

## ***Interactive comment on “Multi-model evaluation of short-lived pollutant distributions over East Asia during summer 2008” by B. Quennehen et al.***

### **Anonymous Referee #2**

Received and published: 29 May 2015

This paper describes the evaluation of a number of global and regional atmospheric composition models over the East Asian region during 2008. It focuses on the summer period, coinciding with the Beijing Olympics, and addresses comparisons with both gas-phase and aerosol observations from a range of different measurement platforms and locations.

The paper is interesting, and is potentially very useful in characterizing the behavior of the models involved. However, the major weakness is that it does not provide sufficient interpretation and attribution of why models differ from observations or each other. Reproducing observations over this highly polluted region requires a good simulation of emissions, chemical processing, local dynamics, and regional meteorology. The analysis presented includes suggestions of the causes of particular differences (e.g.

C3100

that emissions of NO<sub>x</sub> are too high), but these are not firmly supported with clear evidence or convincing argument to back them up. A more thorough attribution of model differences is needed, so that developers and users of the models involved can make some progress towards model improvement, and so that others attempting a similar comparison are aware of the critical factors involved. This would provide a stronger message and make the paper much more useful to a wider audience.

### General Comments

The purpose of the model evaluation is briefly outlined at the end of the introduction, but includes several very different issues (air quality, climate change, long-range transport) which have different evaluation requirements. However, these issues are lumped together in this study. Please provide a clear indication of which parts of the evaluation are relevant to which issue, so that the wider implications are immediately clear.

East Asian emissions are rising rapidly and are less well quantified than in many other parts of the world, and they were temporally reduced by a poorly-known amount during the Olympic period. Given these large uncertainties, it is unclear why this region and period were chosen for the study. Please provide a clear justification for the choice. The impact of the Olympic emission reduction is one aspect of interest in many previous studies, but is not exploited here.

The main discussion sections (e.g. Section 3.5) are insubstantial, major biases are missing, and the links between biases that are described are not clearly and logically laid out. The result is that these sections are inconclusive and provide the reader with little new information about why the models and observations differ. While it may not be possible to diagnose model biases completely with the limited information available, clearer analysis of the data will provide new and useful information. One issue is that the satellites provide estimates of NO<sub>2</sub> (not NO<sub>x</sub>) and that emissions are of NO (not NO<sub>x</sub>). Biases in chemical processing thus strongly influence comparison with measurements, and biases are expected in this region where aerosol loadings can be

C3101

very high. Another issue is that no attempt is made to assess how representative the surface or aircraft measurements are of the wider region, and thus whether it is reasonable to expect the models to match them. A clearer and more detailed discussion of the expected chemistry and transport biases would make the paper much more useful.

#### Specific Comments

Abstract, l.6: "emissions of NO<sub>x</sub> are too high": or that chemical timescales or transport are incorrect. Addressing the contributions of these different biases is the key to strengthening the paper, and simultaneous evaluation of several variables across a number of models in this study should permit this.

Abstract, l.13: I agree with the statement "These results have important implications for accurate prediction of pollution episodes...", but the reader needs to know what the implications are, and how they could be resolved.

Section 2.4: It would be helpful to include a statement on the purpose of the CARE-Beijing flights to provide an indication of how representative they were of the spatial and temporal scales resolved by the models. Were they intended to sample urban outflow, or rural regions? What biases have previous analysis of these flight data identified?

Section 2.5: The meteorological data are highlighted as a source of model differences, but no indication is provided of how good they are (how close are they to observations?) or what differences they introduce. Fig 1 shows that the NorESM fields are very different from the others. What can we learn from pollutant comparisons if the underlying transport patterns are substantially different from those observed?

Section 3.1.1: Are these spatial or temporal statistics/correlations? This should be stated clearly here, and in the captions to the relevant tables. I assume they are spatial statistics, in which case the start of the section (p.11060,l.6) should state clearly that it is the spatial patterns which are being evaluated here.

p.11061,l.14: What does the good agreement actually tell us here? Does it just reflect

C3102

the latitudinal variation in tropopause height, or is there a more interesting story?

p.11062,l.26: observed "spatial" patterns?

p.11063,l.29: Biases are substantially worse at 0-6 km than at 0-20 km, as expected, as they are not dominated by stratospheric influences. What can you say about the source of the biases here? (e.g., overestimation over East China Sea, etc.)

Section 3.5: As noted in the general comments above, a clearer and more detailed discussion of the expected chemistry and transport biases in this section would make the paper much more useful.

Section 4.2: How are the aerosol composition biases identified here likely to influence the optical properties assessed earlier? The large overestimation of BC is a particular issue. Comparing the performance in different variables at the same time should allow stronger conclusions to be drawn.

p.11075,l.21: The short lifetimes aspect is important here (see general comment about Section 3.5), but you don't explain why this might be the case.

p.11075,l.25: Coarse resolution models are not appropriate tools for representing regional air quality, so this sentence simply supports the findings of many previous studies.

p.11076: The final paragraph of the conclusions makes a number of very vague and general statements about model weaknesses and about improvements needed, but there are no specific new conclusions derived from this study. Please sharpen up this paragraph to include some firm guidance on how weaknesses should be addressed.

Table 3: The NMB of the model mean is a lot larger than that of any of the contributing models. This looks odd, please check.

#### Typos

p.11052, l.11/l.26: "Hong Kong"

C3103

p.11052, l.21: missing "the" before industrial

p.11055,l.5-6: "allow to control" would be clearer as "address"

p.11056, l.18: grammar incorrect: perhaps remove "description"?

p.11067, l.6: Monks reference should be 2015.

p.11075, l.13: "climate" not needed.

---

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 11049, 2015.