Review of **"Assimilating WIVERN winds in WRF model: an application to the outstanding case of the Medicane Ianos"** by Stefano Federico et al.
The wind velocity radar nephoscope (WIVERN) represents a pioneering spaceborne mission that aims to deliver in-cloud horizontal wind observations with fine vertical resolution. This paper explores the impact of assimilating WIVERN derived Doppler wind observations into the WRF model for the high-impact case study of medicane Ianos. Results indicated that assimilating WIVERN observations leads to a positive impact on the prediction of the medicane trajectory, reducing trajectory forecast error by 43%. It is also shown that the assimilation of such observations will improve prediction of precipitation and surface winds. This study is innovative, interesting and well-written and I found these results of interest for the scientific community, addressing the value of next-generation wind observations in regional numerical weather prediction. However, there are several aspects that require significant clarification or improvement before the manuscript can be considered for publication. For this reason, I recommend **major revision** before it can be accepted for publication in *Weather and Climate Dynamics*.

We acknowledge the reviewer for the careful and insightful review of the paper. We will answer shortly to this discussion and then we will update the paper according to the comments/suggestions in the review phase.

Major Comments:

1. In L113-114, it is mentioned that the CTRL ensemble is initialized at 12 UTC on 16 September 2022 using the ECMWF-EPS data. First, this appears to be a typo, the correct data should be "16 September 2020", corresponding to the Medicane Ianos case. Second, it should be added information about the approximated horizontal grid resolution of the ECMWF-EPS data used for initialization, to consider the significant resolution mismatch between the coarser ECMWF-EPS fields and the 4 km grid resolution of the WRF model used in this study. More importantly, the current setup does not seem to adequately address model spin-up problem. Since the WRF model is run on a single, relatively small domain at convection-permitting resolution, the atmospheric fields initialized from a much coarser ensemble forecast would require a proper spin-up period to develop balanced mesoscale structures. However, the model begins data assimilation just three hours after initialization. This short spin-up time is likely insufficient for the model to reach a dynamically and thermodynamically consistent state at 4 km resolution, which may degrade the quality of the assimilation and forecast performance. I recommend that the authors reconsider the model spin-up strategy, either by lengthening the spin-up period or by conducting a sensitivity test to justify the chosen approach.
   Yes 2022 is a typo. This will be corrected. Considering the point of the resolution mismatch between ECMWF-EPS (36 km) and our simulations (4 km) we searched for a compromise between the complexity of the model simulations and the computing time. We followed a heuristic approach, and we were guided by the comparison of our results with those reported in Pantillon et al (2024). The evolution of the trajectories of the Medicane Ianos simulated by our ensemble are compatible with those reported in this study and we used this as a proof of a reasonable setting of the model. Considering the point of assimilating WIVERN just after 3h, we agree that this could cause unbalances in the model for the WIV_{3h} simulations. Nevertheless, the

simulations assimilating WIVERN every 3h smoothly converge towards the reference trajectory and, likely, unbalances are eventually mitigated by the frequent DA. However, to consider more in detail the comment raised by the reviewer, some simulations will be performed assimilating WIVERN frequently (3h) after 12h of spin-up time to better investigate this point.

2. Given that the focus of this study is on the assimilation of WIVERN wind derived observations, the lack of a clear and dedicated figure showing the structure of these observations at multiple vertical levels is a significant omission. Please, add a figure that visualizes the WIVERN wind fields (i.e., direction and speed) at different pressure levels or heights. This would greatly enhance the reader's understanding of the observational coverage and characteristics. Indeed, adding a new section in the manuscript focus on the WIVERN data itself, would be very beneficial for the manuscript. It would benefit from a more-in-depth discussion of the WIVERN data itself. For instance, how do the WIVERN winds compare with conventional wind observations, such as from radiosondes? Some sort of consistency check or statistical comparison would help assess data reliability. Also, it should be specified which vertical levels of WIVERN observations are assimilated or if there are any thinning or level selection applied. What is the lowest vertical level at which WIVERN observations are available? Understanding the vertical range of WIVERN is important, especially when discussing its impact on near-surface variables such as wind and precipitation.

   Thank you for noticing this point. We will reorganize and develop more the content about the WIVERN observations. In general, WIVERN will observe the winds along the line of sight (LoS). This LoS varies in time (12 rpm) and the WIVERN observations are a combination of zonal and meridional wind components with sine/cosine of variable angles. In this sense, the WIVERN observations cannot be interpreted easily in terms of zonal and meridional wind components. We will add a figure of WIVERN winds at different levels to better show their density. Considering the error of WIVERN observations it is already full developed in the paper, as it depends on reflectivity plus other factors. In any case the final error is expected lower than 3 m/s. Of course, the error depends on the reflectivity, which is dependent on the three-dimensional structure of the observed cloud; this is, for example, much different from sondes, whose error depends on the height (the higher the height the higher the error). For the case of the Medicane Ianos we reported the WIVERN error (averaged over a single level). We will show the same curve for the wind observations of sondes. For lower levels, the sondes have a lower error; for upper level the situation is reversed. Finally, no data thinning was applied in the vertical direction, and we start to assimilate from 1 km above the surface. This limit is a bit optimistic over the land where WIVERN observations should be available starting from 2km above the surface, nevertheless we assimilate by far over the sea for the case study considered in the paper and this is not an issue.

