**Atmospheric
Chemistry
and Physics**
Discussions

Open Access

Interactive comment on "HTAP_v2: a mosaic of regional and global emission gridmaps for 2008 and 2010 to study hemispheric transport of air pollution" *by* G. Janssens-Maenhout et al.

Anonymous Referee #1

Received and published: 22 May 2015

The paper describes a key dataset that will be widely used. Overall the paper does a good job of describing the data, but there are a number of areas where further details would be useful, and a number of places where the current wording is not very clear and should be revised. Some of the analysis is a bit superficial, and could be easily improved given the data available to the authors.

Specific Comments

Page 12871, Line 17

It should be noted that the Lamarque et al. (2010) data used a similar methodology as

presented here for the latter years in that dataset: using country level inventories (as done here) for most OECD countries, combined with research inventories for Asia, and EDGAR for other regions.

Section 2, general

A table giving the general source and characteristics for the data in each world region discussed in these sections would be useful. Some of the characteristics that could be summarized include: original spatial resolution, monthly disaggregation (since there appears to be a mix of data with and without monthly detail, e.g. P 12879, lines 14-15), type of data source (research estimate, country inventory), and perhaps other relevant details.

It would be helpful to provide further details of the re-gridding procedures, in cases where grid cells overlap between datasets and countries. Were country-level gridded data available so that emissions from countries not included (Russia, etc.) in the HTAP_v2 could be cleanly removed, or did some approximation have to be made? (This is information that could be provided in supplementary text.).

Similarly in the later analysis when gridded emissions were re-aggregated, how were emissions in such "split" grid cells allocated to countries?

It would be of substantial value to the community if summaries of the monthly emission profiles for the HTAP sectors by major region (covering the different data sources, plus the three regions used for the EDGAR data) could be provided in a supplement. (e.g., normalized emissions by month) These assumptions are generally fairly hidden from users, and a variety of different sources for these profiles seem to be used in this data product, so it would be very useful to compare them.

12876, section 2.2.1

It appears that some of the year 2010 data is not actually for 2010, given this text from Pouliot etal 2014 "it was decided to use year 2009 data as the best approximation of

year 2010 for the EU anthropogenic emissions data. However, biogenic VOC emissions, which are driven by year-specific meteorology (see Section 5), were calculated for Europe for the years 2006 and 2010", and "the 2008 National Emission Inventory (NEI) (version 2, released April 10, 2012) produced by the U.S. Environmental Protection Agency (EPA) was used as the basis for both the 2006 and 2010". Given that the other reference for the North American data (Pouliot etal 2015) is not so widely accessible, so it will be important to give more details for this data than provided at present and clarify if actual 2010 emissions were used, or some approximation.

The authors need to carefully examine the sources of other datasets to make sure the "2010" data are actually for 2010 and not an approximation based on extrapolation of emission factors (or driver data) or substitution of another year. This will impact the implied emission factors analysis as well (if this is the case). While this likely does not change overall trends and conclusions, this is an important detail that needs to be better documented in the paper.

12877, line 10

It might be clearer to edit to read "TNO data for countries with only partial coverage in that dataset .... were not used"

12878, line 6

regarding "country-specific and sector- specific data monthly profiles per substance for the EMEP model", where these profiles constant within each sector, and country?

12880, line 6

"overview of the EDGAR emissions database and how it can be used for gapfilling can be found in Balsama et al. (2014)." This reference does not seem to be particularly relevant, particularly regarding gapfilling as done for HTAP V2.

12880, line 19

C2859

More detail is needed regarding how "monthly factors defined for the three different zones" are obtained. Presumably, given this description, these are not based on gridded data, but are the same for all emissions in a given sector in the three regions identified? (If so this should be clarified.) As mentioned earlier, it would be helpful if these were compared across the different datasets.

12882, line 20 "region keeps suffering from" is some unusual scientific language in this context

12883, line 3 International shipping emissions should be compared with the recent IMO projections "Third IMO GHG Study 2014".

pg 12886 - 12887 The point on consumption vs production-based emissions is a good one, but it would be useful to provide a more quantitative indication of how large this might be. There are several studies that compare GHG emission estimates, from production vs consumption perspectives, particularly CO_2. Since CO_2 is closely tied to energy, providing some numbers here for CO_2 would be instructive.

pg 12887, line 3-4 It would be useful to explore the reason for the " slightly lower than the reported value of $11.5 kg SO2 cap-1$ from EUROSTAT" value. The authors have at their fingertips extensive data on European emissions from TNO and EUROSTAT. Is this a grid sampling artifact? A difference between TNO and EUROSTAT numbers? This should be relatively easy to determine.

pg 12888, line 15 This text "the GDP is more difficult to cover with the various inhomogeneous economic activities, which are also influenced by time-dependent inflation and currency exchange rates and which are incomplete with the unrecorded unofficial activities." needs to be re-written to be clearer.

The sentence following is also awkward and should be revised.

