

## ***Interactive comment on “Improving the deterministic skill of air quality ensembles” by Ioannis Kioutsioukis et al.***

### **Anonymous Referee #2**

Received and published: 23 August 2016

The paper analyses the two phases of the AQMEII initiatives to test different techniques for improving deterministic estimates from multi-model ensembles. Even though the paper is generally well written, my opinion is that the scientific novelty is scarce, and most of the conclusions are not solid. Here are my motivations:

1) As stated in the Abstract: Line 5-7 “we demonstrate. . .is far from optimum,”. This has been already proved several times in previous publications. (see Solazzo et al. 2013, Riccio et al. 2007, Galmarini et al. 2013, among others). In these papers, the same concepts and techniques of reducing the dimensionality of multi-model ensembles and optimal combination have been widely and repeatedly presented.

2) Pag 4 line 4-13. The differences between the two experiments are described. The differences in the meteorology (two different years) and stations (amount of observa-

[Printer-friendly version](#)

[Discussion paper](#)



tions and their locations) are those that undermine more the statistical significance of the results. Most of them are presented (see Table 2 and Table 4) without bootstrap confidence intervals or other techniques to assess if the differences between the two phases are statistically significant. The numbers of Phase I and II are often very close, and despite that, the authors build many conclusions on the top of these small differences. Also, most of the differences (if any) could be explained by the meteorology or the underlying changes in the station network. The authors should, at least, have made an attempt to make the two experiments more homogenous, i.e. by keeping a similar kind of stations over the two phases (same amount of urban, background stations).

3) Section 4.1 Forecasting performances. The authors want to prove that the weighting scheme might be used in forecasting mode. There are two issues here that undermine the conclusions of this section. My understanding is that some of the models participating at the inter-comparison are not running in forecasting mode (they use meteorological reanalysis as boundary conditions). While they should run as an operational real-time forecasting model to be considered as realistic forecasts. Running these model in forecasting mode would change the model behaviors and error structures. Hence the conclusions achieved might change as well. How the bias of the models is removed in this test? Using the bias computed over the entire period (as previously mentioned) to correct forecast issued over the same test period would not be possible in real-time forecasting. This simple bias removal technique might not be so effective especially in forecasting mode when data from the future cannot be used.

Minor comments:

The sentence: “In addition, mathematical tools such as ensemble forecasting provide an extra channel for uncertainty quantification and eventually reduction. Such method seems similar to the Monte Carlo approach; in practice, the similarity is only phenomenological since the probability density function of the uncertainty is not sampled in any statistical context like random, latin-hypercube, etc.” is not clear at all. Ensemble forecasting cannot be considered as a mathematical tool in general. What does it

Printer-friendly version

Discussion paper



mean:” Similarity in only phenomenological. . .”?

“benefits from ensemble forecasting arise from the averaging out of the unpredictable components (Kalnay, 2003).” It would be correct to say that benefits arise from averaging estimates with uncorrelated errors.

Pag 3 line 25 “One of the challenges in ensemble forecasting is the processing of the deterministic models”. This is true only if you are talking about a multi-model ensemble.

Eq1 bias, var, cov? Should be presented with a more detailed notation

Eq 2 E is the mean over what?

Line 6 page 8 keep the same stations over the two phases

Line 24 page 8 indirect feedback of what? Some details should be added

Line 19 page 5 I'd say the minimum (what does it mean ideal?)

Section 2.1 86 % is a general value or something related to this paper

The same bunch of authors (or most of them) appears in previous publications regarding AQMEII phase I and II. I have some doubts (but I might be wrong) that they all give an active contribution to this paper or at least original compared to what already provided in the previous publications regarding these experiments. It would be fair to include in detail a description of the contribution of each author to this paper.

---

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-513, 2016.

Printer-friendly version

Discussion paper

