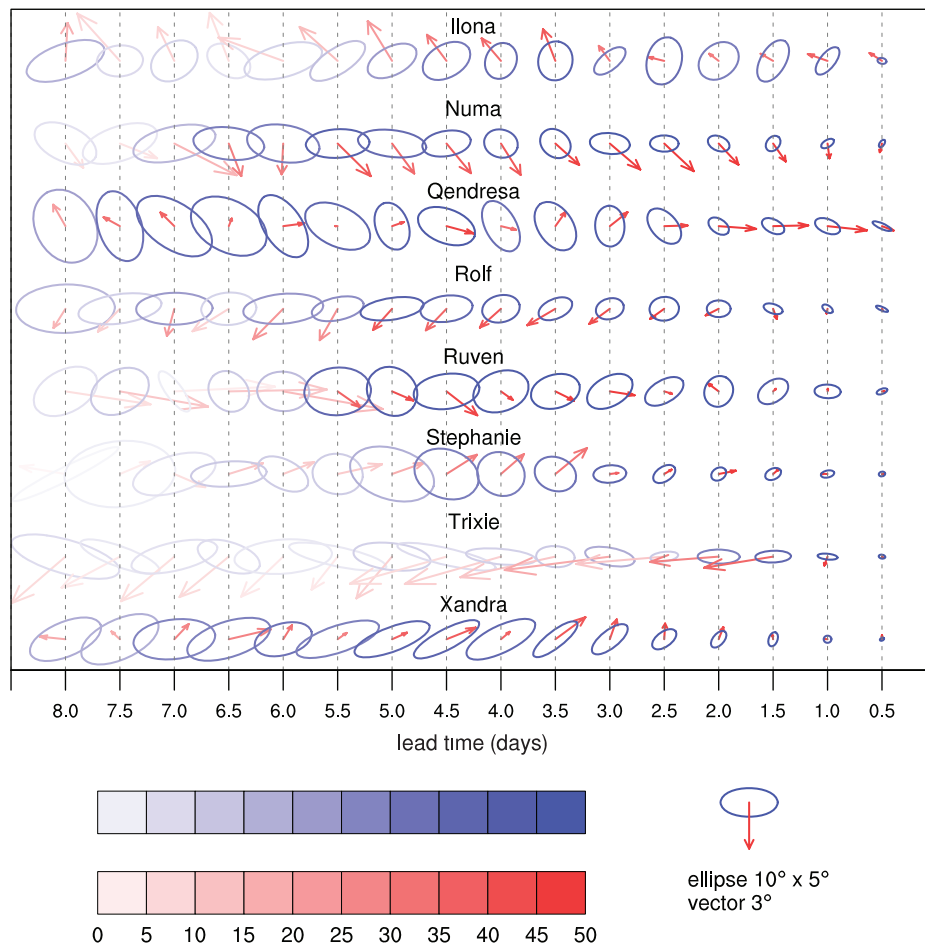


FIGURE 6 Statistics of cyclone position forecasts. For any given forecast, only storm members are considered. The red arrow represents the median of the position errors, its components being longitude (horizontal) and latitude (vertical). The blue ellipse is a bivariate normal distribution fit to the position errors, representing their spread; it is scaled so as to encircle 95% of the error points. The ellipse is oriented along the direction of maximum variability of the error values and its axes are proportional to their variance in the two-dimensional rotated coordinate system defined by the eigenvectors of their covariance matrix (see section 2.6). To improve readability of the plot, two different scales are used for the median (arrows) and the actual errors (ellipses), and each arrow is made to begin from the centre of the corresponding ellipse, even though this point does not correspond to zero error. The more storm members, the stronger the colour of both the ellipses and the arrows [Colour figure can be viewed at wileyonlinelibrary.com]

(although in this case with low occurrence probability until 2 days LT). Similarly, position errors are consistently distributed from west to east for *Numa* and *Trixie*, from northwest to southeast for *Qendresa* and from southwest to northeast for *Rolf* and *Xandra*. In summary, the region where the cyclone is predicted to occur often tends to remain the same between consecutive forecasts. This implies that early forecasts may already contain valuable prognostic information, in that the actual cyclone position may be approximately estimated early on by examining the spatial distribution of position forecasts. One explanation for this may be that certain areas of the Mediterranean Sea are more conducive to Medican development than others (Tous and Romero, 2013; Cavicchia *et al.*, 2014a), so it is more likely that the cyclone will be predicted to spend its mature phase in these regions.

