

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2021-354-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-354

Anonymous Referee #2

Referee comment on "On the impact of dropsondes on the ECMWF Integrated Forecasting System model (CY47R1) analysis of convection during the OTREC (Organization of Tropical East Pacific Convection) field campaign" by Stipo Sentić et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-354-RC2>, 2021

The authors analyze the impact of a large set of dropsondes over the tropical east Pacific and the Caribbean on ECMWF's integrated forecasting system. Data from two IFS experiments (one making use of the dropsonde data and one not) are compared to study the impact on wind, temperature and humidity distributions. Additionally, the analysis is extended to selected convective parameters which are of relevance for studying developing tropical convection. This impact study is of interest to the community as it makes use of a unique data set and shows a model validation in a region where winds are known to be uncertain. However, it would clearly benefit from additional context and a more thorough discussion of the results. Please see the detailed comments below:

Major comments:

1. The introduction does not provide enough context to the study. It is only hidden between the lines why this work is done and what the questions and hypotheses are. So, I suggest to revise and clarify this section.

*L18: I do not see the distribution of the observations. I recommend that dropsonde locations or frequencies of observations per lat-lon bin (bins shown in Fig. 2) are shown on the map. How many flights were conducted in each box on how many days?

*L20ff: The reader would profit from information about the general aims of OTREC (instead of detailed results) and how this study is embedded. How were the flights designed? Did they happen regularly or was there a target process, region or time?

*L25-32: What is the purpose of the paragraph? Please include a clear statement about the scientific aim of the study and why the IFS data assimilation system and the proposed method is suitable to address it. Was there a plan for a dedicated validation in the specific regions? Was there a bias expected?

*L33ff: I am a bit confused about this paragraph as it covers a broad selection of studies making use of additional sonde data "in other regions of the planet" but it remains unclear what this implies for the presented study. The authors are referred to publications of Majumdar (2016, BAMS) and Parsons et al. (2017, BAMS) that summarize the results of targeted observations during THORPEX and beyond.

2. The data and methods sections would benefit from a more thorough explanation of the model experiments and the used data products. It is confusing to me what type of data is used for what type of analysis:

*L73ff: This paragraph is a surprising start as this 3D-var approach was never mentioned before and it is not stated why it is needed.

*L79-84 (and elsewhere): Please specify "model data" and what is used in the departure calculation. Is it an analysis departure or a background (first guess) departure? Please explain how this can be interpreted? Why do only positive values denote a deficit? Please better explain the two experiments (YDPS and NDPS) and what one expects to learn from comparing them? Were these experiments cycled over the entire month or several single cycle experiments for each mission? Please give details about the blacklisted profiles mentioned in L166. I guess the analysis only considers data points that are used in both experiments, correct?

*L83: I guess there is no collocation done in the sense that data is interpolated in space and time, but observation space data is used from the data assimilation system? Is the restriction to significant level data a result of the transmission of thinned data to the GTS? Please explain what data was used and clearly separate observation space data from gridded model data.

*L102: Here, I guess that "operational data" means gridded model space data? Do the $1 \times 1^\circ$ cells mean that data is retrieved at this grid resolution? Is the operational analysis data used here? What is the vertical resolution? L108 says that also ERA5 data is used but it is unclear why? Please be more precise about the data.

3. The section "model analysis" that contains the main results to assess the impact of the dropsondes is well-written and good to follow. The authors identify one particular flight on Aug 18, 2019 providing increased departures which is very interesting. However, it remains unclear whether this is a special case in

terms of the observed profiles (stronger wind, more humidity or warmer?) compared to the other flights and also whether the observed pre-tropical storm environment differs from other flights.

I would like to suggest that the authors enhance this discussion. Difference maps between the analyses of both experiments (or analysis and first guess) could help to understand how the information contained in the dropsondes was spread in the domain. Additionally, the largest departures could be illustrated with respect to the synoptic situation. Did the assimilation also improve the later development of the tropical storm?

*L118ff: This discussion would profit from showing how many observations were made in each box. How many flights contributed to the result in Fig. 2? See also comment on L18.

*L125: I guess "statistically significant" means that in the other boxes, departures from more than one flight are averaged which might have led to reduced departures?

*L128f: "Lack of..." I do not understand this sentence. What does this study tell about a negative impact on the prediction? Does that mean that other obs types are not able to constrain the analysis correctly?

