Review of Baró et al.

Round 2

Regional effects of atmospheric aerosols on temperature: an evaluation of an ensemble of on-line coupled models

for Atmospheric Chemistry and Dynamics, May 2017

The authors have generally addressed some concerns raised during the first review and have conveniently neglected to mention other concerns. The response to the first review is organized by blending comments from each reviewer by theme, since many of the comments were common between reviewers. However, in doing this, it makes it hard for the editor to see what has and has not been addressed. The authors may choose to not revise the manuscript if they disagree with reviewer suggestions, but they should at least say why.

Unaddressed Major Comments from this Reviewer from First Review

The paper could be greatly strengthened by looking at why the model results improve, or at least by providing additional information to help readers gain context. This could be done by looking at the energy budgets. The benefit of having the range of models in the intercomparison is that one can examine if the aerosol-related improvements occur for the same reason in each model, or if there are compensating effects that lead to interesting nonlinearities.

Another way the paper could be improved is by looking beyond daily values. Only looking at daily values hides a lot of model deficiencies. Comparing the models against hourly temperature data, as well as moisture and PM2.5 amounts, would provide much more detail for understanding why the models change when including aerosol feedbacks. This would also bring the plume behavior of local aerosol sources more into play.

Further Comments from Addressed Suggestions/Critiques

The concern was raised that the authors compare the models to one gridded temperature dataset, and differences between datasets could be bigger than the differences shown due to the aerosol impacts within the models. The author confirms this problem, and even goes so far as noting that he wrote a paper showing just this fact. No modifications were made in the manuscript and the response goes on to say that "the main objective of this work is not to rank the ensemble of simulations included in EuMetChem, but to provide a comprehensive comparison between simulations…" This is an OK objective. Unfortunately, the text in the manuscript, at the end of the introduction at lines 18ff, claims an objective that makes the accuracy of the observations paramount: "the objective of this work is to assess whether the outputs of an ensemble of regional on-line coupled models simulations including aerosol radiative feedbacks… improves the prognostic for maximum, mean and

minimum temperature at 2 meters over Europe." If the authors do not wish to add a comparison with a second dataset to take observation uncertainty into account, they should at least add text to the manuscript that puts the observations into context and note the limitation of the current study due to the single dataset.

The authors attempted to clarify the issue of what it means to turn aerosol-cloud interactions on and off within a model. Text has been added on p. 5 stating "Although NRF case does not consider the aerosol effects and feedbacks, there is a standard aerosol assumption of some continental aerosol (250 cm^{-3} used by WRF-Chem in the absence of ACI for estimating cloud droplet number)." However, as phrased this is a bit confusing. Please clarify. The sentence talks about aerosol assumptions and then provides a cloud droplet number concentration. The main issue is that the assumptions regarding physical mechanisms and sources of variability change between the two model configurations. The authors should double check with each modeling group using WRF to identify how they chose to not have aerosol-cloud interactions. This can either be done by not compiling in "chemistry mode" and then one gets the 250 cm^{-3} droplet number concentration for Morrison microphysics. Or, one can compile with the chemistry mode turned on but not use an aerosol module. The latter sets a constant aerosol number concentration (*naer* in the namelist) instead of a cloud droplet concentration. This is important because the physical processes related to activation and cloud formation change depending on the mode used.

p. 13, l. 16–19: The authors claim the following sentence has been corrected, but it still does not make sense: "In general, coefficients of determination are highest for mean temperature (0.60 to 0.78) and lowest for minimum temperature (0.50 to 0.56), presenting the ensemble always maximum values for ρ^2 (0.75, 0.79 and 0.61, respectively for maximum, mean and minimum temperature)." It is unclear what is meant by "presenting the ensemble always maximum values."

Figure 1 is blurred and unreadable.

Minor Comments

p. 4, l. 19: on-line coupled model~~s~~ simulations

p. 5, l. 20: Although the NRF case

p. 10, l. 6: have a notion of the aerosol loading (it would be better to reword to not use the colloquial phrase "have a notion" and replace it with "have an understanding of")

p. 10, l 8ff: The sentence starting with "Despite the work of Palocios-Pena et al." is phrased poorly. It would be better to refer to the other work for full details and to say the current article provide brief details for context.