4.3 | Thermal structure forecasts

After assessing whether a cyclone is going to occur or not and where it is going to occur, the next step is to analyse

its thermal structure. For this reason, we now examine forecasts of upper-level thermal wind, represented by the $-V_T^U$ parameter, which are shown in Figure 7. The evolution of these forecasts with lead time is generally neither gradual nor monotonic, as already noted with regard to forecasts of cyclone occurrence (section 4.1). Overall, the forecast spread does not consistently reduce with decreasing lead time, with some cases exhibiting a smaller (*Qendresa* and *Stephanie*) or comparable (*Ilona*, *Rolf*, *Ruven* and *Xandra*) spread at long lead times compared with the latest forecasts. Similarly, in some cases the forecast distribution increasingly deviates from the analysis value with decreasing lead time, only to get closer again in later forecasts (e.g., *Ilona*, *Numa*, *Qendresa*, *Rolf* and *Stephanie*).

Storms *Rolf* and *Numa* (Figure 7a,h, respectively) show a similar evolution, with the forecast distribution increasingly drifting away from the analysis value and the spread increasing in parallel, until the forecast distribution is mostly or even entirely below the analysis value. The forecast distribution then converges again towards the analysis, while the spread

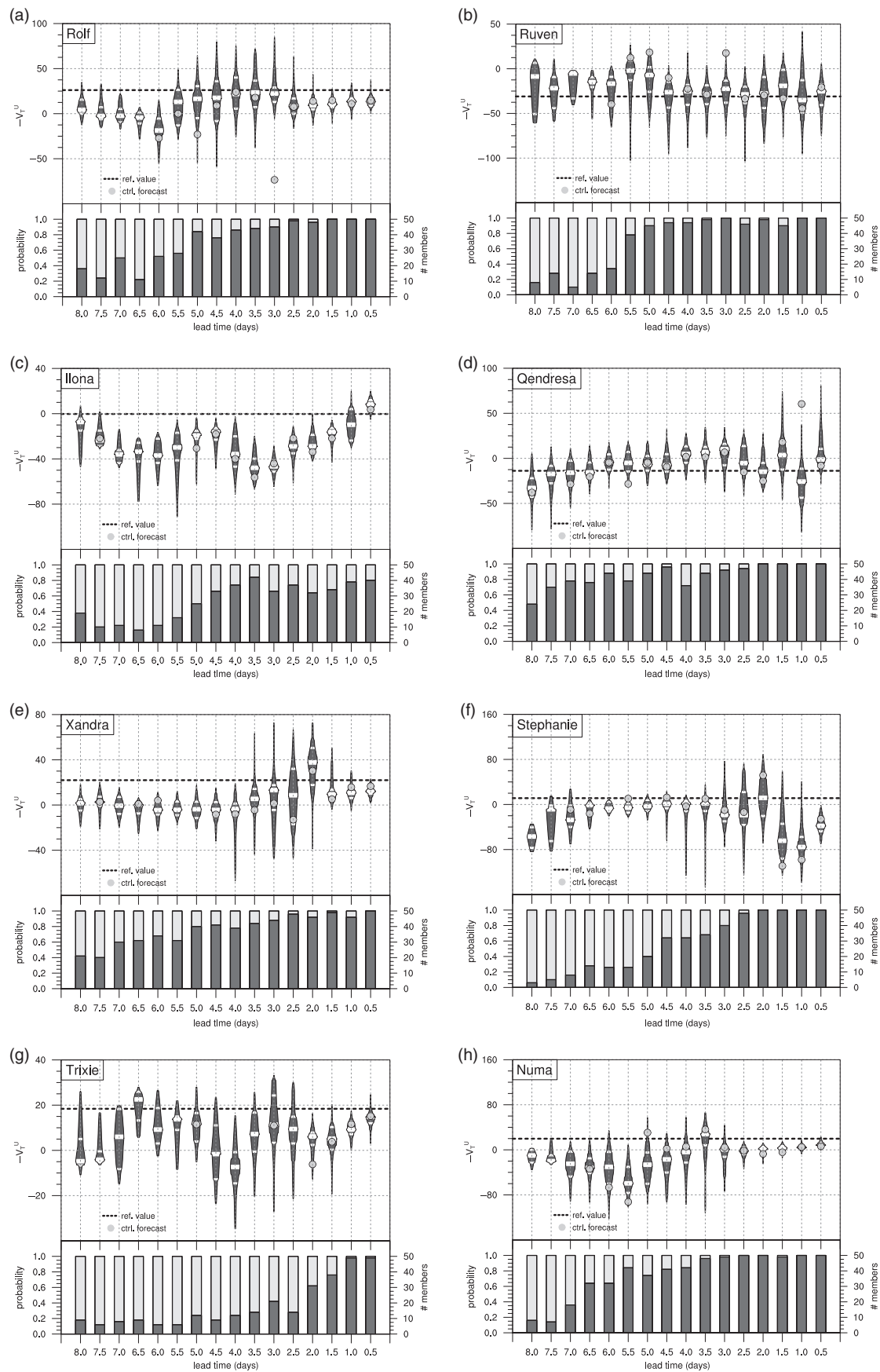


FIGURE 7 The same as Figure 4, but for upper-level thermal wind ($-V_T^U$) forecasts and for all storms

