



Interactive comment on “Global Covenant of Mayors, a dataset of GHG emissions for 6,200 cities in Europe and the Southern Mediterranean” by Kona et al.

Anonymous Referee #1

Referee comment on "Global Covenant of Mayors, a dataset of greenhouse gas emissions for 6200 cities in Europe and the Southern Mediterranean countries" by Albana Kona et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-67-RC1>, 2021

The authors present and describe a dataset named “Global Covenant of Mayors dataset of GHG emissions for 6,200 cities in Europe and the Southern Mediterranean”. The authors do a great job in assembling and maintaining such a huge amount of data and information.

If one would only read the abstract, it would perhaps expect a detailed analysis and description of a GHG emission dataset from cities, in the context of the current Paris Agreement and European emission reduction targets, and this is exactly what I was expecting. However, reading the whole paper, I have the feeling that it misses some substance in the actual meaning and use of this dataset.

Even if I understand that this is a dataset description paper and I should not expect much scientific results, somehow I was a bit disappointed not finding the real aim and scope of this study.

My point is, it is an interesting work but I was looking forward to find something about the actual results reported by these 6,200 cities, what is their current state in terms of GHG emissions, and how are they doing in the reporting process and Paris Agreement targets. This is not more than a methodological/abstract statistical paper which, in my personal opinion, would greatly benefit of a section summarizing the actual data presented by this dataset. It will give the general audience more insight into what actually this database is about – GHG emissions from cities.

For example one could present 1) a classification of top 10 emitters cities/large areas (in CO₂eq or kton gas) and their percentages with respect to the European GHG total emissions highlighting sectors and activities responsible for these emissions (useful for local mitigation strategies); 2) look at emission trends observed between the baseline year (1990) and last available year (2019)? 3) How do local governments make use of this data? Also, authors do not mention other initiatives like e.g. C40 cities.

Overall, I think this paper will be of great interest to all policy/science/industry users and, therefore, I encourage authors to put a little effort and if possible restructure and combine

this methodological abstract description with an overview of scientific results.

The manuscript would also benefit of a native English speaking review.

Therefore, I suggest major revisions and I am really looking forward to an improved version of this manuscript.

Below are my comments which hopefully will be appreciated and help in improving the manuscript, followed by a list of line-by-line specific changes:

- I miss at the end of the title a word, could be cities or countries? I accessed the online database (with more than a click, ESSD data policy states one click to download the data) and I see "countries" after Mediterranean. I would strongly suggest to add it as well to the title.
- In general throughout the manuscript, terminology is not always explained. Readers not familiar with this background need to clearly understand the terms and acronyms.
- Consistency in using acronyms and dataset names is poor. Since the beginning of the paper I stated to get confused by the use of GCoM and CoM, the authors should be consistent and use, in my opinion, only GCoM (online tables as well). I also miss references for many products. Transparency is the key to such data description papers and links to guidelines and regulations need to be provided.
- I strongly recommend the introduction of a Figure 1 to define boundaries and study area. Will be helpful to visualize the countries/cities.
- I somehow can't access the Supplementary files by clicking on the link in the manuscript. I could only view them from the preprint ESSDD website.
- The authors compare GCoM results with EDGAR v5.0 dataset (should be specified everywhere in the manuscript that authors used EDGAR v5.0) They state that "Overall, considering the completely different origin of EDGAR and CoM primary data the agreement has to be considered fully satisfactory". I personally think that a low uncertainty of two completely different datasets could either imply that both are good or that both are bad. Authors should explain better the methodology used to calculate uncertainties (perhaps add it as an Annex).

Specific line-by-line comments:

L3: Title: add "countries" after Mediterranean

L15: Please specify here and throughout the paper what EU are you referring to, I guess it is EU27.

L16: which latest IPCC report? Please add year and reference

L17: should be "emission inventory"

L17: I would reformulate: "To partly address this gap, we present a harmonized..."

