

Interactive comment on “A regional air quality forecasting system over Europe: the MACC-II daily ensemble production” by V. Marécal et al.

Anonymous Referee #2

Received and published: 29 April 2015

This manuscript presents the status of the ensemble air quality forecast system for Europe at the end of the Monitoring Atmospheric Composition and Climate: Interim Implementation (MACC-II) project in the summer of 2014. The ensemble forecast system consists of seven regional air quality models. The median of the ensemble member values is evaluated against observations from a selected set of stations using a number of statistics. The selection of stations is unfortunate and should be revisited – it appears to be eliminating most of the urban stations, which are precisely those where air quality issues are often seen. The performance of ensemble median, which is said to perform better than any of the individual models, should be examined further: how does it perform for high pollutant values, does it capture the peak concentrations as well as an individual model? The robustness of the ensemble median should be examined for the cases when one or two of the best performing models are withheld from the en-

C578

semble. More specific technical comments are listed below, followed by selected edits and corrections as portions of the manuscript require English proofreading/editing.

p. 2743, l. 7: “In PMs, there is no distinction between primary (dust, sea salts, black carbon and organic carbon) and secondary aerosols formed from gaseous precursors such as SO₂, DMS, H₂S, NH₃, NO_x and VOCs.” This should be reworded. There is a distinction – it is ignored when composition is not taken into account, such as when considering mass or number concentration only.

p. 2744, l. 4: “...the approach based on a multi-model ensemble of forecasts has been developed to improve their quality through statistical approaches.” This should be reworded. The ensemble provides better information by combining information from different models - it does not directly improve the quality of forecasts from an individual model (although careful routine evaluation/comparisons may guide model improvement process over time).

p. 2744, l. 27: Introduce individual models, provide references.

p. 2745, l. 1: “Ensemble approach provides on average forecasts and analyses of better quality than any of the individual models” - Add a reference to support this statement. Demonstration in this manuscript is limited to ozone over only one week comparing the “best” individual member vs ensemble median.

p. 2746, l. 2: Do all models have only those 8 (4 earlier) vertical levels, or are these levels the only ones used in the ensemble?

p. 2746, l. 18: “All 7 models do not produce yet the analysis for all the 6 core species.” - How are missing levels/species handled in ensemble construction?

p. 2747, l. 8: “cut-off time at 07:00UTC on Day0 for the dataset covering Day0–1. At this time of the day, more than 90% (on average) of all data are available.” Is 90% of data for Day 0 really available by 7 UTC on Day 0, or should that be 7 UTC on Day 1? Define Day0-1, Day0-2.

C579

- p. 2747, l. 17: quantify “almost complete”
- p. 2747, l. 18-20: Provide specifics on data selection/quality control procedures - what is blacklisted, what thresholds are used, how is representativeness determined?
- p. 2747, l. 22: Is anything in data selection different from Joly and Peuch (2012)?
- p. 2748, l.1: Does station selection essentially eliminate all the urban sites? Are you verifying over rural sites only? This seems to ignore model performance in urban areas, where most serious AQ issues are often seen and weakens all of the subsequent results.
- p. 2748, l.4: How are these 20% of observations that are withheld from the assimilation selected?
- p. 2748, l.7: This needs to be rewritten as the list provided is not complete. Uncertainties in observations and assimilation methods impact analysis uncertainty. Initial condition uncertainties impact forecast and analysis uncertainty.
- p. 2748, l.10: The use of identical emissions and meteorology in all models minimizes the ensemble spread, not necessarily true uncertainty.
- P. 2757, l. 4: Define SIA.
- p. 2762, Figs 2 and 3: While ensemble median might be capturing day-to-day trends, it clearly underestimates observed elevated ozone values during daytime, so why is this statistic used to represent ensemble predictions?
- p. 2763, l. 24-25, Figure 4: It is not clear from Fig. 4 that the median captures the two ozone episodes well. Suggest replacing by a single map with enough resolution to be able to discern the level of agreement between model and observations. In the central European area it is difficult to see the model values among the observations – suggest zooming into that area. At the same time, there are no observations shown in the area of high ozone in Italy.

C580

- P. 2765, l. 12-23: Why are “best” and “worst” models removed simultaneously? This may be misleading as an illustration of the robustness. How robust are the results to loss of one or two of the best models?
- P. 2765, l. 24-25: “Note that we only illustrate here the behaviour of the ENSEMBLE over a short period of time, but these results are still true over longer time periods.” Please support this statement by providing quantitative statistics for a longer period of time.
- Figure 7: For each statistic, use the same vertical range for both years to allow comparison between the years.
- P. 2767, l. 23: Correlations of 0.7 at best, and often lower than 0.5 should not be called “high correlations.
- P. 2767, l. 23: “The changes in scores between 2013 and 2014 are not large enough to be regarded as significant. . .” Please demonstrate the lack of statistical significance to support this statement
- p. 2768, l. 1: “show a specific behaviour that is not analysed since not significant” Please demonstrate the lack of statistical significance to support this statement.
- p. 2768, l. 2-4: “The MB and MNMB both indicate a low bias in the ENSEMBLE. This can be linked to the fact that not all individual models include secondary inorganic aerosols and/or secondary organic aerosols.” – Demonstrate that this is the cause of low PM10 bias by stratifying the statistics by separating the models based on inclusion/exclusion of secondary aerosols.

Selected corrections/edits:

- p. 2741, l. 15: performances of the system are assessed → performance of the system is assessed
- p. 2741, l. 18 and throughout the manuscript: median ensemble → ensemble median

C581

- p. 2741, l. 23: The change of the skills → The change in the skill
- p. 2741, l. 28: new developments on aerosols → improvements in aerosol representation
- p. 2742, l. 22: for the information to people → to inform people
- p. 2742, l. 27: Volatil → Volatile
- p. 2743, l. 4: Particulate matters → Particulate matter
- p. 2744, l. 20: done → accomplished
- p. 2745, l. 6: from → by
- p. 2745, l. 9: give → provide
- p. 2745, l. 11: of → document/evaluate/quantify
- p. 2745, l. 13: short information on recent upgrades done before → a brief description of recent upgrades included before
- p. 2745, l. 15: performances on → performance for
- p. 2745, l. 16: perspective of → perspective on
- p. 2745, l. 22: with → at
- p. 2746, l. 3: The species concentrations available from the forecasts are firstly O3, NO2, SO2, CO, PM10 and PM2.5, called core species hereafter. → The forecast species include O3, NO2, O3, NO2, SO2, CO, PM10 and PM2.5, which are called core species hereafter.
- p. 2746, l. 12: new species are mainly designed for → additional species are provided primarily for
- p. 2746, l. 18: All 7 models do not produce yet the analysis for all the 6 core species. → Analysis for all 6 core species is not yet available from all 7 models.

C582

- p. 2746, l. 18: All models do not provide yet all additional species and vertical levels → All the additional species and levels are not yet available from all models,
- p. 2746, l. 25: processing 24 h per 24 h → processing 24 h segments
- p. 2747, l. 2: set to retrieve → put in place to gather
- p. 2747, l. 6: on → in
- p. 2748, l. 14 : on → in
- p. 2748, l. 15 : are on → stem from
- P. 2757, l. 21: researches → research
- p. 2758, l. 5: dusts → dust
- p. 2759, l. 10: new development → a new approach
- p. 2760, l. 9: firstly interpolated in → first interpolated to

Interactive comment on Geosci. Model Dev. Discuss., 8, 2739, 2015.