The University of Melbourne
Diego Saúl Carrió Carrió
School of Geography, Earth and Atmospheric Sciences and
Centre of Excellence for Climate Extremes
Parkville, 3010
Victoria, AUSTRALIA
diego.carriocarrio@unimelb.edu.au
Melbourne, April 26, 2022

Dr. Joanna Staneva
Editor– Natural Hazards and Earth System Sciences

Dear Joanna Staneva,

Please find attached the revised version of the manuscript **NHESS-2022-58** entitled ***"Challenges assessing the effect of AMVs to improve the predictability of a medicane weather event using the EnKF. Storm-scale analysis and short-range forecast"***.

I have carefully examined the constructive suggestions made by the reviewers and I have taken full account of their comments. Therefore, the main results of the work are now better described and emphasized. The following is a point-by-point response to the comments and inquiries made by the first reviewer. I thank the reviewer for their comments and believe the manuscript is now greatly improved both in terms of clarity and readability. We believe that the new version of the manuscript, which have been improved significantly, will help the reader to better understand the work presented here.

With best regards,

Diego Carrió Carrió

ANSWERS TO THE REVIEWER

Reviewer #1:

The manuscript presents a study to evaluate the impact of assimilation of atmospheric motion vectors and conventional observations with an ensemble Kalman filter to improve forecast of medicane Qendresa.

The manuscript is well written and the numerical experiments well designed.

Major comments:

Main remarks are as follow:

1. *The introduction requires a brief general presentation of numerical weather prediction and data assimilation, before any detailed discussion.*

   I thank the reviewer for pointing out this. I totally agree with the reviewer that a brief general description of numerical weather prediction along with data assimilation will help the reader better contextualize the work presented here.

   To this purpose I have added the following sentences at the beginning of the *Introduction* section:

   "*The Western Mediterranean region is often affected by high impact weather phenomena (e..g, tornadoes, hail storms, tropical cyclones, or flash floods, among others) which produce huge economic losses and fatalities. Although numerical weather prediction models have significantly improved during the last years, the proper prediction of such extreme weather events (timing and location) in the Western Mediterranean remains a challenge. Many extreme weather events affecting the Mediterranean coastlands initiate over the sea, where in-situ observations are scarce. This lack of information impacts negatively on the representation of the initial state of the atmosphere, and consequently on the accuracy of the numerical forecasts (Wu et al., 2013). Over the last years, different sophisticated methods have been designed and implemented to improve the estimation of the atmospheric state in numerical weather prediction models from both in-situ and remote sensing instruments, such as Doppler radars or meteorological instruments on-board satellites (Rabier, 2005; Palmer and Hagedorn (2006); Shen et al., 2016; Geer et al., 2018). These methods are known as Data Assimilation techniques (e.g., Daley, 1991; Evensen, 2009b; Kalnay, 2002), which basically combines information from numerical weather prediction forecasts with all the available observations to create a new set of initial conditions that better represent the current state of the atmosphere*."

   References:

   Xinrong Wu, Shaoqing Zhang, Zhengyu Liu, Anthony Rosati, and ThomasL. Delworth. A study of impact of the geographic dependence of observing system on parameter estimation with an intermediate coupled model. Climate Dyn., 40(7-8), 2013.

   Florence Rabier. Overview of global data assimilation developments in numerical weather-prediction centres. Quart. J. Roy. Meteor. Soc., 131(613), 2005.

Tim Palmer and Renate Hagedorn. Predictability of weather and climate, chapter Observations, assimilation and the improvement of global weather prediction-some results from operational forecasting and ERA-40. Cambridge University Press, 2006.

Feifei Shen, Jinzhong Min, and Dongmei Xu. Assimilation of radar radial velocity data with the wrf hybrid etkf–3dvar system for the prediction of hurricane ike (2008). Atmospheric Research, 169, 2016.

Eugenia Kalnay. Atmospheric modeling, data assimilation and predictability. Cambridge university press, 2002.

Evensen, G. Data assimilation: The ensemble Kalman filter(2nd ed.). Berlin and Heildelberg, Germany: Springer-Verlag, 2009b.

Roger Daley. Atmospheric data analysis, cambridge atmospheric and space science series. Cambridge University Press, 6966(25):809–822, 1993.

