

## ***Interactive comment on “Silicone v1.0.0: an open-source Python package for inferring missing emissions data for climate change research” by Robin D. Lamboll et al.***

### **Anonymous Referee #1**

Received and published: 24 June 2020

#### **1 General assessment**

The methods and software package described in the manuscript will prove useful for the IAM and climate policy analysis community and are definitely worth publishing. The IAM community will benefit from the easy to use and flexible methods to fill missing data in their output. The climate policy analysis community will benefit as well as incomplete data in scenario databases can hamper analyses and influence the results. Filtering for complete scenarios excludes some models and studies from the analyses while creating your own methodology to fill missing data is time consuming and takes away time from the actual analysis.

C1

However, the manuscript needs some work as some aspects are hard to understand and the structure of the manuscript should be improved.

It is very good that usage examples that actually reproduce the paper's results are given including code, however, the jupyter notebook contain some problems and needed fixes (at least on my computer).

Furthermore, there is still room for improvement of the methods especially for splitting gas basket pathways to individual gases.

I have some general comments followed by more detailed comments. I don't ask the authors to implement all of those comments, most are actually just questions and ideas and I think some answers might be interesting to more readers and could be answered in the manuscript.

#### **2 General comments**

- It could be helpful for users to briefly mention the package infrastructure used by silicone and if and how it is connected to the other packages in the IAMC toolbox (e.g. data transfer using a common csv format, or by API calls to other packages)
- For me an overview over the processing steps would be great. This would also help to identify possible further use cases and the options to expand the gas basket splitting functionality such that it can replace the EQW (see also comment below.) (see also comment on line 83)
- It would be good to have more structure in explaining the different crunchers and infilliers (e.g. paragraph headings for each cruncher) The paper will also be used as a reference for silicone and then it's good to easily find information on a specific issue. The information on crunchers and infilliers is somehow scattered through the document with bits in the methods section (text and tables) and some aspects

C2

in the results section. The overview tables are a good idea, but they should be placed in the methods section for the print version of the paper.

- It is a bit unclear how robust the results are and how much work is needed to check them before using the results. It is mentioned in some places, especially the tables, that incomplete databases can influence results. For easy use of the package it would be great if the user is warned in these cases. As one purpose of the package is to fill those gaps it would be a problem if the results are unknowingly influenced by the gaps. Maybe you could add a short section on limitations which gives an overview over such possible problems and references the tables for details. See also comments on lines 99-103 and lines 130-133.
- The gas basket splitting functionality seems to be not fully developed. The package has the potential to replace the EQW and its update used by assessments like the Climate Action Tracker, but in its current state it can not yet do that. It would be great if using QRW or EQW with a KyotoGHG constraint was possible for all gases. Could you add information if/how this can be added e.g. as a multiple infiller or a scaling postprocessing step? This would make the package very useful for the climate policy assessment community.
- RMS closest: please consider referencing the paper "Warming assessment of the bottom-up Paris Agreement emissions pledges" (YR du Pont, M. Meinshausen. Nature Communications, 2018) where a similar method has been used.
- It is very useful that the authors test the correlation between the different variables. The Spearman's rank correlation coefficient does not detect non-monotonic relationships which can be modeled by the Quantile Rolling Window method. Have you tried other methods to detect correlations (e.g. the Hoeffding Dependence Coefficient).
- Did you investigate non-emissions lead variables?

C3

- Have you used the method on other emissions databases (AR5DB, SSPDB, ...). Is this easily possible or does the package need a structure only present in the SR15DB?
- My comments on crunchers and infilling methods are a bit scattered as the comments follow the manuscript which also has information on crunchers and infilling methods in several places.
- Finally a more political comment. Why do you feature a fossil fuel companies' scenario in a climate change paper (Shell sky scenario)? This helps their greenwashing attempts. Please provide the scientific explanation for using this scenario.

### 3 Detailed comments

- Abstract: I think it would be good to mention some of the most important use cases and infilling options in the abstract.
- Line 70: It would be helpful for the reader to have short description of the structure of the paper at the end of the introduction.
- Line 83: Please give an overview over the infilling process at the beginning of the methods overview section or by expanding the introduction of the methods section (i.e. what are the steps in the process).
- Lines 99-103: Does the code have any flags to control behavior in these cases? Does it warn if this occurs or is it up to the user to control the results?
- Line 119: Where do the estimates averaged over come from? Different historical estimates for the same year, or different years as well? Is it possible to use

C4

trends of the last years with data or averages over the last years? If only single last historical data points are used (which is what I understand from table 1) the method is very sensitive to annual fluctuations in emissions data. Thus trending or averaging is needed for robust results that do not drastically change by using one additional year with data that has unusual emissions due to e.g. extreme weather or crises (financial, COVID-19, local crises, ...).

- Line 123: I assume "cases" refers to scenarios and scenarios where the sign for the lead variable does not coincide with the infill are taken out of the average for both lead and follower variable? Maybe revise language to clarify that. A more general point: it's a nice way to avoid numerical problems at transitions to negative emissions. However, when working on small scenario databases and / or with high temporal resolutions you might run into problems with availability of scenarios that use the same time period for the transition to negative emissions as the infill scenario. What do you do in those cases? Does the algorithm allow for a fallback option or does it just fail?
- Lines 130-133: As a lot of scenario databases are not fully consistent regarding data completeness (and sometimes sector definitions) it is dangerous to not automatically check for completeness or at least have an optional filter that removes incomplete scenarios.
- Line 138: Does "with all the same times" mean data availability for the same points in time? The use of "times" seems unusual and is not intuitive to understand for me.
- Line 142: Line 140 states that the average is taken over all points in time, yet line 142 states that "the value of the lead at one time impacts the whole timeline". That sounds contradictory to me. Can you explain? Also, adding a formula would help.