L18: what do you mean with complete and verified dataset, please explain in a footnote "complete and verified" As far as I understood, it only contains the three main GHGs (CO₂, CH₄ and N₂O) what makes it complete?. To be mentioned, since the beginning, that you only report these three GHGs.

L19: "to complement the reported emission data"

L21: I would call them datasets (and not databases)

L22: EU (without `s) progress on the Sustainable Development Goal (SDG) 13 on Climate Action (I would add a reference here)

L22: "The datasets ..." please check the font size and note that this link does not work. Replace it with the one from the Data Availability section or references

<https://doi.org/10.2905/57A615EB-CFBC-435A-A8C5-553BD40F76C9>

Actually, the ESSD data policy states that access to data should be done with one click, I already clicked 3 times to download the data.

L25: I would delete Summary

L27: Perhaps authors could find a more recent reference?

L31: dot at the end of the sentence.

L35: authors state that there is an absence of a global cities datasets, true but other initiatives (e.g. C40 cities) should be mentioned.

L39: "gaps for Europe". In the beginning the authors were talking about EU, now is Europe, please use consistent domain.

L40: Mediterranean countries: on L61 is written cities: the authors should be consistent in using the same name of this dataset e.g. countries everywhere.

L41: now we have again EU, before was Europe. The Southern EU neighborhoods are the same with the Southern Mediterranean countries? Why not simply EU27 +UK and Western Balkans (I guess non-EU)? What countries are part of the EFTA?

As I already mentioned, this paper would really beneficiate of a figure with the study domain, cities/countries.

L49: delete space after neighborhood. It is very abstract, can you please define on the map all these groups? (EFTA, W Balkans, EU27, Eastern and Southern neighborhoods?)

L51: mention that in 2008 we had EU28; please add parties or signatories after 10,000.

L52: delete as of, in March 2020.

L55: please add: reporting based on which methodologies? IPCC same as for country reporting?

L58: The technical report...

L61: here is cities, on L40 is countries, be consistent. Please state which reference years are you talking about

L61 and everywhere in the manuscript: CoM and GCom, please use GCoM everywhere as stated by the title. If otherwise, make it clear in the beginning of the manuscript.

L64: the 4 steps are repeated in the Methods section, I would talk here in more general terms about the dataset and the context and leave the steps to be detailed as done in the methods section

L71: to be referenced which method and where is described e.g. see point 2 Methods or Supplement etc.)

L73: for someone not familiar with this type of terminology, should be explained who are the administrative units in the EU official statistics and link to signatories. Please add footnotes to define each group. Perhaps a list of acronyms would help as well.

L76: EU progress

L77: add (SDG), you use it later in the text.

L78: The datasets (Figure 1): which dataset are we talking about? GCoM? I see now different naming...please explain better the introduction of the "CoM dataset 2019: Emission Inventories" and how it links to the GCoM dataset.

Figure 1: I would actually have it as Figure 2, as mentioned, Figure 1 should present the domain.

Please explain in the figure caption what the acronyms stand for (e.g. NUTS, LAU, GDP, FUA etc.)

L93: perhaps here you can explain these terms: "complete, cleaned, validated and harmonized"

L97: Please add example of scientific and academic communities interested in using the dataset. It is important for the local governments and scientist to connect via such resources.

L99: Therefore, this datasets offers cities...

L100: since the beginning you mention GHG emissions: which GHGs?

L107: please include the comments under L64

L111-L115: Both 2.2 and 2.3 treat the Detection of outliers, 2.2 from large areas and 2.3 from small medium towns, why not call it:

2.2 Detection of outlier from large areas

2.3 Detection of outliers from medium towns

L120: Again GCoM appears, please use everywhere CoM or GCoM.

Perhaps the authors could present in the introduction some more information on the other platform, CDP-ICLEI Unified Reporting System, why these two platforms? what are the commonalities and differences with My Covenant, why cities should report to one or another?. More insight into methodologies governing these two platforms should be mentioned as well.