Geer, A. J., Lonitz, K., Weston, P., Kazumori, M., Okamoto, K., Zhu, Y., ... & Schraff, C. (2018). All-sky satellite data assimilation at operational weather forecasting centres. Quarterly Journal of the Royal Meteorological Society, 144(713), 1191-1217.

2. *When referring to present weather prediction systems/models more details should be included: are these operational forecasts or case studies? Are these global coarse resolution or mesoscale convection permitting models? Do they assimilate observations and how?*

I really appreciate the reviewer's comment about this point. The initial idea was to highlight that even using the most recent numerical weather model versions at high grid resolution we still have problems to accurately predict the intensification and trajectory of this small-scale Mediterranean cyclones, which are mainly initiated over sparse observational regions. To avoid misinterpretation of this sentence to the reader we have rewritten the sentences that appears "current numerical weather models…" as:

Line L6: "*…the chaotic behavior inherent to current numerical weather prediction models*" => "*…the chaotic behavior inherent to* ***numerical weather prediction models***"

L94-95: "*More precisely, the correct prediction of both the northward loop trajectory followed by Qendresa and its intensification still remain a major challenge for most current numerical weather models*." => "*More precisely, the correct prediction of both the northward loop trajectory followed by Qendresa and its intensification still remain a major challenge for most* ***current mesoscale convection-permitting numerical weather models***."

L130-131: "*It is hypothesized that the low predictability of Qendresa is likely associated to the difficulty of current numerical weather models to depict…*" => "*It is hypothesized that the low predictability of Qendresa is likely associated to the difficulty of* ***current mesoscale convection-permitting numerical weather models*** *to depict…*"

L139-140: "*This loop-like trajectory is the main feature that current numerical models do not accurately predict mainly due to a miss representation of the upper-level dynamics*" => "*This loop-like trajectory is the main feature that* ***current convection-permitting numerical models*** *do not accurately predict mainly due to a miss representation of the upper-level dynamics*"

L504-505: "*Specifically, this weather event was selected due to is low-predictable behavior (i.e., poorly forecasted by current numerical weather prediction models) probably because their initial conditions are poorly estimated due to the lack of in-situ observations over the sea." => "Specifically, this weather event was selected due to is low-predictable behavior associated to the poorly initial conditions estimation of the atmosphere over the sea."*

3. *The ensemble was designed to include different sets physical parameterizations among the different ensemble members (if I understood correctly). However, no discussion on how different sets of parameterizations performed is discussed. Can you identify sets of parameterizations that lead to better/worse results in a systematic way? If so, I suggest also drawing them with a different color in the plots.*

The ensemble of forecasts used in this study was generated using initial and boundary conditions from the global EPS-ECMWF and then to add more diversity to the ensemble different combinations of parametrizations were also used for each of the ensemble members. This approach is widely used in the National Severe Storms Laboratory (NSSL-NOAA, Oklahoma,USA) and NCAR (USA), where the author learnt this method among others. At the end of the first paragraph of section 3.1 it is also stated that more details about how the ensemble was generated is explained in the paper by Carrió and Homar, 2016. Table 1, shows the different combination of parametrizations used for each ensemble member.

In this study I focus on the performance of the EnKF to improve the predictability of this extreme and low predictable weather event. To this purpose, the ensemble design plays an important role so we want to avoid having an under dispersive ensemble, which will hamper the data assimilation performance. After testing different sets of parametrizations I ended with an ensemble configuration that presents the statistical properties I was interested in. The ensemble set up is the one described in the manuscript.

In this sense, investigating the performance of individual ensemble members according with its parametrizations is beyond the scope of this study. Adding this information will no contribute to understand better results of this study and could distract the attention of the reader of other crucial aspects of this study. The main aim of this study is to investigate the impact of the assimilation of AMVs to improve pre-convective environment of this high-impact weather event. We do not focus on seeking the best numerical simulation of this event.

Nevertheless, I think that following the reviewer's comment I could improve the section adding the following sentences at the end of section 3.1 (*Numerical Weather Model*):

"Note that the main objective of this paper is not to investigate the sensitivity of grid resolution and the role of different parameterizations to the Qendresa forecast skill.

