Paper wcd-2019-1

"Medicane Zorbas: Origin and impact of an uncertain potential vorticity streamer"

Raphael Portmann, Juan Jesús González-Alemán, Michael Sprenger, and Heini Wernli

Response to the Reviewers' comments:

We thank the reviewers for their many critical and constructive comments that helped improving the manuscript. All points have been carefully considered and led to the following important changes compared to the original revision:

- We better explain the objectives of the study, which focuses on the synoptic-scale processes involved in the formation of this intense cyclone.
- We clearly define what we consider as "medicane" in this study.
- The introduction has been restructured and streamlined, following the suggestions by the reviewers.
- The paper contains a separate data and methods section.
- Large parts of the synoptic overview were rewritten: The cyclogenesis and life-cycle of 'Zorbas' is introduced in more detail, and a discussion of the synoptic situation before cyclogenesis is added.
- The analysis of the error amplification and propagation is improved, and a few aspects of our previous analysis are omitted.
- The final part of the study (section now called: "Cyclone thermal structure and link to upper-level PV and low-level equivalent potential temperature") was completely rewritten, also based on additional analyses.

Below are the detailed replies to the individual comments (in blue).

Reviewer 1 (Ron McTaggart-Cowan)

Background

The authors investigate the predictions of a September 2018 medicane in the ECMWF ensemble system. They identify ensemble members that have differing day-3 representations of a PV streamer involved in the storm's development. They track these differences back to small differences in the initial conditions and show the progression of PV spread in association with an anticyclonic Rossby wave break.
The remainder of the analysis focuses on the interaction between the PV streamer and the developing cyclone: distinct precipitation and storm structures are identified in the different sub-ensembles.
Reductions in predictive skill associated with the development of sub-synoptic systems in the Mediterranean region are an important subject for investigation. Similarly, the limits of predictability imposed by PV streamer evolution and interactions between such features and lower-tropospheric circulations are not well understood. Despite these interesting fundamental underpinnings, the current submission suffers from a large number of flaws in organization, preparation and analysis. Although each of these may not be considered fatal in isolation, a significant amount of effort will be required to address all of them thoroughly. Any revised submission of this investigation will necessarily be heavily modified and constitute a new piece of work. I hope that the authors will find the comments below useful as they pursue this research.

General Comments

1. The manuscript lacks organization and logical flow. This extends from the highest level of structure (sections and subsections), down to the paragraph and even sentence level. It makes the manuscript difficult to read and follow because concepts and details are disjointed, scattered and frequently repeated throughout the text. Ordinarily, I would consider these kinds of organizational issues relatively minor and possibly within the domain of the authors' discretion; however, in this case they seriously detract from the work and make it very difficult for readers to follow the investigation. Addressing these problems will involve the rewriting of much of the text, but will yield a more focused manuscript that will likely be shortened by at least 1-2 pages.

a. High Level

i. The structure of the introduction is ineffective. It begins with a
very cursory review of tropical transition and medicanes (including PV streamers), then switches to a Rossby wave discussion that returns to error amplification twice, and then goes back to a very short summary of Zorbas. The latter also includes thesis questions and an outline of the remainder of the study that lacks section information or internal references.

We agree that the introduction could be better structured. We streamlined and rewrote parts of the introduction and provided section information when presenting the outline of the study (L102-109).

ii. I do not think that the decision to replace the standard "Data and Methods" section with "Operational ECMWF Products" is effective. It means that additional data sources (satellite imagery, lightning detection, etc) have to be described in the text body, descriptions that seriously distract from the analysis when they occur. The same is true of methodological details (e.g. LAGRANTO, which is introduced twice) and the entire discussion of tracking and the CPS (L378-L392). Descriptions of all of these sources and techniques should be centralized in a "Data and Methods" section.

We agree and now include a proper "data and methods" section that introduces most methodological aspects of the study (L110-175)

iii. There are numerous forward-references to section 6 throughout the earlier sections of the manuscript. While appropriate internal referencing is a useful tool, these repeated forward-references are likely an indication of poor manuscript organization, especially when they underpin important elements of the analysis (for example, the CPS referenced in sections 3 and 4 but never shown despite a reference to "Fig. 4a", which does not exist in the submission; L159). I think that the synoptic analysis (ideally shortened from its current length by enhanced focus) should include a discussion of the medicane itself, including the CPS. I understand that the medicane is not the intended focus of the work as is repeatedly stated in the text; however, the reader could be excused for thinking that it is because of (1) the title, (2) the multiple introduction subsections that deal with TC-like features, (3) the statement on L317 that the investigation of the "development of a medicane-like system" is an objective of section 6.2, and (4) the pervasive forward-referencing to the CPS analysis in the text.

We agree that the medicane itself should be more in the focus of the synoptic overview. Apologies, the reference to Fig. 4a should have been Fig. 5a.

We shorten some aspects of the synoptic overview, but add the CPS and the track of Medicane Zorbas to avoid forward references.

b. Medium Level

i. Each section should begin with an introduction of the section contents so that the reader has an idea of how the section fits into the larger narrative of the work. A section should not begin with a description of data used in specific figures, as does section 3. Please revise each section introduction to ensure that the reader is logically guided through the work.

We agree that the text should be written in such a way that it guides the reader as elegantly as possible through the study. We considered this when rewriting the section introductions.

ii. Each paragraph should begin with an introduction to the paragraph contents, and should conclude with a statement that relates back to the material introduced. There are very few paragraphs in the manuscript that follow this basic structure. A particularly clear example of a paragraph that ranges far too broadly occurs on L50-64. The paragraph (and subsection) starts with a description of initial condition uncertainty, moves into ensembles more generally, then into PV error growth, and ends with a discussion of tropical cyclones. This lack of focus makes the study very difficult to follow despite the fact that the investigation itself is relatively straight-forward. In this case, the subject of error growth reappears in the paragraph starting on L73, which further adds to the confusion of the discussion. Please do not simply rewrite the paragraphs noted in this comment: the structure of virtually all paragraphs in the manuscript needs to be reconsidered and revised, an effort that will lead to rewriting of large portions of the work. The readability problems induced by the lack of logical internal paragraph are more than aesthetic in this case, and are serious enough that they significantly reduce the potential impact of this work.

We apologize for the obviously in parts confusing writing style in the original submission. However, we argue that the rule that *every* paragraph should begin with an introduction to the paragraph contents might be too restrictive. Again, we carefully reconsidered the structure of the entire text in order to increase readability.

iii. I do not think that summary paragraphs at the end of a section are useful, particularly given the large number of relatively short sections in this submission. For example, the summary on L163-L167 is redundant with analysis undertaken in the previous page of the manuscript. Please remove summary paragraphs (they also appear at the ends of sections 5 and 6.1) in favour of making the text itself direct, clear and readable (see item 1.c.ii below).

We removed unnecessary summary paragraphs at the end of sections, which just repeated what has been discussed before.

c. Low Level

i. Reference to caption-level figure details within the text is highly distracting. For example, the fact that precipitation is shown in red solid contours in Fig. 1 is referenced three times in section 3 (once erroneously as "blue contours"; L130), while the fact that QG vertical motion is shown in red contours is referenced twice in the same section. These plotting details are described in the caption, and their appearance in the text detracts from the flow of the analysis. Figure and panel references should be enough to guide readers through the discussion. Please consider removing caption information from the text body throughout the submission.

We agree that the original version contained too many caption-level figure details. However, we are still convinced that, in some cases, they can help the reader to quickly grasp the relevant information from the figures. Therefore, we carefully reconsidered the usefulness of caption information in the text and remove them where possible.

ii. The writing style is too informal and lacks the precision required for scientific text. For example, the outline of the manuscript is described as "a journey" on L101, the analyses in section 6.2 "hint" at airstreams (L321), and the first person plural ("we") is used heavily throughout the submission. The introduction of section 6.2 (L320-324) can be summarized as, "we don't do anything thoroughly here, but we show stuff that's different and make some guesses about what that means; then in the next section we do it properly". I don't think that that kind of introduction (or approach to the analysis) will make readers want to continue to invest their time in the rest of the section. Throughout the submission, irrelevant details clutter the text (e.g. does it matter that 1800 UTC 26 September is "in the evening of the same day" on L126?), and ill-defined concepts reappear throughout the analysis (e.g. the "C-shaped" PV cutoff with a "dent" and "dent structure" on L138, 141 and 146, respectively). Every effort should be made to make the text succinct and readable, so that the analysis does not get lost in superfluous details and unnecessary bridging statements.

Here we ask the reviewer to please consider that we are not native speakers and, in some cases, we don't realize that our language becomes too informal. We thought that the "C-shaped" cutoff would be a useful terminology, but obviously it isn't. We improve the text and tried to avoid too informal terminology.

2. I think that cyclogenesis in cluster 3 is really interesting, but that the discussion in the current study misses the opportunity to capitalize on its uniqueness (I do not think that section 6.3 is sufficient in this respect). It looks to me like this is an excellent example of a nonlinear response / bifurcation leading to a real limit on predictability. Clusters 1 and 2 are simply phase shifts of the same cyclogenesis event. From a guidance perspective, both are reasonably useful at least in terms of situational awareness. Cluster 3, however, looks to me like the development of a different cyclone. There's an 850 hPa circulation south of Turkey in all of the groups at 1200 UTC 26 September (Fig. 8, column 2). In fact, a cyclone has already formed in this region in many of the cluster 3 members and one of the cluster 1 members (also shown in Fig. 9a). In groups 1 and 2, the low between Crete and Cyprus disappears as the PV tail promotes development along the African coast. In group 3, the pre-existing cyclone intensifies and fractures the PV streamer as the low retrogresses towards Crete (Fig. 8, column 3). By 1200 UTC 28 September, the medicane lies in the central Mediterranean in groups 1 and 2, but it is a completely different storm that is centered on Crete in group 3 (this differs from the interpretation implied by discussions on L286-L290 and L303-304 of the submission). So the relatively small difference in the location of the PV streamer axis (a linear change from west to east of the observed location) leads to a highly nonlinear response in the form of development of a new cyclone (groups 1 and 2) or intensification of an existing circulation (group 3). (Note that a couple of centers form over northern Africa in group 3 at 1200 UTC 27 Sept – Fig. 8k; these are cases in which the response to the change in PV streamer

position is linear.) The theory that group 3 is fundamentally different from the others is supported by the precipitation patterns and tracks (Fig. 9; noting that the large track jumps between North Africa and Crete are unlikely to be accurate) parcel trajectories (Fig. 10) and parcel properties (Fig. 11). Because a nonlinear response / bifurcation is known to impose strong limits on predictability, identifying and describing such behaviour in this case would be an important outcome of this work. I hope that some of the length reductions achieved by improving the manuscript's focus and organization can be invested in a much more thorough analysis of this possibility.

Many thanks for emphasizing the strongly nonlinear effects seen in cluster 3. We agree that this is an interesting aspect and discuss it in more detail in the revised version of the paper. In particular this aspect now appears in L380-L389, L440-444, L502-504.

3. The values of QG vertical motion seem too small to be very meaningful despite being described as "strong" in the text (L441). Vertical motions of ~0.5 mm/s and 1 mm/s (0.005 and 0.01 Pa/s) are plotted in Fig. 2, while values of up to ~5 mm/s (0.05 Pa/s) are plotted in Fig. 7. These values are all well within the typical rms of QG vertical motion at midlatitudes and mean vertical motions across the globe (Stepanyuk et al. 2017). If these calculations and plots are correct, then the vertical motion forcing from the upper levels is almost irrelevant to the real vertical circulations in most cases. Such weak vertical motions would need to be sustained in-place for many hours/days to have any appreciable impact on moistening or stability. For example, air in the peak ascending region in Fig. 2c ascends <10 hPa in a day in response to QG forcing, an ascent rate that is dwarfed by the 600 hPa ascent in the rising parcels near the centre. If the calculations are correct, then the relevance of the PV streamer to ascent and cyclogenesis needs to be seriously reconsidered in this case, an exercise that will likely lead to conclusions that are completely different from those arrived at by the current submission.

It is true that the values of QG omega shown here are very small compared to the full vertical motion. However, it was not intended and is also not "fair" to try to explain the full vertical motion with the QG omega shown here. Note that we show the QG omega, as forced by levels above 550 hPa (QG omega top) on 850 hPa. These values are expectedly much smaller than the full QG omega on 850 hPa or the QG omega top on higher levels (see also Fig. 1b in Davies (2015), which shows the effect of an isolated forcing region on the vertical velocity field in the surrounding atmosphere). We therefore do not intend to explain the full vertical motion with this variable, but the goal is to show the presence and location of the upper-level forcing.

However, when revising the analyses and streamlining the article we decided that QG omega is not essential to show in order to arrive at the main conclusions. QG omega is therefore not shown anymore in the article.

4. The motivation for the case study approach adopted by the study is weakened by passages that highlight case-to-case variability, and is not supported by a clear statement of the useful aspect of the case study framework. The dominance of case-to-case variability is particularly emphasized

on L38 and L84, with the latter appearing to be a direct criticism of the case study as a useful analytic tool. It is good to identify the limitations of the adopted investigation technique, but this criticism should be balanced with a clear description of what the case study approach can provide that other types of analysis (e.g. climatology) cannot.

We agree and now better explain the motivation for the case study approach, in particular in L87-L90.

5. The analysis of vertical coupling in section 5 is not quantitative enough to be included in the study. Despite significant discussion of Fig. 7e-h (L241-253), the strongest conclusion that is draw is that it is "most likely" that baroclinic instability is active. Even this conclusion appears to over-reach the analysis given that no baroclinic growth rates were computed. Given that the Icelandic low is not the focus of this investigation and that the left-hand column of Fig. 7 shows a convincing evolution of short-wave anomaly growth, I think that the right-hand column of Fig. 7 and the associated discussions should be removed. If this analysis is to be retained, then there needs to be a real quantification of baroclinic coupling and associated growth rates [note that the 12-18h time scale is very rapid for pure baroclinic growth, which typically has a doubling time scale on the order of a day (Hakim, Encyclopedia of the Atmospheric Sciences) and suggests that moist processes are likely to be very important].

We agree that in the original submission it was not clearly shown that baroclinic instability is active. Thanks for pointing out this weak aspect of the analysis.

We changed Fig. 7(e-h) to show (based on the operational analysis and cluster means) the synoptic setting in which the error growth is occurring, without claiming that baroclinic instability is active. Additionally, we add a Figure (new Figure 9) and extend the discussion about upper-level dynamics related to the jet streak, which more convincingly explains the error amplification.

6. The study of PV error growth by Baumgart et al. (2018) is referenced in the introduction, but not in section 5, where the left-hand column of Fig. 7 bears a striking resemblance to Fig. 3 of that work (albeit with a compressed time frame). The discussion of the importance of non-linear upper-level Rossby wave dynamics here follows closely that of Baumgart et al. (2018), so much of this description could be replaced by citations and comparisons. The Torn (2015) normalized difference is a useful measure, so compressing section 5 to focus on that metric in the context of the Baumgart et al. (2018) interpretation of this process would allow for a dramatic shortening of this section and serve to place this submission in the context of investigations by other groups.

We agree that Baumgart et al. (2018) could be referenced and some sentences of this section could be shortened. However, we still think that Fig. 7 needs to be discussed well, also because (a) we show a different measure than Baumgart et al. (2018), and (b) as you point out, the time frame is very different (we show lead times 6-42 h in 12-hourly time steps, whereas Baumgart et al. (2018) showed lead times 2-8 days every 2 days).

We added a reference to Baumgart et al. (2018) (L284, L291) in this section and shortened the text where possible.

7. Assessing the significance of the differences discussed in section 5 is important; however, the technique and in-text descriptions should be revised. Wilks (2016) provides a description of problems with the multiple-testing technique (as adopted in this study), which can lead to over-confident statements about significance. Please consider using the false discovery rate here. Additionally, the level at which the differences are considered significant is not identified in the text, and only appears in the Fig. 7 caption (is 0.05 used throughout?). Note that there is currently a reconsideration of the use of the term "significant", which appears to be leaning in favour of providing p-values rather than definitive statements about significance. I'm not very familiar with that discussion, but it might be of interest to consider during revision.

Thank you very much for pointing out this problem and mentioning the Wilks (2016) paper. We agree that, as we are using multiple-testing, a p-value correction is required to control the false discovery rate.
We corrected the p-values using the Benjamini-Hochberg correction that is suggested by Wilks (2016) (see L160-170) and made sure it is always clear which false discovery rate (alpha) is used to assess significance.

8. I am surprised not to see any references to Wiegand and Knippertz (2014), who study the representation and predictive skill of anticyclonic RWB and PV streamer formation over the Mediterranean region in the ECMWF ensemble (i.e. an earlier version of the same system used here). That work seems so directly relevant to this study (including the conceptual diagram in Fig. 10 of that paper) that it should be leveraged heavily in this investigation, particularly in terms of putting the forecast uncertainty in this case in a broader context.

We apologize for not referencing this important study in the original version. This paper is now included and helps putting our case study in a broader context (see L505-507,L527)

9. The numbering of clusters forces readers to remember the mapping: 1 is centered, 2 is west and 3 is east. Why not call the clusters C, W and E? Then the Fig. 8 rows could be reordered to W, C, E so that there's a progression in the columns rather than having the PV streamer location (and eventual cyclone location) jumping around.

Thanks, very good suggestion, which we adopted in the revised version.

10. Throughout the study, the "surface cyclone" is discussed by the 850 hPa heights are shown. Showing 850 hPa winds is useful, but I don't see anywhere in the manuscript that the 850 hPa heights are essential to the analysis. I think that all plots that currently show 850 hPa heights should be replaced with mean sea level pressure for consistency with the text.

We agree that 850 hPa geopotential heights are not ideal and SLP would be more appropriate. The plots were changed accordingly (see new Figures 3 and 10).

11. Throughout the study, short-range ECMWF forecasts are used to estimate precipitation accumulations. To avoid model biases and potential "twinning", it would be preferable to use an independent product. The GPM IMERG is readily available and would be a better choice for this study than stitched-together IFS forecasts.

No precipitation product is perfect for our analysis. The short-range IFS forecasts have the advantage that they are from the same model as the other data used. GPM data might also not be free of biases. Since the exact precipitation values are not essential for our study, we continued using the short-range IFS forecasts.

12. Advection of cold air over warm Mediterranean waters is identified as a factor that increases latent heat fluxes and promotes convection; however, this effect is not quantified in the current investigation. The OAFlux dataset covers the period of interest and is readily available for this kind of study. Please consider supporting the claims made in the manuscript with an analysis of OAFlux (or equivalent) surface flux estimates. An augmented surface flux analysis may particularly interesting if model-predicted fluxes are found to be very different between groups 1/2 and group 3 (see item 2 above). Such an analysis is essential if the categorical statements about surface fluxes currently found in the conclusions (L429) are to be retained.

We agree that the argument, that latent heat fluxes are active was not well supported in the original submission. Because of the general restructuring and in favour of a clearer focus of the paper, we removed the trajectory analysis in the ensemble members and instead provided a trajectory analysis including surface fluxes based on the operational analysis in the supplementary material, which is briefly discussed and referred to in the synoptic overview (L208-L210)

13. The manuscript really needs to be clear about whether the medicane itself is a focus of the study. In multiple passages, it is stated explicitly that the medicane is not going to be investigated as part of this work (e.g. L97, L161, L320). However, much of section 6 is dedicated to the evolution of the medicane, including trajectory and CPS analyses. The title of the manuscript also emphasizes the storm morphology and will attract readers interested in medicanes. It feels as though the work was initially focused entirely on the PV streamer, and that "mission creep" has led to the introduction of more storm-scale-relevant material. Please reconsider the statements that disavow the relevance of the medicane structure for this work in an effort to remove what seems like a fairly important internal inconsistency in the manuscript.

We apologize for the confusion about the focus of the study, which created the impression of internal inconsistency. The original idea was to submit a two-part paper, where the medicane would be the focus of the 2nd part. Obviously, when we decided to first focus on this paper, we didn't manage to clearly explain the role of the medicane aspect.

We clarified this aspect better. In particular, we clearly defined what we consider as a medicane (a Mediterranean cyclone acquiring a deep warm core) in this study (and distinguish between medicanes that undergo tropical transition and the ones that don't). We make clear that in this study, we focus on the large-scale conditions that influence the predictability of the formation of a medicane. And we also make clear that we do not focus on the meso-scale dynamics that – once a medicane has formed – can lead to tropical transition. See e.g. L34-40, L133-134.

14. Why are the ECMWF data coarsened to 1°, and how is it done? The result is very poor resolution in the graphics, and if it not done carefully, the operation could result in aliasing. Is a conservative remapping used? This is a particularly important question for the precipitation field, where the difference between sampling/interpolation and remapping/aggregation can be enormous when the degradation of resolution is so large.

We downloaded ECMWF analysis data on a 1° grid to be consistent with the resolution of the ensemble data. Such a coarsening of the ensemble data is required to cope with the huge amount of data. Note that, for all ensemble members, we download the 3D fields on all model levels, which is required to accurately calculate PV and trajectories. This data transfer needs to be done within a few hours after completion of the ensemble simulation, because eventually, fields are archived in MARS on a few pressure levels only. All grid interpolations were done with routines available in MARS.

15. Most published works do not consider "medicane" a proper noun (and it is therefore not capitalized). This is analogous to "hurricane", which is only capitalized when a specific storm is discussed (e.g. "Some think that Hurricane Katrina was a category 3 hurricane at landfall"). Consider using lower case "medicane" throughout except in named reference to Medicane Zorbas.

Thank you for the explanation! We changed the use of "medicane" accordingly.

16. The terms "air mass", "airstream" and "parcel" seem to be confused in relation to trajectory analyses (L139 and section 6.2). An "airstream" is a loosely defined concept, but I think that it would be represented by a high density of air parcel trajectories in a limited area. Then the phrase "trajectories of the airstreams" (Fig. 10 caption) doesn't make sense unless the airstream (a feature in storm-relative coordinates) is somehow tracked over time. Similarly, trajectories do not track "air masses" (L139), but parcels. The difference is important, because it is unlikely that all parcels in an "air mass" are ascending near the cyclone centre.

17. The trajectory analysis in section 6.2 is incomplete. The suggestion that moistening is occurring because of surface latent heat fluxes (L345-346) implies that the parcels are in contact with the surface; however, the vertical position of the parcels is never shown. It is also possible for parcels

to be moistened by evaporation of falling precipitation or by turbulent mixing. It is therefore not demonstrated that enhanced surface fluxes are responsible for the moisture changes in groups 1 and 2. The same is true for the potential temperature analysis on L346-348: surface fluxes are only one possible reason for potential temperature increases, and only influence parcels if they are in contact with the surface (even at above-surface levels in the boundary layer, the moistening/heating mechanism would be turbulent flux convergence rather than surface fluxes per se). The lack of information about the trajectories makes it impossible for reviewers or future readers to confirm the validity of the conclusions drawn at the end of this section (L356-370).

18. Section 6.2 ends with a set of suppositions and conjectures based on an incomplete trajectory analysis (see previous item) climatological behaviour. As a result, terms such as “could favour” and “might support” are used instead of definitive statements. If the analysis and descriptions in this section cannot be made robust enough to be able to conclude these statements definitively, then this section should be removed.

Reply to comments 16-18: Thanks for pointing out these weaknesses of the analysis. As mentioned above, we remove this trajectory analysis in favour of focusing on the most important aspects of our study (see reply to comment 12).

19. The description of the CPS (L386-392) is insufficiently detailed to allow independent confirmation of the results (a requirement for publication). Because of the small scales of medicane structures, the hurricane-based radii are usually reduced for studies of Mediterranean storms. Was the same done here, or were the original hurricane-based values used?

Consistently with previous studies (e.g. Gaertner et al., 2018), we have used a radius of 150 km. This is now mentioned in the revised version (L146)

20. I don’t understand the “deep warm core” (DWC) analysis in Fig. 12. Take groups 2 and 3, for example. They have 12 and 18 members, respectively. The average number of DWC in group 2 is 7.2, and 7.0 in group 3 according to Fig. 12. That number is “per ensemble member”, so multiplying by the relevant ensemble size yields 7.2*12=86.4 for group 2 and 126 for group 3. However, the total number of DWC steps for group 2 is given as 43, and that for group 3 is given at just 14 at the bottom of the plot. In the text (L404) the reader is told to consider the group-3 DWC analysis “with caution, due to the small sample size”. However, the average number of DWC steps per ensemble member is as large in group 3 as it is in group 2: why is the sample size so small? There seems to be something about the number of sequential DWC steps (“duration”), but that is never clearly stated in the text or caption. What is wrong with my interpretation of the DWC analysis?

Thank you for pointing out that this analysis has not been straightforward to follow. The missing piece is, that the analysis only includes ensemble members that actually have a deep warm core cyclone. So, the number of DWC steps has to be multiplied by the number of members in the considered cluster that have a DWC cyclone. This number can be read from Table 1. Hence, the sample size for group 2 is small because only 2 members form a deep warm core cyclone.

We replaced the Figure with another one (new Figure 12) that is clearer and additionally provides evidence for conclusions that are so far not very well supported (see your comment 23).

21. Throughout the text, equivalent potential temperature gradients are used to identify both baroclinic zones and moisture gradients [L125, L132, L137 (where the 850 hPa theta-e is inappropriately used to identify a "weak surface cold front") L153 and elsewhere]. Strictly, neither of these is guaranteed by a theta-e gradient, which may arise as a result of either in isolation. If baroclinicity is important, then potential temperature (or temperature on an isobaric surface) should be shown. If moisture is important, then it should be shown. Theta-e is a very useful quantity for assessing convective potential and is a useful way to identify the warm sector for the trajectory analysis, but it does not replace the more basic fields for questions of baroclinicity and moisture.

We agree that it is more appropriate to look at potential temperature and humidity when the focus is on baroclinicity or humidity.

We removed discussions about baroclinicity and moisture gradients as they were not of major importance for the storyline of the manuscript.