L122: The Common Reporting Framework (CRF) is a little bit misleading. CRF is an official terminology used under the UNFCCC reporting, the tables are named Common Reporting Format, can you please specify in a footnote that CRF is not linked to the UNFCCC process? Readers not familiar with it would think are the same tables.

L127: Can you give examples of boundaries?

L128: please reference IPCC guidelines, are the 2006 or the update 2019 Refinement or both?

L130: give examples please of "other city protocols"

L133: replace "they also can report" with "they can report as well.." and who are they? Cities, inventories, protocols?

Table 1: please replace "with IPCC guidance" with "based on the IPCC 2006 guidelines.

In general, always mention IPCC (2006).

In the column IPCC (ref no.) why 1.A.1 appears in all the rows?

In the Description column, please use capitals for ETS, and delete double spaces. Would be useful to add in the caption which GHGs are we talking about.

In general, all captions from this study are very short and not fully explanatory.

Waste: to my knowledge, waste sector and activities should be numbered with 5 (5A, 5B etc.).

L146: The geographical boundaries definition: should be mentioned before, on L127

L147: "regarding the type of gases" they should be detailed before in abstract and introduction

L149-150: there are two GWPs largely used depending to which IPCC AR report you refer to, AR4 uses 25 and 298 for CH4 and N2O respectively (and countries report them to UNFCCC CRFs) and AR5 28/34 and 265/298...which ones are you using??? please define as well GWP in a footnote, you use the GWP 20 or 100 etc.

L152: only ETS

L164: IPCC year and reference

L169: PostgreSQL database, what is it? And where is described? Please add reference to where can be found (Supplement etc.) or add footnote with explanation

L170: Consistency with climate action or Climate Action;

Is it "had been" the correct tense? these complete inventories were submitted and will never be submitted again? Or could be "have been submitted" because they still exist in present times?

L173: SECAP appears already at L69, please describe there what it is (add footnote with a link to it if possible)

L174: add comma after available

L175-178: these two sentences are repeating, can you please merge and make one sentence? Correct with "reported data, GHG emissions"

L178: replace with: Figure 2 shows the frequency distribution of 2019 GHG emissions per capita from emission inventory datasets, .."

L180-183: Could authors please rephrase the explanation for Figure 2? I don't really understand the density here and what exactly is intended with this figure? What is the width of the class, where was defined? Are too many i.e. which interrupts the readability of the sentence.

L200: please add after: waste sector, see Table 1)

L218: 2.2 as mentioned before, this section is also about detection of outliers from large urban areas, for a better flow of information I would:

- rename the section as "detection of outliers from large urban areas"
- add two subtitles Activity data and Emission factors.

L225: Eurostat database, please add version, year, reference.

L225: I think is "are classified as"

L232: please name some EU statistical systems

L233: "In case of reported data that ranges out of"

L234: "accuracy of the platform's reported data"

L237: "have been manually corrected"

L238: "therefore, we assume as valid the data reported in the SECAP document."

L243: IPCC AR4 (and reference) perhaps add a link to a chapter annex etc.

So, this means that you are using AR4 GWP values as well? see my notes on line 149, define it there please.

L249: rename with: detection of outliers from small medium towns (see also comments on lines 108-115) I would add here as well Activity data and Emission factors sub-sections to separate better the discussion.

L264: I think CAP is with capitals and please define it. Again consistency needed throughout the entire manuscript with EU-ETS, or ETS only or acronym versus EU-no acronym etc.

L265: could you please name these few exceptions in ()?

L269-274: Does this method is mentioned before on L202 as the rule to treat the outliers? Of yes, please explain it there.

L277 and everywhere else: should Supplementary files 1 and 2 be accessible with one click from the text?