decreases, first slowly and then much more quickly, to eventually level off at short lead times. Storms *Ilona* and *Ruven* instead exhibit a contrasting evolution. For *Ilona* (Figure 7c) the forecast distribution twice drifts away from the analysis value with decreasing lead time, to eventually approach it in the latest forecasts; the spread oscillates considerably between consecutive forecasts throughout the period considered. For *Ruven* (Figure 7b) the distribution changes little in shape and position, with the spread remaining almost constant throughout all forecasts and the median always somewhat close to the analysis value.

A peculiar evolution is exhibited by storm *Xandra* (Figure 7e). Early forecasts have very little spread and the analysis value lies consistently beyond the upper end of the forecast distribution. The spread then increases considerably between 4 and 2 days LT as the distribution shifts to higher $-V_T^U$ values. The spread finally decreases again rapidly in the latest forecasts, while the forecast distribution remains slightly below the analysis value for the most part. We interpret this behaviour as follows: with relatively high storm occurrence probability (0.6 and higher) and little spread at longer lead times, ensemble forecasts indicate the development of a weaker warm core or a cold core. The increase in spread with decreasing lead time, which is associated with a slight rise in the occurrence probability, indicates that new information available in the initial conditions allows the development of a warmer upper-level core to occur in some ensemble members, that is, a higher probability of a Mediane developing.

Forecasts of cyclone thermal structure do not appear to be consistently linked to occurrence probability. However, some cases show interesting behaviours: for instance, for *Rolf* (Figure 7a) the forecast distribution closely approaches the analysis value only when the probability is higher than 0.8; for *Stephanie* (Figure 7f) the increase in occurrence probability around 4.5 days LT appears to be associated at first with a broadening of the $-V_T^U$ forecast distribution and later with its shift towards lower values; for *Trixie* (Figure 7g) the rapid increase in occurrence probability at 2 days LT is associated with a reduction in the $-V_T^U$ forecast spread.

In all cases, forecasts initialised when the cyclone has already developed have a much lower spread of upper-level thermal wind than previous forecasts, and their distribution also tends to be closer to the analysis value. This is probably explained by the inherently low probability of Mediane occurrence and the fact that the development of a warm core depends on a variety of factors, including small-scale ones such as surface fluxes, so that a pre-existing cyclone constitutes a marked improvement in the initial conditions. We observe that the latest (0.5 day) $-V_T^U$ forecast is more accurate than earlier ones in most cases, in terms of the ensemble distribution being closer to the analysis value and its spread being lower, while the analysis value lies within the distribution in all cases. Overall, this is evidence that the ECMWF ensemble

model can adequately reproduce warm-core cyclones despite its relatively low horizontal resolution.

4.4 | Kinematic structure and intensity forecasts

The last step in our analysis of the ensemble forecasts for the eight storms is to assess how their kinematic structure and intensity are predicted by examining forecasts of symmetry, compactness and CP. Overall, these forecasts also show a non-gradual evolution with lead time, as previously observed for occurrence, position and thermal structure forecasts. Specifically, the forecast distribution often does not converge gradually or monotonically towards the analysis value, the spread does not always decrease gradually or monotonically, and forecast jumps occur at some lead times for most cases. However, the evolution of these forecasts is more gradual than that of the forecasts previously examined. For this reason, we focus here on the overall performance of these forecasts rather than on their evolution with lead time. Full forecast statistics are only shown for two representative cases, namely *Numa* for compactness (Figure 8a) and *Stephanie* for symmetry (Figure 8b).

It is apparent that the forecast distribution of both compactness and symmetry consistently lies mostly, if not entirely, below the analysis value in these two cases, though with a clear convergence towards the analysis value at short lead times. These two forecasts are representative of compactness and symmetry forecasts for other cases, in that a similar behaviour is seen for most cases. One could naturally expect compactness to be systematically underpredicted to some extent, given the low resolution of the ECMWF ensemble prediction model. However, the clear convergence of forecasts at short lead times, with a markedly reduced spread and an overall much closer distribution to the analysis value, as well as the fact that the distribution tails reach or exceed the analysis value even at long lead times, indicates that the model is capable of producing high values of compactness. Moreover, compactness and symmetry forecasts appear to be well correlated with each other, so that high values of either parameter are associated with high values of the other. We conclude that the underdispersion arises at long lead times because the occurrence of a very symmetric and compact storm is a highly unlikely event and, as such, it is by its nature near the tail of the forecast distribution (especially at long lead times), as observed by Majumdar and Torn (2014). Late forecasts thus tend to converge at short lead times as they benefit from improved initial conditions.