22. There are a lot of very specific geographical references throughout the text, probably more than there need to be. I'm a geographer, but I still found myself having to look for specific place-names on maps. It would be very useful to have a new Fig. 1 that shows (at least) the storm track and labels for all place names referred to in the text.

We agree with this point and included a new Fig. 1 that shows the storm track and labels for relevant places.

23. The conclusions of the study are not supported by the evidence provided in the text:

a. The "clustering" technique is not rooted in a mathematic definition and fails to guarantee the separation of the members into distinct "scenarios" as stated in the text (e.g. at L481). Is it true that there are three "distinct scenarios"? I agree that there are two (see General Comment 2), but I don't see why there are three. Groups 1 and 2 are distinguished only by the fractional overlap of the PV streamer, and there was no demonstration that there is any sort of heterogeneity in overlap space. This is a weakness in the analysis that results from the failure to use a true clustering analysis, and the decision to rely on a classification heuristic. There is no guarantee that group 1 and 2 events are separate from each other in any kind of meaningful way, and selection of a different overlap threshold (70%, for example) would result in the progressive reclassification of members from one group to the next. To demonstrate the presence of different scenarios, a true cluster analysis should be performed, and the optimal number of groups should be identified (e.g. using the "elbow method).

We agree that generally, and for climatological or operational purposes, an objective mathematical clustering is clearly preferable. However, the beauty of this specific case is that it is very clear to the reader that the uncertainty of the PV streamer prediction is large in terms of the meridional position of its tip and its shape (both are somewhat connected) and that it is meaningful to separate the ensemble into a group where the meridional position of the PV streamer is more or less correct, another group that has the streamer too far to the east, and one that has it too far to the west. If you agree that groups 1 and 3 are different, then you could also agree that group 2 is different, because group 3 and group 2 are distinguished from group 1 in the exact same way (with group 3 having the PV centre of mass to the east and group 2 to the west of the analysis). We are very thankful that you pointed out the fact that cyclogenesis in group 3 shows a non-linear response to the PV streamer position (General comment 2), and that it shows also a clearly distinct surface cyclone scenario in response to the distinct PV streamer scenario. However, from the fact that the surface cyclone scenarios of group 1 and 2 are not as distinct as of 1 and 3 (albeit clearly shifted) we cannot conclude that there are no distinct PV streamer scenarios. We therefore argue that, in the present case, this kind of ad-hoc clustering is useful and suits the purpose of this case study. It is not at all guaranteed that an objective clustering provides more appropriate scenarios for this kind of investigation. For example, if we assumed that the overlap space is fully homogeneous and the PV streamers are shifted by constant distances from west to east, an objective clustering would not necessarily provide us with useful information how many clusters to choose. However, if we are interested in what the shift of the PV streamer does to the predictability of the surface cyclone, it would still be meaningful to split the ensemble into equal bins with one containing the westernmost PV streamers, one the more central PV streamers, and the third the easternmost PV streamer.
In the Zorbas case, we argue that the heterogeneity in the PV streamer distribution is captured very well by the ad-hoc clustering and an objective clustering technique is not required for this study.

We now added an inset to Figure 4 of the initial submission (new Figure 5) which shows a histogram of longitude of the maximum PV value within the latitude band 36°-37°N. The three maxima in this histogram, representing each one of the clusters, hopefully convince you that – from the PV streamer perspective – there are three clearly distinct scenarios. Note that a clustering according to this measure would put a few borderline members into another bin (see green and red "outliers"). However, we stick to the original clustering as this accounts for the uncertain positon/shape of the PV streamer on a larger domain.

b. There was no analysis of the near-surface flow induced by the PV streamer, so how is the conclusion about induced advection (L432-434) supported by the evidence provided in the submission? Particularly given the limited spatial extent of the streamer immediately prior to cyclogenesis, it is possible that the induced near-surface flow is very strong. For the arguments regarding air parcel modification by surface fluxes, the parcels approaching the centre in groups 1 and 2 must be in contact with the surface, putting them as far as possible from the upper-tropospheric streamer.

We agree that for this statement to be fully supported, a PV inversion would be necessary. From our general understanding however it is very sensible to assume that the PV streamer, as a strong upper-level PV anomaly, affects the lower-level circulation. Note that this conclusion is not about the air parcels modified by surface fluxes, but the ones that end up in the warm sector (which are never argued to experience surface fluxes, on the contrary they move almost adiabatically and without increase in specific humidity). However, as mentioned above, in response to above comments and for the sake of enhancing the focus of our study, we remove the trajectory analysis (see reply to your comment 12).

c. I cannot see what part of the analysis is used to conclude that the group-1 PV streamer was better able to "maintain the cyclonic circulation" (unclear whether this refers to the upper- or lower-level flow) than the group-2 or group-3 features (L434-438). There appears to have been a rigorous analysis behind this statement (something that determines the number of members that meet a "condition"), but I don't know what section this analysis was described in.

We agree that composites may not be fully ideal/fair to draw this conclusion because spatial shifts could also result in low PV values in the composite.
We therefore provide an additional analysis (new Figure 12) that better supports this statement and make sure the formulation is clearer. We also rephrase the conclusion. This aspect is now not a major conclusion anymore.

d. The increase in the amplitude of the cyclonic PV anomaly from about -0.5 PVU to beyond -2.5 PVU (combined with a rapid areal expansion) over the 24-h period ending at 1800 UTC on 25 September (Fig. 7b and d) is "rapid" as stated on L442. However, as noted in item 4 above, this growth rate appears to exceed that expected for typical midlatitude baroclinic growth. It is highly likely that moist processes are involved, but because no estimates of growth rates are made in this study, it is impossible to know. It is therefore also inappropriate to conclude that the observed growth is "as expected from baroclinic instability" (L442) because the expected value remains unknown in the context of this work.

We agree that this conclusion has been a bit shaky. Note that the argument about baroclinic growth was (at least this was our intention) mainly made for the 6/12-h period from 12-18 UTC 24. Sept. After this time, we argued that non-linear barotropic dynamics is mostly responsible for the growth and downstream development. Still we agree that this is very rapid for (dry) baroclinic growth.

The reviews sparked further analyses of this aspect. As stated in the reply to General Comment 5 we revised in Figure 7 and made sure the conclusions drawn from it are well supported.

e. It is unclear to me what part of the analysis demonstrates that "the contributions of diabatic airstreams ... were negligible for the uncertainty amplification in this case" (L444-445). The non-conservative evolution of the PV streamer was remarkable in this case (Fig. 2), and the impact of diabatic PV reduction in WCB outflow on ridge amplification during the upstream RWB (Fig. 7a-d) was not analyzed in the study, as far as I can tell. This statement about the role of diabatic process on forecast uncertainty (L444-445) is very strong, inconsistent with previous work, and needs to be clearly supported by the presented analysis.

We meant that, for the uncertainty amplification shown in Figure 7 (which is before the PV streamer forms), WCB outflow was not present in the vicinity of the region with strong error growth. This was based on a careful analysis of WCB trajectories, but not shown, because of the few WCB air parcels identified in the domain. Of course, you are right that the PV streamer evolution was non-conservative, but with our statement about the contribution of diabatic effects we did not aim to characterize the later times but the timesteps shown in Figure 7. Also, we argue that it is not at all inconsistent with previous work, that direct modification by diabatic processes and diabatic airstreams are not very relevant for the amplification of forecast uncertainty in individual cases (see e.g. Baumgart et al. 2018). This of course, does not mean at all that we consider diabatic airstreams as generally not important for the amplification of forecast uncertainty. The attribution of the amplification of forecast uncertainties in Rossby waves to individual dynamical processes is a rather new research subject, but it seems that pure non-linear barotropic Rossby wave dynamics is very important (Davies and Didone 2013; Baumgart et al. 2018). Of course, diabatic airstreams can influence these dynamics, for example by placing a negative PV anomaly close to the wave guide. The absence of warm conveyor belts in the phase of the rapid error growth shown in Fig. 7 is a very strong indication that diabatic airstreams are not relevant for the uncertainty amplification in this case. However, this has not been discussed in the initial submission.

We now include the intersection points of WCB air parcels with 325 K in a revised version of Fig. 7 and in the synoptic overview. With this, we show that, even if the WCB was present at the later stage of the wave breaking, the few intersection points present in the early stage are far away from the region of the amplification and highlight this aspect when we discuss Figure 7, see L293-294, and L301-305.

Minor Comments

There are a relatively large number of grammatical errors in the submission, which I have not itemized here because of the major reworking of the text that will be required to address the issues identified above.

1. [L50] It is not clear why the introductory reference to parameterization uncertainty is useful here, where initial condition uncertainty is described in the subsequent passage. I would suggest starting this paragraph with "A major source ...".

2. [L51] I don't think that "slight uncertainties in initial conditions typically grow" (my emphasis), because the majority of uncertainties in any given analysis project onto decaying modes in the atmosphere (Privé and Errico 2013). I think that it would be more precise to say something like, "Slight uncertainties in the initial conditions that project onto the growing modes of the atmosphere can increase in amplitude during the forecast and potentially ...". You could also just replace "typically" with "can" in the current phrase.

3. [L87-L94] Suggest dropping this subsection in favour of the analysis in section 3.

4. [L95-L97] Having a clear set of objectives is a good idea, but these questions are framed in a way that is too complex to make them useful for the reader (e.g. "what is a and what of b leads to c and d in e"). Suggest simplifying or removing these questions.

5. [L99-L105] Provide a standard outline with section references.

Reply to comments 1-5: We adopt the suggestions in 2 and 5 (L102-109), and considered 1, 3 and 4 when rewriting the introduction (see your General Comment 1.a.i)

6. [L108] How are the ensemble members "perturbed": initial conditions, stochastic physics, SPPT, etc?

It is the standard ECMWF operational ensemble forecast. Because the details of how the ensemble is created are not relevant for the conclusions in this analysis, we will not go into detail here, but add a short statement clarifying this point (L114-116)

7. [L111 and elsewhere] The word "data" is plural, so "data are available", etc.

Thanks, we change it accordingly.

8. [L115] What climatology is used for the ACC calculation?

It is daily mean Z500 values from ERA-Interim from 1979-2014. We add this information in the corresponding section (L120-123)

9. [L122] Reference to a URL is inappropriate. At the very least, an access date needs to be provided. Consider including lightning strike information on the plots, rather than making reference to external information that may not be permanently available.

We added a map of lightning strikes to the supplementary material.

10. [L152] How is conditional instability identified in this analysis?

It is not directly identified. We remove “conditional”.

11. [L159] Figure 4a does not exist and Fig. 4 is not the CPS.

Apologies for this typo, it should have been Figure 5a. We corrected it.

12. [L207] Reference to Fig. 7 is out of order.

We do not see that this reference was out of order. We make sure that the reference is in order in the new version as well.

13. [L221] At what level are the differences significant?

Always 0.05 is used in the initial submission. We make sure that (now using false discovery rate alpha) it is always clear which level is used for assessing significance (L168)

14. [L231-L232] This sentence doesn’t make sense: does the amplitude “propagate” at a different speed from the difference? Are you differentiating between phase speeds and group velocities here? Please rephrase to make this clearer.

15. [L263] The section title should be much clearer, and not read like a news headline.

16. [L274] Three different time references begin this sentence. Please determine whether it is the time relative to streamer extension, Gregorian date/time, or forecast time that is most relevant here and stick to this description of the first column of Fig. 8.

Reply to comments 14-16: Thank you for pointing out these unclarities. We considered them in the revisions.

17. [L278] I don’t see that cluster 3 trough is “clearly” shifted to the east of the analysis at 1200 UTC 25 September (Fig. 8i). Instead, I see a trough that is too narrow, notably on the upshear flank over Germany.

Thanks for pointing this out. We agree and changed the text accordingly. (L363-364)

18. [L281] Why isn’t significance plotted here as in the first column?

For visibility reasons. In the later plots, large areas would be covered by the significance shading. The significance for all timesteps is shown in the supplementary material.

Using false discovery rate (as response to your general comment 7) gives a less “spotty” field and we now plot significant regions in all panels.

19. [L304-L305] This looks like more than just smoothing of the ensemble mean. Because averaging is a linear operation, the area-averaged ensemble-mean precipitation should match the observed values if the ensemble does not under-predict rainfall.

We do not agree, that area-averaged ensemble-mean precipitation should match the observed values. As long as the "observation" lies within the range of the ensemble, we don't think that there is an under-prediction. The area-averaged precipitation can be highly different between the ensemble members and as long as there are (even only very few) ensemble members that have equal or higher amounts of area-averaged precipitation than the "observed", the ensemble is fine (unless this is the case for most ensemble forecasts over many cases). We computed area-averaged accumulated precipitation for each member and the analysis over the study region and it shows that the value for the analysis is around the 90^{th} percentile of the ensemble.
We now add additional panels to Figure 9 that show the members with the highest and the lowest accumulated precipitation in each cluster in a box over the study region to illustrate the variability among members.

20. [L297] Are these SLP changes computed from the central pressures of the ensemble members, of from the ensemble mean? The search for a minimum central pressure is not a linear operation, so the results will likely be sensitive to the method. Particularly given the broad spatial distribution of group-2 centres, some/much/all of this apparent weakening may simply be the dilution of the ensemble mean if the ensemble averaging is done first.

Thanks for pointing this out. We computed it from the cluster mean. We agree that this is not a very good way to assess the SLP evolution in the clusters.
We now include boxplots of minimum SLP values of the cyclones in each cluster and add it to the cluster composites (new Figure 10).

21. [L312-315] The four lines of hypothesis here would be much better invested in the actual analysis rather than this forward-referenced supposition (I recommend the removal of this whole paragraph as noted in item 1.b.iii above).

22. [L317 and L320-L321] There seems to be an internal inconsistency here. On L317 the objective of the section is stated to be "to investigates ... subsequent development of a medicane-like system". However, on L320-321 you state that you "do not identify low-level warm cores directly and do not investigate their formation in detail". Because the warm core is one of the primary structural ingredients that distinguishes medicanes from typical Mediterranean cyclones (considering the CPS), these two statements seem to be in direct conflict.

23. [L321] What do you mean that you don't identify warm cores directly? The CPS-based warm-core detection is the basis for a large part of section 6.3.

24. [L344] Do you mean a larger increase in specific humidity in groups 1 and 2, than in group 3? This sentence suggests the opposite, likely because the intended target of the pronoun "this" is unclear (although the construction suggests group 3).

25. [L353-L355] What is the physical relevance of this comparison?

26. [L393-L399] Why bother with a set of conjectures right before performing the actual analysis? A far more direct approach would be to explain why the fractions of medicanes in each group differ, based on the analysis presented in earlier sections. The conjectures do nothing to build suspense for the big "reveal" of Table 1, and just serve to consume five lines of text unnecessarily.

27. [L399-L401] This text contains every number shown in Table 1, without offering any physical insight. Choose to present these numbers either within the text, or in a table, but not both.

Reply to comments 21-27: Thanks for these helpful comments that highlight parts of the analysis and the text that require revisions. This part of the paper was completely revised in the new version, considering these comments.

28. [L413-415] How is it concluded that "the detailed interaction between the surface cyclone and the upper levels become limiting factors" for predictability? Why can't internal storm processes or air-sea exchanges be the limiting factors? Those processes have not been investigated or ruled out as limits on storm structure predictability in this analysis, as far as I can tell.

We conclude that "sub synoptic-scale processes including the detailed interaction between surface cyclone and upper-levels become limiting factors". This does not exclude other sub-synoptic scale processes but maybe puts too much emphasis on the interaction of the surface cyclone and upper levels. We improve this sentence and also mention that internal storm dynamics/convection can be relevant. L480-482

29. [Fig. 6] Can the map resolution be increased a bit? (Similar in Fig. 7 zooms.)

Yes, this was done.

30. [Fig. 6] At what level are the contours significant?

Again, 0.05. As mentioned above, we made sure this is always clear.

31. [Fig. 6] The means are too similar to be usefully distinguished on the plot. Consider plotting the full ensemble mean only rather than solid and dashed contours.

We agree and changed the Figure accordingly (now showing the analysis field)

32. [Fig. 7] Is that a reference vector between (d) and the colour bar? If so, it should be highlighted and described in the caption. If it isn't, then one should be added.
Yes it is. Thanks, we made the reference vector better visible and mention it in the caption.
33. [Fig. 10] Use a fixed domain to ease comparison between panels.

34. [Fig. 10] What does the colour-coding of the trajectories represent (the last sentence of the caption is not clear about what is indicated "in colors")? Are different members assigned random colours? Why are there fewer cyclone positions in the groups than members within the groups? Are there multiple trajectories ending at the same point because of the degradation of the grid resolution? If so, there should be some way to represent the number of overlapping triangles (potentially the size of the triangle).

Each ensemble member has a different color.

35. [Fig. 10] How does the maximum "percentage of ensemble members with an airstream occurring at the specific grid point" occur outside of the trajectory envelope? For example, the maximum departure frequency in Fig. 10b occurs poleward of any trajectory. Is it because these trajectories are actually averages of many trajectory calculations? If so, then there must be some unusual spatial distributions to obtain density maxima away from the means. How many trajectories are computed in each member?

The "average trajectories" represent means of several trajectories (~12, depending on the location) for each member. Therefore, it is possible that there is a density maximum (of the all actual trajectories) away from the starting position of the "average trajectory". We realized that this analysis is a bit confusing.

36. [Fig. 11] Why are radii the best way to identify the blue and green lines? It would be clearer to label the blue line "center" and the green line "warm sector" because the radii are technical details rather than relevant features.

Reply to Comments 33-36: As mentioned above, we removed the trajectory analysis to achieve a better focus of the paper and instead provide a trajectory analysis based on the operational analysis in the supplementary material (which we discuss in the synoptic overview).

References
Privé, N. C. and Ronald M. Errico (2013) The role of model and initial condition error in numerical weather forecasting investigated with an observing system simulation experiment, Tellus A: Dynamic Meteorology and Oceanography, 65:1, DOI: 10.3402/tellusa.v65i0.21740

Stepanyuk, O., Jouni Räisänen, Victoria A. Sinclair & Heikki Järvinen (2017) Factors affecting atmospheric vertical motions as analyzed with a generalized omega equation and the OpenIFS model, Tellus A: Dynamic Meteorology and Oceanography, 69:1, 1271563, DOI: 10.1080/16000870.2016.1271563.

Wilks, D.S., 2016: "The Stippling Shows Statistically Significant Grid Points": How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It. Bull. Amer. Meteor. Soc., 97, 2263–2273, https://doi.org/10.1175/BAMS-D-15-00267.1

Wiegand, L. and P. Knippertz, 2014: Equatorward breaking Rossby waves over the North Atlantic and Mediterranean region in the ECMWF operational Ensemble Prediction System. Quart. J. Roy. Meteor. Soc., 140, 58-71.

Reviewer 2 (Florian Pantillon)

The paper investigates the large-scale dynamics that led to the formation of a tropical- like cyclone over the eastern Mediterranean in late September 2018, which was characterized by high forecast uncertainty in the operational ECMWF ensemble prediction system. A potential vorticity streamer issued from an anticyclonic Rossby wave breaking over eastern Europe was key in the Medicane dynamics. Two clusters of ensemble members with zonal shift of the streamer can be tracked along the Rossby wave guide back to initial conditions over North America. The evolution of the streamer into an upper-level cut-off low then controls the surface cyclogenesis, the stability and the advection of warm moist air that all support the Medicane formation.

Hybrid cyclones in general and Medicanes in particular are current sources of vivid discussions in atmospheric dynamics and objects of broad interest in the Mediterranean community. Contributions to better understand their dynamics and predictability are thus welcome and the paper presents interesting new material based on sound methods and high-quality figures. However, it suffers two major shortcomings: possible contributions from small-scale dynamics are largely ignored, although they at least partly explain Medicane formation, and the manuscript needs reorganization, as already pointed out by Referee 1. These shortcomings are linked somehow, as the tropical transition of the cyclone is actually assessed at the very end of the paper only. They are described below, as well as (many) specific comments.

The paper thus requires substantial revision before it can be considered for publication in Weather and Climate Dynamics.

General comments

Scales: as stated in the introduction, "the relative role of positive upper-level PV anomalies and air-sea interaction for the intensification of Medicanes is currently debated, as well as the question to which degree they are dynamically similar to tropical cyclones". The paper focuses on the synoptic scale and is based on model forecasts that do not explicitly resolve convection. This is fine but (1) the focus should be explicitly stated, (2) the limitation should be kept in mind throughout the paper and (3) the results should contribute to the current debate.

Organisation: as already pointed by Referee 1, the structure of the paper is unsatisfactory. Please better organize the Sections, make sure important concepts are introduced early in the paper (then stick to the terminology) and methods are described in the appropriate section, and avoid referring to later sections. In particular, show the warm core structure early in the paper, and comprehensively, based on the analysis for instance; in the present form the reader must wait until the last subsection of the last results section to learn the cyclone actually developed a warm-core structure.

Thanks for these comments. We obviously failed to clearly state the focus and intention of the study and the organisation of the Paper was not as appealing as we thought.

When revising the manuscript, we made sure that the focus is clearly stated and kept in mind throughout the study. In particular we state a definition of a medicane based on current literature, even if it is debated. In particular, we mention that there are cyclones classified as medicanes that do not seem to have real tropical dynamics while others undergo proper tropical transition. We then define how we identify medicanes in this study (which is by the presence of a deep warm core). This allows for a clearer terminology, which we make sure is consistently used throughout the text. We state clearly that the focus of the study is to investigate the predictability of the medicane in the early stage (the formation of the deep warm core) and not the later phase, when Zorbas acquires more tropical-like appearance (see e.g. L183-188).

We now show the CPS of Zorbas based on the operational analysis early in the paper (new Figure 1) and make clear which part of the life cycle we are looking at and why (see lines mentionned above, but also e.g. L417 -419.

We keep in mind limitations of the model data and approach used in this study. L482-484

We also provide a standard data and methods section (L110-175), and extend the synoptic overview section.

Specific comments

Title: the position and depth of the PV streamer exhibit some uncertainty in the ensemble forecast but the streamer itself is not uncertain; the link between PV streamer and Medicane could be more explicit.

Thanks for pointing this out. We changed the title of the paper as follows:

Medicane Zorbas: Origin of an uncertain potential vorticity streamer position and impact on cyclone formation.

Abstract
l. 3-4 This statement is not clearly supported. l. 5 “uncertain” is not properly used here (see comment on title).

We change the wording and make sure it is clear that mainly the PV streamer position/shape was uncertain (L5)

l. 7-8 “demonstrated”, “the dominant source”: not necessarily. See comments below.

Thanks for this comment.
We changed the wording to reduce the strength of this statement (e.g. L6-9)
See also replies to your comment below.

l. 9 Twice "strong(ly)".

l. 12 More details about the two air streams and their key role?

Reply to above two comments: Thanks, we rephrased and reconsidered the abstract content after the revisions.

1 Introduction
l. 19-25 All references relate to the North Atlantic, which should either be explicitly stated or extended to other oceanic basins.

We add two more references of studies in the North Pacific at this location (L24-25), and further above a reference to the Australian region L22.

l. 19-20 ET could also be mentioned here.

We agree that ET is an important process, but not really essential for this study. Therefore, we decided to not mention ET.

l. 26-33 It is unclear what is the difference between subtropical, tropical-like and hybrid cyclones, if there is one at all. And do not they by definition undergo tropical transition?

We agree that the wording is unclear here. We now try to make a better distinction between the different terms in the introduction. L21-31

l. 29 "air-sea feedback" is not precise enough.

We make this more precise by mentioning the WISHE mechanism that becomes active, when tropical transition occurs. L30-31

l. 32 This "may" result in high damage, as Medicanes often remain over sea.

l.35-39 Confusing what "they" and "their" refer to.

l. 41-42 Forecast uncertainty and the link with process understanding and ensemble forecasting needs better introduction.

l. 42-49 This appears too early and several keywords are not introduced yet (Zorbas, warm core, practical predictability, . . .).
l. 50 The transition should be smoother between 1.1 and 1.2.

l. 63 Lamberson et al. studied "extratropical" cyclone Joachim.

l. 64-65 The link between the predictability of breaking Rossby waves and Medicanes is far from obvious and need more details; it was extensively explored for a case study in September 2006:
Chaboureau, J. , Pantillon, F. , Lambert, D. , Richard, E. and Claud, C. (2012), Tropical transition of a Mediterranean storm by jet crossing. Q.J.R. Meteorol. Soc., 138: 596-611. doi:10.1002/qj.960
Pantillon, F., Chaboureau, J., Lac, C. and Mascart, P. (2013), On the role of a Rossby wave train during the extratropical transition of hurricane Helene (2006). Q.J.R. Meteorol. Soc., 139: 370-386. doi:10.1002/qj.1974
Pantillon, F. P., J. Chaboureau, P. J. Mascart, and C. Lac, 2013: Predictability of a Mediterranean Tropical-Like Storm Downstream of the Extratropical Transition of Hurricane Helene (2006). Mon. Wea. Rev., 141, 1943–1962, https://doi.org/10.1175/MWR-D-12-00164.1

l. 67 Who are "they"?

l. 67-69 The mentioned studies do not clearly attribute forecast busts to initial uncertainties rather than to the representation of diabatic processes. Both error sources would thus better be described together here.

l. 79 Which process?

l. 86 Even if only few case studies exist, there are certainly more than the two cited here.

l. 89 Some basic information about the case study are needed here (what, where, when). And where does the name "Zorbas" come from?

l. 91 Please explicit "short lead times".

l. 94 Which PV streamer? Either detail or remain general; referring to Section 3 does not help. l. 96 "leads" or "led" l. 100-106 Detailing Sections 2, 3, 4, . . . might be required.

Reply to Comments above: Thanks for all the above comments that point out unclarities, inconsistency, and missing depth in the introduction. We completely revise the introduction and consider these valuable points.