L282: Now we have a new name of the dataset: "CoM dataset 2019: Emission Inventories" Is this the same with GCoM datasets of 6200 cities from the title? Is this study about this dataset? Is yes, please clarify this in the beginning, as it creates really creates confusion. We have now CoM and GCoM and "CoM dataset 2019" only and "CoM dataset 2019-Emission Inventory" etc.

It actually appears on L78 but in this paragraph you state that hereafter you will name it

as such.

L285: approaches, Similar to

L286: .."is motivated by the fact that observations..."

L292: IPCC 2006 and reference

L294: be consistent in the paper with space between the value and %, ± 50 %

L295: "the reference value should be (IPCC 2006/JRC)"

Table 2: Please add a more explanatory caption. The baseline year refers to 1990? should be mentioned.

L305: submitting their data **to** the platform

L320: Please introduce SDG acronym before (line 22) and be consistent again with the Climate Action in capitals or not, also L322

L322: now is Global Covenant of Mayors, why not as before, GCoM?

L324: "*we extracted the national values of GHG emissions per capita from EDGAR for the corresponding CoM activity sectors (Table 3).*" Please add v5.0 to EDGAR. When you say you extracted the national values of GHG emissions per capita I would expect to see emission values reported in the table, you only present source categories which are not clearly defined in the columns, not knowing which ones belong to which dataset.

Table 3: this table needs a better caption and explanation as one would not understand where is EDGAR v5.0? and what is the aim of this comparison? This mapping of emission source categories belong to which dataset?

L335: I would write: 4. "Comparison with ancillary emission inventories"

L337: This paragraph introduces already the uncertainty subject before the comparison is actually discussed, I would first present the comparison and after discuss uncertainty.

Everywhere: add EDGAR v5.0 please and references.

L345: when comparing GCoM to EDGAR, how about proxies used by GCoM compared to those in EDGAR? Could you summarize them in a table? This will make a clear discussion on why these datasets are completely different.

L349: explain LAU please; delete "s" from emissions (EDGAR emission grids)

L350: "and two source categories" delete for

L351: "energy in buildings (RCO)". Where was this defined? I do not find this sector in Table 3, for comparison purposes. Please explain where RCO is coming from, it is an EDGAR code? Why do you use code only for RCO and not for road transportation? I would delete RCO

L351: This general sentence: "EDGAR includes emissions from a variety of sources" it really needs a reference, examples etc.

L352: superscript for km²

In this paragraph also subscript for CO₂

L358: before on line 352 you used RCO, please be consistent, also with capitals (energy vs Energy)

L367: here you use " " for the Energy in building(s) sector, and singular for building

L367: is there a methodology for uncertainty calculation and what exactly this uncertainty represents? Very low values given that datasets are so different

L368: whereas without capital W

L372: "emissions in the Covenant framework cover"

L375: "emissions"

L376: Further limitations are discussed in the next section

L377 paragraph: I would still like to know more about the methodology used to calculate uncertainties

L390: I would use Secondly,

You state that "there is a limited knowledge on the methods by cities in determining the emissions..." this is exactly why methodology should become mandatory to GCoM submission system, parties must offer transparency when calculating their emission stating the guidelines, methodology and protocols they use.

L393: Can you please explain here the word "modeled" ? I am a modeler and I am familiar with the EDGAR database since a long time, I would not name it modeled, It calculates emissions as $EM = AD \times EF$ a simple global Tier 1 consistent approach for all countries, but not modeled.

"whereas EDGAR calculates emissions following a consistent Tier 1 approach based on AD and EF country specific information"

L393: replace suppose with assume

L393: based on collected activity data

L394: Replace Indeed with: "For these areas, there is a good match ..."

L397: Please replace modeling exercise with "emission calculation methodology"

L408: delete "also"

L408-410: Will be of great interest to find out more about CDP-ICLEI platform and I am still interested to know why are two systems and how do they differ.

L421: our colleague; replace deep reviewing with "in-depth review"