Jan 23, 2020

Ref.: WAF-D-19-0244

Dear Mr. Short,

I am now in receipt of all reviews of your manuscript "Verifying Operational Forecasts of Land-Sea Breeze and Boundary Layer Mixing Processes." On the basis of these reviews and my own evaluation, it appears that the manuscript must undergo Major Revision before it can be considered for publication in Weather and Forecasting. These reviews are enclosed below or attached.

There is notable diversity in the reviews. Reviewer #2 (Minor revisions) notes that the paper is excellent and appreciates a paper that tackles the issue of how humans might improve upon model guidance. His/her main comment pertains to the figures and converting UTC to local time as well as better geographic arrangement along the ordinate.

Reviewer #1 (Major revisions) also thinks the paper is well written, but does have four major comments for you to consider. For example, he/she wants to see (1) a better explanation and more information about what “edits” are being performed and (2) more information/better statistics about the performance of various systems.

Finally, Reviewer #3 (Reject) notes that your paper contains some interesting concepts and does identify the importance of verifying human-edited forecasts (similar to Reviewer #2). However, this reviewer also has concerns about the methods, which affect the interpretation of results and subsequent conclusions; the varying size of the spatial domains is of particular relevance. And like Reviewer #1, he/she notes the methods could be better described. Most notably, this reviewer finds the paper difficult to read (in contrast to the other two reviewers). All of this is detailed in his/her three major comments and 15 minor comments.

Given that the comments from Reviewers #1 and 2 are relatively short, I have opted for Major revisions instead of Reject. I do believe that the comments from Reviewer #3 can be adequately addressed in the allotted time for revisions—perhaps with a minor extension granted if necessary.

Please upload your revised manuscript by Mar 23, 2020. If you anticipate problems meeting this deadline, please contact me to discuss an extended due date for your revision.

Along with your revision, please upload a point-by-point response that satisfactorily addresses the concerns and suggestions of each reviewer and the Editor. To help the reviewers and Editor assess your revisions, our journal recommends that you cut-and-paste the reviewer and Editor comments into a new document. As you would conduct a dialog with someone else, insert your responses in a different font, different font style, or different color after each comment. If you have made a change to the manuscript, please indicate where in the manuscript the change has been made. (Indicating the line number where the change has been made would be one way, but is not the only way.) Although our journal does not require it, you may wish to include a tracked-changes version of your manuscript as part of your response to the reviewers. Should you disagree with any of the proposed revisions, you will have the opportunity to explain your rationale in your response. No separate cover

letter to me is needed unless it contains essential information that does not appear in your reply.

Before submitting your revision, please carefully review the AMS Guidelines for Revisions found at http://www.ametsoc.org/PUBSrevisions to be sure you have complied with all instructions for quick processing of your revised manuscript.

To submit your revision, go to https://www.editorialmanager.com/waf/ and log in as an Author. Click on the menu item labeled "Submissions Needing Revision" and follow the directions for submitting your revision.

Please contact me at Bunkers.WAF@ametsoc.org if you have any concerns or questions. I look forward to receiving your revised manuscript.

Best regards,

Matthew J. Bunkers

Editor

Weather and Forecasting

Comments:

Reviewer #1: Drawing of the weather maps used to be one of the best example of human expertise and intervene in weather forecasts. From the time of super computers and high performance computers the role of human was limited to the up-gradation of initial and boundary conditions and further execute the model to improve the model forecasts. However, at the time of Artificial Intelligence (AI) and Deep learning Machine, how human intervene proved to lower the mean absolute error, is a quite challenging one??!! Perhaps, one can look on "Edits", as one of the other method where error and biased of the model can be reduced over a particular time but at coarser resolution. Some of my comments are as follows:

1. From the explanation on "Edits", I hardly understand what kind of "edit" is this to improve the model performance! The author completely missed to explain clearly, what are Edits? Edits is only choosing the base model forecast or there are many more rules and regulations to guide the daily forecasts. How does Edits differ from Data assimilation?

2. Similar to Fig 6, how does Official vs ACCESS perform for Airport station and City Stations? Official Vs HRES shows very low confidence over most of the locations. DAE (Difference in Absolute error) is a very simple skill score and been obsolete. What about the skill scores of forecasts based on probabilistic forecasts, e.g Ranked Probability Skill Score (RPSS)?! It is observed that most of the time probabilistic skills are superior than dynamical skill.