2 Operational ECMWF products
l. 110 Why 46 members only? What is the operational short-term forecast?

For technical reasons, four ensemble members were missing in the data we downloaded operationally. And as this forecast is not in the core of the study, we thought that 46 members were sufficient. The operational short-term forecast are based on the first 12 hours of forecasts started at 00 UTC and 12 UTC each day and are used to get an estimate of actual precipitation. We now include all ensemble members of this forecast and explain better what we mean with operational short-term forecast. L114

l. 112 Forecast data is actually available at higher frequency.
This is true, but not with the full vertical resolution which is needed to compute potential vorticity appropriately. Note that we download ensemble data on full model levels right after forecast completion because much fewer (pressure) levels are in MARS.

l. 115 What is the reference for computing ACC?

It is daily mean Z500 values from 1979-2014. We add this information in the corresponding section. L121-123

3 Synoptic overview
Figures 1-3 Zooming in on the region of interest would be very helpful to follow the discussion. Large-scale dynamics play an important role but, e.g., the Irish and Red Seas are not relevant. Consider then merging the three figures to avoid jumping from one to the other.

l. 135-136 The spiral-like structure is hardly seen. Or do you mean the frontal band? Spiral often refers to a tropical structure. Again, zooming in would help.

l. 139 Fig. 2C

l. 140-141 Move to the methods section.

Thanks, we now put most methodological aspects into the methods section.

l. 144, 146 What are “they”?

l. 139-149 Is it convection and/or large-scale ascent? The 600 hPa in 24 h criteria suggests the latter, while lightning suggests the former. The ECMWF model cannot actually resolve convection but you could check whether the precipitation is issued from the convection scheme or not.
We agree that the 600 hPa in 24h ascending air parcels are not ideal to show here.

We quantify the contribution of convective precipitation and mention it in the text (L221-227). Additionally, we show lightnings in the supplementary material.

l. 135-149 There is a general confusion in the paragraph between what was stated by previous studies and what happens here.

l. 152 Fig. 3d

l. 152 Clarify the precipitation is from model data; using different colors would help distinguishing the −11, 8, 15 and 21 mm (6h) contours on Fig. 1.

l. 153 How can you know it is due to conditional instability?
We do not check directly what type of instability occurs.
We now don't specify that it is conditional instability.

l. 156-157 How do you discern a warm seclusion from a warm core?

l. 157-160 and 166-167 There is not enough evidence at that point to claim a tropical transition. Relying here on Sec. 6.3 is not a good idea and there is no Fig. 4a. Either show more details or keep for later.

As commented above, we now show the CPS early in the paper and make sure we have a consistent wording. The tropical transition, that likely occurred in the later stage of Zorbas life-cycle, is mentioned now on L186-188, but it is clearly stated that it is not the focus of this paper.

l. 162 This would be worth showing!

This paper doesn't aim to deal with the tropical transition of Zorbas and discussing the eyewall formation would be clearly beyond the scope of this study. For the sake of focus, we don't show this aspect of Zorbas and leave it for further studies.

Reply to all above comment for section 3 with no direct replies: Thanks, these comments are all valuable to improve the text and/or figures. We considered them when revising and rewriting the synoptic overview.

4 Ensemble clustering according to position of PV streamer
l. 170-172 The sentence presents essential information but needs more support: why 00 UTC 27 Sep? Why 00 UTC 24 Sep? Is it perhaps the combination of valid time and initialization time resulting in largest spread? Can we see this somewhere? "Shown in Fig. 2 six hours earlier" is not too convincing.

These times were chosen because the PV steamer position showed this "nice" tri-modal behaviour just before cyclogenesis. We added a better motivation for these initial and valid times. L239-244.

l. 172-174 Referring to a later Section to motivate the present one is surprising.

l. 174-176 Please move technical details to the Methods section.

l. 178 Average of 320 K and 330 K levels or are there additional levels in between?
It's every 5 K, so we average 320,325,330 K. This should be now clear enough from the "data and methods" section (L119).

l. 178 Why set PV < 2 pvu to zero?

This was done to "mask" out the troposphere and just use stratospheric PV air when averaging PV vertically. This gives a field that is sensitive to the depth and PV values within the PV streamer. For example, we get the same values if the PV streamer is present just at one level with a PV value of 6 PVU or at three levels with each PV values of 2 PVU. Another way to look at it is that we want to cluster the ensemble members according to the PV streamer, and therefore remove the contribution of the variability of tropospheric PV values to the averaged field.

l. 180 Remind there are 50 members?

l. 190-191 Again, referring to a later Section is surprising.

l. 192-193 "Decrease" rather than "drop"? Are these values of ACC particularly low? And why not color clusters 2 and 3 in green and red on Fig. 5 as on Fig. 4?

Thanks for this suggestion. ACC values are not particularly low. We did not color clusters 2 and 3 because the focus here is to show the better performance of cluster 1. If clusters 2 and 3 are added the plot is less easy to read. We adopt the suggestion to use "decrease" (L265) but decided to not colour clusters 2 and 3.

l. 198 Errors in the shape of the PV streamer have not been discussed yet, only the zonal shift.

Thanks for this comment. The shape is somewhat included in the sense that if the shape of the PV streamer is completely wrong, the overlap would be too low to satisfy the criterion for cluster 1. The overlap can be low because of a shift or because of a different shape, or of course, a combination. The member that has been excluded from the analysis because it did not fit into a cluster actually shows no zonal shift at the tip of the streamer but a very special shape, such that the overlap is not large enough.

We now make sure, that it is mostly the zonal position of the PV streamer that is accounted for (L 237-238)

l. 199 Which characteristics are relevant?

l. 200-204 ACC may not account for the cyclone at all, at least the link is not showed yet. The link with the PV streamer is not obvious either. Consider adding Z500 on one of Figs 1-3.

Thanks for this comment. We agree that Z500 may not account for the cyclone. But it is reasonable to assume that Z500 is linked to upper-level PV, which is what is needed for the argumentation in this section. We now add Z500 in the synoptic overview (new Figure 3) and it is now nicely visible that first, Z500 accounts for the upper-level PV streamer/cutoff and later, the cyclone.

Reply to all above comment for section 4 with no direct replies: Thanks, these comments all contain valuable suggestions to improve the text and/or figures. We considered them when revising the paper.

5 PV streamer scenarios emerge from initial condition uncertainties and baroclinic amplification
l. 205-206 The jet streak has not been mentioned before. Consider either adding the large-scale dynamics leading to the PV streamer in Section 3 or, at least, shortly describing these dynamics here and motivating why they are the focus of the following analysis.

This is a good idea. We now include a Figure that provides an overview of the large-scale dynamics leading to the PV streamer (new Figure 2) and discuss it (L190-201)

l. 209-220 Please move to the Methods. Can you say some words about delta PV values, e.g., is there a threshold that indicates bi-modal distribution?

Delta PV values are just normalized differences and say something about how different the clusters are relative to the ensemble spread. We move this part to the methods and make sure the reader understands the meaning of delta PV values (L150-159).

l. 221- 223 While the normalized PV difference is clearly highlighted, PVU and wind contours barely differ between clusters and are not discussed at all.

We agree and use analysis wind contours now (see new Figure 7).

l. 230 The separation between clusters is hardly seen at that point.

We agree and only mention the PV difference in the text at this point. L285-290

l. 238-239 Remind Fig. 7d; better “stronger anticyclonic wave breaking” than “westward phase shift and larger amplitude”?

.

l. 242- 254 The description of Fig. 7e-h is difficult to follow and Fig. S1 suggests that differences in omega and Z850 are hardly significant in the region of interest. Consider removing altogether.

Thanks for these two comments. We also received comments regarding this Figure from Reviewer 1. We completely revised Figure 7, especially the right hand panels, where we present analysis and cluster mean fields with the goal to show the synoptic situation in which the error amplification takes place. And we now provide a better explanation for the uncertainty amplification (new Figure 9 and L306-348).

l. 244, 250 Either show or omit potential temperature.

l. 255- 263 This interpretation is meaningful and consistent with displayed material overall but (1) the formulation partly sounds speculative ("strong QG forcing", "uncertain low-level wave", "exponential growth", "strong vertical coupling") and (2) ensemble members differ not only in their initial conditions but also in their physical parametrisations or any other perturbations implemented by ECMWF to increase ensemble spread.

Thanks for this comment. Regarding point (2) we agree that the ensembles also differ in their physical parametrisations and errors might come from there but we can still see an initial condition uncertainty (at lead time 0, Figure 6 (new Figure 7)) that seems to propagate into the North Atlantic and amplify there. We therefore argue that some of the uncertainty most likely comes from initial condition uncertainty. We weakened the strong statements regarding the initial condition uncertainty (e.g. we remove it from the title of section 5 (L274), and more carefully describe its role (L280-281). We thoroughly revised this section and the analysis in Figure 7, add a new Figure (new Figure 9) and analysis (L306-349) to arrive at conclusions that are less speculative and are clearly shown in the section (e.g. L349-351).

Reply to all above comments for section 5 with no direct replies: Thanks, these comments point out unclarities or suggest an alternative wording. We considered them when revising the paper.

6 Diverging synoptic development impacts Medicane predictability
l. 268 What is a "Medicane-like" cyclone?

We will make sure our terminology is clearly defined in the introduction and consistently used throughout the text. See reply to your initial comment.

l. 270 The subsection provides a synthetic summary of the dynamics of all clusters, but what is the variability between members within each cluster?

We agree that the variability between members is not shown (except for the cyclone locations). We understand the goal of making clusters to reduce the information from all members into scenarios. The statistical significance test allows to draw the conclusion that the member within a cluster are really different from the members in another cluster. We now add information regarding the variability in cyclone intensities (in new Figure 10 and L381-390), precipitation (new Figure

11, L399-409) . The variability of upper-level PV and low level equivalent potential temperature is now discussed in the new section 6.2 (L415-483)

l. 275-276 Mention the analysis PV is depicted by the black contour.

l. 277 What is meant by "exactly the ones"?

l. 278-279 There is no visible difference in PV between clusters in Fig. 8 a, e, i.

On the western side of the trough (marked by the p-values) slight differences are discernible. The point here is that the differences are maybe still small when comparing the full PV fields, but (as shown with the PV difference plots) they are significant and they propagated into the trough from upstream.

l. 288 Fig. 2c; slightly different time.

l. 288-291 Mention the cyclone in individual members is depicted by dots.

l. 295 Differences are substantial but not necessarily due to latent heat release (only).

We agree that the differences in the PV cutoff evolution are not necessarily all due to latent heat release. However, if PV is eroded in cluster 1 and not in cluster 3, and as erosion of PV cutoffs is known to be related to substantial latent heat release, this is a clear indication that cluster 1 experiences more latent heating or at least, a different one (i.e. one in the vicinity of the PV cutoff).

l. 297-298 Clarify these are mean values.

l. 305 Fig. 9d

l. 306-306 The smoothing effect due to averaging makes the comparison difficult for precipitation; how do individual members look like? You could e.g. compute PDFs of accumulated or instantaneous precipitation for each member and the analysis.

Thanks for this comment. We included new panels in Figure 9 (new Figure 11) and added a paragraph that provides information about the internal variability within the clusters to make them better comparable to the analysis (new Figure 11, L399-409, see comment above).

l. 314-316 These arguments are too speculative and are better left to Sec. 6.3.

l. 318 Again, what is a Medicane-like system?

l. 318-325 This paragraph is confusing and must be rewritten/streamlined. How do you define a low-level warm core, a warm seclusion and a Medicane-like system, and why do you focus on two air streams?

l. 326-327 The analysis tool Lagranto belongs to the methods and is already mentioned above.

l. 330 Is the warm core formation shown somewhere?

l. 345 "weaker" not "stronger" increase.

l. 346 In clusters 1 and 2.

l. 347 The Mediterranean Sea is not an ocean.

l. 353-356 This would likely better fit at the beginning of the paragraph.

l. 363 Closer to the coast but the region remains the eastern Mediterranean.

l. 363-371 The discussion is speculative so far, as the cyclone thermodynamics have (still!) not been documented yet. There is also a general confusion between warm core, warm seclusion, warm sector and tropical structure.

l. 377-394 This all belongs to the Methods. What radius is used to compute CPS metrics?

l. 394-399 I expect clusters 2/3 to show more favourable low-level/high-level forcing but not necessarily to produce a stronger/weaker Medicane.

l. 400 Avoid introducing an additional name ("DWC"), which adds confusion, better stick to the terminology used up to that point.

l. 400-402 What about the two other CPS metrics, symmetry and low-level warm core? The upper-level warm core metric might be contaminated by the presence of the PV streamer/cutoff. And what about the cyclone intensity?

Thanks for this comment. Although it could be worth looking at these additional aspects, the focus of this study is to investigate the factors affecting the formation of the deep warm core in Zorbas, which is a major characteristic of medicanes. The low-level warm core is indirectly included in the sense that it is a necessary requirement for a deep warm core. However, as the upper-level warm core is a distinguishing factor that separates so-called "medicanes" from subtropical cyclones, it is in the main focus. Regarding parameter B, this is a measure of the frontal nature of the cyclone, but analysing the frontal structures of Zorbas is beyond the scope of this paper.

In order not to extend the analysis, we have decided not to analyse those parameters. However, we now additionally analyze the maximum cyclone intensity in each cluster (in new Figure 10 and L381-390)

l. 403-404 But cluster 3 produces stronger upper-level warm cores than cluster 2, which contradicts the other results and interpretation.

This is true, but cluster 3 produces only 2 cyclones with an deep warm core, so this average has to be taken with caution. We completely revised and extend the deep warm core analysis and now show individual members rather than box plots to make the clusters better comparable (see new Figure 12) .

l. 407-409 The three-day long sustained deep warm core (Fig. 5a) appears unprecedented. Can you provide CPS diagrams for the analysis?

Yes, we provided the CPS of the analysis in the introduction (new Figure 1).

l. 412-413 Why?

l. 414-416 What about convection?
We agree that convection can also be a relevant sub-synoptic scale process. This sentence was not meant to exclude other factors. We make this sentence clearer and mention that also internal storm dynamics/convection can be relevant (L481-483).

Reply to all above comment for section 6 with no direct replies: Thanks, these comments point out unclarities or suggest small structural or content changes of the text. We considered them when revising the paper. Especially we made sure that the terminology is clearer.

7 Conclusions
l. 427-418 Again, what is a subtropical cyclone, a tropical-like system or a Medicane?

l. 431 More details about this “first case”?

l. 432 Which process?

l. 435-438 This is not shown here.

l. 445-446 How do you know this?

It has also been pointed out by reviewer 1 that it is not clear how we arrive at the conclusion that diabatic air streams were not relevant for the uncertainty amplification. We supported this point better with the analysis (new Figures 2 and 8) by providing more information about warm conveyor belts and precipitation. See L293-294, L301-305.

l. 446-447 Which process, baroclinic instability?

l. 455-457 Ensemble forecasts are computed with different perturbation methods thus the error growth cannot be attributed to initial conditions only here.

Thank you for this comment. As already commented above, we agree that model errors can contribute to errors in this case. However, we show that a patch of uncertainty is present at initial time of the forecast that then propagates over the North Atlantic where the amplification takes place in its vicinity. Even if we cannot exclude model error here, we argue that the analysis shows strong indication, that initial condition uncertainty was very relevant in this case.

We reduce the strength of the statement such that it is clear that we argue that initial condition uncertainties contributed substantially to the forecast uncertainty, but not exclusively. L494-497

l. 458-460 . . .and convection and its organisation.

l. 462 As the used data is from ECMWF essentially, it could be stated how to access it.

The data used is not available long term from ECMWF with this high vertical resolution (which was required to compute PV and trajectories). We downloaded this data immediately after the event occured

Reply to all above comment for section 7 with no direct replies: Thanks, these comments point out unclarities or inconsistencies in the conclusions. We carefully considered them when revising the paper.

References
Providing DOIs or URLs for all papers would be helpful
Thank you for this comment. We provided DOIs (see reference list)

Figures
Moving all figures to the end of the paper would ease the review.

In the submitted document for discussion they are all at the end of the paper.

Fig. 5: it is unclear which date relates to which tick mark.
We changed the Figure to make this more clear (new Figure 6).
Fig. 7 appears before Fig. 6.

Fig. 6: consider changing the color scale to [-1,5; 1,5] and plotting coast lines at higher definition. We changed the color scale to [-2,2] for all PV difference plots (new Figure 7 and 8)

Fig. S1: the title should refer to Fig. 7 not 6.

Reply to all above comments regarding the Figures: Thanks, for these comments. We considered them when revising the paper.

Reviewer 3

This interesting paper deals with the predictability of the Medicane Zorba, which affected the central Mediterranean in September 2018. ECMWF ensemble forecasts are used in this effort. The limit in the predictability of the cyclone is analyzed and discussed in connection with the upper-level PV, which also affected the low-level evolution.

This is one of the first paper that clearly identifies the relevance of PV features in the predictability of Medicanes. The results are a relevant contribution in the field; however, the analysis should be substantially improved. Some points are indicated hereafter.

Major points:

Line 27, Line 154-157: I would like to see some clarifications about the definition of Medicanes. Although there is no general consensus, in most of the literature (e.g., Miglietta et al., 2011; Picornell et al., 2014; Cavicchia et al., 2014), a Medicane is considered as an extra-tropical cyclone that acquires a symmetric, deep warm core in the Mediterranean region. At the same time, the presence of a deep, warm core is not always an indication of tropical-like processes going on: as discussed in Fita and Flounas (2018) and Mazza et al. (2017), a deep warm core is not necessarily associated with a WISHE mechanism, but it can also be induced by a warm seclusion. However, Miglietta and Rotunno (2019) have shown that the intensification of the same cyclones discussed in the latter two papers cannot be explained without considering the sea surface fluxes and the latent heat release, in analogy with the WISHE mechanism typical of tropical cyclones. For these reasons, I suggest to remove "occasionally" (Line 27) and the sentence "Warm seclusion have been previously linked to Medicane formation" (Lines 155-156). Also, at Line 317 and 321 you "investigate potential pre- cursors of a low-level warm core": however, the low-level warm core is not relevant for the following development of the cyclone in itself (see also Line 365, 396), but because of the high values of equivalent potential temperature that are responsible for potential instability and favor the development of convection at later times.

Thank you for these explanations. We made sure that in the revised paper we have a clearer definition of medicanes and that the subsequent argumentations are consistent with this definition (see e.g. L32-44, L134-135)

Figure 10: I found the understanding and interpretation of Figure 10 quite difficult; in particular, I did not understand if the back trajectories you show are averages over all the ensemble members, since this is not mentioned in the Figure caption and not clearly reported in the text (Line 334); also, the presence of high percentages far from the plotted trajectories (purple shading) is counterintuitive.

Line 350-355: for a more comprehensive analysis of the trajectories in Fig. 10, some information should be included about the change of height along them.

Thanks for pointing out these unclarities in Figure 10 and the related analysis. As the other reviewers were also critical regarding this analysis and because it is not the centrepiece of the paper, we removed this analysis in favour of enhancing the focus of the paper. We provide a trajectory analysis based on the operational analysis in the supplement and an additional analysis of low-level THE and and upper-level PV evolution in each member (see new Figure 12)

Minor points:
Line 115: please provide the definition of ACC

There is now the definition of ACC in the supplementary material.

Line 130: red instead of blue

Lines 143-145: the role of upper level PV anomaly in the generation of Medicanes is also discussed in Miglietta et al. (2017)

Thanks, we now include this reference. L44

Line 201: were instead of is

Line 202: severely instead of severly

Line 207: please can you provide an approximate indication of the height the isentropic = 325 K corresponds to?

Isentropic surfaces are generally not horizontal, they can intersect the surface towards the equator and be in the stratosphere at the pole. We provided information about the approximate pressure level of the 325 K in the specific location over the North Atlantic that is discussed in this part of the text in the caption of new Figure 9. Additionally we provide this information for the Mediterranean region (see caption of new Figure 3).

Line 214: why do you give less weight to the regions of strong gradients?

The idea of standardized anomalies is to quantify how different two clusters are, relative to the ensemble spread at a specific location. Just looking at absolute PV differences can result in high values just because the spread is high. But in this case, we are interested in regions where the clusters start to separate within the ensemble. As a result, regions of strong gradients are usually given less weight, because that’s where the ensemble spread is usually large.

Line 221: explain why “stratospheric” side

This is because we use the 2 pvu contour to separate stratospheric and tropospheric air masses. The PV difference is located poleward of the 2 pvu contour and of the jet streak which is a region of stratospheric air.

Line 311: favorably instead of favourable

Line 317 and elsewhere: Medicane or tropical-like, not Medicane-like!

We avoid the term Medicane-like and provide a better definition of the terms used in the introduction (L21-L44)

Figure 5 caption: why do you use only 46 members for the second ensemble?

Because this is what we downloaded operationally and for technical reasons, 4 members were missing in this forecast.

We now include the missing 4 members.

Figure 6: change contour line colors to facilitate interpretation

Figure 7 caption: please indicate that the black contour refers also to captions (a-d)

Figure 8 caption: (e,i) instead of (e,f); the black contours around the teal patches create confusion

Reply to all above comments with no direct reply: Thanks, for these comments that show wrong spelling or suggest Figure improvements. We considered them when revising the paper.

Supplement material, Line 9: "The results show that significant differences of QG are located in the region of strong QG on 1800 UTC 24 Sep 2018": it does not seem to be the case, at least at that time.

There are some patches of significant differences of QG omega in the region where QG omega is high. But we agree that this argumentation and analysis are a bit shaky. Also as response to comments from the two other reviewers, we revised the analysis shown in Figure 7 completely and it does not require QG omega anymore.

REFERENCES:

Cavicchia, L., von Storch, H., Gualdi, S. (2014). Mediterranean tropical-like cyclones in present and future climate. J. Clim. 27, 7493–7501.

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., Cerrai, D., Laviola, S., Cattani, E. and Levizzani, V. (2017). Potential vorticity patterns in Mediterranean "hurricanes". Geophysical Research Letters, 44, 2537–2545.

Picornell, M. A., J. Campins, and A. Jansà (2014). Detection and thermal description of Medicanes from numerical simulation, Nat. Hazards Earth Syst. Sci., 14, 1059–1070.

Medicane Zorbas: Origin ~~and impact~~ of an uncertain potential vorticity streamer position and impact on cyclone formation

Raphael Portmann[1], Juan Jesús González-Alemán[2], Michael Sprenger[1], and Heini Wernli[1]

[1]ETH Zurich, Institute for Atmospheric and Climate Science, Zurich, Switzerland

[2]University of Castilla-La Mancha, Environmental Sciences Institute, Toledo, Spain

Correspondence: Raphael Portmann (raphael.portmann@env.ethz.ch)

Abstract. Mediterranean tropical-like cyclones (~~Medicanes~~medicanes) can have high societal impact and their accurate forecast remains a challenge for numerical weather prediction models. They are often triggered by upper-level potential vorticity (PV) anomalies, such as PV streamers and ~~cut-offs~~cutoffs. But knowledge is incomplete about their detailed formation processes and factors limiting their predictability. This study exploits a European Centre for Medium-Range Weather Forecast (ECMWF) operational ensemble forecast with an uncertain PV streamer position over the Mediterranean, ~~which~~that, three days after ~~initialisation~~initialization, resulted in an uncertain development of Medicane Zorbas in September 2018. ~~Using an~~ An ad-hoc clustering of the ensemble members according to the PV streamer position ~~, it is demonstrated that~~ is used and cluster differences show a region of enhanced uncertainty in the initial conditions ~~near an upper-level jet streak~~ in the upper troposphere on the stratospheric side of the jet stream over the Gulf of Saint Lawrence~~is~~. This feature is advected over the North Atlantic where it amplifies rapidly on the stratospheric side of an emerging upper-level jet streak and further propagates into the Mediterranean. Uncertainties in the ~~dominant source of the subsequent uncertainty in the position of the PV streamer over the Mediterranean. The initial condition uncertainty strongly amplifies baroclinically after 18 h in a region of strong quasi-geostrophic forcing for ascent in the left exit of a jet streak~~ ageostrophic circulation associated to the jet streak contribute substantially to the initial amplification over the North Atlantic. The further amplification and downstream propagation of the tropopause-level PV uncertainty leads to a large spread in the position of the PV streamer over the Mediterranean after three days, directly limiting the predictability of the position, thermal structure and evolution of ~~Zorbas. Two~~ the Mediterranean cyclone. In particular, the eastward displacement of the PV streamer in more than a third of the ensemble members results in a very different cyclone scenario, where low values of low-level ~~airstreams possibly play a key role in linking the uncertainties of the large-scale~~ equivalent potential temperature and the missing erosion of the high upper-level ~~flow with meso-scale uncertainties in the cyclone structure~~PV prevent the development of a deep warm core that is typical for medicanes. Overall, this study ~~is an illustrative example~~ illustrates that uncertainties in large-scale initial conditions and uncertainty amplification over the North Atlantic can determine the practical predictability limits of a high-impact weather event in the Mediterranean.

1 Introduction

1.1 Medicanes: high impact, limited understanding, and uncertain forecasts

In the last 15 years, ~~there has been increasing~~ research on cyclones ~~placed~~ at the interface between the ~~the~~ classical concepts of tropical and extratropical cyclones ~~. Several studies focused on~~ has been increasing. Sometimes, these cyclones are called hybrid cyclones (e.g. **?**). They are usually characterized by a lower-tropospheric warm core and an upper-tropospheric cold core. Frequently occurring hybrid cyclones are so-called subtropical cyclones~~(STCs;~~ , which mainly occur over ocean basins (e.g. ~~**??**). STCs are low pressure systems that~~ **????**). Often, they initially form from ~~baroclinic processes, but they later acquire tropical characteristics due to convective processes, while fronts dissolve. In the second phase convectionbegins to build a low-level warm core~~elongated equatorward intrusions of stratospheric high potential vorticity (PV), so-called PV streamers, that subsequently detach from the main stratospheric high PV reservoir, forming a PV cutoff. The large-scale ascent and destabilization by the high upper-level PV favors convection, which starts to form a lower-tropospheric warm core. Fronts dissolve and the environment becomes more barotropic. ~~STCs~~ Subtropical cyclones have gained attention because they can have substantial ~~impacts~~ impact on society (**?**), but also because they ~~potentially convert into fully-tropical cyclones, a process called~~ can undergo tropical transition (**?**)~~.~~, i.e. they can amplify through the so-called wind induced surface heat exchange mechanism (WISHE; **?**) and convert into tropical cyclones.