However, if one is interested in the potential effect of these aspects to Qendresa forecast, please see recent work from Pytharoulis et al., 2018 and Mylonas et al., 2019."

4. *The use of the global model for initial and boundary conditions is mentioned in the manuscript, however, no discussion is provided regarding the improvement that I supposed was achieved by the WRF model and WRF-DART simulations, both due to better resolution and assimilation at high resolution. This should be discussed in the manuscript and lines depicting the global model results should be added in the plots.*

   The comparison and discussing that the reviewer is making about the global model is fair and indeed, that discussion is already present in the paper. To investigate the potential of assimilating the different kind of observations I was interested in (i.e., conventional and AMVs) I designed 4 experiments: (a) NODA, in which no observations are assimilated, (b) SYN, only conventional observations are assimilated, (c) RSAMVs, only AMVs observations are assimilated and (d) CNTRL, both conventional and AMVs are assimilated. These experiments share the same spatial grid resolution, so I can avoid attributing benefits of the results to such aspect and I can focus more on the observations assimilated.

   It is also important that it would not be fair to compare the DA simulations against the global model, which is the EPS-ECMWF that runs at ~31 km for the period when Qendresa took place. Our NODA experiment is the equivalent to global model downscaled at the resolution of the rest of experiments (i.e., 4 km). That was the main reason behind the design of the NODA experiment, to be able to compare its performance with the DA experiments and be able to assess the impact of assimilating different observations, all of them using the same grid spacing resolution.

Minor remarks:

Line 8: with operational models? At what resolution? Lower than your resolution?

As far as the author is aware, accurate prediction in terms of intensity and trajectory of Qendresa, especially the observed loop in the trajectory eastern of Sicily, was very challenging. Operational models that tried to predict such event fail and other researchers in the community (some of them exposed at conferences but never published) that run their own models, such as WRF, also failed. In terms of spatial grid resolution, there were simulations with higher and lower grid resolution compared to the one we used in this study, and none of them succeeded.

This issue was solved in the above answer to point number 2.

Line 16: Is your high resolution the main factor of improvement or the data assimilation?

Results of this study, after comparing the DA experiments (i.e., SYN, RSAMV and CNTRL) against the simulation where no data were assimilated (i.e., NODA), all of them using the same grid resolution, reveal that the main factor of improvement comes from the data assimilation procedure.

Line 18: A short introduction of NWP and data assimilation should be provided before directly jumping into detailed review of data assimilation literature

I agree with the reviewer and the revised version of the manuscript contains the following lines at the beginning to the introduction section:

"The Western Mediterranean region is often affected by high impact weather phenomena (e..g, tornadoes, hail storms, tropical cyclones, or flash floods, among others) which produce huge economic losses and fatalities. Although numerical weather prediction models have significantly improved during the last years, the proper prediction of such extreme weather events (timing and location) in the Western Mediterranean still remains a challenge. Many extreme weather events affecting the Mediterranean coastlands initiate over the sea, where in-situ observations are scarce. This lack of information impacts negatively on the representation of the initial state of the atmosphere, and consequently on the accuracy of the numerical forecasts (Wu et al., 2013). Over the last years, different sophisticated methods have been designed and implemented to improve the estimation of the atmospheric state in numerical weather prediction models from both in-situ and remote sensing instruments, such as Doppler radars or meteorological instruments on-board satellites (Rabier, 2005; Palmer and Hagedorn (2006); Shen et al., 2016; Geer et al., 2018). These methods are known as Data Assimilation techniques (e.g., Daley, 1993; Evensen, 2009b; Kalnay, 2002), which basically combines information from numerical weather prediction forecasts with all the available observations to create a new set of initial conditions that better represent the current state of the atmosphere"

Line 36: This part of the sentence is not clear. May be: Due to the limited oceanic and maritime coverage of meteorological radars,....

I thank the reviewer for pointing out this. We have modified the sentence according with the suggestion provided by the reviewer.

Line 39: May be: observational coverage of

I Agree. We have applied the reviewer suggestion.

Line 50: These are not so recent. Are there any more recent references?

I thank the reviewer's suggestion. Yes, there are more recent references. The revised version of the manuscript contains the following recent references about bias correction.

