Atmospheric
Measurement
Techniques

Open Access

EGU

Discussions

Interactive comment on "Uncertainty Quantification for Atmospheric Motion Vectors with Machine Learning" *by* Joaquim V. Teixeira et al.

Joaquim V. Teixeira et al.

joaquim.p.teixeira@jpl.nasa.gov

Received and published: 11 August 2020

"Using Machine Learning to Model Uncertainty for Water-Vapor Atmospheric Motion Vectors" Teixeira et al. Responses to Referee 1

We would like to thank the referee for the careful read of the paper and for the detailed comments. Please see our responses below:

1. I have a problem with the present title. Reading it the first time I thought that paper was about improving error/quality of AMVs during the extraction process, and not during the assimilation process. From my understanding a title like: 'Use of Machines

Printer-friendly version

Discussion paper

Learning to improve Uncertainty Quantification of Atmospheric Motion Vectors assimilated in NWP models', would certainly match better the real content of the paper and be less confusing.

We certainly understand the reviewer's outlook on this; uncertainty quantification can often be a confounding term with different interpretations across subject areas. The title has been modified to "Using Machine Learning to Model Uncertainty for Water-Vapor Atmospheric Motion Vectors" to reflect this.

2. The test presented in this paper is limited to water vapour AMVs extracted on specific layers. This potentially corresponds to extraction of 3D winds from hyperspectral sounders, as mentioned in the introduction. However, there is actually no evidence that the results can be generalised to the common AMVs extracted from clouds tracking in infrared or visible channels. If the method is limited to hyperspectral winds, this must be clearly specified in the text and probably also in the title of the paper, and not let the reader supposed that it works for all types of AMVs. If the method is not limited to hyperspectral AMVs authors have to present results also with common cloud motion winds extracted from satellite imagery. I understand from the text that another paper is upcoming (line 325), but there is no description or information that can actually let me assume that common AMVs have been used, and that the results are positive.

Referee #2 expressed similar concerns, and we can understand the reviewers' perspective. We have qualified our statement that the approach may be globally applicable to any measurements, and have stated more specifically that it is likely to be useful for other sources of AMVs (especially those obtained by tracking gradients in trace gases). In paragraph 3 of the introduction we have included additional mention of the height assignment errors known to be an issue in tracking cloud features from radiances. This source of error is expected not to be as great of an issue when tracking retrieved trace gases (as shown in Posselt et al. 2019), as it is when tracking cloud features or radiance images. In addition, there are sources of error that are expected to be common to any feature tracking algorithm (e.g., regions without strong gradients

in the field being tracked, or regions in which the wind is oriented parallel to contours in the field being tracked). We have modified our conclusions to include this in the last paragraph of discussion.

3. The algorithm seems to be too dependent on the user's choice of the number of clusters, and the paper does not discuss the dependence of the algorithm on the chosen training dataset. It is also very unclear if the different clusters identified could refer to kind of physical or geographical AMVs properties, or if they are only blindly resulting out of the numerical tests. Authors must clarify/discuss if the results may depend on the AMV extraction model used (Mueller 2017). It is not clear if the same clusters can be used for operational AMV extracted from other schemes too (NOAA, EUMETSAT, C2 JMA. . .Etc). If it is not the case I guess this study must be repeated individually for every different AMV extraction schemes and maybe after every releases of these codes, which should represent an important limitation for operational use in NWP models. Although the authors promise the possibility to distinguish different geophysical regimes, the application ultimately presented by the paper comes down to discriminating the AMVs that are null because they are tracking the ground radiance, which is much too simple to showcase the real benefits of the algorithm.

This study is meant to be a proof of concept – to show how a combination of random forest, plus a Gaussian mixture model, can be used to learn error structures found via comparison of simulated measurements with a reference "truth" dataset (as was done in our previous work). Naturally, the particular algorithm developed in this paper is wholly dependent both on the nature run and the AMV extraction method. However, it is not intended to be an algorithm that can be immediately used in NWP models. Instead, we aim to present a model that can be reproduced (and tuned) for use in specific contexts of AMV methods and data assimilation frameworks. The computational costs of training the algorithm ($\sim$1 day on a single processor, per pressure level) and even the computational costs of running the AMV extraction on the nature run (an average of 3 days per pressure level, on a non-optimized cluster network), are not outside the

usual demands when updating parameters of NWP models.