*L131: "This suggests..." If so, are there other examples or might that be just very case sensitive? This also affects the next argument on the sensitivity to the boxes. Was this flight designed differently or focusing at a different met. situation? Was it the only pre-storm flight? Was this case unique, e.g., in terms of wind speeds? See also comment on L20ff.

*L143: I do not understand this sentence. What is the relative error?

*L156: What does "issues" mean? Were large FG departures leading to a withdrawal of the data? Is the used data rejected during the data assimilation process?

*L169: Please explain more carefully how this is done as it remained unclear what kind of data is used. Please specify "point data to a regular grid". Is this observation space or model space data? Is this related to the methods briefly introduced in the first paragraph of section 2? I guess this method uses observation data and the two analyses from the experiments? A bit more information is needed for readers that are not aware of results shown in Fuchs-Stone et al. (2021). What if the results were compared with gridded model space data? Is it possible to derive reliable vorticity data from the relatively coarse observation grid?

4. The section on cloudiness effects is entirely convincing to me, primarily as I missed a hypothesis about what is expected and why this is done. I do not get why x^* should be higher in clouds especially for winds. The presented analysis doesn't tell anything about whether a cloud also exists in the model and does not consider the vertical extent of the clouds and cloud layering. Why did the authors not consider a point by point analysis and use something like a dew point depression or RH as an indicator for clouds? Could it be that highest departures occur where no cloud was simulated but observed or vice versa?

*L223: I do not understand this sentence as I thought the difference between the experiments is only the information contained in the dropsondes (plus a potential cycling effect that should also be discussed in the final section)? Are there other obs types that do not perform well? What types of wind observations do exist in clouds?

*L232: The Instability Index seems also to show some overestimation in both experiments. Although this is also the case for DCIN it is stated that the convection parameters are well-represented. Later departures are considered "small" and an "excellent agreement" is seen. What is the underlying accuracy requirement for these statements? I guess a 1 g kg^{-1} or a 1 K might strongly impact a process like convection.

5. I asked myself what section 5 has to do with the topic of this paper? What is the reason for adding an analysis of the diurnal variability of these parameters in ERA5. It would be more interesting to know to what extent the assimilation of dropsondes contributed to the diurnal cycle representation in the forecasts, e.g., when using the 12 UTC FCs over 24 h and compare it to the analyses.

*L236: Why is ERA5 data used here instead of the experiments? Please clarify how observations and simulations are combined and whether this 3D var method is applied to the ERA5 gridded data to calculate a time series of convective parameters.

*L237: The separation of non-convective vs. convective regions in the domain is done only for the dropsonde releases, right? It would be good to give an overview about when the flights are conducted and dropsondes are released in the overview section. Does the difference in release time lead to an average moisture convergence below 2 kW m^{-2} as shown in the time series? How many flights or individual days did contribute to this analysis?

*L260: The decomposition is interesting but to understand the inter-box differences it would be good to know how many days are investigated here.

6. The **summary and conclusion** section misses a discussion of the results in the context of previous literature and the aims of this study. Although the study addresses the IFS model performance by comparison of model results with observational data, I wonder why the authors submitted the paper to GMD. Therefore, I suggest that the authors comment on how an NWP center can profit from these results. I would also encourage the authors to discuss their results in the context of the quality of tropical winds.

Minor:

*Title: I suggest using "ECMWF" instead of "EC"

*L5 and L59: Please check the abbreviation "EC" as typically ECMWF is used.

*L9: Please specify "precursor"

*L16-17 and elsewhere: correct citation brackets

*L29: reanalyses?

*L37: Please specify "affected"

*L105: What is the top pressure? The uppermost model level?

*L109: Please find a different and more specific title.

*L113 i.e. 700 hPa?

*L121: "all" seems not to be correct.

*L123: Box 3?

*Fig.4: At what vertical resolution is this analysis performed and what is the data availability? Use Theta symbol for potential temperature instead of t.

*L230: Please specify "considering the previous section"

*Fig.8 and Fig.9 would profit from a fit to the data (Fig.9) and some statistical numbers

*L272: "evaluates the performance" isn't it rather an evaluation of the impact of sondes as stated in the title.

*L282: "variable departures" is ambiguous

*L284: Please rephrase "benefit by ... smaller departures"

*L293: Change to "more scatter, i.e. more variable departures"

*L297: please mention that this was done using ERA5 data

*L307: "whether improvements could be made"