Kumar, S. V., Reichle, R. H., Harrison, K. W., Peters-Lidard, C. D., Yatheendradas, S., & Santanello, J. A. (2012). A comparison of methods for a priori bias correction in soil moisture data assimilation. Water Resources Research, 48(3).

Otkin, J. A., Potthast, R., & Lawless, A. S. (2018). Nonlinear bias correction for satellite data assimilation using Taylor series polynomials. Monthly Weather Review, 146(1), 263-285.

Ma, C., Wang, T., Jiang, Z., Wu, H., Zhao, M., Zhuang, B., ... & Wu, R. (2020). Importance of bias correction in data assimilation of multiple observations over eastern China using WRF-Chem/DART. Journal of Geophysical Research: Atmospheres, 125(1), e2019JD031465.

Line 54: It should be stated that one of the weaknesses of retrievals is the need for a first guess or additional atmospheric information, usually coming from short model forecasts that may be in turn too coarse and inaccurate. Some references should be included here.

I thank the reviewer for pointing this valuable information out. I have added the following sentence in the revised manuscript with some new references:

L53: "Retrievals are easier to assimilate and interpret because they provide information that can be directly related with atmospheric variables, and its assimilation avoids the use of relatively complex RTM. However, one always should consider that atmospheric retrievals are sensitive to many factors, including the first guess that usually comes from a coarse global numerical weather model that could result too inaccurate (Hannon et al., 1996; Li et al., 2000; Zhang et al., 2014)."

References included:

Hannon, S. E., L. L. Strow, and W. W. McMillan, 1996: Atmospheric infrared fast transmittance models: A comparison of two approaches. Proceedings of SPIE, 2830, 94–105.

Li, J., W. Wolf, W. P. Menzel, W. J. Zhang, H. L. Huang, and T. H. Achtor, 2000: Global soundings of the atmosphere from ATOVS measurements: The algorithm and validation. J. Appl. Meteor., 39, 1248–1268.

Zhang, J., Li, Z., Li, J. et al. Ensemble retrieval of atmospheric temperature profiles from AIRS. Adv. Atmos. Sci. ***31,*** *559–569 (2014). https://doi.org/10.1007/s00376-013-3094-z*

Line 61: any more recent reference?

Walker, E., Mitchell, D., & Seviour, W. (2020). The numerous approaches to tracking extratropical cyclones and the challenges they present. *Weather*, *75*(11), 336-341.

Choy, C. W., Lau, D. S., & He, Y. (2020). Super typhoons Hato (1713) and Mangkhut (1822), part II: challenges in forecasting and early warnings. *Weather*.

Dorian, T., Ward, B., & Chen, Y. L. (2018). Tropical Cyclone Amos (2016) Forecasting Challenges: A Model's Perspective. *Tropical Cyclone Research and Review*, *7*(3), 172-178.

Line 99: These were not assimilated in previous studies? Are these assimilated in operational forecasts that issued real-time forecasts of Qendresa? If so, they did not improve forecasts? If so, for what reason? Coarse resolution, not efficient data assimilation technique?

As far as the author is aware AMVs have been not assimilated at this grid resolution for the Qendresa case or any other extreme weather event in the Mediterranean region. Global models, such the ones form the ECMWF tend to assimilate such observations among others. However, the grid resolution for such models is not comparable with the grid resolution used in this study, in which the model is able to resolve much more complex physical processes providing more realistic atmospheric fields. In this study I want to assess the potential of assimilating AMVs. Are the AMVs able to produce a significant improvement to high resolution simulations?

It is difficult to answer if the assimilation of AMVs improve global models or, even more difficult, to what extent they improve the analysis. Global models assimilate AMVs among millions or other types of observations, so one cannot answer this question without performing additional sensitivity experiments. However, in my opinion, for extreme weather events it is crucial to have a high-resolution model running together with an efficient data assimilation method.

In any case, here I want to emphasize that I am not interested in competing with the global model. The motivation of this study was that running a high-resolution model we were not able to reproduce the intensity and track of Qendresa. Due to the limitation of observations available over the region where the medicane took place one appealing observation alternative was to assimilate information about the wind field. According with *Carrió et al., 2017*, improving the upper-level dynamics could contribute to improving the forecast of Qendresa. The assimilation of AMVs offers us this possibility.