Section 3.4

The analysis of implied emission factors is quite interesting and useful. Further de-

tails on how these are calculated, however, are needed. In particular, it is not clear how process and fossil fuel emissions are combined? The equation seems to imply that emissions and driver data are available in quite some detail, but my impression on reading the text was that this is not the case for the most part (with the exception of EDGAR data). Instead what appears to have been done is that total sector emissions were divided by energy consumption in TJ (with the exception of agricultural emissions). If this is true, it is this that should be represented in the equation. Also, if this is true, then some of these EFs will be skewed due to the inclusion of process emissions, since these are not necessarily proportional to energy consumption. For that reason, it might be useful to indicate where this is the case in the appendix tables using EDGAR data, since the process vs combustion split by country and sector is known to the authors for this dataset. (Even though the numerical values will differ from inventories from other sources, this will still give a useful indication of where the implied EF may be less relevant/accurate.)

The statistics (and ranges) presented could well be misleading since many small countries will have emission errors due to grid-resampling and general data issues (energy consumption statistics, particularly by country, are not very good for many of these countries, for example). Further, I suspect there are few independent data points for many of the small countries, instead these will be assumptions from EDGAR, likely similar emission factors, and extrapolations of more aggregate energy data. To deal with this, I suggest presenting present the same figures but only for the countries that contribute, for example, 80-90% of global world emissions (in order to determine if the statistics are skewed by the poor quality data for many small countries). This could be done instead of the current figures, or in addition (with the less relevant version in the supplement). Otherwise this analysis is a bit suspect.

This text "It should be noted that the aggregation of the country cells, taking into account the relative areal fraction of that 15 country in cross-border cells, needed to be corrected with country-specific reporting, in order to allocate point sources (e.g. power

C2861

plants) at borders (e.g. waterways) to the responsible country. " needs further clarification. (as noted above)

It is rather odd to see "The high SO2 implied emission factor might indicate the use of lower quality fuels in sea transportation, especially in international waters." since this data is from EDGAR and the authors know exactly what assumptions were used here. It also is well known through fuel sampling that shipping fuels are high in sulfur, e.g. see IMO report referenced earlier. Therefore, "might indicate" is not necessary.

Section 3.5

As mentioned earlier, the reality of the change in data from 2008 to 2010 should be further investigated. Its not clear if all these data actually represent estimates of emissions from these years.

Again, it seems odd to see wording "possibly due to the impact of increasing coal use". The authors should be able to easily determine if this is "possibly due" or not.

Section 3.6

"and lead to inconsistencies over borders," needs to be further discussed. Is this true for all borders, or just for boarders (relatively few, globally speaking) between where different data sources (EDGAR vs TNO) are used?

clarify text: "bottom-up methodology with activity data and emission factors is applied to calculate emission totals and distribute these on the grid." My reading of this manuscript is that the gridded emissions are from each original data source, not distributed to the grid as part of the HTAP dataset construction, so this was confusing.

Table 3 is interesting, but the justification for the uncertainty ranges should not be squeezed into the caption of the table. Instead these need to be further discussed (either in the text or a supplement) with a more complete discussion of the selection of uncertainty ranges.

C2862

Some of the details here should be re-considered. The quality of statistical infrastructure that is relevant to uncertainty in CO2 (as discussed by Marland et al, etc.) is primarily concerned with total fossil fuel consumption. For most of these pollutant emissions much more detailed information is needed, including fuel consumption by technology and well sampled and measured emission factors covering conditions appropriate for that country. Some discussion of sectoral differences in uncertainties would also be useful. (Traditional bio fuel use in many countries, for example, is not well quantified, and emissions from the use of this fuel even less so. So this is a substantial source of uncertainty in sectors/emissions where these sources are important.)

pg 12891, line 14 What does "as base year for the HTAP scenarios." mean? Clarify.

In this section it should be mentioned that these datasets are not independent (as GAINS work is informed by country-level inventories, common EF assumptions, etc.), so differences do not necessarily indicate uncertainty. (The comparison is useful, but needs to be put into better context.)

pg 12892, line 15

clarify "Another type of inconsistency occurs when speciation of a substance is done with gridmaps of different data sources."

clarify "Thereto a re-allocation of the emissions of some point sources (industrial facilities) was needed within Europe (e.g. Poland) and performed in consultation with TNO."

This sentence "Another check was to estimate per grid cell the change in emission from 2008 to 2010 and allowed to find missing sources." is unclear. The entire following paragraph is very difficult to understand and should be re-written.

pg 12893, line 6 clarify: "which is also regionally accepted as reference."

pg 12893, line 15

C2863

I'm not sure that "international standards" is quite the correct wording. Perhaps what is meant is "consistent international statistics"?

Figure 2

The captions for this figure could be usefully compressed by not repeating duplicate text.

The sector names here are not the same as used in Table 1, are they the same? Make consistent if so.

It would be useful to place the species name within each sub-figure (for example, perhaps on top of the center of the antartica region, which otherwise contains no information)

_____