A similar phenomenon occurs over the Mediterranean Sea~~. So-called Medicanes (Mediterranean Hurricanes; **??**~~, and the resulting cyclones are sometimes called medicanes (Mediterranean hurricanes; **???**) or Mediterranean tropical-like cyclones (**?**)~~are (most often) hybrid systems, i. e. between extratropical and tropical cyclones, which occasionally~~. There is not yet a clear definition of medicanes. A common property seems to be the development of a (symmetric) warm-core structure throughout the whole troposphere, often called deep warm core. The deep warm core of medicanes is not necessarily produced directly by convection but can, especially at lower levels, be promoted by horizontal advection and the seclusion of warm air in the cyclone center, which is why it has been argued that some medicanes are rather subtropical than tropical-like cyclones (see e.g. **?**). Medicanes can undergo tropical transition ~~. If the latter occurs, they can amplify through air-sea feedback (**?**)~~ and acquire the typical appearance of a hurricane (**?**) with convective cloud bands wrapped around a cloud-free central eye and a typical size of their associated cloud clusters on the order of 300 km in diameter. In this case, convection forms ~~a robust warm core that reaches the upper troposphere and is~~ and maintains a robust deep warm core and the medicanes are associated with strong horizontal pressure gradients, wind, and rainfall. This ~~results~~ may result in high damage, although ~~Medicanes~~ medicanes rarely attain hurricane intensity. ~~Medicanes can form when positive upper-level potential vorticity (PV) anomalies (e.g. PV streamers and cut-offs that result from Rossby wave breaking) reach the Mediterranean Sea. These PV anomalies trigger extratropical cyclogenesis and if they are accompanied by substantial deep convection they can further develop into a Medicane (**????**).~~ The relative role of positive upper-level PV anomalies and air-sea interaction for the intensification of ~~Medicanes~~ medicanes is currently debated~~, as well as the question to which degree they are dynamically similar to tropical cyclones (see e. g. **?**). There~~. In most cases the positive upper-level PV anomaly seems to be ~~a high case-to-case variability (**?**), and, given their infrequent occurrence, the systematic study of Medicanes, their formation dynamics, and sources of forecast~~

~~uncertainty remains a challenge (?). In addition, in~~ important for the initial intensification (?).

In situations with the formation of ~~Medicane~~a medicane, ensemble prediction systems often show ~~a large spread even at short lead times~~ large uncertainties in cyclone occurrence, cyclone intensity, and intensity of the upper-level warm core even at few days lead time (e.g. ?). ~~This study investigates the uncertain three-day ensemble prediction of Medicane Zorbas in September 2018, with the aim of better understanding the effect of large-scale upper-level dynamics on the formation of a warm-core Mediterraneancyclone. Other studies have shown that a detailed analysis of ensemble forecasts, especially of different scenarios offered by a particular ensemble prediction, can be highly rewarding for better understanding the involved dynamics and the practical predictability limits. For example , such scenarios have been used to identify key dynamical elements~~ For a medicane case in 2006, it was shown that beyond 36 h lead time ensemble members that were missing a PV streamer over the Mediterranean had a lower probability of forecasting a medicane (??). This was linked to uncertainties in an extratropical transition event over the North Atlantic, which rapidly propagated downstream into the Mediterranean. However, even from the ensemble members that formed a PV streamer over the Mediterranean, only about 9% actually forecasted a medicane. This indicates that even if the PV streamer is forecasted, other synoptic-scale aspects related to the PV streamer, for example uncertainty in its position, shape, or intensity may be limiting the predictability of ~~tropical cyclones (????) and atmospheric blocking (?).~~

1.2 ~~Origin and amplification of forecast uncertainties in Rossby waves~~

~~Besides uncertainties inherent to the forecast model (e. g. due to parametrisations), a major source of uncertainty in numerical weather forecasts is~~ the ~~limited accuracy of initial conditions. In fact, even slight uncertainties in initial conditions typically grow with increasing forecast lead time and potentially result in very different large-scale weather patterns a few days to a week after initialisation (e.g. ??). However, forecast uncertainty does not grow homogeneously but depends on the flow itself. To quantify this flow-dependent~~ medicane. Also in this study about Medicane Zorbas, the position of a PV streamer in the Mediterranean varies between ensemble members and it is decisive for the evolution of the medicane. PV streamers result from the non-linear development of Rossby waves, which is known to be susceptible to the growth of forecast ~~uncertainty, operational ensemble predictions are conducted since the early 1990s and have continuously improved since then (??). The uncertainty can be quantified with the ensemble spread that serves as a proxy for the expected forecast error of the ensemble mean if the forecast is reliable (which is not true in all flow situations, see ?). Understanding~~ uncertainties. Therefore we now briefly review the relevant literature on the origin and amplification of forecast ~~uncertainty and error~~ uncertainties in Rossby waves.

1.2 Origin and amplification of forecast uncertainties in Rossby waves

Understanding forecast error and uncertainty is crucial to reveal the limits of and potential for improving weather forecasting systems. In recent years, research increasingly focused on the origin and amplification of forecast errors along the upper-level wave guide, i.e. the near-tropopause band with a high isentropic PV gradient (??), because there they tend to ~~amplify and propagate downstream due to Rossby wave dynamics~~ be largest (e.g. ???). ~~Such near-tropopause errors in Rossby waves can~~

~~limit tropical cyclone predictability (e.g. **?**) and they can also become relevant for Medicanes related to breaking Rossby waves that form strong upper-level PV anomalies over the Mediterranean.~~ Many studies used the PV framework to investigate forecast errors along the Rossby wave guide (for example **?????**). ~~They~~

Forecast errors in Rossby waves can *originate* from errors in initial conditions ~~and have shown to~~ or model errors (e.g. the misrepresentation of diabatic processes). In certain situations they can result in so called "forecast busts" over Europe, i.e. periods of anomalously low predictability (**???**). ~~Also, the~~ The misrepresentation of diabatic processes ~~in the forecast model can induce PV errors, as for example in~~ within warm conveyor belts can be an important error source for Rossby waves (**??**). Warm conveyor belts affect the Rossby wave guide via their low-PV outflow in the upper troposphere (**??**). PV errors near the tropopause ~~, due to errors in initial conditions and /or model physics, translate into Rossby wave errors~~ represent errors in the structure and amplitude of Rossby waves that then propagate downstream.

To become relevant, forecast errors ~~in Rossby waves~~ and uncertainties must *amplify*. It is well known that even slight uncertainties in initial conditions can grow with increasing forecast lead time and potentially result in very different large-scale weather patterns a few days after initialisation (e.g. **??**). However, the growth of forecast uncertainty is not homogeneous but depends on the flow itself. To quantify this flow-dependent growth of forecast uncertainty, operational ensemble predictions are conducted since the early 1990s and have continuously improved since then (**??**). The uncertainty can be quantified with the ensemble spread that serves as a proxy for the expected forecast error of the ensemble mean if the forecast is reliable (which is not true in all flow situations, see **?**). Different case studies have shown ~~that this can be~~ than forecast error in Rossby waves can amplify due to a range of processes. **?** ~~point~~ pointed out the importance of non-linear (barotropic) dynamics near the tropopause and to a lesser extent baroclinic interaction and effects of upper-tropospheric divergent winds. The contribution of non-linear tropopause dynamics to amplification (and downstream propagation) of forecast errors can be understood by the mutual interaction of negative and positive PV errors near the tropopause ~~(**?**)~~(**???**). **?** identified warm conveyor belts in a forecast bust case as key for amplifying forecast errors in the tropopause region. ~~This process includes error amplification in the baroclinic growth of a cyclone, possibly enhanced by model error in the representation of diabatic processes in the warm conveyor belt, resulting in errors in the divergent wind and size and amplitude of the negative PV anomaly in the upper troposphere.~~ Together, these studies indicate that forecast errors in near-tropopause Rossby waves can grow just due to their internal non-linear dynamics, but baroclinic ~~coupling can, via rapid baroclinic growth or~~ growth and warm conveyor belts and their associated divergent low-PV outflow ~~,~~ can enhance this amplification. The relevance of these processes to the amplification of forecast errors and uncertainty is ~~likely very case dependent~~ case and location dependent (**?**) and their interplay potentially complex. ~~However, only few case studies exist that investigate the origin, amplification and propagation of forecast errors in Rossby waves and their impact on the uncertainty in the prediction of high-impact weather events.~~Therefore, more case studies are useful to better understand the involved mechanisms.. In the context of high impact weather, a better understanding of the synoptic-scale elements controlling the practical predictability (i.e. the predictive skill of state-of-the-art forecasting systems), can help to improve situational awareness.

1.3 ~~Zorbas: An uncertain Medicane~~

~~In this context , we exploit the recent Medicane Zorbas, which rapidly acquired fully tropical-like characteristics. Operational ensemble forecasts by the European Centre for Medium Range Weather Forecast (ECWMF) did not agree on the development of Zorbas, even at short lead times.Zorbas, from the moment of its formation until it acquired full tropical-like characteristics, led to considerable damage through severe winds, torrential rainfall, major flooding and even tornadoes. The main affected region was Southern Greece, especially Crete, Peloponnese, Evia, and the region around Athens.We investigate the origin and amplification of~~

1.3 Zorbas: An uncertain medicane

This study investigates the ECMWF ensemble prediction initialised 84 h before cyclogenesis of Medicane Zorbas in September 2018, which is characterized by an uncertain position of the precursor PV streamer over the Mediterranean after 72 h lead time. The main objectives are: (i) to identify the ~~forecast uncertainty associated with the PV streamer that triggered Zorbas (see Section 3) and identify the relevant large-scale conditions limiting the Medicane's predictability. The key questions of our study are: What is the~~ origin of the forecast uncertainty ~~in the PV streamer and which sequence of dynamical processes lead to its amplification and propagation into the Mediterranean? Was the uncertainty in the PV streamer the direct cause of the uncertain Medicane prediction and why? This study does not aim to analyze the details of the Medicane dynamics but rather focuses on the~~ and analyze the synoptic environment in which the ensemble spread amplifies and propagates; and (ii) to investigate the different large-scale ~~processes prior to~~ scenarios over the Mediterranean offered by the ensemble and understand how they are related to the ~~Medicane formation.~~ formation of a medicane. Other studies have shown that a detailed analysis of ensemble forecasts, especially of different scenarios in a particular ensemble prediction, can be highly rewarding for better understanding the involved dynamics and the practical predictability limits. For example, such scenarios have been used to identify key dynamical elements limiting the predictability of tropical cyclones (**????**), medicanes (**?**), and atmospheric blocking (**?**).

The remainder of this article is structured as follows. ~~After describing the data used in this study we give an overview on~~ A description of the data and methods used (section 2) is followed by an overview of the synoptic evolution of Zorbas and ~~introduce the main method that builds the basis for all subsequent analyses~~the large-scale situation previous to cyclogenesis (section 3). Then, ~~the next Sections are organized as a journey from the origin of the large-scale forecast uncertainty to its effect on the uncertainty of the cyclone development and precipitation impacts. This is then connected to uncertainties in the low-level moisture and warm-air advection prior to the Medicane formation~~a pragmatic clustering is introduced that uses the uncertainty in the PV streamer position to separate the ensemble forecast into three distinct PV streamer scenarios (section 4). Section 5 discusses the origin, amplification, and propagation of the forecast uncertainty before the formation of the PV streamer. sectionion 6 presents the synoptic evolution of the three scenarios in the Mediterranean and investigates how the scenarios influence the resulting Mediterranean cyclone. Finally, ~~uncertainties in the evolution of the vertical thermal structure of Zorbas are diagnosed and links to the relevant precursors discussed. We close with highlighting the~~ in section 7, we draw the main conclusions and discuss implications for further research on ~~Medicane dynamics and on~~ uncertainty in Rossby wave forecasts and on medicane dynamics.

2 ~~Operational ECWMF products~~Data and methods

2.1 Data

The basic data for this study ~~is~~ are from the ECMWF Integrated Forecasting System (IFS, Cycle 45r1; **?**). ~~We use the operational~~ Operational ensemble forecasts with 50 perturbed members initialized at 0000 UTC 24 Sep 2018 and 0000 UTC 27 Sep 2018~~(46 perturbed members available)~~, the operational analysis, and operational short-term (12 -hourly accumulated, initialized at 0000 and 1200 UTC) forecast of 6-hourly accumulated precipitation, are used. The ECMWF operational ensemble forecast is based on perturbed initial conditions as well as stochastic perturbations of model physics (for details see **?**). The spectral resolution of the operational ensemble is TCO639 (about 18 km) on 91 model levels, and the resolution of the operational analysis TCO1279 (about 9 km) on 137 model levels. The data ~~is~~ are available every 6 h and ~~has~~ have been interpolated to a regular grid with a horizontal resolution of 1° ~~x 1°~~. From the standard variables we additionally compute PV ~~,~~ on isentropic surfaces (every 5 K), and equivalent potential temperature ~~and quasi-geostrophic omega (QG ω) on 850 hPa as forced by levels above 550 hPacomputed the same way as in **?**~~ (θ_e) on pressure levels (every 25 hPa). As a measure for forecast skill, anomaly correlation coefficients (ACC) are calculated for geopotential height at 500 hPa for each ensemble member of the forecasts initialized at 0000 UTC 24 Sep 2018 and 0000 UTC 27 Sep ~~2018.~~

3 ~~Synoptic overview~~

~~Figures **??**, **??** and **??** provide an overview of the atmospheric processes before and during the formation of Zorbas based on ECMWF operational analysis and infrared images from~~ 2018 using the daily mean ERA-Interim climatology from 1979-2014 as a reference (for details see supplementary material). Anomalies of 900 hPa θ_e were computed with respect to the September/October ERA-Interim climatology from 1979-2017.

Additionally, satellite data (infrared channel 9 (10.8 μm) of MSG SEVIRI) provided by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) ~~. At 0000 UTC 26 Sep 2018, about a day before cyclogenesis, a few convective clouds (white patches)and some precipitation (red contours) can be identified in the Southern Mediterranean and Lybia (Fig. **??**a). Convective activity is confirmed by the occurrence of lightning in this region (as inferred from www.lightningmaps.org, not shown). At this time, a large-scale trough is present over Eastern Europe (Fig. **??**a). In the Mediterranean region, the low-level air masses (850~~are used.

2.1 Cyclone phase space and cyclone tracking

The cyclone phase space (CPS; **?**), is a useful tool to diagnose the thermal structure of cyclones throughout their life-cycle. The CPS is a descriptor of the three-dimensional thermal structure of cyclones at a given timestep in terms of three parameters: lower-tropospheric horizontal thermal asymmetry (B), which measures the across-track 900-600 hPa thickness gradient, i.e. frontal nature, and thermal winds in the lower ($-V_T^L$; 900-600 hPa) ~~have very high equivalent potential temperatures whereas in Eastern Europe the values are much lower (Fig. **??**a) indicating substantial baroclinicity and moisture gradients~~

~~between these two regions.~~ and upper troposphere ($-V_T^U$; 600-300 hPa), which measure the vertical thermal structure. In this three-dimensional parameter space, cyclones can be classified as frontal ($B > 0$) or non-frontal ($B \leq 0$), cold-core ($-V_T^L < 0$ and $-V_T^U < 0$), hybrid ($-V_T^L > 0$ and $-V_T^U < 0$), or deep warm-core ($-V_T^L > 0$ and $-V_T^U > 0$). In this study, cyclones that at least once in their life-cycle fulfill the deep warm-core criterion are classified as medicanes. For simplicity, the symmetry parameter B is not considered.

~~In the evening of the same day, at 1800 UTC 26 Sep 2018, cloud formation and precipitation over the Mediterranean is enhanced (Fig. ??b). The upper-level trough has elongated into a narrow PV streamer on 325~~Cyclone tracks at 6-h temporal resolution are obtained for each of the 50 ECMWF ensemble members and the operational analysis using the cyclone detection and tracking method described by **?**. This method was specifically designed to study meso-scale cyclones in the Mediterranean Sea, including medicanes (**?**). More specifically, 6-hourly SLP fields are used to identify pressure minima after applying a Cressman filter (radius of 200 km; **?**) to smooth out noisy features and small cyclonic structures. Weak cyclones are then filtered with a SLP gradient threshold of 0.5 hPa per 100 km. Cyclone tracks are identified with the aid of the horizontal wind field at 700 hPa, which is considered the steering level for cyclone movement. For one member (member 32) the cyclone track showed an unrealistically large jump from the first to the second timestep and therefore the first timestep was removed from the track. Note that mean sea level pressure data of the ensemble forecast is only available in 6-hourly resolution until 144 ~~K that extends towards the Central Mediterranean (Fig. ??b). At the streamer's downstream side (to the east), a large area of enhanced QG forcing for ascent by the upper levels (QG ω, red contour in Fig. ??b) provides favorable conditions for the relatively strong convective precipitation off the Lybian coast local maximum of~~ h lead time. Therefore, all cyclone tracks in the ensemble initialized at 0000 UTC 24 ~~mm (~~Sep 2018 end at the latest at 0000 UTC 30 Sep 2018. The CPS is calculated every 6 ~~h)$^{-1}$ indicated by the blue contours in Fig. ??b. High wind speeds on 850~~h based on the track positions and the CPS values at each time step are smoothed using a running mean filter with a 24-h window. Due to the small size of medicanes, a radius of 150 ~~hPa over the Aegean Sea (arrows in Fig. ??b) indicate strong low-level cold air advection from Eastern Europe and~~ km is used to calculate the CPS values, consistent with previous studies (e.g. **?**) .

2.2 Normalized PV differences

To compare PV of two ensemble clusters at different lead times, it is useful to compute normalized cluster-mean differences (see e.g. **?**):

$$\Delta \mathrm{PV}_{\mathrm{AB}} = \frac{\overline{\mathrm{PV}}_{\mathrm{A}} - \overline{\mathrm{PV}}_{\mathrm{B}}}{\sigma_{\mathrm{PV}}} \qquad (1)$$

where $\sigma_{\Delta \mathrm{PV}_{AB}}$ is the standard deviation of all ensemble members and the subscripts A and B denote any two clusters of an ensemble forecast. Hence, $\Delta \mathrm{PV}_{\mathrm{AB}}$ becomes large when the cluster-mean difference of PV between A and B at a given location is much larger than the ensemble standard deviation at the same location, i.e. when the two clusters contain the members of the ensemble that are most different from each other. Large absolute differences in regions of strong gradients, particularly at the tropopause, are given less weight. Additionally, it allows different lead times to be easily compared. For example, if $\Delta \mathrm{PV}_{AB}$

increases with lead time, the cluster differences grow faster than the ensemble standard deviation, which means that the clusters become increasingly distinct from each other relative to the full ensemble.

2.3 Statistical significance

In order to be confident that the differences between two ensemble clusters are robust, a two-sided Wilcoxon rank-sum test (?) for each cluster pair and considered field was applied. With such a test, the null hypothesis is investigated, that it is equally likely that at a certain grid point the value of a randomly picked ensemble member in one cluster is larger or smaller than in a randomly picked ensemble member of the other cluster. When applying such a statistical test to a field, the false discovery rate should be controlled in order to avoid over-interpretation of the results (?). This can be done by correcting the p-values of the statistical test taking into account the number of tests. For this study we are only interested in a domain covering the North Atlantic and the Mediterranean and consider a box from 30 to 70°N and 80°E to 30°W. Therefore, a number of 4400 tests are used for the ~~Black Sea region towards Greece and the Central Mediterranean. Cold and dry air masses that are advected over a warm ocean surface can enhance sea-surface latent heat fluxes and become relevant for Medicane formation (?)~~correction. As suggested by ?, a Benjamini-Hochberg correction is used in this study and the false discovery rate is set to a rather conservative value of $\alpha_{\mathrm{fdr}} = 0.1$ for all analyses. Regions where the null hypothesis is rejected on this level of α_{fdr} can then be used to identify where and when robust differences in the clusters emerge in the ensemble forecast.

2.4 Trajectory calculations

Computing trajectories provides insight into the Lagrangian history of air parcels. In this study, the Langrangian analysis tool LAGRANTO (??) is used to identify warm conveyor belt trajectories (ascent rate larger than 600 hPa in 48 h, see e.g. ?) and to compute backward trajectories from the cyclogenesis region (see supplementary material).

3 Synoptic overview

This section provides an overview of the life-cycle of Medicane Zorbas, the large-scale synoptic situation over the Euro-Atlantic region prior to its genesis, and of the synoptic elements that accompanied the first stage of Zorbas' life-cycle during which it acquired a deep warm core.

~~A day later, at 1800~~ Medicane Zorbas formed at 1200 UTC 27 Sep 2018 ~~, the clouds have formed a spiral-like structure with a weak frontal cloud band extending from the Mediterranean over Lybia~~ close to Benghazi, moved into the Central Mediterranean Sea and then sharply turned eastward and moved over Greece into the Aegean Sea, where it finally decayed four days after its formation (Fig. ~~??c). Surface cyclogenesis has taken place off the Lybian coast close to Benghazi (closed yellow contours in Fig. ??c)and increased gradients of equivalent potential temperature at this location~~ 1a). Zorbas led to considerable damage through severe winds, torrential rainfall, major flooding and even tornadoes. The main affected region was Southern Greece, especially Crete, Peloponnese, Evia, and the region around Athens. According to the CPS (Fig. 1b), Zorbas formed as a cold-core cyclone and within 18 h acquired a deep warm core that was sustained for more than three days.

Zorbas reached its maximum intensity (992 hPa) already 12 h after cyclogenesis. Satellite images indicate the formation of an eye-like feature shortly before Zorbas reached Greece on 29 Sep, suggesting that Zorbas likely underwent tropical transition (not shown). However, this aspect of the life-cycle is not in the focus of this study. Instead, we focus on the synoptic aspects influencing cyclogenesis, the initial intensification, and the formation of a deep warm core.

The period before cyclogenesis of Zorbas was characterized by the transition from a relatively zonal jet in the Euro-Atlantic region at 0000 UTC 24 Sep 2018 (Fig. **??**c) confirm the presence of a weak surface cold front. The PV streamer has broken up into segments (**?**) . Its tip has formed a C-shaped PV cut-off (shading in Fig. **??**) . To the east of it2a) with a jet streak (J1) over the Gulf of Saint Lawrence and a moderate ridge (R) and trough (T) pattern over the eastern North Atlantic/Scandinavia to a very wavy jet at 0000 UTC 26 Sep 2018 (Fig. 2b) with an elongated PV streamer (S) over the central North Atlantic, a large region of enhanced QG ω (red contours) is present and strongly ascending air masses (black crosses, ascent rate larger than 600 hPa in 24 h) as identified from trajectory calculations using the Lagrangian analysis tool LAGRANTO (**??**) are located in the dent of the cut-off, above the cyclone centre. This is where also the precipitation maximum occurs 88 mm $(6\,\mathrm{h})^{-1}$, red contours in Fig. **??**c.The presence of rapidly ascending and strongly precipitating air masses are indicative for substantial diabatic effects cf. the concept of warm conveyor belts; **??**. They not only heat the cyclone centre and potentially form a low-level positive PV anomaly but favor the direct erosion of ridge (R) and a north-eastward directed jet streak (J2) over the western North Atlantic, and a narrow, elongated trough (T) over eastern Europe . During this transition, the ridge and the associated surface high pressure system (H) amplified and the latter moved from the eastern North Atlantic to eastern Europe. The PV streamer over the North Atlantic then interacted with Hurricane Leslie (L) resulting in substantial warm conveyor belt activity. Note that the anticyclonic wave breaking (i.e. the amplification of the ridge and the non-linear elongation of the trough over eastern Europe) preceded the North Atlantic warm conveyor belt activity, indicating that in this case the warm conveyor belt was less relevant for the ridge amplification than for typical events with the formation of strong Mediterranean cyclones (**?**). Subsequently, the narrow trough over eastern Europe further elongated into a PV cut-off due to diabatic PV modification and entrainment (**?**). Furthermore, they transport low-PV air to the 325 K level, which causes the dent structure in the upper-level PV and further affects the evolution of the cut-off. The ascent seems to be strongly convective, as confirmed by the high lightning activity in this region (as inferred from www. lightningmaps. org,not shown). At 1800 streamer reaching the Central Mediterranean Sea at 0000 UTC 28 27 Sep 2018, almost no strongly ascending air masses are present on 325 K, 2018 (see Fig. 3a).

The formation of the PV cut-off has almost completely decayed on 325 K, and streamer over the Mediterranean was followed by cyclogenesis of Zorbas at the PV streamers' south-eastern flank and its rapid intensification (see Fig. 3a-c). At the same time, the enhanced QG ω has vanished PV streamer broke up resulting in the formation of a PV cutoff, which was rapidly eroded during the intensification of the surface cyclone. Shortly before cyclogenesis, the large mean sea level pressure gradients between the surface high pressure system over eastern Europe and a surface low pressure area over the Levantine Sea resulted in strong low-level advection of air with low θ_e across the Aegean Sea (Fig. 3a,d). Air with anomalously high θ_e was present over Lybia and at cyclogenesis in immediate proximity of the cyclone center (hatched regions in Fig. 3d,e). At a first glance, it seems that the air with high θ_e was advected cyclonically into the cyclogenesis area. However, a detailed trajectory analysis (see supplementary material) shows that the low-level air in the 850-950 hPa layer and within a radius of 250 km around the cyclone centre originated from the Aegean and Black Sea and was substantially moistened by sea surface fluxes as it traveled across the Aegean Sea. This is consistent with the direction of low-level winds (Fig. ??d3d,e). The surface cyclone has further intensified and moved over the Central Mediterranean . The satellite data shows a cyclone without clear frontal cloud bands and still substantial precipitation in its centre relevance of a similar process for medicane formation in the western Mediterranean has already been pointed out by ?. A day after cyclogenesis, the cyclone developed into a barotropic system with a deep warm core structure. The surface cyclone, the low-level warm core, the minimum of geopotential height at 500 hPa, and the upper-level PV maximum were vertically aligned (Fig. ??d), which is, however, not associated with strong QG ω but likely due to conditional instability. The region with enhanced gradients of equivalent potential temperature previously identified as a weak cold front has moved eastward and become disconnected from the cyclone centre. A local maximum of low-level equivalent potential temperature is present in the cyclone centre, indicating the formation of a warm seclusion (?). Warm seclusion events have been previously linked to Medicane formation (?). All this suggests that the cyclone underwent the transition from an extratropical to a subtropical-like or even tropical-like system between 1800 3c,f). The upper-level PV cutoff decayed and the remaining, small-scale PV maximum with values around 2 PVU above the cyclone centre at 1200 UTC 27 Sep and 1800 UTC 28 Sep 2018. As diagnosed from the cyclone phase space CPS; ?, for more details see Sect.2.1, a 2018 was most likely diabatically produced, as stratospheric upper-level positive PV anomaly would imply an upper-level cold core. The low-level warm core is already present at 1800 UTC 27 Sep and a maximum of θ_e was lower at this stage than at cyclogenesis and not anomalous with respect to climatological values.