3. A major concern with the current study is the design of the WIVERN24h assimilation experiment, which only assimilates WIVERN wind observations at 12 UTC on 17 September, when the medicane is already fully developed, followed by a 24 h free forecast until 12 UTC on 18 September. From a predictability and forecasting perspective, this approach raises questions about the broader relevance of the

findings. Since the cyclone structure is already well stablished at the time of assimilation, the potential for WIVERN winds to meaningfully influence the genesis or early intensification phase is not tested. As such, the experiment does not provide insight into whether WIVERN observations can enhance forecast skill in the more critical lead-up phase, when predictability in inherently lower and guidance is more valuable for early warning. I strongly recommend that the authors design an additional experiment where WIVERN observations are assimilated prior to cyclone development, such as during the early stages on 15 September. This would allow the authors to assess whether WIVERN winds improve the characterization of the pre- convective environment, and whether that leads to improved prediction of the cyclone's formation, track, or intensity. Such an experiment would greatly increase the impact and relevance of the study by demonstrating the added value of WIVERN observations in a more operationally realistic forecasting context.

As stated into the paper, the time for data assimilation was chosen when the storm was well formed to fulfill these two requirements: a) we are able to have at least one member that simulates well the real Ianos trajectory (this occurs, to our knowledge, only starting the forecast on the 16 September and we need this simulation for generating pseudo-observations); b) we are enough far from the landfall to ensure at least few hour of alerting time. Of course, as suggested by the reviewer, it would be interesting to investigate the WIVERN potential in the early stages of the storm, and we will design an additional experiment starting on the 15 September (12 UTC) using a similar approach of that used in the paper. Nevertheless, we cannot guarantee to have a useful member to generate pseudo-observations to fully explore the point.

4. To improve clarity and reader understanding of the experimental set-up, add a schematic diagram or timeline that illustrates the configuration and sequencing of the different simulations performed in this study (CTRL, WIVERN24h, WIVERN3h). Include initialization times, assimilation windows, duration of the free forecasts, timing of observation ingestion, among others. Such a visual aid would be particularly helpful in understanding how the experiments differ in terms of when and how WIVERN observations are assimilated, and it would complement the textual descriptions in the methodology section.
   Ok. We will add this schematic diagram to be clearer in the methodology section.

5. Another aspect that needs to be improved from this study is related to the lack of standard data assimilation diagnostics in the performance of the WIVERN3h cycling DA experiment. Typically, in cycling DA, it is standard practice to include observation-space diagnostics to evaluate how the DA system is performing. These diagnostics often include: i) root-mean-square innovation, ii) total spread, iii) mean innovations and iv) consistency ratio, which compare the ensemble spread against the observation error (see Dowell et al., 2004; Yussouf et al., 2013 or Jones et al., 2016). These diagnostics not only help in understanding whether the system is behaving as expected but also provide confidence and robustness to the DA system. In particular, the "sawtooth" plots (e.g., Fig. 2 in Jones et al., 2016) are useful for showing the temporal evolution of these quantities across assimilation cycles, helping to assess whether the ensemble has sufficient spread and whether the assimilation is consistent with the assumed error statistics. I strongly encourage the authors to

include these diagnostics to demonstrate that the assimilation of WIVERN winds is working properly and to validate the underlying assumptions of the ensemble DA framework.

Thank you for the comment. We will introduce some analysis as those suggested by the reviewer. Specifically, the sawtooth plots (Figure 2 Jones et al., 2016) seems relevant for the purpose of this paper.

6. A key omission in the current manuscript is the lack of discussion regarding observation error characterization, which is essential in any data assimilation study. Specifically, the study does not clarify i) what observation error variance was assigned to the WIVERN-derived wind observations, ii) whether this error was estimated, assumed or tuned and, iii) if the same error was applied uniformly across vertical levels or observation types. Given that WIVERN represents a novel observing system, it is particularly important to justify the error assumptions used in the assimilation process. This includes specifying the source of the observation error (e.g., instrument noise, representativeness errors), and how it was implemented in the assimilation system.
We will clarify better the comments raised above. The material, however, is already into the paper. The WIVERN simulator starts from a scene of member 42. The observation errors depend on the observed signa to noise ratio (ultimately the reflectivity, see Eqn. 7), which is derived from the hydrometeors simulated by WRF member 42 at the assimilation time (the reflectivity in the W band is simulated by the WIVERN simulator starting from the hydrometeor content). Moreover, we need to add a some error to that computed in Eqn. (7) to take into account for mis pointing of the radar antenna and for partial beam filling. This is accounted in Eqn. 8. The equations (7) and (8) show the formulation of WIVERN winds along the LoS error. The error of Eqn. 8 is applied to WIVERN winds pseudo-observations along the LoS. Observations errors are assumed uncorrelated, and a data thinning of 10 km is applied (this corresponds to take one pseudo-observation every 2 pseudo-observations). The average value of the error as a function of the vertical levels is shown in Figure 4b (black curve). Incidentally, we use the term “WIVERN simulator” to indicate the forward observation operator. The name “simulator” is preferred in the radar community.