In regards to the physical and geographical properties of the identified clusters, we have added a section in lines 329-345 and Figures 8-11 discussing this. They illustrate that the clustering algorithm manages to generally discriminate among geophysical regimes. Regarding the choice of number of clusters, this is a tuning parameter that is highly specific to application. We note that having one or more tuning parameters is not uncommon in many data analysis methods (e.g., k-means, PCA, self-organizing network, random forest, neural nets, regularized regression, smoothing splines, wavelets, etc.). Here, our method requires only 1 major tuning parameter (the random forest model also has tuning parameters, but that process, being a supervised regression, can be guided by cross validation). We note that the search for the 'optimal' number of clusters should be guided by expert knowledge, although this process should be greatly simplified by including an information criterion (e.g., the Bayesian Information Criterion) in the Gaussian Mixture Modelling algorithm. We have updated the end of the last paragraph of Section 3.4 to include this discussion.

Specific Comments:

1) Everywhere I would change the denomination "true wind" to "G5NR wind" throughout the text. No matter the quality of any dataset relating to physical quantities, it does not deserve to be called "true".

We understand that the term 'true' can often be controversial even when referencing a simulation. The denomination has been changed to 'Nature Run Wind' throughout the text. Thank you for the comment.

2) Line 144 It would be good to recall that this Figure relates to the first 1.5 months of the dataset, in the caption of the Figure.

Thank you. The distinction between training and test dataset has been made throughout the figure captions.

Printer-friendly version

Discussion paper

3) Lines 144-145 This is disappointing. Given the use of a powerful tool like GMM and the possibility of identifying "geophysical regimes" (line 132), I expected far more than just discriminating two groups, one being functional AMVs, and the other merely being the AMVs tracking the ground radiance, when the water vapour layer is too thin.

Figure 8-11, and lines 329-345 show that the clustering algorithm performs adequately in capturing consistent geophysical regimes. We focus in this paper on the 'skillfull' vs 'unskillfull' distinction because it is the most straightforward analysis for our purposes. More specific regime dependent uncertainties (as discussed in response to reviewer 2) is certainly a forward step after scaling this methodology beyond proof of concept.

4) Line 270 This parts misses a "is" between "xi" and "the".

Thank you for catching this. It was been corrected.

5) Section 4 The term Continuous Ranked Probability Score should be mentioned at least once before the formula at line 278. The two acronyms CPRS and CRPS are used in this section. Please correct.

The typo has been corrected. We mentioned the full name for CRPS immediately preceding its equation in (4), and we added a reference to a paper (Gneiting and Katzfuss, 2014) immediately after the equation.

6) Line 309 You are referring to Figure 13, and not Figure 12 as written.

Thank you for catching this. It has been corrected.

7) Lines 329-330 I find your conclusion a little daring, knowing that you had to try different numbers of clusters before actually managing to discriminate the null AMVs.

We apologize for the ambiguity. Our intention in these lines was different from what came across. We meant to say that our algorithm is able to 'find' or separate geophysically meaningful clusters without requiring domain knowledge expertise or prior information on the distribution of the variables. Granted, the algorithm requires the

users to slide the number of clusters across some scales, but this process is vastly simplified since there is only 1 scalar parameter to vary. As we noted before, having tuning parameters is par-the-course for the majority of data analysis methods such as k-means, PCA, self-organizing network, random forest, neural nets, regularized regression, smoothing splines, wavelets, etc.

We understand the referee's concern, however. Therefore we have removed the afore-mentioned lines in the Conclusion, and we have included a note about the need to optimize over the number of clusters in 2nd paragraph of the Conclusion.

————————————————————