The 12 hours after cyclogenesis were characterized by very intense precipitation north-west of the cyclone centre according to the IFS short-term forecasts (red contours in Figure 4a, maximum value of 101 mm in 12 h). The dense cloud patch and lightning activity (see supplementary material) in this area are indicative of strong latent heating and the presence of deep convection. According to the IFS model output, 38% of the area-averaged precipitation was produced by the convection scheme. This indicates that additionally strong large-scale ascent occurred. Such a situation with strong ascent and latent heating is expected to result in vorticity generation at lower levels, which is consistent with the cyclone path in this period. It also helps to explain the rapid diabatic erosion of the PV cutoff (see e.g. ?). In contrast, at the time when the deep warm core is reached only 12 hours later at 0600 UTC 28 Sep (see Fig 1a). Consequently, Zorbas also visually acquires tropical characteristics including the formation of an eye on 29 September 2018 as it moves over Greece (not shown). However, this later period of the cyclone evolution is not in the focus of this study. was well established, the precipitation was substantially weaker (Fig. 4b) and mostly produced by the convection scheme (67%). The cloud structure indicates that well-defined fronts were absent, i.e. Zorbas acquired a more tropical-like appearance.

In summary, Zorbas forms via extratropical cyclogenesis forced by a breaking-up PV streamer over a baroclinic zone in the Central Mediterranean. As low-level moisture and temperature over the Mediterranean are very high The evolution of Zorbas in the first day of its life-cycle agrees well with the climatological evolution found for the strongest Mediterranean cyclones around their time of maximum intensity (?). However, remarkable in this case was the rapid erosion of the upper-level PV cutoff, the anomalously high values of 900 hPa θ_e at cyclogenesis and north-easterly cold air advection at low levels occurs, cyclogenesis is accompanied by strong convection, cloud formation and intense precipitation, first at the PV streamers downstream flank and later in the cyclone centre.Zorbas then transitions into a subtropical and later even tropical-like system: It loses its frontal structures and becomes a circular cyclone with a warm seclusion, which is where most precipitation occurs. the very intense precipitation in the early stage of its life-cycle. Values of 900 hPa θ_e at cyclogenesis reached above 330 K in this case. This value provides the approximate maximum isentropic level to which these air masses can ascend by the release of latent heat. In this case, the PV cutoff was located at 325 K, and therefore diabatic ascent of the low-level air present at cyclogenesis had the potential to erode the stratospheric upper-level PV anomaly.

4 Ensemble clustering according to position of PV streamer

The In the following, a procedure is presented that allows separating the ensemble members of the 0000 UTC 24 Sep 2018 initialization into meaningful clusters based on position of the Mediterranean PV streamer at day 3 of the forecast. This is motivated by the previous section, which showed that the PV streamer is crucial for the cyclogenesis of Zorbas. At 0000 UTC 27 Sep (shown in Fig. ?? six hours earlier) 2018, the position of the PV streamer varies strongly in the ensembleforecast initialized at 0000 UTC 24 Sep 2018, with about equal shares of ensemble members where the PV streamer is roughly correct, too far west and too far east, respectively. This offers the opportunity to use this ensemble forecast to study the dynamical processes that lead to this significant forecast uncertainty in the upper troposphere, which subsequently affected Medicane formation (see Sect.6). Therefore a position of this upper tropospheric PV structure. The ensemble members are clustered based on the PV streamer position at 0000 UTC 27 Sep 2018 (shown in Fig. 3a), which is immediately before the PV streamer breaks up cyclonically and cyclogenesis occurs in the operational analysis. A pragmatic clustering method was designed specifically for this situation to separate the strongly diverging PV streamer evolutions in these ensemble membersinto clusters with a similar evolutiondiffering PV streamer positions in the ensemble members. The three identified PV streamer scenarios, i.e. clusters, are the basis for all remaining analyses. For the clustering, a box is defined around the PV streamer as identified at 0000 UTC 27 Sep 2018 in the operational analysis (Mediterranean box, 5-30°E, 30-45°N, see black box in Fig. 15). The clustering uses PV vertically averaged from between 320 to and 330 K, where all tropospheric PV values (hereafter called PV_{av}. Before the averaging, all PV values with PV<2 PVU) are set to zero , hereafter called PV_{av} to remove the contribution of the variability of tropospheric PV values. Hence, PV_{av} is high in areas where the PV streamer is strong and deep, and low where it is weak and shallow. The clustering is based on two different steps: First, all ensemble members from all 50 ensemble members the ones are identified for which the region with $PV_{av} \geq 2$ PVU in the box has more than 75% overlap with the corresponding area in the analysis. In these 19 members (cluster 1C), the streamer has a similar shape and location as in the analysis (see blue shading in Fig. 15). The remaining members are separated into two clusters depending on whether the maximum PV_{av} is shifted to the west (cluster 2W, 12 members, green shading in Fig. 15) or east (cluster 3E, 18 members, green red shading in Fig. 15) relative to the analysis. There is one ensemble member that can not cannot be attributed to one of the three clusters because its overlap is less than 75% and but the maximum of PV_{av} is located at the same longitude as in the analysis. The histogram of the longitude where the maximum of PV_{av} occurs between 36-73°N (inset in Fig. 5) shows three clearly distinct peaks, one for each cluster, supporting the simple clustering approach. There are a few borderline members but they don't affect the main results of this study.

The meaningfulness of this clustering for ~~understanding predictability is~~ studying the predictability of this case is further supported by the fact that it helps explaining the temporal development of the ~~ACC~~ anomaly correlation coefficient (ACC, for details see supplementary material) in the Mediterranean box. To this aim, Fig. ~~??~~6a summarizes the synoptic sequence of the case, i.e. the formation of the PV streamer over the Mediterranean, the break-up resulting in a PV ~~cut-off~~ cutoff (grey boxes), the time evolution of the minimum sea-level pressure of Zorbas (solid line), and its thermal structure as diagnosed from the CPS (colors~~, for details see Sect.2.1~~). As shown in Fig. ~~??~~6b, the ACC of geopotential height at 500 hPa in the Mediterranean starts to ~~drop (median and many members in clusters 2 and 3)~~decrease in the majority of the ensemble members, when the PV streamer reaches the Mediterranean on 26 Sep 2018 and cyclogenesis occurs, while it remains high (close to 1) until 29 Sep 2018 for most members of cluster ~~1~~ C (blue lines in Fig. ~~??~~6b). After the ~~drop~~ decrease from 1 to around 0.8 the median ACC remains fairly constant until 29 Sep 2018. In comparison, for the ensemble forecast initialized at 0000 UTC 27 Sep 2018, i.e. at the time when the PV streamer has developed, the ACC remains high in all members during the intensification and deepest phase of Zorbas, ~~dropping~~ decreasing only after 29 Sep 2018 (Fig. ~~??~~6c), likely due to errors associated with a second PV streamer reaching the Mediterranean in the northern part of the box (not shown). ~~We conclude~~ It can be concluded that errors in the position ~~and shape~~ of the PV streamer limited the large-scale predictability as measured by the ACC of geopotential height on 500 hPa in the Mediterranean, and that ~~the clustering incorporates the relevant characteristics of the PV streamer well. However, the ACC of geopotential height on 500 hPa does not fully account for errors in the vertical structure of the cyclone and its intensity and exact position, aspects that are potentially relevant for predicting the cyclone's impact. Hence, even if the large-scale predictability related to the PV streamer is high, still there can be relevant errors in the details of the cyclone evolution and its interaction with the upper levels, which may severly limit meso-scale predictability~~cluster C contains the members with the most accurate forecasts.

5 PV streamer scenarios emerge from ~~initial condition uncertainties and baroclinic~~ uncertainty amplification in a North Atlantic jet streak

We now investigate how the diverging PV streamer scenarios identified in ~~Sect.4 emerge~~ section 4 evolved from differences in ~~initial conditions that amplify baroclinically at the left exit of~~ the initial conditions and a jet streak over the North Atlantic. To this end, ~~we analyze the differences of the means of clusters 2 and 3~~ the normalized PV differences (see section 2.2) between clusters E and W (ΔPV_{EW} are analyzed, as these are the clusters that deviate the most in terms of the PV streamer ~~evolution.We investigate the upper-level development using differences in PV and winds on 325 K (Fig. ??a-d), and the baroclinic interaction with the lower levels using QG ω and geopotential height on 850 hPa (Fig. ??e-h). For upper-level PV and winds, we use~~

normalized cluster-mean differences (?):

$$\Delta PV = \frac{\overline{PV_3} - \overline{PV_2}}{\sigma_{PV}}$$

where σ_{PV} is the standard deviation of all ensemble members. Hence, ΔPV becomes large when the cluster-mean difference of PV at a given location is much larger than the ensemble standard deviation at the same location, i.e. when the two clusters contain the members of the ensemble that are most different from each other. Large absolute differences in regions of strong gradients, particularly at the tropopause, are given less weight. Additionally, it allows us to easily compare different lead times. For example, if ΔPV increases with lead time, the cluster differences grow faster than the ensemble standard deviation, which means that the clusters become increasingly distinct from each other, relative to the full ensemble. Further, in order to make statistically robust statements, regions where cluster-mean PV values significantly differ are identified using a two-sided Wilcoxon rank-sum test (see supplementary material). By design, significant PV differences tend to be co-located with high values of ΔPV. position.

Already at 0000 UTC 24 Sep 2018, i.e. the initialization timeof the forecastAt forecast initialization time, a relatively large area of significantly positive positive normalized PV differences is discernible on the stratospheric side of a the jet streak over the Gulf of Saint Lawrence (teal contour Fig. 7, jet streak is marked as J1 in Fig. ??2a). The normalized PV differences in this area are between 0.5 and 1.5 standard deviations (red shading in Fig. ??). With and not statistically significant. However, it seems that with increasing lead time this PV difference in the initial conditions moves eastward along the 2-PVU contour, amplifies over the North Atlantic and produces results in a dipole of significant negative and positive PV differences further downstream in the trough over eastern Europe (T) that ultimately result in the differences in the shifted PV streamer formation over the Mediterranean (Figs. ??Fig. 8a-d). In the following, this development is discussed in more detail.

The initial PV difference as shown in Fig. ?? moves eastward and after 6 h is located northeast of Newfoundland, with an amplitude still between 0.5 to 1.5 standard deviations (Fig. ??a). After The evolution of the normalized PV differences shows similarities with the evolution of the PV error in ?, but here it occurs on a much shorter time scale. The first significant PV difference appears after 18 h (Fig. ??b) , the PV difference is largest 8b) over the central North Atlantic, has amplified, reaching values above 1.5 standard deviations, and covers a larger area. Also, the mean 2-PVU contours of clusters 2 (dashed) and 3 (solid) start to separate at this location and a . A clear cyclonic difference wind field (arrows) is present. Additionally, downstream of the positive PV difference , a negative Subsequently, during the non-linear development of the Atlantic ridge and the trough over eastern Europe, the PV difference propagates downstream in a wave-like manner giving rise to alternating positive and negative PV differences resulting in a more progressed anticyclonic Rossby wave breaking in cluster W compared to cluster E (Fig. 8a-d). Ultimately, this results in a zonally shifted tip of the narrow trough, and later, the PV difference north of the British Isles emerges. In the following 24 h (Figs. ??c,d)both PV differences propagate along the wave guide and the maximum amplitude propagates from the positive to the negative PV difference. Another positive PV difference occurs further downstream (red patch at 43N and between 20E and 30E in streamer (consistent with Fig. ??d5). This gradual downstream development of the original PV difference can be explained by non-linear (barotropic) Rossby wave dynamics : The cyclonic difference wind field leads to stronger northward advection of tropospheric low-PV air in cluster 3 compared to

~~cluster 2 downstream of the positive PV difference, leading to a negative PV difference there. This negative PV difference then propagates downstream and induces a positive PV difference by the same mechanism. At~~ (see **?**). However, the initial rapid amplification of the positive normalized PV difference at 1800 UTC ~~25~~ 24 Sep is not straightforward to explain. A contribution of warm conveyor belts can be excluded, as no intersection points of warm conveyor belt trajectories with the 325 K isentropic level (light green crosses in Fig. 8a-d) occur in the vicinity of the investigated PV difference.

The synoptic environment of the amplification of this positive normalized PV difference is characterized by a jet streak (Fig. 8e-h). At 0600 UTC 24 Sep 2018 the jet streak is further upstream and much weaker, but at 1800 UTC 24 Sep 2018, when the rapid amplification takes place, the positive PV difference is located on the stratospheric side of the jet streak. Subsequently, the positive PV difference propagates faster downstream than the jet streak maximum. The positive PV differences are associated with a slight shift in the jet streak between clusters E and W at 1800 UTC 24 Sep 2018 and thereafter (solid and dashed white contours). There is some precipitation in the region of the positive PV difference at 1800 UTC 24 Sep 2018, but the values are low and the ~~2-PVU contours are clearly separated, showing a westward phase shift and larger amplitude of the Rossby wave in cluster 2 (dashed contour) compared to cluster 3 (solid contour). Hence, this amplification and downstream development of the PV differences result in the diverging formation~~ precipitation area small compared to the later time steps. Furthermore, the positive PV difference is located in a region that is clearly stratospheric. Hence, a contribution of direct diabatic PV modification to the amplification is very unlikely and it seems to be the upper-tropospheric dynamics associated to the jet streak that drives the amplification at 1800 UTC 24 Sep 2018.

Jet streaks are regions with strong ageostrophic winds. In order to understand the effects of the geostrophic and ageostrophic wind components to the amplification of the positive PV difference we qualitatively apply the equation for the local tendency of the PV ~~streamer, which forms earlier and more to the west in cluster 2 and later and more to the east in cluster 3 (consistent with Fig. 1). Figures **??**e-h show how the diverging~~ error following **?** and **?**:

$$\frac{\partial \mathrm{PV}^*}{\partial t} = -\mathbf{v}^* \cdot \nabla_\theta \mathrm{PV} - \mathbf{v}^* \cdot \nabla_\theta \mathrm{PV}^* - \mathbf{v} \cdot \nabla_\theta \mathrm{PV}^* + \mathrm{DIAB}^* + \mathrm{RES}^* \tag{2}$$

where PV^* and $\mathbf{v}^*$ denote the error PV and wind field, respectively. The first three terms of Eq. 2 represent the contributions of adiabatic upper-level ~~wave propagation interacts with the lower troposphere. Upper-level PV differences project to the surface via QG ω (shown in red contours in Figs. **??**e-h). Six hours after initialization (Fig. **??**e), cluster-averaged geopotential height (purple contours) and potential temperature (not shown) on 850 hPa are still almost identical. The regions of upper-level induced QG ω on 850 hPa (red contours) have slightly different magnitudes but are well aligned. As the upper-level PV differences intensify, QG ω becomes stronger in that region (red contours with a maximum around 58N and 35W in Fig. **??**f), which is associated to~~ dynamics to the PV error tendency and the fourth term the contribution of diabatic processes. The last term contains the residual contribution (e.g due to frictional effects, numerical diffusion etc.) and seems to be systematically reducing forecast error (**?**). As in our case a substantial direct contribution of diabatic processes to the amplification of PV^* at 1800 UTC 24 Sep 2018 is very unlikely, we focus on the first three terms of Eq. 2. In our case the *-terms denote differences between clusters E and W, and the first term represents the advection of the PV field in cluster W by the difference wind and the second term the advection of the PV difference field by the difference wind. The second term is likely much smaller as the PV differences are still very small so early in the forecast. The third term is the advection of the ~~left exit of an upper-level jet streak (blue and green shading).The jet streak and the resulting forcing for ascent in cluster 2 (green shading and dashed contours) are slightly shifted to the northwest compared to cluster 3 (blue shading and solid contours). These differences are partially significant (see supplementary material). In the region of the forcing, a wave-like stucture in geopotential height (purple contours) and potential temperature (not shown)occurs, indicating the initiation of a baroclinic wave. QG ω and the baroclinic wave propagate further in both clusters,but in cluster 2 the center of action is more to the west (Figs. **??**g,h). At 0600~~ PV difference by the wind in cluster W, which advects individual difference features along the flow in cluster W. Note that strictly speaking, Eq. 2 only applies to the PV difference between forecast and analysis (the PV error) or between two forecasts. Here, we are looking at the PV difference of the means of two ensemble clusters, but material PV conservation does not hold necessarily for average fields. Therefore this equation has to be applied with caution and is used here only for qualitative argumentation. However, it helps to explain the relevant processes responsible for the amplification of the PV difference at 1800 UTC ~~25~~ 24 Sep 2018.

The above argumentation implies that a substantial part of the amplification of the PV difference at 1800 UTC 24 Sep 2018 is due to the first and the third term on the r.h.s. of Eq. 2. If the difference wind is additionally separated into a geostrophic component $\mathbf{v_g}^*$ and an ageostrophic component $\mathbf{v_a}^*$ we arrive at the following approximation for the tendency of the not yet normalized PV difference for this specific time step and location:

$$\frac{\partial \mathrm{PV}^*}{\partial t} \approx -\mathbf{v_g}^* \cdot \nabla_\theta \mathrm{PV_W} - \mathbf{v_a}^* \cdot \nabla_\theta \mathrm{PV_W} - \mathbf{v_W} \cdot \nabla_\theta \mathrm{PV}^* \qquad (3)$$

where the subscript W denotes the fields in cluster W. Figure 9 shows the normalized PV differences at the beginning and the end of the 6 h time interval during which the amplification takes place. Additionally, the geostrophic and the ageostrophic difference wind fields together with the PV field and the jet streak in cluster W are shown. As discussed above, it is reasonable to assume that, for the given synoptic situation, the positive PV difference between 40°W and 50°W at 1200 UTC 24 Sep (Fig. 9a,c) is advected by the wind field in cluster W (see third term in Eq. 3) and amplified (by the first two terms in Eq. 3), resulting in the positive PV difference 6 h later with a larger amplitude and spatial extent further downstream (Fig. 9b,d). The magnitude of the amplification can be estimated as about 1 standard deviation when using the normalized differences and about 1 PVU when using absolute differences (not shown). The geostrophic difference wind field shows a cyclonic pattern around the positive PV difference (Fig. ~~**??**h), closed contours in geopotential height indicate the presence of a cyclone in both clusters, and the shifts in all fields between clusters 2 and 3 become more evident. To sum up, the uncertainty of the position of the Mediterranean PV streamer can be traced back to initial condition uncertainty on the poleward side of an~~ 9a,b), essentially leading to a retrogressive propagation of the PV difference relative to the flow in cluster W, i.e. counteracting the advection of the PV difference. This is comparable to the effect of the error wind associated with the upper-level ~~jet streak over the Gulf of Saint Lawrence. This uncertainty amplifies as it moves into the left-exit region of a jet streak over a strongly baroclinic zone 18 h after the forecast is initialised. Via the strong QG forcing for ascent in this cyclogenetic region~~ PV anomalies described in **?**. However, the cyclonic geostrophic difference wind field is not fully symmetric and results in larger PV advection on the upstream side of the PV difference than downstream and, hence, contributes to the ~~upper-level~~

uncertainties couple with the lower level, grow, and give rise to an uncertain low-level wave in a baroclinic zone. Hence, most likely, the upper-level uncertainties amplify due to the exponential growth inherent to baroclinic instability. However, after this about 12-18amplification of the positive PV difference. Larger in this case is the contribution of the ageostrophic wind (see Fig. 9c,d). At 1200 UTC 24 Sep large ageostrophic difference winds almost perpendicular to PV isolines are present just downstream of the positive PV difference, near the left exit of the emerging jet streak are . Six hours later, ageostrophic difference winds with magnitudes up to 5 h lasting period of strong vertical coupling and co-location of upper- and lower-level uncertainties, $\mathrm{m\,s^{-1}}$ coincide with the region where the positive PV difference is largest, but also with the jet streak and large PV gradients. Also further east large ageostrophic difference winds are present within the emerging negative PV difference. The effect of the ageostrophic wind is contained in the amplified upper-level uncertainties propagate via non-linear Rossby wave dispersion faster downstream than the low-level wave. They reach the Mediterranean region while the low-pressure system resulting from this interaction becomes stationary east of Icelanddivergent wind in **?**. However, they attributed it mainly to the divergent outflow above a region with intense latent heat release, whereas in this case the divergent wind is strongly related to the ageostrophic circulation in the jet streak.

It can be concluded that the initial amplification of the positive PV difference at 1800 UTC 24 Sep 2018, especially the peak values, can be well explained by differences in the ageostrophic circulation in combination with the strong PV gradients associated with the emerging jet streak.

6 Diverging synoptic development impacts Medicane predictabilityEffect of PV streamer position on cyclone formation and evolution

Having elucidated the reason for the In the previous section we identified the dynamical pathway leading to the uncertainty in the position of the PV streamer , we now have a closer look at at day 3 of the ensemble forecast, and here we investigate how this uncertainty affects the subsequent cyclone development formation and precipitation patterns in the Mediterranean. In addition, we analyze how the PV streamer scenarios affect the potential precursors for the formation of a low-level warm core, which is crucial for the meso-scale dynamics of the transition of an extratropical to a Medicane-like cyclone. Finally, we diagnose the predicted number of Medicane-like systems Finally, the occurrence of medicanes in each cluster and discuss the relevant large-scale precursorsis diagnosed and links to upper-level PV and low-level θ_e discussed.

6.1 Synoptic development over the Mediterranean

To examine the diverging synoptic development we analyze for all three clustersof the three clusters, the evolution of mean cluster-mean upper-level PVas well as geopotential height on 850 hPa , sea level pressure, and surface precipitation . Regions where these fields significantly differ between two clusters are identified using a two-sided Wilcoxon rank-sum test (see supplementary material). Figure **??** shows the three is analyzed. Figure 10 shows the upper-level PV streamer scenarios from clusters 1-3 and scenarios for clusters W, C and E and how they translate into distinct low-level cyclogenesis scenarios. Before Prior to the formation of the narrow PV streamer , at 1200 UTC 25 Sep 2018, i.e. 36 hours after forecast initialisation

~~(Figs. ??~~PV streamer from the narrow trough over eastern Europe (Fig. 10a,e,i), the differences between the three clusters are ~~significant in the region between 8-15E and 48-55N~~ still small, but significant on the upstream side of the narrow trough (indicated by the teal ~~patches in Figs. ??e~~contours in Fig. 10a,i). ~~These differences are exactly the ones investigated in the previous Sect.(compare Figs ??~~This is the same region where the largest negative normalized PV differences occur (Fig. 8c,d). While the 2-PVU contour ~~of cluster 1~~ representing the trough in cluster C is very similar to the analysis (black contour), the contour is ~~clearly~~ slightly shifted to the west compared to the analysis in cluster ~~2 and to the east in cluster 3. At 1200 UTC 26 Sep 2018, after~~ W, and the trough is too narrow on the upstream side resulting in an eastward displacement of the 2-PVU contour in cluster E at this location. After the narrow PV streamer has formed, the differences between the clusters become more obvious (~~Figs. ??~~Fig. 10b,f,j). The shape and position of the PV streamer in cluster ~~1~~ C is still very close to the analysis, whereas in cluster ~~2~~ W the tip of the streamer is thinner and extends more to the west, and in cluster ~~3~~ E it is shifted to the east. In these regions, clusters ~~2 and 3~~ W and E significantly differ from cluster ~~1 (see supplementary material)~~C. This is not surprising as the clustering was specifically designed to focus on these differences. ~~After the PV cut-off formation, at~~ At 1200 UTC 27 Sep 2018, i.e. the time of cyclogenesis in the analysis, a PV cutoff has formed in all clusters (Figs. ~~??~~10c,g,k) and the differences in the scenarios over the Mediterranean are very prominent. While in cluster ~~1, the cut-off~~ C the cutoff is located south of Italy in the Central Mediterranean (as in the analysis), cluster ~~2~~ W exhibits a much weaker ~~cut-off shifted~~ cutoff further to the west over Tunisia~~and cluster 3 a stronger cut-off~~, and cluster E shows a stronger cutoff shifted to the east~~over the Eastern Mediterranean~~. In all clusters ~~surface cyclogenesis occurs~~ the developing surface cyclones are located slightly east of the ~~cut-off (closed purple contours and black circles), which is where we expect the strongest QG forcing for ascent (as visible in Fig. ??).~~ cutoff (cyclone centres of individual ensemble members are shown in black dots). Hence, in cluster ~~1 the cyclone forms in northeastern Lybia~~ C the cyclones are located close to Benghazi (as in the analysis, indicated by the teal star), in cluster ~~2 in northwestern Lybia~~ W close to Tripoli, and in cluster ~~3 in the Eastern Mediterranean~~ C over Crete. ~~At this stage, cluster 3 exhibits the strongest, most developed surface cyclone.~~

One day later ~~, at 1200 UTC 28 Sep 2018 (Figs. ??~~(Fig. 10d,h,l), the ~~cut-off~~ cutoff in the analysis has decayed ~~already~~ into smaller patches~~due to the effects of strong latent heat release in rapidly ascending and precipitating air masses, as discussed in Sect. 3. In cluster 1 the cut-off~~. In cluster C the cutoff has clearly weakened (PV values < 3 PVU), in cluster ~~2 it~~ W it has fully decayed, and in cluster ~~3~~ E it is still very prominent and strong (PV values > 6 PVU), indicating substantial differences in latent heat release~~.~~, which is a major factor contribution to the erosion of PV cutoffs (e.g. ?). In both clusters ~~1 and 3~~ C and E, the vertical structure of the system has become more barotropic, i.e. the ~~cut-off~~ high upper-level PV and the surface cyclone are vertically aligned.~~In cluster 1 the cyclone has further intensified with the minimum cluster-mean sea-level pressure dropping from 1010 hPa to 1007 hPa (not shown)~~

The surface cyclones in cluster E show a very different behavior than in clusters W and C (see box plots in the individual panels of Fig. 10). First, cyclogenesis occurs earlier and takes place in a pre-existing low pressure area over the Levantine Sea. At 1200 UTC 26 Sep 2018, 9 out of 18 members have a cyclone identified in the Levantine Sea close to Cyprus, whereas in clusters ~~2 and 3 it has weakened from 1011~~W and C most cyclones form later in the southern part of the Central Mediterranean Sea. Second, cyclones in cluster E are on average much weaker than in clusters W and C. The pre-existing cyclones over the

Levantine Sea deepen slightly when they interact with the PV cutoff but – hPa to 1013with the exception of two cases – hPa and from 1009 hPa to 1011 hPa, respectivelyweaken again when the system becomes more barotropic. Cyclones in clusters W and C intensify more strongly and reach their highest intensities later. It can be concluded that the eastward shift of the PV streamer (cluster E) leads to a non-linear response of the evolution of the surface cyclones, whereas the westward shift (cluster W) just results in slightly weaker surface cyclones with a similar cyclogenesis and intensity evolution than in cluster C.