C5

- I have quite a lot of comments on the quantile rolling window cruncher because it was not straight forward to fully understand your algorithm. I rephrased parts of the algorithm in my own words to understand it and to check if I understood correctly. Maybe this can also give you hints on possible rephrasing to ease understanding for the reader:
  - If I understand correctly, the windows are not windows with clear boundaries but created by the weighting functions. I think the phrasing "Five windows are drawn" is a bit misleading because to me it suggests that the data-points are binned. However, my understanding is that you actually create a smooth distribution from each discrete data point such that at every lead emissions level chosen to calculate a quantile you have all data follower emissions data-points available, just weighted by the weighting function.
  - Please introduce an index for the data points used for the analysis. E.g. the formula in line 160 appears to use  $p$  as the sum index, but I assume it actually sums over  $w_p(e_l(i))$  for all scenarios  $i$  used at evaluation point  $p$  (lead emissions  $e_p$ ).
  - If I understand correctly, you calculate known quantile levels  $q_i$  through the summing of the ordered (by ascending  $e_{f,i}$ ) normalized weights. For each of the five lead emission levels you analyze and each point in time, you obtain a quantile level  $q_i$  for each follower emissions level  $e_{f,i}$  in the scenario database. This creates a lookup table where you can search for a given target quantile (usually 0.5, I assume), obtain the adjacent higher and lower quantiles and use these results to find the suitable follower emissions level by interpolation between the points.
  - As the quantiles are calculated for 5 lead emissions levels per point in time the above is done for two lead emissions level  $e_p$  and then linear interpolation is used to obtain the follower emissions for the actual lead emissions level. This leads to a further question: why do you use the approach of in-

C6

interpolating between only 5 points. When calculating the quantiles each time you use the method, you could just calculate them for the exact value of the lead emissions. If you pre-calculate the quantiles and use a lookup table to fill emissions (as done in the old EQW), you could use a higher resolution than just 5 points. I get the impression that you kept some aspects of the Generalized Quantile Walk that are no longer necessary with your improved method, because you don't have the binning.

- Line 149: Maybe add "for each time step" to clarify that the operation is carried out at each time step individually. This information could also be added in the description of the process in the following lines.
- Line 149: If I understand correctly you don't interpolate between quantiles, but between follower emissions associated to a given quantile at given lead emission values. If that's correct, please rephrase the sentence.
- Line 155: There is an additional full stop.
- Line 162: The formula uses  $r_i$  as sum index, but neither  $r_i$  nor  $r_p$  are introduced in the text. (I assume  $r_i$  are the lead emission data points in the scenario database ordered by the associated follower value and  $r_p$  is the highest one to be taken into account at quantile  $q$  (so should it not be  $r_q$  instead of  $r_p$ ??))
- Line 181: Why did you choose 0.03?
- Line 215: Did you analyze if the RMS closest cruncher chooses scenarios coming from the same model or study such that chosen pathways are very similar?
- Line 252: The sentence starting with "that" is incomplete.
- Line 256: Does the EQW use interpolation between quantiles and smoothing before calculation of the quantiles?

C7

- Lines 259 - 262: Did you test the influence of the weighting function on the conservativeness of the QRW cruncher. I assume that in the extreme case of constant weighting functions this would always return the median of the scenario database. So the smoothing through the weighting function has to be used carefully.
- Line 275: See also comment on line 123. How many scenarios are left when restricting to scenarios with same sign. This will especially be problematic for extreme stylized scenarios with early negative emissions.
- Line 279: a similar approach has been taken in "Warming assessment of the bottom-up Paris Agreement emissions pledges" (YR du Pont, M. Meinshausen. Nature Communications, 2018)

#### 4 Comments on testing the examples repository

- There are problems with directory names. The input data directory in the repository is called "input", while in the code is is referenced as "Input" with upper case "I". This leads to the code failing on unix systems.
- Splitting up basket: The notebook fails in cell 7, line 8 with: "TypeError: convert\_unit() missing 1 required positional argument: 'to'"
- Stylized Path: the legend of the plot resulting from cell 9 shows 4 cases of MES-SAGE scenarios, yet in the plot only two can be distinguished. Is that correct?
- Is it possible to store calculated quantiles for the QRW and EQW method to increase calculation speed by using a lookup table? itemize

C8

## 5 Comments on tables and figures

- Table1: latest time ratio: see comment on line 119. Please explain what re-harmonizing means in this context.
- Table1: Time dependent ratio: when using the optional filtering for same sign, how do you ensure that there is a sufficient number of scenarios available? See also comment on line 123.
- Table 1: Linear interpolation: Is it possible to use scenario binning here to not select two very similar pathways (e.g. same model and storyline with small parameter variation)? I think this could be made more stable by using more than two scenarios, but the problem will not vanish. Can you give an example, where using this cruncher makes sense?
- Table1: EQW: Does the EQW cruncher use smoothing through a weighting function when calculating the quantiles?
- Tables 1 and 2: please repeat the table header after a page break
- Fig 2b: there are some artifacts around the subfigures (partly visible plot boundaries?)
- Fig 3: The spike in the linear interpolation pathway shows that this cruncher is very problematic.
- Figs 4-6: please add a legend

---

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-138>, 2020.