Comment on essd-2021-67

Anonymous Referee #2

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-RC2, 2021

Review of ESSD-2021-67: "Global Covenant of Mayors, a dataset of GHG emissions for 6,200 cities in Europe and the Southern Mediterranean"

Overall, this is a useful dataset and the work deserves a place in ESSD. I have a number of comments.

Firstly, there are numerous problems with language, which I'm sure ESSD's copy-editor will correct. However, passing the manuscript by someone with good English before submission would make the reviewers' job much easier. There are many situations where the language leaves the meaning unclear.

While it is several times mentioned in the introduction that this dataset is useful, and that the lack of such data has previously been noted (not "denoted"), the introduction does not say how these data would be useful. The fact that data do not exist is insufficient argument for their utility, and it is insufficient to say the data are "extremely useful" or "of great value and interest" without saying how and why. It seems that I wait until section 2.4 on page 11 to find what some potential uses would be, namely "support further research on investigating drivers of climate action at city level and the development of urban policy design." But that only appears to apply to the 'ancilliary data', leaving this reviewer with the belief that the authors don't actually see a use for the main data presented in the article. This cannot be the case. Please make the case for the utility of this dataset much clearer in the introduction of the article.

The dataset presented excludes some important components. This is stated in the article, but this must be made much more salient, since many readers do not read every word of the data description article. The first time this is mentioned appears to be on page 4, line 131: "the emission inventory is not meant to be an exhaustive inventory of all emission sources in the territory." Then later, at lines 152-3 the reader learns that emissions from "industrial plants involved in Emissions Trading Systems" are excluded. This in fact substantially reduces the utility of the dataset, because it must be combined with other datasets before it can present a comprehensive picture of emissions within a municipality. If these datasets are not going to be combined by the authors, then some clues on how to do this would be of significant use to the potential users of this dataset.

When presenting the histogram of per capita emissions, the text implies that there is at

least one municipality with zero emissions (line 170 "observations that range from 0..."). Some comment on how it is possible that a municipality can have zero emissions is warranted. Or if this is simply rounded down from some non-zero number, then please provide more precision to the number. It seems highly unlikely that a municipality could have zero emissions.

Lines 204-5: This is unclear. The statement is that the authors are making performance indicators on the impact and the contribution of climate actions planned and implemented by CoM signatories." But in what follows I cannot see such performance indicators. It seems rather than the assessments being made are informed by the potential uses of the dataset in assessing policy impact. Please clarify.

The headings of sections 2.2 and 2.3 do not seem appropriate to their contents.

Various statements are made about corrections to the data, but only in specific case (lines 279-80) is it made clear that such corrections are fed back to the reporting parties. One is left with the impression that all other 'errors' discovered are not reported back. Please clarify.

Line 306: A centroid is a weighted centre based on all points within an area. It seems very unlikely that each of the 6000 municipalities would report a geographic centroid, but rather the coordinates of their town square or post office or similar. Many will not know what a centroid is. Were parties really required to report their 'centroid' and do the authors really believe that centroids were reported?

Line 362: While the R-squared is unitless, the RMSE is not. What is the unit of the RMSE here?

Lines 367-70: I cannot make sense of these sentences. You say uncertainty is low, and then say this could be because EDGAR does it one way while CoM does it a different way, which could only be a supporting argument for uncertainty being high.

Line 371: Why have the R-squared values been ignored in this analysis, and only RMSEs used? The R-squared for transport is quite poor at 0.66, while the RMSEs (with unknown units) seem both to be relatively good. While 11% sounds low, 0.66 does not sound like low uncertainty, nor does 0.66 seem only "slightly" lower than 0.92.

Line 372: "the Covenant framework covers only the urban fraction", but given that you've used GIS to isolate EDGAR emissions in the urban areas, so does the EDGAR emission you're comparing with cover only the urban fraction. Please clarify.

Lines 377-8: The authors conclude from this analysis that the agreement between the CoM dataset and EDGAR is "fully satisfactory", which is for me very difficult to agree with. The poor reported correlation for transport emissions actually suggests that one potential use of this dataset might be to indicate that EDGAR's simplistic methods for spatializing transport emissions need to be improved. A 0.66 correlation does not seem 'fully satifactory'.

Line 396: Additional detail is provided here that should have been presented in the analysis in the previous section, not introduced in the conclusions.