

Interactive comment on “Variations of China’s emission estimates response to uncertainties in energy statistics” by Chaopeng Hong et al.

Anonymous Referee #3

Received and published: 28 September 2016

Overall Quality

Although this paper seeks to address an important topic and is well written, it lacks sufficient scientific merit for publication. It repeats the general strategy of an earlier paper by one of the co-authors (Guan et al. 2012) that sought to repackage the existence of large inconsistencies between different official Chinese datasets concerning energy as an analytical research finding. Those inconsistencies are important to understanding China’s air pollution and greenhouse gas emissions, but their existence is not newly recognized (Sinton 2001; Akimoto et al. 2006) and, more importantly, processing them with pre-packaged emission estimation protocols does not yield findings that should be considered publishable original research.

The paper essentially does the following. First, it assembles a set of publicly available

C1

energy datasets. Second, it processes these datasets using the pre-existing MEIC model for calculating atmospheric emissions. And third, it uses statistically questionable comparisons of the resulting disparities in both energy and emission datasets to draw inferences about the scale and sources of emission uncertainty. The mechanical processing of existing datasets using preexisting (and opaque) research tools is neither innovative nor novel. Importantly, it is also not reproducible, at least as currently presented. Last, the inferences about uncertainty are speculative, as no rationales for use of the metrics defined and employed in the paper are presented.

The extent to which the results are interesting is derived from the scale of the inconsistencies of the underlying data, not from the analysis itself. While the authors appear positioned to undertake a more rigorous assessment of their important topic, the current paper is too formulaic, unsupported, and speculative to justify publication.

Individual Questions/Issues

1. The paper is irreproducible, as it does not describe the methods of estimating emissions applied to different energy consumption datasets. It instead refers the reader to the website of the MEIC model, which does not present all of the underlying data and assumptions of the emission estimation model. To be reproducible, methods and assumptions must be described for each category of energy use (industrial subsector, for example, or vehicle type) treated uniquely in the assessment. Other researchers therefore cannot replicate the emission estimation as currently presented, except by blind trust in the same MEIC model.

2. The paper draws inferences about uncertainty based on two values defined in lines 9-10 of page 3: “We defined the apparent uncertainty as the maximum discrepancy among different datasets and the apparent uncertainty ratio as the ratio of the maximum discrepancy to the mean value from the different datasets.” These two concepts sound attractive but the rationales for their use to draw inferences about statistical uncertainty are currently lacking in the paper.

C2

Both concepts appear problematic. Regarding “apparent uncertainty,” a case has been made that a rough estimate of the uncertainty of energy data might be based on differences in values of subsequently revised data in the same official series (Marland et al. 2009). The rationale rests on a reasonable expectation that revisions represent increasing accuracy in the data and/or calculations, or “learning and convergence.” In the current paper, however, any connection to this rationale is lost because the authors simply compile datasets from different series (national, provincial, and IEA) and seek a maximum differential. Some sort of conceptual rationale for readers to find meaning in the value defined as apparent uncertainty is required for this calculation to be interpretable.

The “apparent uncertainty ratio” is problematic first because the numerator is apparent uncertainty, with the conceptual concern just noted, but then compounded by a denominator that is also hard to rationalize because of autocorrelation. Taking the mean of all datasets, including sequential revisions of the same dataset (CT-CESY-Ori, CT-CESY-1C, CT-CESY-2C, and CT-CESY-3C), implicitly assumes that they are independent. Without a defensible justification of this assumption, the calculations should recognize that revisions represent improving accuracy and should not be treated equally (as in a mean) in assessment of uncertainty.

The paper requires a more rigorously conceived statistical basis to draw the sort of inferences about uncertainty that it seeks as its primary conclusions.

Technical Corrections:

The authors need to revisit the above fundamental issues first before they (and reviewers) put time into other issues and technical corrections that this paper needs.

References:

Citations are referenced in the manuscript except:

Marland, G., Hamal, K., and Jonas, M.: How uncertain are estimates of CO₂ emis-

C3

sions?, *J. Ind. Ecol.*, 13, 4–7, 2009.

Interactive comment on *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-459, 2016.

C4