3. HRES and ACCESS carry lesser error and higher confidence (Fig 7) and compared to Edits. A method which is inferior in the forecast, in general been rejected. This research work shows the forecasts skills are very random in nature. Thus, with the facts shown in the illustrations, it is a bit hard to trust in these forecasts.

4. Knowing the "edits" are constrained to and much depended on the expertise of the human, who performs the edits. I think, Machine learning and AI will be superior options to choose a base model! However, the error of model bias and random variability in the forecast by edits will retain questionable?

Reviewer #2: In short, this paper is excellent, and provides a very interesting analysis of how model forecast guidance and "official" forecasts (via human edits to model guidance) are able to represent diurnal wind variations various at coastal sites across Australia. The analysis methods are rigorous, well executed, they are novel, and employ considerable insight into diurnal variability in the near-surface wind field. The manuscript is well written and organized, and the rationale for the various methods and analyses are clearly explained. Overall, the figures clearly display the pertinent information and are easy to follow, and provide some new and novel ways in which to view the performance of the various forecast products. I commend the author for this fantastic piece of work, and believe that this paper is an excellent and welcome addition to the Weather and Forecast and the wider body of research model and forecast verification.

I have one (very) minor comments for the author to consider.

In Figs. 2, 6-8, 10, it might be very helpful to convert the abscissa to Local Time, rather than UTC. Additionally, the groupings along the ordinate should be arranged from geographic locations in the east near the top, those in central Australia in the middle, and those in western Australia at the bottom. This follows the natural march of the sun across the sky during the day.

In its present form, the figures require the reader to mentally translate UTC to local time for each of the geographical groupings, particularly for those readers who reside outside of Australia. And the choice or order of geographic groupings is seemingly random, with Brisbane (far east) plotted 1 row above Perth (far west). This must be contributing to the "noisiness" in the plots, which the author frequently notes throughout the text.

I am motivated by the desire for the reader to able to instinctively determine the points on those plots where sunrise and sunset occur, and for the rows along the ordinate to be logically arranged by time zone, following the march of the sun. Doing so will make these plots far more interpretable, and far more accessible. Also, I have the sense that by re-ordering the abscissa to local time is likely to reduce the "noisiness" in the plots.

Reviewer #3: Review of paper submitted to Weather and Forecasting

Title: Verifying Operational Forecasts of Land-Sea Breeze and Boundary Layer Mixing Processes

Author: Ewan Short

I would like to thank the author for submitting his manuscript for review. The manuscript contains some interesting concepts, but at this stage, I do not believe it is ready for publication in Weather and Forecasting. My reasons for this recommendation are outlined below.

a. Does the paper fit within the stated scope of the journal?

Yes

b. Does the paper identify a gap in scientific knowledge and add new knowledge to the overall body of scientific understanding?

The paper does identify the importance of verifying human edited forecasts.

c. Is the paper free of errors in logic?

Yes.

d. Do the conclusions follow from the evidence?

I have some concerns about the methodology. Until these concerns are addressed, I do not know how much weight to put on the results from which the conclusions are drawn.

e. Are alternative explanations explored as appropriate?

Yes.

f. Are biases, limitations and assumptions clearly stated, and uncertainty quantified?

I am not sure the author is aware of the limitations in his methodology. For example, the spatial domains he defines are of widely varying sizes, and I think this may well make comparison of the results from different locations difficult.

g. Is the methodology explained in sufficient detail so that the paper's scientific conclusions could be tested by others?

I think the methodology probably could be reproduced by others. But the author could have made his descriptions of the method much easier for readers to understand.

h. Is previous work and current understanding cited and represented correctly?

Mostly. But the author either ignores or is unaware of the current forecaster practise of using a gridded consensus forecast as the starting point for manual edits.

i. Is information conveyed clearly enough to be understood by the typical reader.

I think this is a major weakness in the paper. The method and results could be conveyed much more clearly. This paper is difficult to read.

j. Are all the figures and tables necessary, appropriate, legible and annotated (as appropriate)?

No. I have provided a minimum set of improvements for the figures in the comments below.

