

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

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.

Specific comments:

L1, GNSS-RO should be spelled out.

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).**”

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

Fig. 1,

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

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?

L133: Looking again at 2 -> Fig. 2

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.

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?

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.

L 236, spell out DWD please.

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

L348, adding a reference could be helpful.

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