The shifts in the PV streamer position and resulting differences in cyclone evolution also result in different precipitation scenarios. The accumulated cluster-mean precipitation during the period when most members exhibit a cyclone track (three-day period after the formation of the PV streamer (between 1800 UTC 26 Sep –2018 and 0000 UTC 30 Sep 2018) clearly differs between the three clusters (shading in Fig. ??11a-c; differences are statistically significant, see supplementary material). In all clusters the precipitation mainly occurs in the immediate surrounding of the cyclone tracks (red lines), with maximum values in the Central Mediterranean Sea near 20°W in cluster C and about 500 km further southwest and northeast in clusters 2 and 3W and E, respectively. This, of course, is consistent with the positions of the PV streamer and cut-off. Cluster 3 Cluster E is associated with the smallest precipitation area, whereas the areas of clusters 1 and 2 C and W are of similar size. The precipitation pattern in cluster 1 C matches best with the short-term forecastforecasts (Fig. 11d), but the precipitation amounts are smaller, also due to the smoothing effect from averaging. The cyclone tracks in cluster 1 also fit best the observed track in the first half of the life cycle. However,there is no cyclone track in . This shows that the precipitation amounts derived from the short-term forecasts are substantially higher than in most ensemble members. In fact, the ensemble that closely follows the observed one in the second part of the life cycle, when the cyclone moves over Greeceand leads to substantial precipitation over the Peleponnes, the Athens region, and Eviavalue of the area-averaged accumulated precipitation in the box from 30 to 40°N and 5 to 30°E derived from the short-term forecasts is close to the 90^{th} percentile of the ensemble forecast. The variability among members within each cluster is large, as seen from accumulated precipitation for the members with the highest (Fig. 11 e,f,g) and the lowest (Fig. 11 h,i,j) area-averaged precipitation in each cluster. Member 28 (Fig. 11f) shows the best agreement with the precipitation derived from short-term forecasts and also the cyclone track shows good agreement with the track in the analyses. Note that in all ensemble members the tracks stop at 0000 UTC 30 Sep at the latest due to limited data availability (see section 2.1), which explains why no ensemble member follows the track in the analyses after reaching Greece. The members with lowest precipitation in clusters W and C have very short cyclone tracks and substantially less precipitation than the members with highest precipitation, indicating large variability within these clusters. In cluster E, there is less variability, as the precipitation maximum for both ensemble members shown is located over eastern Greece and of similar amplitude.

We conclude This section has shown that the uncertainties in the PV streamer's zonal position over the Mediterranean , as a consequence of the processes discussed in Sect.5, directly lead to directly resulted in uncertainties in the location of cyclogenesis and the amount and location of precipitation. For cluster 1, where The eastward shift of the PV streamer location most closely matches with the analysis, the predicted cyclogenesis , cyclone tracks, PV cut-off evolution, and precipitation patterns also compare most favourable with the operational analysis and short-term forecasts. The uncertainty led to a different type and timing of cyclogenesis and substantially weaker cyclone intensities, whereas the westward shift mainly shifts the location of cyclogenesis and peak precipitation.

6.2 Cyclone thermal structure and link to upper-level PV and low-level equivalent potential temperature

In this section, the occurrence of medicanes in each ensemble member is diagnosed and connected to differences in the evolution of ~~the PV cut-off likely directly affects the development of the cyclone and might be a crucial factor determining if a Medicane finally forms or not: If the cut-off is weak and decays early (as in cluster 2), the cyclonic circulation, destabilization, and quasi-geostrophic forcing for ascent are also weak and might not be sufficient to produce a strong cyclone~~ upper-level PV and low-level θ_e. Medicanes are identified as cyclones with a deep warm core ~~. We will elaborate further on this hypothesis in Sect.2.1.~~

6.3 ~~Low-level airstreams relevant for Medicane formation~~

~~To investigate potential precursors of a low-level warm core and subsequent development of a Medicane-like system, we now focus on two different low-level air masses in the cyclone around cyclogenesis time: The air directly located in~~ in the CPS (see section 2.1). Subsequently, for each cyclone and timestep of its life-cycle θ_e between 850 and 950 hPa is averaged within a 250 km radius (as shown for the analysis in Fig. 3e), hereafter referred to as LLTHE. The motivation to look at this quantity comes from the analysis in section 3, which showed anomalously high θ_e values at 900 hPa in the cyclogenesis region immediately before strong convection and large scale ascent occurred, the cyclone rapidly intensified and the ~~cyclone centre and the air with the highest equivalent potential temperature in the warm sector of the cyclone, which could potentially be advected into the cyclone centre to form a warm seclusion. Note that we do not identify low-level warm cores directly and do not investigate their formation in detail. However, we hint on two airstreams that are potential precursors of a low-level warm core and are dominated by the large-scale situation in each cluster. We highlight the important differences between the clusters and use them to hypothesize about the relevance of the large-scale situation for~~ upper-level PV cutoff eroded. High θ_e values at low levels indicate the potential for strong cross-isentropic upward transport associated with latent heating, which is relevant for the erosion of the ~~probability of transition into Medicane-like systems in each cluster. This probability is then quantified in the subsequent section. For each ensemble member, 48-hour backward trajectories are calculated using the Lagrangian analysis tool LAGRANTO (**??**) (i) from all grid points on 850~~PV cutoff. In addition, the 50% grid points with the highest PV values on 325 ~~hPa~~ K within a radius of ~~200~~750 km around the cyclone centre ~~, and (ii) from the 15% of grid points with the highest equivalent potential temperature within 500 km from the cyclone centre. With the latter, we aim to select the warmest and moistest air parcels in the warm sector of~~ are averaged to account for the positive upper-level PV anomaly, subsequently referred to as ULPV. The larger radius is chosen because the highest values of upper-level PV are usually expected west of the cyclogenesis area. Each cyclone can now be positioned in a LLTHE-ULPV diagram. Figures 12a,b show the geographical location of cyclogenesis in each ensemble member and the position in the LLTHE-ULPV diagram, respectively. The markers provide information about the vertical thermal structure of the cyclones, i.e. all cyclones that once fulfill the medicane criterion have a white centres and the ~~developing cyclone~~ marker size is proportional to the maximum intensity of the upper-level warm core. ~~The backward trajectories are started at 1200 UTC 27 Sep 2018,which is at cyclogenesis in the operational analysis and at or shortly after cyclogenesis in most ensemble members and , hence, before cyclone dynamics starts to dominate air mass~~

advection.It is also before a warm core forms in the operational analysis and some ensemble members. Therefore, this timing allows identifying the large-scale airstreams prior to cyclogenesis and discussing their potential role as “predictors” for the development of a Medicane-like cyclone.

Figure 2 shows, for each ensemble member and separate for the three clusters, the 48-hour backward trajectories averaged for all selected grid points (coloured lines) . For each cluster, shading indicates the distribution of the air parcels at 1200 UTC 27 Sep 2018 (red) and 48Cluster C (blue markers) produces most medicanes (see also Table 1) and the strongest upper-level warm cores. Cyclogenesis occurs mostly at the Lybian coast around Benghazi. In the LLTHE-ULPV diagram, the medicanes in cluster C are characterized by high LLTHE values up to 332 K and ULPV values between 1.5 and 3 h before (purple). The air parcels in the cyclone centre (Fig. 2a-c) originate mainly from the western part of the Black Sea and Eastern Europe in all three clusters (see purple shading and coloured diamonds). On the contrary, the air parcels with the highest equivalent potential temperature in the warm sector of the cyclone (Fig. 2d-f) originate from PVU. One medicane (with a weak upper-level warm core) and three other cyclones in this cluster exhibit a late cyclogenesis (blue diamonds) over the Central Mediterranean in clusters 1 and 2 but from the Black Seain cluster 3. Sea. The non-medicane cyclones are characterized by very low LLTHE and high ULPV values.

The evolution of equivalent potential temperature (Fig. 3a-c), potential temperature (Fig. 3d-f) and specific humidity (Fig. 3g-i) along the backward trajectories reveals additional differences between cluster 3 and clusters 1 and 2. In cluster 3, the air parcels in the cyclone centre (blue shading) have an average equivalent potential temperature more than 7 K lower than in clusters 1 and 2, even though 48 h before the equivalent potential temperature was very similar in all clusters. To a large part, this can be attributed to a stronger increase in specific humidity (Fig. 3g-i) especially in the period between -36 h and -12 h. The air parcels are, in contrast to cluster 3, transported over the Mediterranean (Fig. 2a-c)where they most likely moisten due to ocean evaporation. In the same period, potential temperature also increases more strongly in clusters 1 and 2, likely due to surface sensible heat fluxes. The rapid increase in potential temperature and drop in specific humidity in clusters 1 and 2 during the last 6 h clearly indicates strong latent heating due to cloud formation. This signal is almost entirely missing in Half of the cluster 3. The air parcels selected in the warm sector of the cyclone have very different properties in clusters 1 and 2 compared to cluster 3 already at their origin, when they are much moister and warmer (green shading in Fig. 3) W cyclones (green markers) are medicanes that have a substantially weaker upper-level warm core than medicanes in cluster C. Cyclogenesis is generally shifted to the west, but all medicanes in this cluster occur close to medicanes in cluster C, showing that there is a preferred region for the formation of medicanes in this case. In the LLTHE-ULPV diagram, the medicanes in cluster W are placed similarly to the medicanes in cluster C but with a tendency to lower ULPV values. The non-medicanes in this cluster have either low LLTHE or low ULPV values, with the exception of two members, one of them is member 12 (see below). However, they experience only a slight warming and moistening as they are transported into the warm sector, whereas in

Most of the cyclones in cluster E form close to Crete and earlier than in the other clusters. There are only two medicanes in this cluster and the stronger one forms close the medicanes in cluster 3 they are significantly moistened and heated. Nevertheless, the air parcels in cluster 3 are eventually still cooler and drier. In cluster 3, the temporal evolution of all three variables C. In the LLTHE-ULPV diagram, most members are characterized by low LLTHE values and all members with an early cyclogenesis (squares) by very low ULPV values. The reason for this is simply that the upper-level PV anomaly is still far away at cyclogenesis in these members. Cyclones with later cyclogenesis have higher ULPV values.

Two days after cyclogenesis, the 34 cyclones that are still existing are mostly located in the northern part of the Central Mediterranean Sea (clusters W and C) or the southern part of the Aegean Sea (cluster E) and the positions of the medicanes vary strongly (Fig. ~~3c,f,i)is very similar for the air parcels selected in the cyclone centre (blue) and the ones in the warm sector (green), with the main difference that the latter are warmer and moister. In clusters 1 and 2, however,~~ 10c). In the LLTHE-ULPV diagram (Fig. 10d) the ~~air parcels are part of airstreams with clearly distinct properties. We conclude that in the ensemble members of clusters 1 and~~ cyclones cluster into two distinct groups. All medicanes (with one exception), including the one in the operational analysis, are positioned in a group with low ULPV values (well below 2~~, the air which originates from the Black Sea / Eastern Europe region and later constitutes the cyclone centre substantially moistens as it is transported rapidly over the Central Mediterranean (wind speeds over the Bosporus reach above 20 m~~ PVU) and LLTHE values between 320 and 330 ~~s^{-1}~~ K. Overall, they have experienced a reduction in both, ULPV and LLTHE within the first 2 days after cyclogenesis (with very few exceptions). This is a strong indication, that high LLTHE values favored substantial latent heat release and cross-isentropic upward transport that led to the erosion of the upper-level PV anomaly, i.e. a reduction of ULPV similar to what was observed in the operational analysis ~~, see Fig. ??). Previous studies have already pointed out the importance of strong surface fluxes for Medicane formation due to dry and cold winds reaching the Mediterranean (?). Together with the destabilization and quasi-geostrophic forcing for ascent by the upper-level PV streamer/cut-off, this favours the strong latent heating of the air masses eventually constituting the cyclone centre. In cluster~~ (see section 3). This is also consistent with the cluster mean evolution shown in Fig. 10. The second group contains only members of cluster E and is characterized by low LLTHE and high ULPV (between 3 ~~, this process is lacking as the cyclone formation occurs in the northern part~~ and 6 PVU). This indicates that latent heating did not reach high enough to erode the PV cutoff, which can be connected to the low LLTHE values. These low values can be understood because cyclogenesis occurred closer to the northern coast of the Mediterranean ~~. Hence, in clusters 1 and 2, moister and warmer air is present in the cyclone centre that could favour the transition into a Medicane-like cyclone. Moreover, the PV streamer is located such that very warm and moist air from the Central Mediterranean is advected cyclonically into the warm sector, which might support the formation of a warm seclusion and transition into a Medicane-like cyclone. Again, in cluster 3 this process is lacking. Hence, from the perspective of low-level processes, clusters 1~~ and ~~2 favour the~~ as a result, the low-level air parcels were less exposed to the sea surface in order to be moistened strongly by latent heat fluxes, as it occurred in the operational analysis (see supplementary material). Hence, the erosion of the upper-level PV anomaly appears as a necessary ingredient for the formation of a medicane in this case. For most medicanes this erosion can be connected to high low-level ~~warm core, whereas this is not the casefor cluster 3. As caveat, we mention that this analysis neglects that the cyclone formation does not occur exactly at the same time in all members, and the stage of the cyclone may be slightly different in different members. Nevertheless, it provides a basic understanding of the fundamental differences in the low-level processes between the clusters.~~

6.3 ~~Explaining Medicane predictability~~

So far, the uncertain forecast of the PV streamer was shown to be directly linked to uncertainties in the position and magnitude of the PV cut-off, and the location of cyclogenesis and precipitation. Further, the eastward displacement of the PV streamer in cluster 3 leads to substantial differences in the low-level flow and, likely, air-sea interaction, compared to clusters 1 and 2. We now argue that this helps to explain major differences in θ_e at cyclogenesis.

There are a few members, where the cyclones do not follow this storyline. Some of them are now briefly discussed. In member 10, which belongs to cluster C, the vertical thermal structure of the cyclone in the three clusters and the extent to which it acquires tropical-like characteristics. To this aim, we consider the cyclone phase space (CPS; **?**) cyclone follows the pattern observed for most cyclones in cluster E. This is explained by the fact that member 10 is a borderline member of cluster C (as mentioned in section 4, visible as the easternmost blue member in the histogram in Fig. 5). In fact, the borderline member in cluster E (westernmost red member in Fig. 5) is member 43, which is a useful tool to diagnose the thermal structure of cyclones. Cyclone tracks at 6-h temporal resolution are obtained for each of the 50 ECMWF ensemble members and the operational analysis using the cyclone detection and tracking method described by **?**. This method was specifically designed to study meso-scale cyclones in the Mediterranean Sea, including Medicanes (**?**). More specifically, 6-hourly SLP fields are used to identify pressure minima after applying a Cressman filter (radius of 200 km; **?**) to smooth out noisy features and small cyclonic structures. Weak cyclones are then filtered with a SLP gradient threshold of 0.5 hPa per 100 km. Cyclone tracks are identified with the aid of the horizontal wind field at 700 hPa, which is considered the steering level for cyclone movement. Then, also a special case in Fig. 12. Another special case is member 12, which appears in a region in the LLTHE-ULPV diagram that is favorable for medicanes, but does not develop a medicane. A detailed investigation of the upper-level PV evolution indicates, that the PV cutoff is rapidly advected away from the cyclone towards the southwest immediately after cyclogenesis, indicating high vertical wind shear. Low vertical wind shear has already been identified as a synoptic condition of medicane formation (**?**). On the contrary, the CPS is calculated every 6cyclone in member 28 forms at low LLTHE values compared to other medicane members, but nevertheless the ULPV is eroded and the cyclone develops into a medicane. The upper-level warm core becomes about as strong as in the analysis and the medicane shows a similar track and precipitation pattern (member with most precipitation in cluster C, see Fig. 11f). The maximum intensity of the surface cyclone however remains 5 h based on the track positions and the CPS values at each time step are smoothed using a running mean filter with a 24-h window. CPS is a descriptor of the three-dimensional thermal structure of cyclones at a given timestep in terms of three parameters: lower-tropospheric horizontal thermal asymmetry (B), which measures the across-track 900-600 hPa thickness gradient, i. e. frontal nature, and thermal winds in the lower ($-V_T^L$; 900-600 hPa) and upper troposphere ($-V_T^U$; 600-300 hPa), which measure the vertical thermal structureabove the value in the analysis. In this three-dimensional parameter space, cyclones can be classified as frontal (B $>$ 0) or non-frontal (B $\leq$ 0), cold-core ($V_T^L < 0$; $-V_T^U < 0$), hybrid ($-V_T^L > 0$; $-V_T^U < 0$), or deep warm-core ($-V_T^L > 0$; $-V_T^U > 0$). Cyclones that at least once in their life cycle fulfill the deep warm-core (DWC) criterion are classified as Medicane-like systems. Based on the results obtained in Sections 6.1 and **??**, we expect cluster 1 to produce most Medicane-like systems as a strong upper-level PV cut-off is present even after cyclogenesis and there is supply of very moist and warm air in the lower levels. Cluster 3, on the other hand, is expected to produce less Medicane-like systems, as the conditions for the formation of a low-level warm coreare much less favorable. Finally, we expect cluster 2 to be placed between clusters 1 and 3, because conditions are favourable for the formation of special case, the reduction in ULPV is not connected to a low-level warm corebut likely the PV cut-off is too weak to maintain cyclonic circulation, destabilization, and forcing for ascent after cyclogenesis. For each cluster we identify how many members form a DWC cyclone . As shownin Table 1, this is reduction of high LLTHE values. The last special case discussed here is member 14. In this member, the case for 15 out of the 19 members (79%) in cluster 1. In cluster 3, only cyclone in general follows the pattern in cluster E but becomes a medicane, albeit with a very weak upper-level warm core. At the time of the formation of the upper-level warm core, the PV cutoff exhibits a ring-like shape, with PV values around 2out of 18 members (11%) develop a DWC cyclone, and, as expected, cluster 2 shows an intermediate scenario with 6 out of 12 members (50%) producing a DWC cyclone. Cluster 1 not only produces significantly more DWC cyclones but also shows stronger PVU above the cyclone centre and much higher PV values around it (not shown). Hence, it is possible that the PV cutoff was only substantially eroded in the centre, enough for the formation of a weak upper-level warm cores (indicated by the higher $-V_T^U$ values in Fig. **??**) and longer duration of the DWC stage (indicated by the number of DWC steps in Fig. **??**). Note that , especially for cluster 3, the number of DWC steps has to be considered with caution, due to the small sample size. Nonetheless, these results show that cluster 1 tends to produce not only more but also more robust DWC cyclones. Interestingly, the Medicane in the operational analysis has an upper-level warm core that is on the weaker side of what members in cluster 1 forecasted and is about as strong as the upper-level warm cores produced in clusters 2 and 3. But it maintains a deep warm core about twice as long as the average Medicane in all clusters. core. These special cases illustrate the variety of systems that can emerge in this specific synoptic situation, the complexity of the cyclone dynamics and the limits of the LLTHE-ULPV framework to understand medicane formation.

Overall, these findings suggest that cluster 1 provides the best synoptic environment out of the three clusters for a Medicane-like system It can be concluded that the uncertain position of the PV streamer is a major factor that determines the uncertainty in the cyclone formation and evolution in this case and that the central PV streamer position provides the most favorable synoptic environment for a medicane to form. However, the strength of the upper-level warm core in cluster 1 is overestimated compared to the analysis and the variability among the cluster members is large. This indicates that the synoptic setting in cluster 1 has the potential for much stronger Medicane-like systems than the one that actually occurred. On the contrary, storm-internal processes like convection, which is parametrized in the duration of the deep warm-core stage is strongly underestimated. We conclude that, once the PV streamer is forecasted well, sub synoptic-scale processes including the IFS model, and the detailed interaction between the surface cyclone and upper levels become limiting factors to accurately predict the cyclone evolution including its vertical structureupper and lower levels have not been discussed in this study. They may further limit predictability of the cyclone and its meso-scale characteristics, including surface winds and precipitation.

7 Conclusions

The basis of this study was an ECWMF operational ensemble forecast that showed large uncertainties in the position of a PV streamer over the Mediterranean and the subsequent development of Medicane Zorbas in September 2018. The ensemble members were clustered into three distinct scenarios according to the position and shape of a PV streamer over the Mediterranean

Table 1. Number of ~~members with a deep warm-core (DWC) cyclone characteristic~~ medicanes in each cluster (bold font).

	cluster W	cluster C	cluster E
# medicanes	~~**15**~~ **6**	**15**	**2**
out of	~~19~~ 12	19	18

~~that lead to the development of Zorbas~~of the PV streamer at day 3 of the forecast. The differences between these scenarios were used to provide new insight into, on the one hand, the ~~link between uncertainties in the large-scale flow and the meso-scale development of the Medicane, and, on the other hand, the~~ origin, amplification and propagation of forecast uncertainties in Rossby waves~~. It is clear that the uncertain position of the PV streamer was the dominant factor limiting the predictability of Zorbas for both its location andits vertical thermal structure. The first aspect, the direct influence on the cyclone formation is straightforward, as the PV streamer and cut-off provided the main forcing for cyclogenesis. Regarding the second aspect, the cyclone structure, we identified two possible~~ , and, on the other hand, the link between uncertainties in the large-scale ~~key ingredients relevant for the transition of the extratropical cyclone into a sub-tropical cyclone and, consequently, a tropical-like system, i.e. a Medicane. First, low-level advection below the eastern side of the streamer had to be such that dry and cold air masses from Eastern Europe and the Black Sea region were transported rapidly over the Mediterranean. They took up moisture by surface fluxes, experienced diabatic heating and ended up in the cyclone centre, where they helped to form a strong low-level warm core similar to the first case in **?**) . The location of cyclogenesis and hence the extent to which this process could be active was directly linked to the position of the PV streamer. Additionally, the PV streamer had to be far enough west for its induced circulation to reach the region with very warm and moist low-level air over the Central Mediterranean and advect it cyclonically around its tip into the warm sector of the cyclone. As a second ingredient, the upper-level PV cut-off had to be strong enough to maintain the cyclonic circulation and destabilize its immediate surrounding to favour deep convection even after cyclogenesis. This is reminiscent of the second Medicane case in **?** and the case discussed in **?**. In most members of cluster 1, both conditions were fulfilled, whereas in cluster 2 the second and in cluster 3 the first condition was mostly missing.~~ flow and the potential formation of a medicane.

~~The uncertainties in the position of the PV streamer after~~ The uncertainties in the position of the PV streamer after 72 h forecast lead time could be ~~clearly linked~~ clearly linked to the amplification of uncertainties in a jet streak over the North Atlantic in the first 18 h of the forecast. Uncertainties in the ageostrophic circulation combined with the presence of large isentropic PV gradients associated with the jet streak contributed strongly to this amplification. To some extent, the uncertainty could be traced back to relatively large-scale uncertainties in the initial conditions on the stratospheric side of an upper-level jet streak over the Gulf of Saint Lawrence. They ~~propagate~~ were advected along the dynamical tropopause ~~and strongly amplify in the left exit of a~~ before they amplified in the jet streak over the North Atlantic~~. At the same time, the strong QG forcing for ascent in this region enables a coupling of upper and lower levels and initiates a baroclinic wave. This wave and the associated rapid growth, as expected from baroclinic instability, are crucial for the in situ amplification of the uncertainties in this case. Non-linear tropopause dynamics then leads to the rapid downstream development of the uncertainties~~ , propagated into the

Mediterranean, eventually resulting in the uncertain position of the PV streamer and~~the~~ , as a result, the uncertain development of Zorbas. The contributions of diabatic airstreams, such as warm conveyor belts, or the direct diabatic PV modification were negligible for the uncertainty amplification in this case. The described amplification ~~process could be an important element to better understand the amplification of forecast uncertainties also in other flow situations~~due to the ageostrophic circulation in a jet streak could be also relevant in other cases for the initial uncertainty growth that can then further amplify and propagate downstream, especially in ~~the storm~~ strom track regions. Further case studies as well as more climatological analyses are needed to quantify ~~its relevance~~ the relevance of this process.

