

Interactive comment on “Evaluation of the United States National Air Quality Forecast Capability experimental real-time predictions in 2010 using Air Quality System ozone and NO₂ measurements” by T. Chai et al.

Anonymous Referee #1

Received and published: 26 June 2013

1 General comments

The manuscript describes an evaluation of forecast skill of the "US National Air Quality Forecast Capability" for the year 2010. The evaluation is very well described, with a clear division into the relevant sub topics: spatial properties per region or per station type, and temporal patterns within year, month, or day. How certain statistics are computed is immediately clear from the text. For example the test on time shifts present between the archives is well described, even if it is not too important for the

C878

main results. The figures are well chosen, with only a limited amount of information shown; the notable features in each figure are then addressed in the text, which helps the reader to pick up the main messages. In summary, the manuscript provides a good overview of how to evaluate an ozone/no₂ forecast system, and in particular the system discussed here; therefore, the manuscript deserves to be published in GMD.

My main comment on the manuscript is the intended goal, which is not too clear from the beginning. The best description of the goal might be present in the very last paragraph, which starts to say that "The type of analysis presented here has guided recent updates". This is also mentioned as last sentence in the abstract, but throughout the paper I hardly find it back. Another goal might be that the paper will serve as a reference for future evaluations, or in papers focussing on the model improvements; if this is the case, it should be addressed more clear. Will this type of analysis be done every year, with some overview after 10 years or so? The paper also mentions at some occasions the difference between the operational system and the experimental system; results are solely for the experimental system. Isn't one of the goals that the experimental system in the end should replace the current operational system, when it is proven to have similar or preferably better skills? This would require a more detailed comparison between the skills of the two systems, which is probably left for other publications; if this is planned, it should be mentioned.

2 Specific comments

- Section 3
How many sites are actually used, and how distributed over rural and (sub)urban classes? Might have missed that while reading. From figures 8-9 one can see that some regions have hardly any stations; could this have an effect on the conclusions?

C879

- Section 3 and Figure 2
Why could be (in summary) the reason for differences ? The time shift is mentioned, but what about other causes (other correction factors, conversion to standards, etc) ? The use of circles in Figure 2 makes the spread much bigger than it is, given the very high correlation coefficient. A density plot would be more useful here. Correlation coefficients could be put in the figure too.
- Section 3
Very elegant method of testing on time shifts that might be present. This takes a large part of the section, might be useful to put this in a subsection.
- Section 3, p 2619, line 6
What is meant with "Without NO₂ measurements from AIRNow for the extra checking ...". Isn't only AQS used for the evaluation ?
- Section 4, p 2619, line 13-14
Is it possible to explain shortly the observation representation methods referred to?
- Section 4.1
For some types of NO₂ monitors it is suggested that these actually observe NO_y, see for example www.atmos-chem-phys.net/7/2691/2007/ . Is this taken into account ? Might increase the bias even more, but will help to judge how well the model represents the observations.
- Section 4.1
Pictures instead of tables might help to quickly interpreted the statistics. This holds for tables 1-4 and 7-8.
- Section 4.2
Why showing pictures for 2 summer months? Differences are not large.

C880

- Section 4.3
The acronyms HIT, CSI, etc are only expanded in the introduction; might be useful to have them here too. Also an intuitive explanation of what each measure tells you would be be useful, not this is only done for ETS.
- Section 4.5, p 2626, line 9
Is local time (LT) including the daylight saving regime, or is it actually a local standard time based on the longitude ?

3 Technical corrections

- p 2616, line 2
Isn't the second 48 hour forecast produced for 00Z instead of 06Z ?
- Fig 13, caption
Figure does not show the bias, but just the concentrations I guess.

Interactive comment on Geosci. Model Dev. Discuss., 6, 2609, 2013.

C881