Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2017-257-RC1, 2017

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Advanced error diagnostics of the CMAQ and Chimere modelling systems within the AQMEII3 model evaluation framework" *by* Efisio Solazzo et al.

Anonymous Referee #1

Received and published: 25 April 2017

In this work authors pursue their efforts in devising more sophisticated methods for error analysis in air quality models. In the light of experience gained during the AQMEIIs phases, they analyze results for ozone simulation for North America (CMAQ) and Europe (CHIMERE).

Given that a statistical analysis of model results can only be the first step toward a more-in-depth diagnosis of model deficiencies (it is difficult at this level to disentangle the impact of NOx/VOC chemistry, radiation, boundary layer dynamics and biogenic emissions, dry deposition, etc.), they report some interesting results which would be help in orienting modelers during model development or gaining confidence in using model predictions.

They suggest a combination of spectral (wavelet), time series (ACF, PACF and Kolmogorov-Zurberko filter) statistical tools and simple linear regression analysis to apportion model errors, applied to the decomposition of mean square error into its components (square bias, variance and covariance).

In my opinion, an intelligent approach is to subdivide the AQMEII domains into subregions and highlights the differences in model performances for each domain. For example a striking feature is the CHIMERE bust for the Po valley (EU3), associated with the diurnal component, opposed to the better behavior for EU1 and EU2 domains. Why CHIMERE performs better for north Europe? In a certain sense this work opens more questions than answers, but (I think) this is exactly the aim of authors. Though their analysis does not provide a solution to the problems raised during their evaluation, the combination of these statistical tools allows a better understanding of model deficiencies.

My major remark is that authors condense a great mass of information, difficult to assimilate without rereading back and back again. Moreover, they present results for different sensitivity scenarios, e.g. zero BC, const BC, 20% red, . . . Given a so large mass of information and different statistical analysis, I suggest, if possible, to re-organize the paper. I prefer a more in depth discussion of what may be the physical reasons for model deficiencies. A suggestion may be to try to highlight the role of physical mechanisms (this has already be done here and there, throughout the paper, but I prefer more emphasis). The logical course could be to start from model components (dry deposition, PBL dynamics, etc.), show what are model deficiencies and how your analysis is able to highlight these deficiencies. In this manner your ideas could are introduced as a "proof of concept" applied to a concrete example.

What is the role of dry deposition? Could the PBL dynamics better analyzed? The mean square error decomposition into its component and spectral components could identify where model need a deeper analysis? A plainer analysis of these aspects, moving secondary results to the supplementary section, would help the reader to track

the pieces of information better.

Overall, I recommend the publication of this work, since it summarize the efforts made during AQMEII3 phase and suggest useful statistical analysis, well beyond the 'standard' statistical metrics, often used to qualify model results.

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-257, 2017.