Finally, we note that CP forecasts (not shown) are overall the most gradually evolving forecasts with lead time, albeit with a tendency for the distribution to lie mostly (or, more rarely, entirely) above the analysis value at long lead times for many storms, which is associated with the low probability of cyclone occurrence in early forecasts. This hypothesis is supported by forecasts of *Qendresa*, the most intense of the eight storms (see section 3), for which the analysis value lies

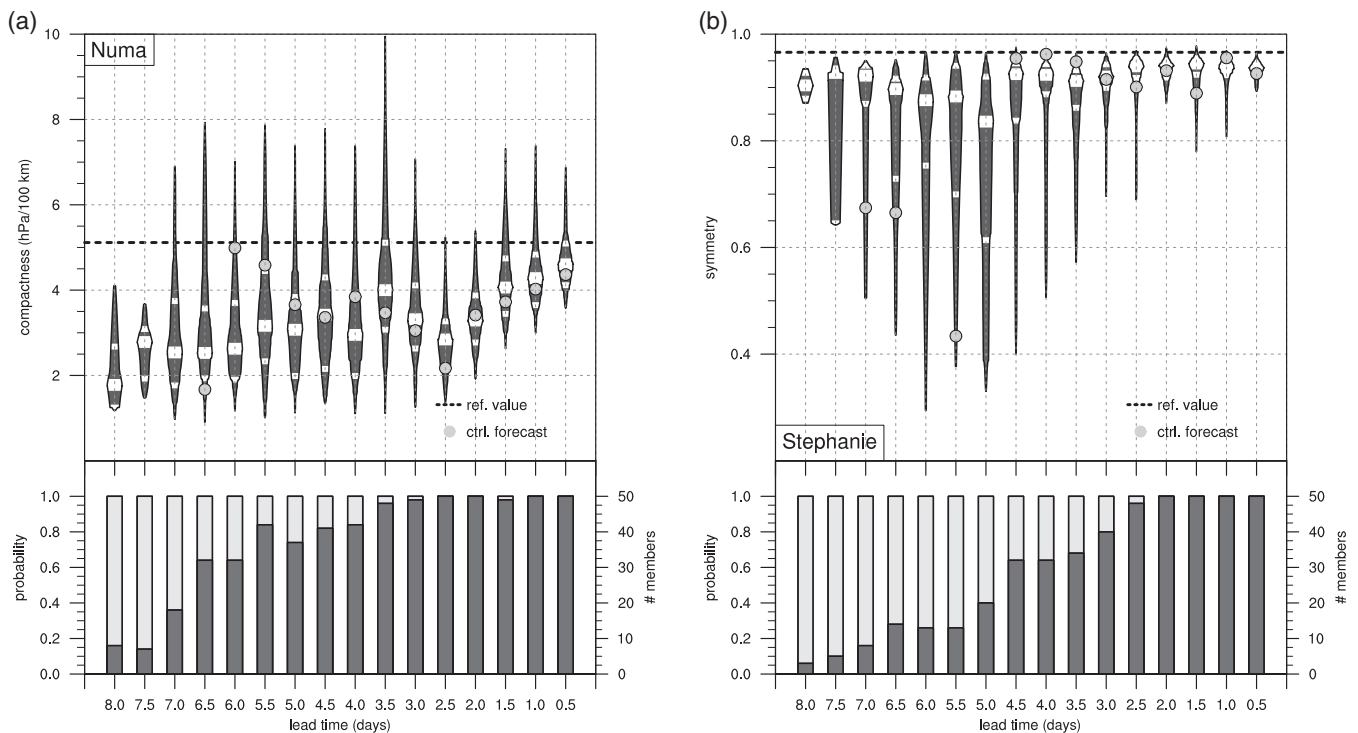


FIGURE 8 The same as Figure 4, but for (a) the compactness forecast for *Numa* and (b) the symmetry forecast for *Stephanie*

consistently at the far lower tail of the forecast distribution at long lead times (Figure 4a). *Qendresa* indeed underwent an extremely rapid development (with a pressure drop of more than 15 hPa in 18 hr; see Cioni *et al.*, 2018) which appears as highly unlikely especially in early forecasts, although the probability of cyclone occurrence is high from 7 days in advance (Figure 4). The fact that the underdispersion is more evident for symmetry and compactness than for CP supports the idea that Medicanes are more distinctly characterised by their high compactness and symmetry than by their intensity, with the former far less predictable than the latter, especially at long lead times.