7. There is a significant ambiguity in the Methodology section (L103-105) regarding the type of data assimilation approach employed. The manuscript states that a 3DVar is used, yet it also notes that the background error covariance matrix is computed from the CTRL ensemble, which is inconsistent with traditional 3DVar frameworks that typically rely on climatological, static covariance estimates. Are the authors using a pure 3DVar approach with a static covariance matrix or is this a form of ensemble-based 3DVar (En3DVar) or a hybrid 3DVar-ensemble method? If an ensemble is being used to estimate flow-dependent covariance, this must be explicitly stated and clearly explained. The choice of DA techniques is crucial to interpreting the system’s ability to adjust to observational input, especially with novel data like WIVERN. Additionally, further clarification is needed regarding the NMC method (used in traditional 3DVar) mentioned for estimating background errors. What period or set of simulations was used to generate the perturbations for the NMC covariance matrix? Was any covariance inflation used? These details are essential for understanding the structure

and realism of the background error covariance, which strongly affects how observations influence the analysis.
We are sorry for not being clear about the terminology. In the paper there are two experiments presented. The experiments WIV_12h and WIV_3h use the En3DVar approach. In our setting of the En3DVar, the background error matrix is entirely determined by the ensemble. In the experiment NMC the background error matrix is static, and it is computed for the month of September 2020 (Lines 316-323). We apologize for the confusion and we will correct this point in the revised version of the paper.

8. Results and conclusions drawn in this study are based solely on a single case study. While the results are promising and the case is certainly relevant, relying on one event significantly limits the generalizability and statistical robustness of the findings. medicanes are highly variable in structure, environment, and predictability, and it remains unclear whether the demonstrated benefits of WIVERN wind assimilation would hold across other cases with different dynamical regimes or forecast challenges. In this sense, a larger number of cases is necessary to draw robust conclusions. In addition, it would be very interesting to know how useful are the WIVERN observations when the baseline model performance (CTRL) is poor?

   Ok. We will clarify better this limitation in the conclusion section. Of course, a much larger number of case studies is required to better define the usefulness/limitation of the WIVERN winds data assimilation.

Jones, T. A., Knopfmeier, K., Wheatley, D., Creager, G., Minnis, P., & Palikonda, R. (2016). Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast system. Part II:

Combined radar and satellite data experiments. Weather and Forecasting, 31(1), 297-327.

Minor comments:
Thank for the careful review of these minor comments. We will consider all the points and correct the paper accordingly.

The following are some suggestions that could help to improve the quality of the manuscript:
Introduction Section:

1) L25: "*… The purpose of data assimilation is …*" => "*The purpose of data assimilation (DA,* **add references***) is …*"

2) L31: "*… using a Doppler Wind Lidar*." => Please add references.
3) L36: "*The Wind Velocity Radar Nephoscope (WIVERN) Illingworth et al. (2018); …*" => Add parenthesis before references => "*The Wind Velocity Radar Nephoscope (WIVERN)* **(***Illingworth et al. (2018); …*"

4) L45: It is introduced for the first time the acronym (DA). Remove from here (see my previous comment 1) above).

5) L45-46: It is stated that an ensemble-based data assimilation framework is used in this study. This is very confusing because later it is stated that a 3DVar, which is based on static climatological background error covariance matrix, is used. This is inconsistent. Please, clarify this point and be more specific about the DA methodology used in this study.

1. 6) L58: Please, be more specific on the dates and hours in which medicane Ianos took place.
2. 7) L75: After introducing acronym 3DVar, please add some references. Again, note that in L46 it was mentioned the use of ensemble-based DA method. However, here it is mentioned the use of the standard 3DVar, which is NOT an ensemble- based DA technique. Please, clarify this point carefully.

Data and Methods Section [**Section 2.1**]:

8. 8) L79: "*WRF model 4.1 with ...*" => "*WRF model* ***V****4.1****,*** *with ...*"
9. 9) L79: I suggest removing (WE) and (SN).
10. 10) L80: Why the authors use such a reduced number of vertical levels (55)?. Add information about how vertical levels are distributed through the atmosphere (e.g., denser vertical levels near surface, ...).