It was shown that the uncertain position of the PV streamer was a major factor limiting the predictability of Medicane Zorbas for both its location and its vertical thermal structure. An eastward shift of the PV streamer in the ensemble forecast led to a non-linear response in the surface cyclone evolution. This is particularly interesting as a comprehensive analysis of the predictability of PV streamers over the North Atlantic and the Mediterranean in the ECWMF ensemble forecasts showed that there is a tendency for eastward displacement in the forecasts, compared to the analysis (**?**). The central (i.e. correct) position of the PV streamer provided the best synoptic conditions for the formation of a strong medicane. These conditions were characterized by high values of low-level equivalent potential temperature and high upper-level PV in the cyclogenesis region. The subsequent rapid erosion of the high upper-level PV was identified as a necessary condition for medicane formation in this case. However, it has to be kept in mind that other potentially relevant factors were not analyzed, such as vertical wind shear and mid-tropospheric humidity (as identified by **?**), or the details of the convective processes. In a subsequent study we plan to investigate how these factors limit the predictability of Zorbas' life-cycle after cyclogenesis.

Since the seminal work of **?**, the growth of very small uncertainties on convective scales to large-scale uncertainties, ~~so called~~ so-called upscale error growth, has been discussed as ~~theoretical limit of~~ a process limiting atmospheric predictability (e.g. **?**). ~~The picture of the flapping wings of a butterfly influencing the development of a storm much later has become well known outside research.~~ However, recent studies suggested that the practical limits of atmospheric predictability often come from uncertainties on much larger scales, even if they are very small compared to the average kinetic energy on that scale (e.g. **?**). This study provides an illustrative example that large-scale ~~, but relative to the background kinetic energy small,~~ initial uncertainties in an upper-level jet streak over North America~~can dominate~~ , which are of small amplitude relative to the background kinetic energy, can influence the forecast uncertainty of a ~~storm~~ very intense cyclone in the Mediterranean.

~~Finally, we note that many other factors may be also relevant for the exact evolution of the Medicane even if the PV streamer is at the right location. The details of the interaction between lower and upper levels, for example, likely influence the formation the warm core, the intensification, and the track of the cyclone. This in turn determines if the cyclone remains over the sea and is able to intensify, or, if it makes landfall and decays. These aspects are subject of a follow-up study.~~ The results and approach of this study are relevant for understanding uncertainties in high-impact weather in the Mediterranean region beyond the relatively rare phenomenon of medicanes. For example, PV streamers are known to be typically connected to Alpine lee cyclogenesis and heavy precipitation on the Alpine south side (**??**), and it has been shown that the predictability of such heavy precipitation events is highly sensitive to the substructure of the PV streamer (**??**), which is well in agreement with the results of this study. Studying the sensitivity of high-impact weather to the structure and position of PV streamers gains additional relevance considering that the ECMWF ensemble prediction system is strongly underdispersive for PV streamer situations over the North Atlantic and the Mediterranean (?). Further research could therefore investigate how uncertain PV streamer positions come about and affect the predictability for other high-impact weather events.

Data availability. All data is available from the authors upon request

Author contributions. JJGA sparked this work by discovering the uncertain forecast of Zorbas. RP and JJGA designed the basic idea for the study. JJGA carried out the analyses that required the computation of the CPS and RP all remaining analyses. MS and HW helped with the data access and handling and gave important guidance and useful inputs during the whole project. RP and JJGA prepared the manuscript and all authors gave critical feedback that helped to improve the article.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the reviewers Ron McTaggart-Cowan, Florian Pantillon, and one anonymous reviewer for their detailed feedback that helped to improve the manuscript. RP acknowledges funding from the ETH research grant ETH-0716-2. JJGA has been funded through the ~~PhD-grant~~ PhD-grants BES-2014-067905~~and grant~~ , EEBB-I-18-12841 for short research stays, ~~both by~~ and by grant CGL2017-89583-R by the Spanish Ministry of Science, Innovation and University, and co-funded by the European Social ~~Fund~~and Regional Development Funds. RP thanks Emmanouil Flaounas ~~(ETH Zurich)~~ for insightful discussions about ~~Medicanes~~medicanes.

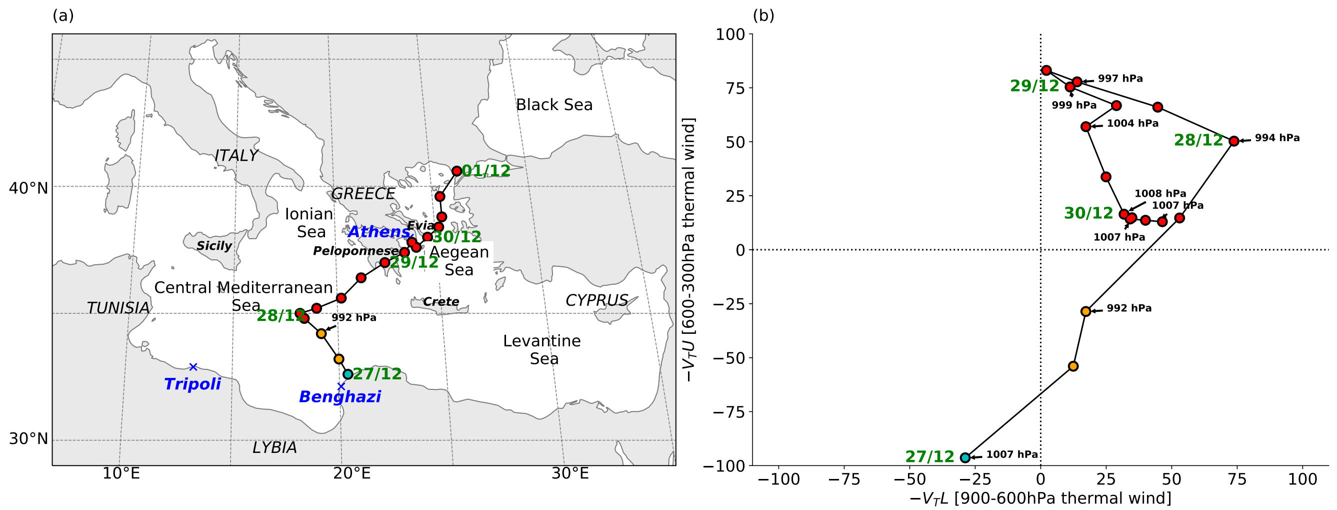

Figure 1. Infrared channel 9 (10.8 μma) Track of MSG SEVIRI provided by EUMETSAT Medicane Zorbas (grey shadingcircles and black line) and total precipitation accumulated over (b) cyclone phase space diagram derived from the previous 6 hours based on ECMWF operational short-term forecasts analyses. Cyclone positions are colored according to the quadrant in the CPS diagram (red contoursblue: cold core, 1orange: shallow warm core, 8, 15 and 21 mm red: deep warm core). Black numbers indicate the minimum sea level pressure (6hhPa) $^{-1}$) of the cyclone at this particular time of its life cycle and green numbers the day (a) 0000 UTC 26 in Sep , (b2018) 1800 UTC 26 Sep, (c) 1800 UTC 27 Sep, and time (d) 1800 hours UTC28 Sep 2018.) .

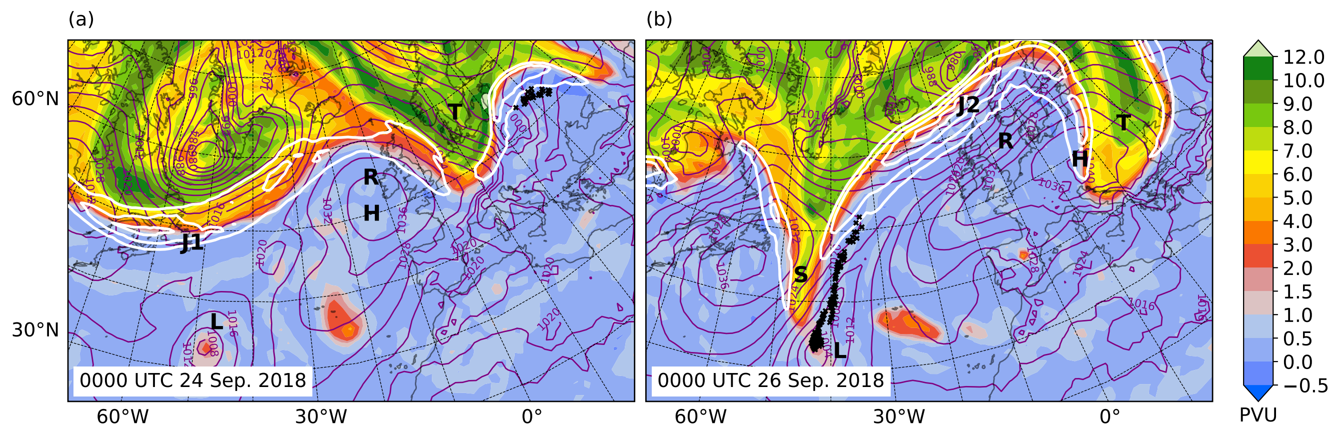

Figure 2. Synoptic situation over the Euro-Atlantic sector before the formation of the PV ~~on 325 K~~ streamer over the Mediterranean. PV (shaded, in PVU) and wind speed (white contours, in m s^{-1} on 325 K, intersection points of ~~air parcels with an~~ warm conveyor belts (ascent rate of more than 600 hPa in ~~24~~48 h) with the 325 K isentrope (black crosses), and ~~red contours show QG ω~~ sea level pressure (~~-0.5 and -1 10^{-2} Pa s^{-1}) on 850~~ purple contours, in hPa~~as forced from levels above 550 hPa~~) at (a) 0000 UTC ~~26~~ 24 Sep ~~,~~ and (b) ~~1800~~ 0000 UTC 26 Sep ~~, (c) 1800 UTC 27 Sep, and (d) 1800 UTC 28 Sep~~ 2018.

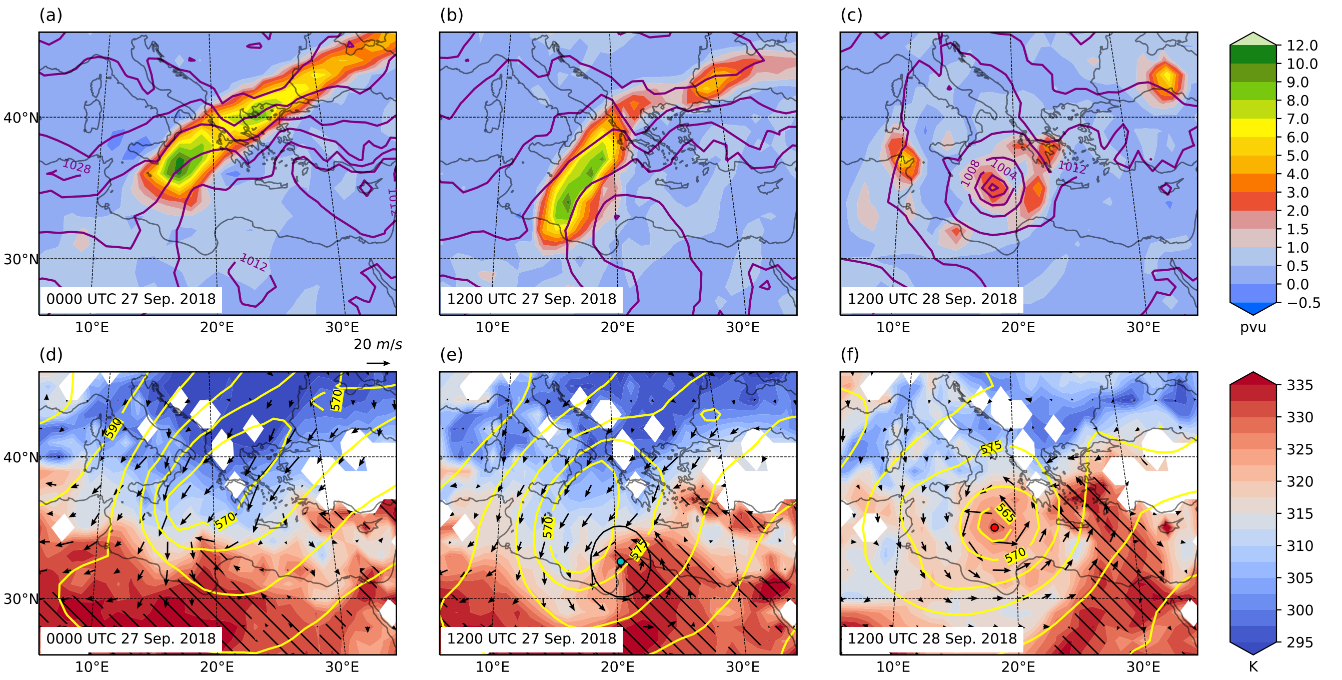

Figure 3. ~~Equivalent potential temperature~~ Synoptic situation over the Mediterranean after the formation of the PV streamer. (a-c) PV on 325 K (shaded, in ~~K~~PVU; this level corresponds approximately to the 300-350 hPa pressure levels in this region) ~~, geopotential height~~ and sea level pressure (~~yellow~~ purple contours, in ~~gpdm~~hPa) and (d-f) equivalent potential temperature (θ_e, shaded, in K) and wind vectors (black arrows) on ~~850~~900 hPa, and geopotential height on 500 hPa at (a,d) 0000 UTC ~~26~~ 27 Sep, (b,e) ~~1800~~ 1200 UTC ~~26~~ 27 Sep, (c,f) 1200 UTC 28 Sep 2018. The hatched areas in (d-f) show regions where θ_e on 900 hPa is anomalously high (at least one standard deviation larger than climatology) with respect to the Sep-Oct ERA-Interim climatology for the period 1979 – 2017. The 325 K level

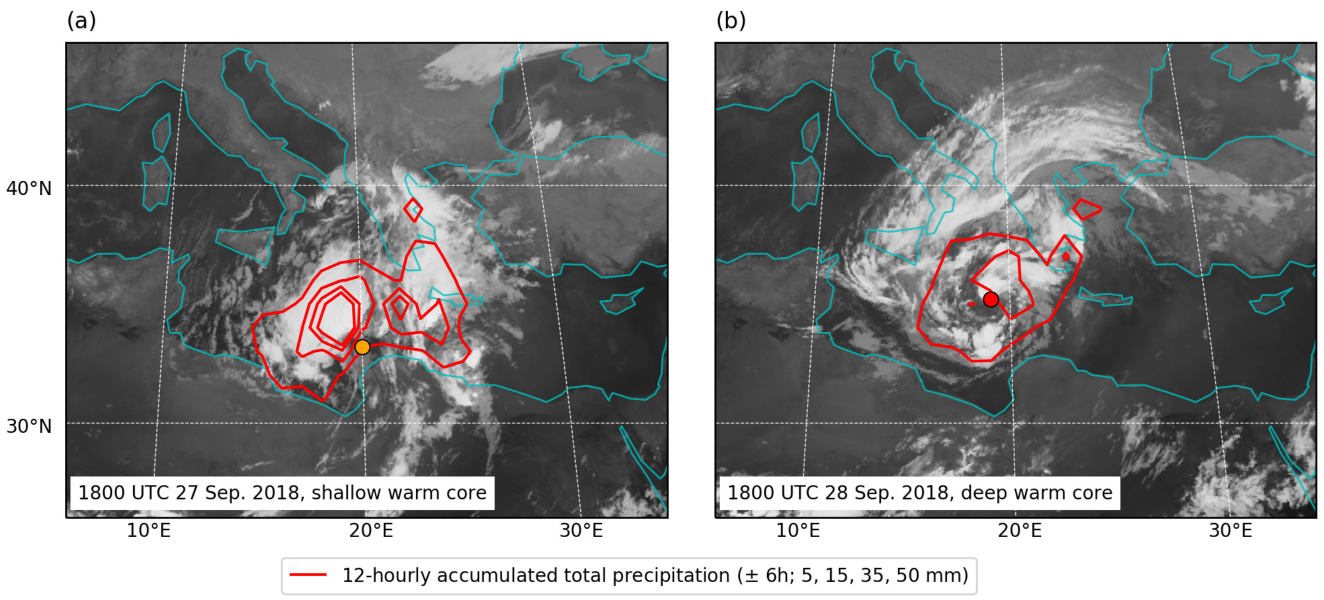

Figure 4. Infrared channel 9 (10.8 μm) of MSG SEVIRI provided by EUMETSAT (grey shading) and total precipitation accumulated over the previous and following 6 hours based on ECMWF operational short-term forecasts (red contours, 5, 20, 35 and 50 mm) at (a) 1800 UTC 27 Sep , and (db) 1800 UTC 28 Sep2018. . Cyclone positions are marked with circles and colored according to the thermal structure of the cyclone (as in Fig. 1).

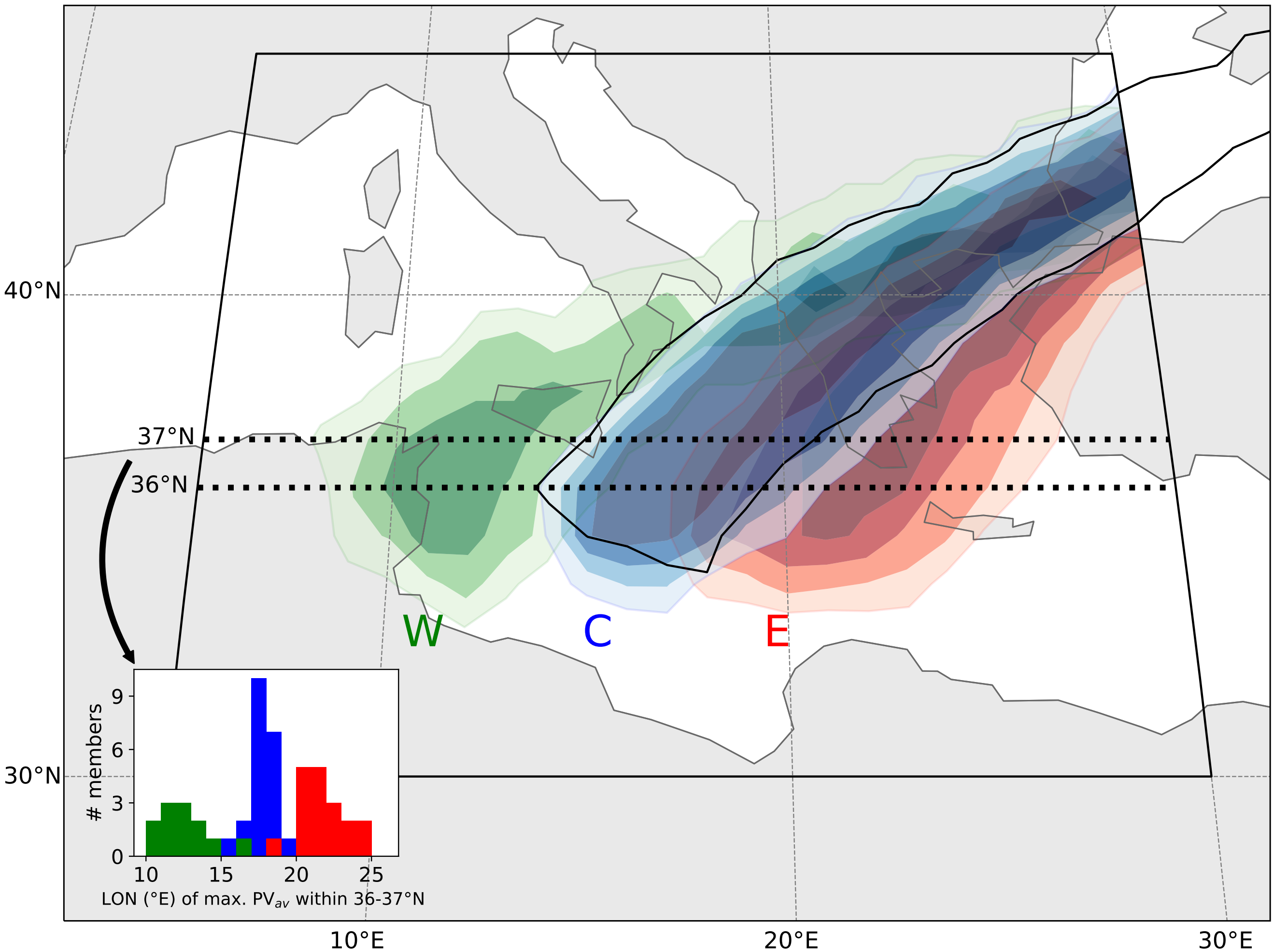

Figure 5. Clustering of ensemble members (initialized at 0000 UTC 24 Sep 2018) according to the position of the upper-level PV streamer in the Mediterranean at 0000 UTC 27 Sep 2018. Colors show frequencies of $PV_{av} \geq 2$ PVU (shading, every 20%) for each cluster (blue: cluster ~~1~~C, green: cluster ~~2~~W, red: cluster ~~3~~E) and the black line the $PV_{av} = 2$-PVU contour in the operational analysis. The region considered for the clustering is shown by the black box (see text for details).

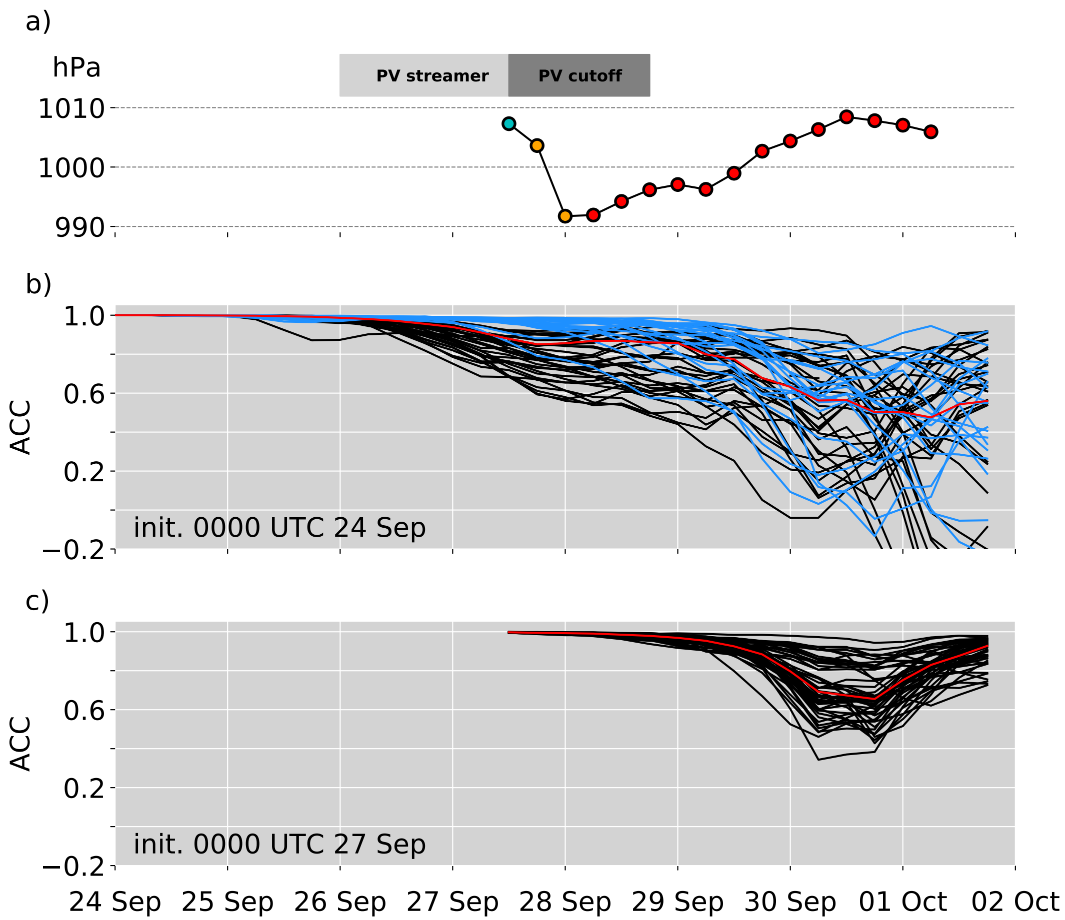

Figure 6. (a) Temporal evolution of synoptic elements discussed in this study. Grey boxes indicate times when the PV streamer or ~~cut-off~~ cutoff on 325 K is present in the analysis and the solid line shows the evolution of the minimum ~~sea-level~~ sea level pressure of the cyclone, colors indicate the cyclone stage as identified from the CPS (cold-core: blue, shallow warm-core: orange, deep warm-core: red). (b) and (c) Temporal evolution of the anomaly correlation coefficient of geopotential height at 500 hPa in the Mediterranean box (see Fig. ~~1~~5) ~~is shown~~ for each ensemble member (black lines for clusters W and E, ~~cluster 1:~~ blue lines for cluster C) and the median (red line) of the ensemble ~~forecast~~ forecasts initialized at (b) 0000 UTC 24 Sep 2018 ~~(50 members)~~ and (c) 0000 UTC 27 Sep ~~2018 (46 members).~~ 2018.