5 | SUMMARY

Medicanes have been gaining increased attention in the research community in the last two decades. These storms constitute a major threat in the Mediterranean region, due to intense winds and rainfall. Although the pathway leading to the formation of Medicanes is by now well known, they remain elusive characters of the Mediterranean climate, in that their frequency of occurrence is low and an objective definition has not yet been established. Predicting Medicanes also poses a considerable challenge, due to scarce observations over the sea and the interplay between the numerous factors influencing their entire life cycle at multiple spatial and temporal scales.

In this article, the predictability of eight southern European tropical-like cyclones – seven Medicanes and the first-ever documented case of such a storm in the Bay of Biscay

– is analysed by evaluating ECMWF operational ensemble forecasts against operational analysis data. We apply an object-based approach that allows us to focus on specific storm features, while tolerating their shifts in time and space to some extent. Each storm is then treated as an object and its forecasts are evaluated using suitable parameters: CP, symmetry, compactness and $-V_T^U$, which give a measure of the cyclone's intensity, symmetry, pressure gradient and upper-level thermal structure, respectively. The tropical-like traits attained during the mature phase of the cyclone's life cycle are therefore well represented. This object-based approach has shown strength in extracting the most relevant information from the data, and value in condensing it into intuitive quantities: for these reasons, it could easily be applied to other types of forecasts and atmospheric features. The DTW technique in particular appears promising for further application in the atmospheric sciences due to its intuitiveness and flexibility in providing a meaningful space–time matching of time series.

Findings reveal that the evolution of ensemble forecasts with lead time is far from gradual, generally differing from steady convergence towards the analysis value that may be expected (Buizza, 2008). In particular, rapid increases in the probability of cyclone occurrence (forecast jumps) are seen in most cases. This behaviour is consistent with the existence of predictability barriers, which are overcome only, as we hypothesise, when initial conditions adequately represent the variety of factors playing a role in Medicanes development. Cyclone thermal structure forecasts also exhibit a non-gradual evolution in some cases, with the forecast distribution drifting away from the verifying analysis value and the spread

increasing with decreasing lead time. However, late forecasts which have been initialised when the storm has already developed tend to be more accurate than earlier forecasts. This supports previous findings of high sensitivity of Mediane simulations to initial conditions (Cioni *et al.*, 2016).

On the other hand, forecasts of cyclone position exhibit a visible tendency to a consistent spatial distribution of cyclone position uncertainty and bias (i.e., a non-zero median position error) between consecutive forecasts, which may be explained by the fact that some regions of the Mediterranean Sea are more conducive to Mediane development than others (Tous and Romero, 2013; Cavicchia *et al.*, 2014a), thus favouring the occurrence of the cyclone in the same region between consecutive forecasts. This implies that early forecasts may already contain valuable prognostic information on the cyclone's position during its mature phase.

Unlike other parameters, forecasts of compactness and symmetry consistently exhibit underdispersion at long lead times in most cases. A marked improvement of these forecasts is however seen at short lead times. In light of these contrasting behaviours, we exclude the presence of any systematic bias that could be expected due to the relatively low resolution of the ECMWF ensemble model. We instead deem the underdispersion to be a result of the intrinsically low probability of the occurrence of a Mediane (that is, a highly axisymmetric and compact storm) in early forecasts, which causes it to be found near the tail of the forecast distribution, as observed by Majumdar and Torn (2014). We interpret in the same way the underdispersion tendency appearing in CP forecasts at long lead times in some cases. Considering all parameters, forecasts indicate that the ECMWF ensemble model can adequately reproduce Medianes in terms of their tropical-like traits, albeit only at relatively short lead times.

The present work paves the way towards an in-depth investigation of the physical mechanisms underlying the features revealed by our analysis, in particular the non-gradual evolution of forecasts with lead time, forecast jumps and the consistent spatial distribution of cyclone position forecasts. A future study will examine the complex interplay between processes at different spatial and temporal scales leading to the formation of a Mediane and its impact on their predictability and the evolution of ensemble forecasts with lead time.

ACKNOWLEDGEMENTS

The European Centre for Medium-Range Weather Forecasts (ECMWF) is acknowledged for providing the ensemble forecast and operational analysis datasets. The research leading to these results has been carried out within the subproject C3 “Multi-scale dynamics and predictability of Atlantic Subtropical Cyclones and Medianes” of the Transregional Collaborative Research Centre SFB/TRR 165 “Waves to Weather,” funded by the German Research Foundation (DFG). We thank two anonymous reviewers for their constructive comments,

which helped improve the quality of the manuscript, as well as Florian Pantillon and other colleagues for their useful feedback.