11)L80: "*The model horizontal resolution is ...*" => "*The model horizontal* ***grid*** *resolution is ...*"
12) L80-81: Remove "*in both WE and SN directions*"
13) L82: Remove "*with the SW and NE corners ..., respectively.*"
14) L82: I suggest your 1st figure in the paper to be your numerical domain.
15) L82-85: Add a justification about the choice of these parameterizations used in this study. What are the main reasons the authors use this configuration?
16) L86: Again, at this point it is confusing why it is used the EPS if you are going to use 3DVar, which only requires a deterministic field.

17) L87: Are you using "analysis" or "forecasts" from the ECMWF-EPS? Please, add this clarification. Also, add further information about spatial resolution of the fields provided by the ECMWF.

18) L89: A direct downscaling from global model to a single domain at 4 km is used here. Why do the authors not make use of a nested approach?

Data and Methods Section [**Section 2.2**]:

19) L92: Be consistent with the notation. In the previous section, it was used 12:00 UTC, instead of 12 UTC.
20)L93: "*data assimilation cycles are considered*" => "*data assimilation* **experiments/simulations** *are considered*"

21. 21) L94: What do you mean by "*a longer repetition cycle*"? Please, clarify.
22. 22) L98: "*In the 24h cycle*" => "*In the* **24-hourly** *cycle*"
23. 23) L98: Why do the single assimilation is performed at 12 UTC? Which was the state of Ianos at this time (initiation phase, fully developed, decaying, ...)? Please add this information.
24. 24) L99: What do you mean by "*WIVERN overpasses a mature storm system*"? It is not clear to me. Please, improve this sentence.
25. 25) L103: Add more references to the 3DVar system. Again, this is not an ensemble-based system. If you are referring to an ensemble-based 3DVar, please specify.

26)L104: "*background error matrix*" => "*background error* **covariance** *matrix*". Replace along entire manuscript.
27) L104: Typo: "form" => "**from**"
28)L105: Why the background error covariance matrix is computed at 12 UTC?
Provide justification.
29) Equation 2: According with the notation, XNens should be replaced by X**b,**Nens

30) L113-114: The model should spin-up first using 6-12 hours. Direct downscaling from global resolution to 4 km without considering spin-up time will lead to imbalance physical fields.
31) L117: What do you mean by "*simulator*"? Do you mean "forward observation operator"? Please clarify.
32) L114: The assimilation of WIVERN observations in this study is performed in stratiform areas, where no convection is present. What do you obtained if you assimilate observations where vertical velocity W is not negligible?

33) L127: Is this assumption, right? How do you know that the medicane is well sampled by WIVERN? A two-panel figure comparing the model wind field of the medicane (left) and the observed wind field from WIVERN (right) at different vertical levels and times is missing in the manuscript and should be added.

Data and Methods Section [**Section 2.3**]:
34) L138: "*The red line in 2*" => "*The red line in* **Figure** *2*"
35) Figure 2: Remove title and add this information in the caption of the figure. Labels in x- and y- are not consistent with notation used in Fig.1. Please, use same notation.
36)Figure 3: a) Remove title of figure. Add blank space between AVG and parenthesis in y-label; b) Use same x- and y-labels as the rest of figures. Could you make larger the trajectory points. They are too small.
Results Section:
37) L197: "*has a very important impact*". Subjective comment. Please, rephrase.
38) Figure 4: Remove titles. Add blank space between Height and [cm] in y-labels.
39)Figure 5: Remove title and add this information to the caption of the figure. Consistent x- and y-labels with the rest of figures.
40) Figure 6: Use same y-axis limits for both panels. Use black colour to x- and y- labels.

41) Figure 8: Use black colour to x- and y-labels.
42) Figure 9: To small figure. Add units to panel d). Modify x- and y-labels according to the rest of figures. Use same notation.

Results Section [**Section 3.1**]:

43) Figure 10: Remove panel titles. Modify x- and y-labels according to the rest of figures. Use same notation for latitude and longitude. Add units to colorbar.

44) L289-290: What was the observed precipitation? It seems you are assessing the performance of your simulations without comparing with the observations. If not, please rephrase sentence to clarify this point.
45) L295: Add reference to Table 2, where the RMSE is shown.

46) L303: Table 2 or Table 3? Table 3 is not mentioned in the text.
47) Figure 11: Adjust the colorbar height to match the figure panels. Remove decimal places from colorbar labels.

Conclusions Section:
48) L347: “*a constellation of many (4)*”. This sentence is unclear. Please clarify what the “*(4)*” stands for.

Appendix A Section:
49) L391: Should **Bx** and **By** been replaced by **Ux** and **Uy**, respectively?

50) Equation A5: Should n, *U* and *H* been written in bold notation?