

Reviewer Comment 2

Review of “Experiments with large number of GNSS-RO observations through the ROMEX collaboration in the Met Office NWP system”

By Neill Edward Bowler and Owen Lewis

Thank you for your review. The original text is maintained in black, whereas our responses are marked in red. Any changes to the revised manuscript are also indicated in red text.

This manuscript presents a study on the impact of a large number of GNSS-RO data provided by the ROMEX project on the Met Office’s NWP system. The introduction is concise and very well written. It provides a clear and easy-to-follow overview of the GNSS-RO observations and the ROMEX project. The authors have done substantial work on this topic and cover many aspects in the assimilation of GNSS-RO observations from ROMEX. The results are solid and inspiring.

I feel that the presentation could be improved so that it can be more easily followed by readers from a broader community. Particularly, the authors use “to test the hypotheses” quite a few times to introduce their new experiment. This makes the presentation flow hard to follow sometimes. I suggest that the authors include a table listing all the experiments and their major configuration parameters before discussing the results, or alternatively, place the table in an appendix. Otherwise, I have only a few minor comments below.

Adding a table of experiments is a very good suggestion. We have added such a table at the start of the section discussing the results. In this table we have included the headline scores (i.e. the average change in the RMSE for that experiment).

Specific comments:

L1, GNSS-RO should be spelled out.

Done.

L 29, 39, 44,47, 56, 156, and many others....

I do not know the specific formatting requirements of AMT, but having references outside the parentheses seems weird. For example, in L29: “NWP forecasts have been gradually improving in quality over time **Bauer et al. (2015).**”

This error derives from the history of the manuscript (it was originally written in a format which only allowed the use of the \cite command). This has now been fixed.

L 40-42, I’m uncertain about the relationship between WMO and CGMS, but this phrasing looks questionable to me. Please verify.

The CGMS is an independent organization, and WMO members may make use of its recommendations. This paragraph has been updated to:

The Coordination Group for Meteorological Satellites (CGMS) provides recommendations on the number of observations that should be made each day

by the various observation platforms. Whilst the CGMS is not able to mandate the number of observations to be made, it does provide guidance to national meteorological centres and world meteorological organisation (WMO) members on the number of observations that should be made.

Fig. 1,

- a) why are TR Z forecasts missing in the scorecards? Is there any reason for this?

Geopotential height is not a useful quantity in the tropics. Therefore, it is omitted from all Met Office scorecards.

- b) The right panel of the scorecard figures shows verification against observations. Could the authors provide more details about this? Specifically, what types of observations are being referred to?

The scorecards are calculated against radiosondes. The word radiosonde has been added to the caption of Figure 1.

L133: Looking again at 2 -> Fig. 2

Fixed

Fig. 6, The figure was wrong in the original submission. While I can access the new figure attached in the Editor's comment, the authors need to include the new one in the revision. It is true that "The experiment which adjusts the observations by 0.05% approximately halves this reduction, and the experiment adjusting by 0.1% eliminates it entirely. This seems to be the main reason that the adjusted experiments perform better than the initial experiment". However, the forecast bias of the adjusted experiments increases with lead time. Please explain.

The error with Figure 6 has been corrected in the latest revision. We were perhaps imprecise in our explanation of this figure – it is the short-range forecast bias in 500 hPa height which is approximately halved by the 0.05% experiment. We have adjusted the text to hopefully be clearer:

The experiment which adjusts the observations by 0.05% approximately halves the negative bias in the short-range forecasts of this quantity. The experiment adjusting by 0.1% entirely eliminates the negative bias replacing with a slight positive bias, similar to the control NWP system. With increasing lead time, the forecast tends towards a positive 500 hPa geopotential height bias. The change of bias in the short-range forecast} seems to be the main reason that the adjusted experiments perform better than the initial experiment – they are able to remove the large negative bias in the geopotential height forecasts.

Fig. 7, Could the authors adjust the y-axis range for the humidity panel to improve visualization, given that there is nothing significant above 10–12 km?

Thank you for the suggestion – we have revised the humidity panel to only plot up to 15km for Figs 7 and 8.

Fig. 9, the paper presents many scorecard figures to show the overall impact of a test. As stated from L194 "There is evidently strong similarity between the results of this experiment and the ones shown in Figure 5, although not exactly equivalent." I

do not think Fig. 9 is necessary to be shown. A summary sentence could be enough. The authors can consider remove Fig.18 for similar reason too.

We have removed Figure 9, and updated the text to give the following summary:

To run an experiment equivalent to the ones above, k_1 was reduced by 0.1%. The results are very similar to those shown in Figure 5, and are therefore not included here. Changing the k_1 value gives slightly better forecasts of extratropical temperature and wind, but slightly worse forecasts of extratropical geopotential height. This highlights that a bias correction to the observations can have a very similar effect to an adjustment to the observation operator.

We feel that the differences between Figure 18 and Figure 10 are sufficient to justify its inclusion in the paper.

L 236, spell out DWD please.

We have added definitions for DWD and ECMWF.

Fig. 16a, please provide more explanation on “COSMIC-2 observations are an exception to this...”

Unfortunately, a limitation has been discovered in the ROM SAF monitoring software (which is written by the Met Office and used to produce graphs on the ROM SAF NRT monitoring pages). This software, as originally written, does not account for the drifting tangent point within a GNSS-RO profile. This only has a small effect on the calculated biases for most satellites, but is important for the COSMIC-2 satellites, due to their low-inclination orbit. Therefore, Figure 16 has been updated to use the corrected software which accounts for the drift in the tangent point of the observation. However, we have added an additional figure to demonstrate the difference that this limitation has in the case of COSMIC-2, as the experiments were run based on biases which were calculated with the original system.

L348, adding a reference could be helpful.

Unfortunately, we don't have a reference for this – we are told that a paper is in preparation on the upgrade. We have added (Yan Liu, personal communication) to note this.

Eqs. 3-5, I do not think these equations are necessary given these are standard statistics, but I am fine with them included.

We feel that they're worth including. One justification of this is to note that the standard deviation of forecast error is derived from the RMSE, rather than the square of differences from a mean. These are equivalent in a machine with infinite precision, but can lead to issues when accumulating very large amounts of data.