ORCID

Enrico Di Muzio  <https://orcid.org/0000-0002-5964-1908>

Andreas H. Fink  <https://orcid.org/0000-0002-5840-2120>

Michael Maier-Gerber  <https://orcid.org/0000-0003-3677-4414>

REFERENCES

- Berndt, D.J. and Clifford, J. (1994) Using dynamic time warping to find patterns in time series. *10*, 359–370.
- Buizza, R. (2008) The value of probabilistic prediction. *Atmospheric Science Letters*, 9, 36–42.
- Buizza, R. and Hollingsworth, A. (2002) Storm prediction over Europe using the ECMWF ensemble prediction system. *Meteorological Applications*, 9, 289–305.
- Buzzi, A. and Tibaldi, S. (1978) Cyclogenesis in the lee of the Alps: a case study. *Quarterly Journal of the Royal Meteorological Society*, 104, 271–287.
- Cavicchia, L. and von Storch, H. (2012) The simulation of Medianes in a high-resolution regional climate model. *Climate Dynamics*, 39, 2273–2290.
- Cavicchia, L., von Storch, H. and Gualdi, S. (2014a) A long-term climatology of Medianes. *Climate Dynamics*, 43, 1183–1195.
- Cavicchia, L., von Storch, H. and Gualdi, S. (2014b) Mediterranean tropical-like cyclones in present and future climate. *Journal of Climate*, 27, 7493–7501.
- Chaboureau, J.-P., Nuissier, O. and Claud, C. (2012a) Verification of ensemble forecasts of Mediterranean high-impact weather events against satellite observations. *Natural Hazards and Earth System Sciences*, 12, 2449–2462.
- Chaboureau, J.-P., Pantillon, F.P., Lambert, D., Richard, E. and Claud, C. (2012b) Tropical transition of a Mediterranean storm by jet crossing. *Quarterly Journal of the Royal Meteorological Society*, 138, 596–611.
- Cioni, G., Malguzzi, P. and Buzzi, A. (2016) Thermal structure and dynamical precursor of a Mediterranean tropical-like cyclone. *Quarterly Journal of the Royal Meteorological Society*, 142, 1757–1766.
- Cioni, G., Cerrai, D. and Klocke, D. (2018) Investigating the predictability of a Mediterranean tropical-like cyclone using a storm-resolving model. *Quarterly Journal of the Royal Meteorological Society*, 144, 1598–1610.
- Davolio, S., Miglietta, M.M., Moscatello, A., Pacifico, F., Buzzi, A. and Rotunno, R. (2009) Numerical forecast and analysis of a tropical-like cyclone in the Ionian Sea. *Natural Hazards and Earth System Sciences*, 9, 551–562.
- Ebert, E.E. and McBride, J.L. (2000) Verification of precipitation in weather systems: determination of systematic errors. *Journal of Hydrology*, 239, 179–202.
- Emanuel, K. (2005) Genesis and maintenance of “Mediterranean hurricanes”. *Advances in Geosciences*, 2, 217–220.
- Ernst, J.A. and Matson, M. (1983) A Mediterranean tropical storm? *Weather*, 38, 332–337.
- Esty, W.W. and Banfield, J.D. (2003) The box-percentile plot. *Journal of Statistical Software*, 8, 1–14.
- Fita, L. and Flaounas, E. (2018) Medianes as subtropical cyclones: the December 2005 case from the perspective of surface pressure tendency diagnostics and atmospheric water budget. *Quarterly Journal of the Royal Meteorological Society*, 144, 1028–1044.
- Fita, L., Romero, R., Luque, A., Emanuel, K. and Ramis, C. (2007) Analysis of the environments of seven Mediterranean tropical-like storms using an axisymmetric, nonhydrostatic, cloud resolving model. *Natural Hazards and Earth System Science*, 7, 41–56.
- Giorgi, F. and Lionello, P. (2008) Climate change projections for the Mediterranean region. *Global and Planetary Change*, 63, 90–104.
- González-Alemán, J.J., Evans, J.L. and Kowaleski, A. (2018) Use of ensemble forecasts to investigate synoptic influences on the structural evolution and predictability of Hurricane Alex (2016) in the midlatitudes. *Monthly Weather Review*, 146, 3143–3162.
- Hamill, T.M., Whitaker, J.S., Fiorino, M. and Benjamin, S.G. (2011) Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Monthly Weather Review*, 139, 668–688.