Line 135: can you please further describe how this was done?

I simply took infrared imagery, where it is easy to identify the eye of the cyclone, and thus, its center. One can "approximately" identify the center of the cyclone for each time stamp infrared imagery is available. Thus, the observation track showed in Fig. 2 only shows the trajectory of the cyclone when the center of the cyclone was identifiable trough infrared imagery.

Following the reviewer suggestion, we have modified that sentence as follows:

"Due to the lack of in-situ observations present over the maritime region where the medicane took place, infrared imagery was used as a proxy to visually identify the center of the cyclone and estimate the track followed by Qendresa (Fig. 2)."

Line 140: which models? Operational? which ones? Other studies? Please cite as necessary.

I have rephrased the sentence as follows (adding some additional references):

"This loop-like trajectory poses a clear challenge in terms of numerical predictability, even for current most advance mesoscale numerical weather models (see Carrió et al., 2017, Pytharoulis et al., 2018 and Mylonas et al., 2019)."

Line 158: why was one-way nesting chosen?

I initially tested two-way nesting and I did not find any significant difference from one-way nesting apart from the computational time required to run the two-way nesting in comparison with the one-way nesting (at least for this case). For this reason, all the results shown in the manuscript were obtained using the one-way nesting.

Line 193: I think that this sentence should appear earlier in the manuscript (where the aim of the study is described) and in the abstract

We thank the reviewer for noticing this. Following this suggestion, we have deleted this sentence from Line193 and added to the abstract and the introduction sections.

Line 308: I think that only a short paragraph with aim and main results should be here and the rest as supplementary material.

I agree with the reviewer that this section could be shortened

Line 351: Why? It has no data assimilation

I thank the reviewer for pointing out this. I agree that this sentence is not clear and needs further clarification. The reason behind that sentence is because the initial and boundary conditions from NODA are obtained from the analysis EPS-ECMWF at 00 UTC 7 November, which are obtained from the assimilation of all the available observations (including radiances). Thus, one could expect that the initial conditions at that time could be better than the initial conditions obtained from the rest of DA experiments after cycling 12 hours and only assimilating conventional and AMV observations. Here it is showed the importance of assimilating observations using a higher grid spacing.

To avoid misunderstandings, I have rewritten the sentence in the following way:

"~~Among the entire set of analysis obtained from the different numerical experiments, NODA was initially expected to produce an accurate representation of the medicane trajectory.~~ F*orecast results from NODA show that most ensemble members do not properly simulate the observed trajectory, particularly the loop-ending on the eastern coast of Sicily, which is produced when the cyclone made landfall and its dissipation phase begin (Fig. 11a).*"

Line 392: This requires further description. Do the results using this technique add further value? I wonder whether this is really needed. As written now it is unclear to me.

I agree with the reviewer that this section is not adding new information to the conclusions of these experiments. For this reason, I have deleted the results from the *Kernel Density Estimation*.

Line 401: All of these are very low probabilities and the kernel method was not explained, therefore the reader cannot understand the importance of these figures

Following the above comment, I have deleted these results.

Line 457: Are RSAMV assimilated too?

As far as the author is aware, only AMVs are assimilated, which are available hourly. RSAMVs are not considered in ERA5 analysis. RSAMVs are the same type of observations but instead of being available hourly, they are available every 20-min. As ERA5 analysis are obtained hourly, there is no need to assimilate them every 20-min.

I have rewritten the sentence to clarify this point raised by the reviewer in the following way:

"*… ERA5 provides high-temporal resolution fields (i.e., hourly) obtained from the assimilation of vast amounts of observations (most of them satellite radiances)…*" =>"*… ERA5 provides high-temporal resolution fields (i.e., hourly) obtained from the assimilation of vast amounts of observations (most of them satellite radiances), although they are found on a 30 km grid resolution mesh. It is important to note that to obtain ERA5 analysis AMVs have been*

assimilated. However, the RSAMVs observations assimilated in this study, which are available every 20-min, are not used to produce ERA5 analysis (Hersbach et al., 2020).”

References:

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Thépaut, J. N. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146(730), 1999-2049.