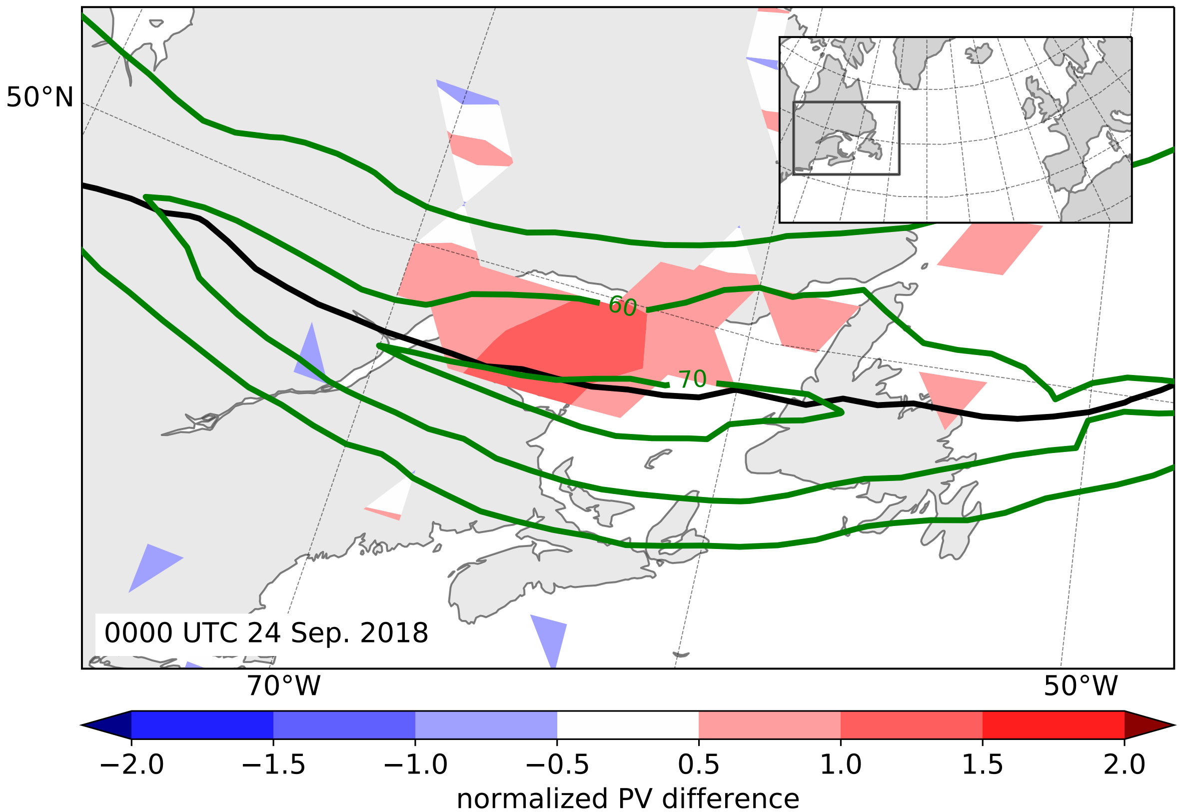

Figure 7. Normalized differences of PV ~~(shaded) between~~ on 325 K betweens cluster ~~3~~ E and ~~cluster 2~~ W (~~$\Delta PV_{3,2}$~~shaded), ~~regions where they are statistically significant (teal contour),~~ 2-PVU contour (black~~; cluster 3: solid, cluster 2: dashed~~) and ~~60 and 65 ms^{-1}~~ wind speed ~~contours~~ (~~cluster 3: blue solid, cluster 2:~~ green ~~dashed~~contours for 50, 60, and 70 m s^{-1}) on 325 K derived from the operational analysis centered over the Gulf of Saint Lawrence at 0000 UTC 24 Sep 2018, i.e. ~~initialisation~~ at the initialization time of the ensemble forecast. Normalized PV differences are not significant on the $\alpha_{\mathrm{fdr}} < 0.1$ level.

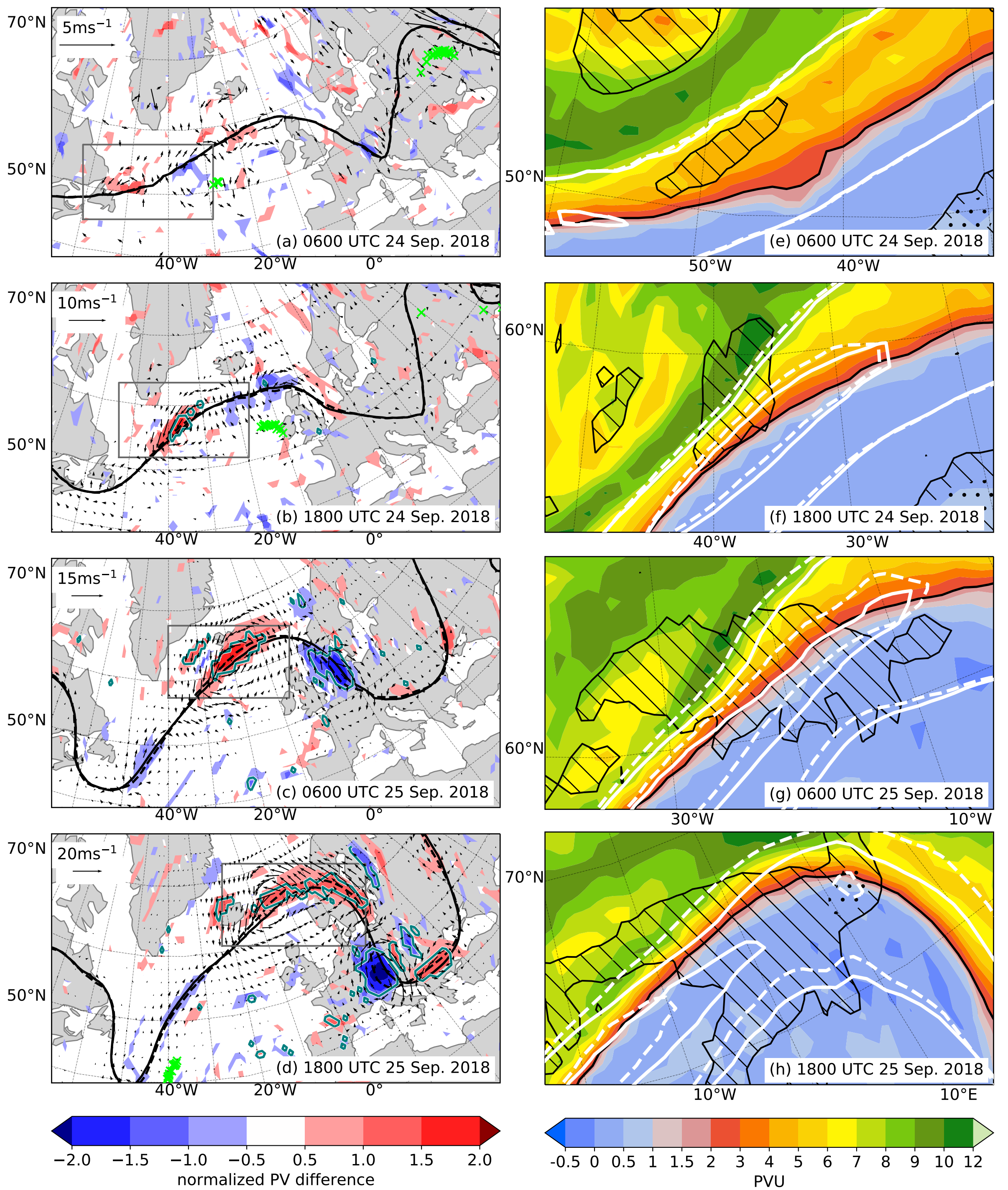

Figure 8. (a-d) Normalized ~~differences of~~ PV differences between clusters E and W (shaded)~~and~~, full difference winds (arrows, only if larger than 1 ~~standard deviation~~m s^{-1}, reference vectors is in top left of panels), 2-PVU contour (black lines) of clusters E (solid) and W (dashed), and intersection points of warm conveyor belt air parcels on 325 K ~~of cluster 3 and cluster 2~~ (green crosses) from 0600 UTC 24 Sep ~~2018~~ to 1800 UTC 25 Sep 2018, every 12 hours. Regions with statistically significant PV differences (~~p-value < 0.05~~α_{fdr} <0.1) are marked with teal contours. (e-h) ~~For~~ PV (shaded, in PVU) and 2-PVU contour (black) on 325 K and 6-hourly accumulated total precipitation (2-10 mm in hatches, >10 mm in stippling) from the operational analysis at the same ~~timesteps~~times as (a-d). Additionally, ~~high~~ average wind speeds on

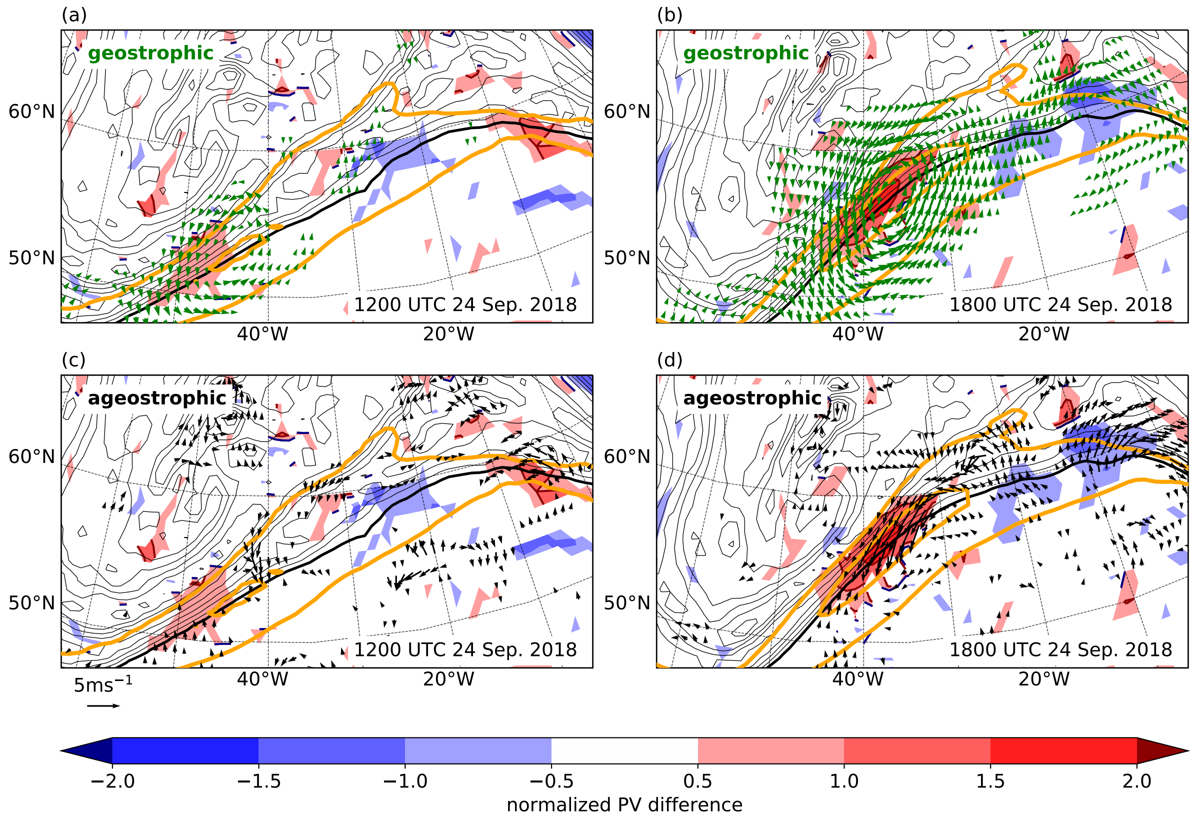

Figure 9. Normalized differences of PV on 850325 hPa forced from above 550 hPa K between clusters E and W (red contours; -0.01shaded, -0.03 as in Figs. 7 and -0.05 Pa s^{-1}8) and the contributions of (a,geopotential height b) the geostrophic (green arrows) and (c,d) the agestrophic wind (black arrows) to the difference winds on 850250 hPa (purple contours; every 10which corresponds approximately to the pressure level of the 325 gpdmK isentrope at this location) for cluster 3 at (solida,c) 1200 UTC 24 Sep 2018 and cluster 2 (dashedb,d) and 2-PVU contours 1800 UTC 25 Sep 2018. Additionally, wind speeds on 325 K (black orange contours; cluster 3: solid, cluster 2: dashed50 and 70 m s^{-1}) derived from the operational analysis are shown in each panel.

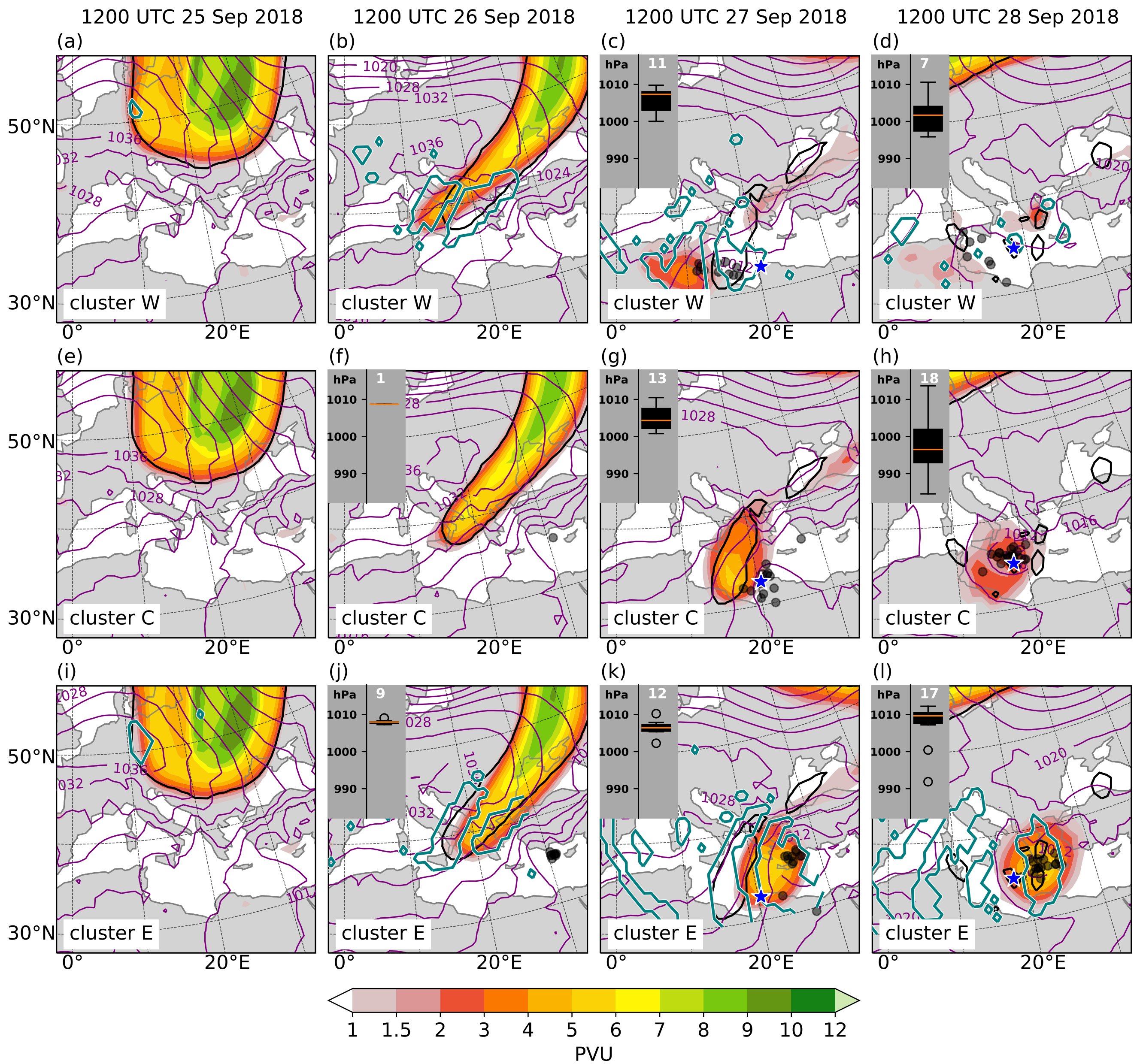

Figure 10. Cluster-mean PV on 325 K (shaded, in PVU), cluster-mean geopotential height on 850 hPa sea level pressure (purple contours, every 24 gpdmhPa), analysis 2-PVU contour on 325 K (blackcontour), cyclone positions (as identified with the method described in Sect.section 2.1) in each ensemble member (black dots), and in the operational analysis (teal blue star) for each cluster clusters W, C, and E (panels a-d, e-h, and i-l)) from 1200 UTC 25 Sep to 1200 UTC 28 Sep 2018 every 1224 h. Insets at the top left of the panels show box plots of minimum sea level pressure of the cyclones in each cluster and white numerals indicate the number of cyclones. Additionally, at 1200 UTC 25 Sep 2018for clusters 2 and 3 , regions where the differences to cluster 1 C of the PV field on 325 K are statistically significant on the α_{fdr}=0.1 level are shown for clusters W and E as teal patches contours in (ea-d, fi-l)(see supplementary material).

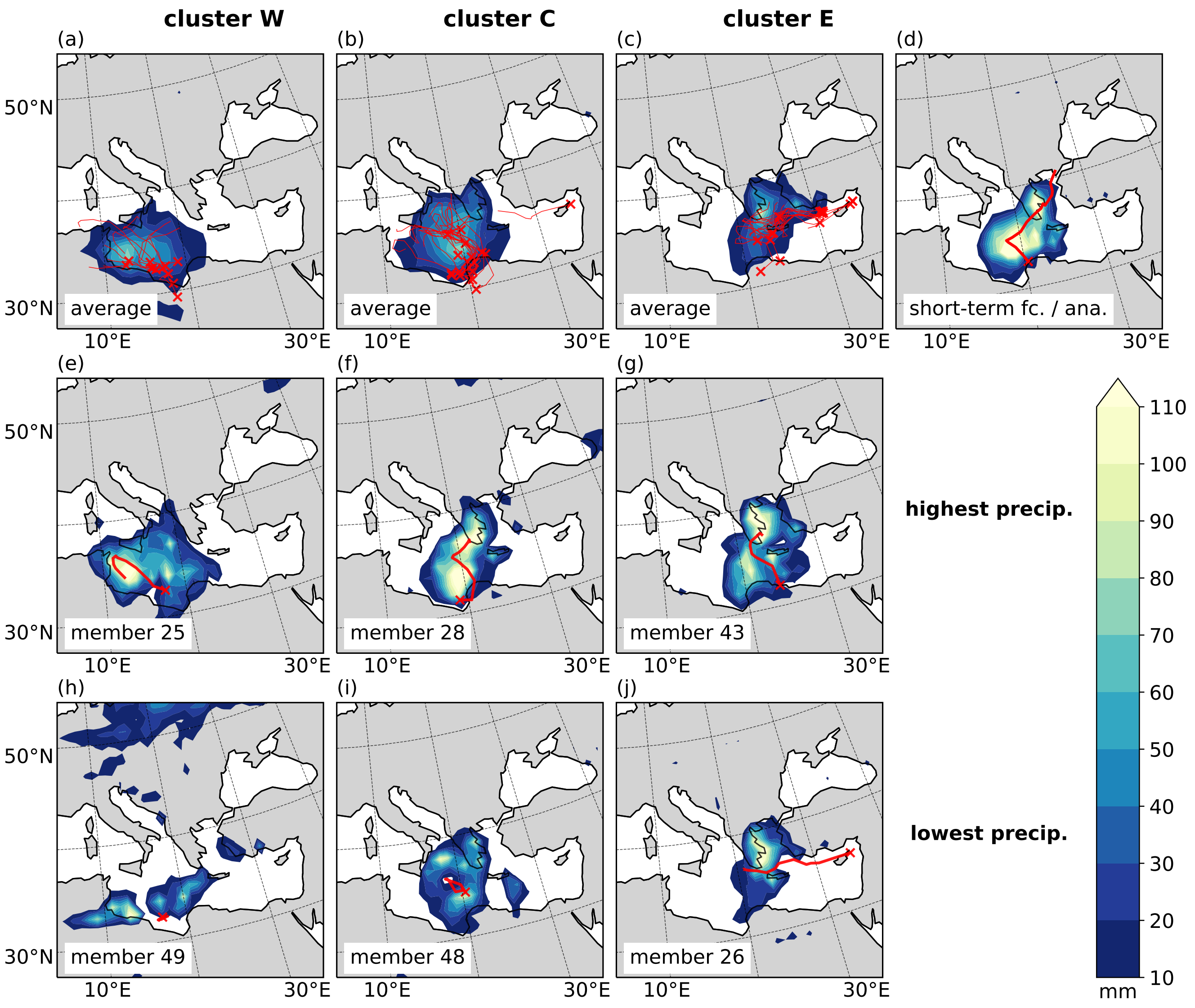

Figure 11. Accumulated precipitation (shading, in mm) from 1800 UTC 26 Sep to 0000 UTC 30 Sep 2018 for (a-c) cluster means of clusters ~~1-3~~ W, C, and E; (d) the short-term forecasts; (e-g) the members in each cluster with the highest and (h-j) the lowest area-averaged accumulated precipitation (average over 30-40°N and 5-30°E). Additionally, red crosses indicate cyclogenesis and red lines cyclone tracks for (a-c) each member ~~and~~ within the cluster, (d) the operational analysis and (e-j) the corresponding ensemble member.

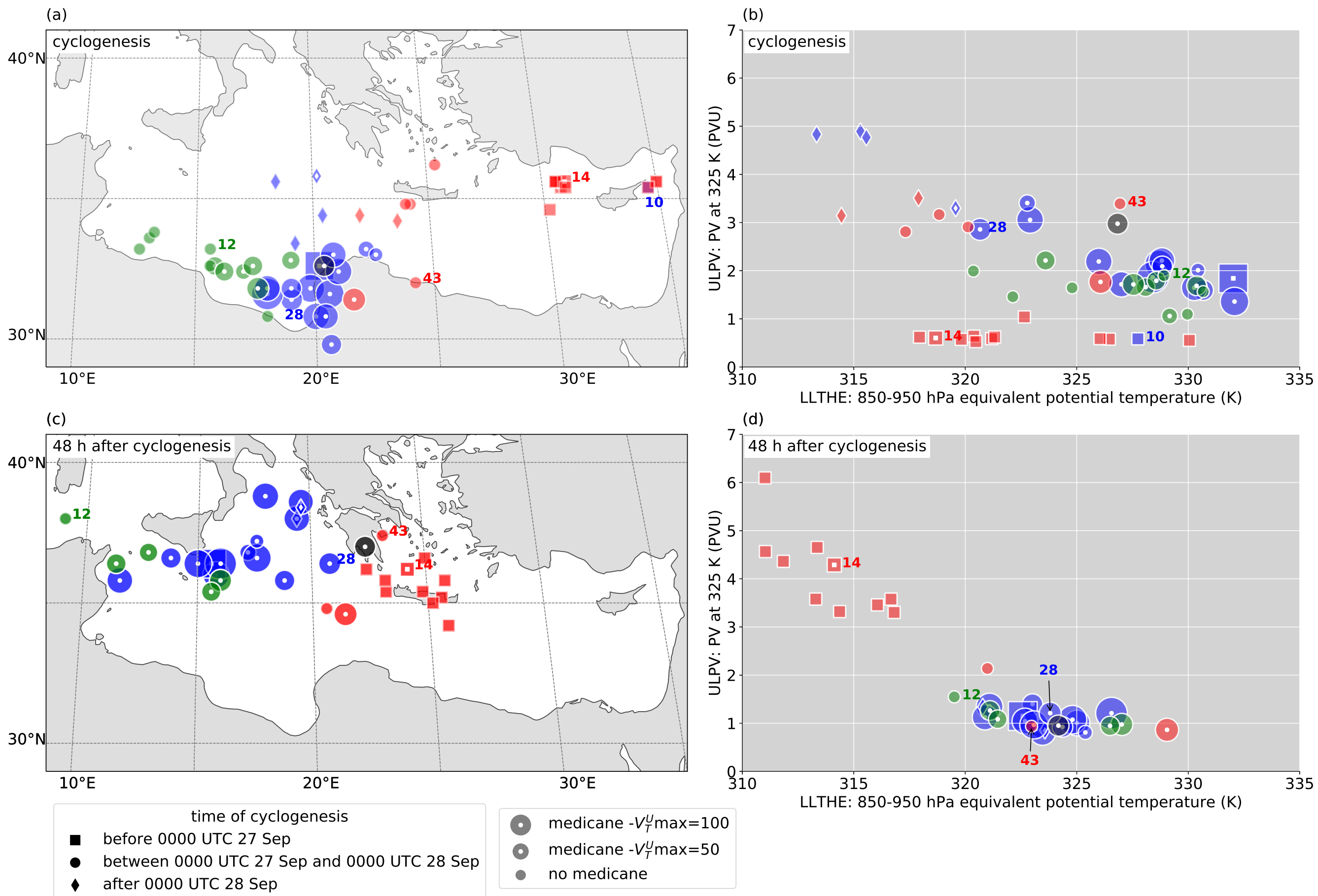

Figure 12. Position of two specific low-level airstreams in the three clusters at 1200 UTC 27 Sep 2018 (red shadinga,c) and two days earlier (purple shading). Values (in %) indicate the percentage Geographical maps of ensemble members with an airstream occurring at the specific grid point. The considered airstreams constitute the center of the cyclone on 850 hPa (a-c) positions and the region with maximum equivalent potential temperature in the cyclone's warm sector on 850 hPa (d-f)b,see text for details. For each member, the cyclone centres (trianglesd) as well as average 48-h trajectories of the airstreams LLTHE-ULPV diagram for each member cyclogenesis (linesupper panels) with their origin (diamonds) and end point 48 h after cyclogenesis (dotslower panels)are indicated in colors. Median (solid line) and Marker shape corresponds to the interquartile range (shading) time of equivalent potential temperature cyclogenesis (a-esee legend below panel c), potential temperature colors to the cluster (d-fW: green, C: blue, E: red, analysis: black), andspecific humidity , for medicanes (g-imarkers with white centre)of two specific low-level airstreams, identified as 48-h backward trajectories started at 1200 UTC 27 Sep 2018, from the center of size to the cyclone on 850 hPa (blue) and the region with maximum equivalent potential temperature in the cyclones warm sectors (green). Values indicate averages intensity of all trajectories of the considered airstreams in each member. Numbers and black dots mark the cluster-average values at 1200 UTC 27 Sep 2018. Box plots of all positive $-V_T^U$ values (i.e. time steps with a deep upper-level warm core) in all ensemble members, for the three clusters, and in the operational analysis.Numbers on top of each box indicate the average number of deep warm-core (DWC) steps per ensemble member and numbers at the bottom the total number of DWC steps in each box plot.