- Hart, R.E. (2003) A cyclone phase space derived from thermal wind and thermal asymmetry. *Monthly Weather Review*, 131, 585–616.
- Hewson, T.D. (2009) Diminutive frontal waves—a link between fronts and cyclones. *Journal of the Atmospheric Sciences*, 66, 116–132.
- Hewson, T.D. and Tittley, H.A. (2010) Objective identification, typing and tracking of the complete life-cycles of cyclonic features at high spatial resolution. *Meteorological Applications*, 17, 355–381.
- Homar, V., Romero, R., Stensrud, D.J., Ramis, C. and Alonso, S. (2003) Numerical diagnosis of a small, quasi-tropical cyclone over the western Mediterranean: dynamical vs. boundary factors. *Quarterly Journal of the Royal Meteorological Society*, 129, 1469–1490.
- Keogh, E. and Ratanamahatana, C.A. (2005) Exact indexing of dynamic time warping. *Knowledge and Information Systems*, 7, 358–386.
- Lalauette, F. (2003) Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Quarterly Journal of the Royal Meteorological Society*, 129, 3037–3057.
- Limbach, S., Schömer, E. and Wernli, H. (2012) Detection, tracking and event localization of jet stream features in 4-D atmospheric data. *Geoscientific Model Development*, 5, 457–470.
- Magnusson, L., Bidlot, J.-R., Lang, S.T.K., Thorpe, A., Wedi, N. and Yamaguchi, M. (2014) Evaluation of medium-range forecasts for Hurricane Sandy. *Monthly Weather Review*, 142, 1962–1981.
- Magnusson, L., Thorpe, A., Buizza, R., Rabier, F. and Nicolau, J. (2015) Predicting this year's European heat wave. *ECMWF Newsletter*, 145, 4–5.
- Maier-Gerber, M., Pantillon, F.P., Di Muzio, E., Riemer, M., Fink, A.H. and Knippertz, P. (2017) Birth of the Biscane. *Weather*, 72, 236–241.
- Maier-Gerber, M., Riemer, M., Fink, A.H., Knippertz, P., Di Muzio, E. and McTaggart-Cowan, R. (2019) Tropical transition of Hurricane Chris (2012) over the North Atlantic Ocean: a multi-scale investigation of predictability. *Monthly Weather Review*. <https://doi.org/10.1175/MWR-D-18-0188.1>.
- Majumdar, S.J. and Torn, R.D. (2014) Probabilistic verification of global and mesoscale ensemble forecasts of tropical cyclogenesis. *Weather and Forecasting*, 29, 1181–1198.
- Mayengon, R. (1984) Warm core cyclones in the Mediterranean. *Mariners Weather Log*, 28, 6–9.
- Mazza, E., Ulbrich, U. and Klein, R. (2017) The tropical transition of the October 1996 Medicane in the western Mediterranean Sea: a warm seclusion event. *Monthly Weather Review*, 145, 2575–2595.
- Miglietta, M.M., Moscatello, A., Conte, D., Mannarini, G., Lacorata, G. and Rotunno, R. (2011) Numerical analysis of a Mediterranean ‘hurricane’ over south-eastern Italy: sensitivity experiments to sea surface temperature. *Atmospheric Research*, 101, 412–426.
- Miglietta, M.M., Laviola, S., Malvaldi, A., Conte, D., Levizzani, V. and Price, C. (2013) Analysis of tropical-like cyclones over the Mediterranean Sea through a combined modeling and satellite approach. *Geophysical Research Letters*, 40, 2400–2405.
- Moscatello, A., Miglietta, M.M. and Rotunno, R. (2008) Numerical analysis of a Mediterranean “hurricane” over southeastern Italy. *Monthly Weather Review*, 136, 4373–4397.
- Munsell, E.B., Zhang, F. and Stern, D.P. (2013) Predictability and dynamics of a nonintensifying tropical storm: Erika (2009). *Journal of the Atmospheric Sciences*, 70, 2505–2524.
- Neu, U., Akperov, M.G., Bellenbaum, N., Benestad, R., Blender, R., Caballero, R. and Wernli, H. (2013) IMILAST: a community effort to intercompare extratropical cyclone detection and tracking algorithms. *Bulletin of the American Meteorological Society*, 94, 529–547.
- Palmer, T.N. (2002) The economic value of ensemble forecasts as a tool for risk assessment: from days to decades. *Quarterly Journal of the Royal Meteorological Society*, 128, 747–774.
- Pantillon, F.P., Chaboureaud, J.-P., Mascart, P.J. and Lac, C. (2013) Predictability of a Mediterranean tropical-like storm downstream of the extratropical transition of Hurricane Helene (2006). *Monthly Weather Review*, 141, 1943–1962.
- Pantillon, F.P., Chaboureaud, J.-P. and Richard, E. (2016) Vortex–vortex interaction between Hurricane Nadine (2012) and an Atlantic cut-off dropping the predictability over the Mediterranean. *Quarterly Journal of the Royal Meteorological Society*, 142, 419–432.