Major comments

1. The author correctly identifies the procedures that Bureau of Meteorology forecasters use to prepare their official forecasts. First gridded guidance is loaded into the Graphical Forecast Editor (GFE), and then the forecaster makes manual edits to the guidance. It is correct that the forecasters have access to several NWP models for initial guidance, including the ACCESS-R and ECMWF models. However, additionally, there is a consensus forecast available called Gridded OCF (Operational Consensus Forecasts). This is the guidance which forecasters are now expected to use as their first guess in most situations. The Gridded OCF system is described in several "Operations Bulletins" available at:

http://www.bom.gov.au/australia/charts/bulletins/nmoc\_bulletin.shtml

In particular, Bulletin 91 may be of interest to the author:

http://www.bom.gov.au/australia/charts/bulletins/apob91.pdf

For the period of the author's study (mid 2018), forecasters may have used ACCESS-R and ECMWF guidance, but they will also have utilised the Gridded OCF guidance. Unfortunately, it is not clear if there are records of which guidance the forecasters used for any particular official forecast. Given that forecasters now mainly rely on Gridded OCF guidance, it is disappointing that the author didn't include this data in his study.

2. The difference of absolute errors metric (pages 7 and 8)

This metric is interesting. However, I have some concerns about its applicability to a vector wind field. In meteorology, wind fields commonly contain sharp discontinuities at fronts. It is not uncommon for the wind direction to change by very large amounts in a period of minutes.

The author attempts to identify diurnal cycles by subtracting a twenty four hour centered running mean from the observed or modeled wind. But consider this hypothetical situation: the first twelve hours of winds are northerlies at a constant speed, and the next twelve hours are southerlies with the same speed. The twenty four hour vector mean will be zero, and the perturbations at each hour will be quite large. I question whether in this situation the perturbation is of use for identifying diurnal cycles.

The above considerations aside, the difference of absolute errors metric may still be of value. The author may find that it would be possible to produce a shorter paper focusing solely on the applicability of this metric for weather forecast verification.

3. Coarser spatial scales (city station groups and coastal station groups)

The author looks at a number of spatial scales. The city station groups comprise the ten stations closest to each capital city. The coastal station groups comprise all stations within 150 km of the nearest coastline. I have the following concerns:

\* The area of the city spatial groups varies quite a lot. The largest city spatial group is three or four times the area of the smallest group. Is comparison of statistics from these differing sized groups valid?

\* It is debatable if Canberra can be considered a coastal area. It is too far from the sea to be subject to sea breezes.

I believe the above concerns need to be adequately addressed before the results can be published.

Minor comments and typos

There are numerous minor concerns with the paper. Some are listed below:

1. Line 98 and elsewhere: I believe the ACCESS model the author is using is ACCESS-R. There are other configurations (a global ACCESS-G and high resolution ACCESS-city models as well, so it is important to specify which model configuration is being used.

2. Lines 110-111: Although the ACCESS time steps may be in the order of 5 minutes, forecasters only get to see hourly data. It may pay to note this point.

3. Line 128: I think upscaled should be downscaled

4. Line 131-132, last sentence: A reference describing the standard approach the BoM takes would be helpful.

5. Pages 7 and 8: A diagram may help readers more easily understand how the perturbations are calculated.

6. Line 147 and elsewhere: u appears to refer to the wind vector. But in meteorology, the standard usage of u is for the west-east component of the wind vector. Perhaps a different symbol could be used to avoid confusion?

7. Line 151: "means" → "arithmetic means".

8. Page 10 and elsewhere: A map with place names would be helpful. Remember that the audience of Weather and Forecasting is international, so international readers will be less familiar with Australian geography.

9. Multiple places in the results section: Rather than mention UTC after all dates and times, at the beginning of the section you could mention that all dates and times are in UTC.

10. Figure 1:

1. No scale for the height colours

2. Inadequate place name labelling

3. State name abbreviations (WA, ACT etc) need defining.

11. Figure 2: The right hand side panels are difficult to read. The black numbers are hidden by the dark blue colouring.

12. Various figures and other places in the text: Wind speed units of knots are used. Although knots are commonly used in aviation (where the ICAO abbreviation is kt, i.e. different from the ISO abbreviation of kn), perhaps these units should be converted to m/s?

13. Figures 6-8 and elsewhere: Same dark colour problem as in Figure 2.

14. Figure 9 (and elsewhere): Where panels in a plot refer to different places, rather than labelling the panels a, b, c and d and then mentioning what these refer to in the caption, put a heading on each panel (e.g. Northern Territory, South Western Australia and so on).

15. Figure 11: The colour scale here is not very helpful; most of the boxes are the same red colour.