- Pappenberger, F., Bogner, K., Wetterhall, F., He, Y., Cloke, H.L. and Thielen, J. (2011) Forecast convergence score: a forecaster’s approach to analysing hydro-meteorological forecast systems. *Advances in Geosciences*, 29, 27–32.
- Petterssen, S. (1956) *Weather Analysis and Forecasting*. New York: McGraw-Hill.
- Picornell, M.A., Campins, J. and Jansá, A. (2014) Detection and thermal description of Medicanes from numerical simulation. *Natural Hazards and Earth System Sciences*, 14, 1059–1070.
- Pytharoulis, I. (2018) Analysis of a Mediterranean tropical-like cyclone and its sensitivity to the sea surface temperatures. *Atmospheric Research*, 208, 167–179.
- Pytharoulis, I., Craig, G. and Ballard, S. (2000) The hurricane-like Mediterranean cyclone of January 1995. *Meteorological Applications*, 7, 261–279.
- Pytharoulis, I., Matsangouras, I.T., Tegoulías, I., Kotsopoulos, S., Karacostas, T.S. and Nastos, P.T. (2017) Numerical study of the Medicane of November 2014. *Perspectives on Atmospheric Sciences*, 115–121.
- Rasmussen, E. and Zick, C. (1987) A subsynoptic vortex over the Mediterranean with some resemblance to polar lows. *Tellus A*, 39, 408–425.
- Reale, O. and Atlas, R. (2001) Tropical cyclone-like vortices in the extratropics: observational evidence and synoptic analysis. *Weather and Forecasting*, 16, 7–34.
- Riemer, M. and Jones, S.C. (2014) Interaction of a tropical cyclone with a high-amplitude, midlatitude wave pattern: waviness analysis, trough deformation and track bifurcation. *Quarterly Journal of the Royal Meteorological Society*, 140, 1362–1376.
- Rios-Berrios, R., Torn, R.D. and Davis, C.A. (2016) An ensemble approach to investigate tropical cyclone intensification in sheared environments. Part I: Katia (2011). *Journal of the Atmospheric Sciences*, 73, 71–93.
- Romero, R. and Emanuel, K. (2013) Medicane risk in a changing climate. *Journal of Geophysical Research: Atmospheres*, 118, 5992–6001.
- Romero, R. and Emanuel, K. (2017) Climate change and Hurricane-like extratropical cyclones: projections for North Atlantic polar lows and Medicanes based on CMIP5 models. *Journal of Climate*, 30, 279–299.
- Sakoe, H. and Chiba, S. (1978) Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26, 43–49.
- Torn, R.D. and Cook, D. (2013) The role of vortex and environment errors in genesis forecasts of Hurricanes Danielle and Karl (2010). *Monthly Weather Review*, 141, 232–251.
- Tous, M. and Romero, R. (2013) Meteorological environments associated with Medicane development. *International Journal of Climatology*, 33, 1–14.
- Tous, M., Romero, R. and Ramis, C. (2013) Surface heat fluxes influence on Medicane trajectories and intensification. *Atmospheric Research*, 123, 400–411.
- Walsh, K., Giorgi, F. and Coppola, E. (2014) Mediterranean warm-core cyclones in a warmer world. *Climate Dynamics*, 42, 1053–1066.
- Wernli, H., Paulat, M., Hagen, M. and Frei, C. (2008) SAL—a novel quality measure for the verification of quantitative precipitation forecasts. *Monthly Weather Review*, 136, 4470–4487.
- Wiegand, L. and Knippertz, P. (2014) Equatorward breaking Rossby waves over the North Atlantic and Mediterranean region in the ECMWF operational ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 140, 58–71.
- Wilks, D.S. (2011) *Statistical Methods in the Atmospheric Sciences*. Oxford: Academic Press.
- Yamaguchi, M. and Majumdar, S.J. (2010) Using TIGGE data to diagnose initial perturbations and their growth for tropical cyclone ensemble forecasts. *Monthly Weather Review*, 138, 3634–3655.
- Zhang, F. and Tao, D. (2013) Effects of vertical wind shear on the predictability of tropical cyclones. *Journal of the Atmospheric Sciences*, 70, 975–983.
- Zsoter, E., Buizza, R. and Richardson, D. (2009) “Jumpiness” of the ECMWF and Met Office EPS control and ensemble-mean forecasts. *Monthly Weather Review*, 137, 3823–3836.

How to cite this article: Di Muzio E, Riemer M, Fink AH, Maier-Gerber M. Assessing the predictability of Medicanes in ECMWF ensemble forecasts using an object-based approach. *Q J R Meteorol Soc.* 2019;145:1202–1217. <https://doi.org/10.1002/qj.3489>