

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2021-164-RC2>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Review of “CLIMFILL: A Framework for Intelligently Gap-filling Earth Observations”

Rene Orth (Referee)

Referee comment on "CLIMFILL v0.9: a framework for intelligently gap filling Earth observations" by Verena Bessenbacher et al., Geosci. Model Dev. Discuss.,
<https://doi.org/10.5194/gmd-2021-164-RC2>, 2021

Review of Bessenbacher et al., gmd-2021-164
“CLIMFILL: A Framework for Intelligently Gap-filling Earth Observations”

This study introduces a sophisticated procedure to gap-fill Earth observation time series while benefitting from independently and concurrently observed related variables. The authors showcase the method with reanalysis data where some parts are intentionally masked, and the reconstructed estimates are finally compared with the original data. Thereby, they consider ground temperature, terrestrial water storage, surface layer soil moisture and precipitation and discuss the results both in terms of reconstructed individual time series, and for the interactions between reconstructed variables compared with respective estimates from the original data.

Recommendation:
I think the paper requires major revisions.

This is a useful and timely contribution for the Earth science community, and interesting for the readership of the Geoscientific Model Development. Benefitting from a growing suite of Earth observations, complex statistical tools and machine learning applications are increasingly employed in Earth science research.

Mostly, these analysis tools require gap-free data which is often derived through gap-filling procedures.

In this context, improving the quality of the gap-filling by exploiting the relationships between the independent Earth observations is a promising avenue.

However, I have some concerns regarding the description of the method and the benchmarking of the results, as detailed below.

General comments:

(1) Comparing the results from the plain interpolation with that at the end of all four steps of the gap-filling procedure is interesting to understand the method and the relevance of the various steps. However, it is not a suitable benchmarking exercise as it is to be expected that the results after four steps are closer to the original ERA5 data than the result after the first relatively crude interpolation step. Instead, an established univariate gap-filling technique should be employed here as a benchmark to illustrate under which circumstances the presented methodology offers benefits over previous approaches. Also, this could reveal to which extent the gap filling can be improved by (i) complete exploration of uni-variate time series beyond neighbors, versus (ii) a multivariate approach.

(2) I think it would be useful for future CLIMFILL users to give more guidance on the methods to use in each step of the algorithm. Table 2 offers many possible choices, but in addition some recommendations would be needed on when to use which method and why. Also, the selection of employed variables is important as their inter-relations are a key source for the gap reconstructions, so also some additional advice on this would be helpful.

(3) I think that the feature selection is a bit arbitrary and dependent on expert knowledge. To somewhat address this issue, maybe several features could be used by default, such as the 34 features used in the presented example and maybe even additional time lags and windows. Then, the random forest model

can be employed to rank the features by their importance (e.g. using SHAP value importance) to make a more informed decision on the useful features. Finally, the gap-filling could be re-run with only retaining relevant features.

(4) There is advanced statistical and data science language used across the manuscript and I recommend to clarify this with additional information to allow a broader geoscientific audience to follow this manuscript. Please see my respective suggestions in the specific comments below.

I do not wish to remain anonymous - Rene Orth.

Specific comments:

line 2: estimates for what?

line 5: remove "up"

line 7: I agree that technically the algorithm does not require a gap-free donor variable; however if all variables have gaps at the same time and if this period is longer, then the final gap-fill estimate will naturally have a low quality

line 15: "profit", maybe rephrase as "are improved by"

lines 45, 144 & Table 1: Jung et al. 2019 and O & Orth 2021 are relevant studies in this context and could be mentioned here

line 46: please clarify "scale somewhere between"

line 84: please clarify "difficult observational record"

lines 108/109 and 111 are in contrast to each other

line 151: this is unclear, please rephrase

line 154: "another" should be "other" I guess

Table 2, caption: "other" should be "another" I guess

Table 2, right column: "or more complex interpolation methods", "Guided by ...", these are not exactly examples as the column title suggests

line 170: remove "on"

line 171: feels a bit random which letters are capitalized here and which are not

line 173: "the highly structured nature", please explain

Figure 2, caption: The framework is divided into four steps, not three.

line 178: Abbreviation CLIMFILL is mentioned earlier and should be explained at the first occasion

line 181: please clarify "correlation structure"

lines 203, 311: please clarify "constant"

line 216: quotation marks not needed

lines 229: please clarify "stabilising the results"

line 231: please clarify "terminal clusters"

line 243: I think this should be "to overwrite the former estimates"

lines 250/251: "learns different weights", please clarify

Figure 3, caption: replace "substracting" with "subtracting"

line 272: How are deserts defined and detected?

line 311: It should be 4 and not 3 additional features I guess?

line 314: please clarify "non-normality"

line 316: How does this add up to 34?

line 319: "respectively" should be added after "clusters" I guess

line 326: I wonder if and how different spatial resolutions can affect the accuracy of the gap filling,
it would be great if the authors could shortly discuss this.

line 326: "where one fold is one year", please clarify

Figure 7, caption: what is "CLIMPUTE-RF"?

line 351: please clarify "det"

Figure 8, caption: sentences should not end with "with" and "create".

line 361: "This" should be added before "leads".

line 367, section 3.4: I very much like the idea of studying the performance of the gap-filling across missingness patterns and different severity of the gaps.

Figure 10, caption: the B-distance is not actually displayed in this figure

line 373: How exactly are the satellite swaths imitated?

line 401: I do not quite understand the point on the bias correction.

line 427: similar in "remotely sensed" data but underestimated in "satellite observations", this should be the same thing?

Figure 2: The figure is rather small now and should be enlarged to make it easier to see all details.

Figure 4: The months axis should not go to 12.5

References:

Jung, M., et al., The FLUXCOM ensemble of global land-atmosphere energy fluxes, Sci.

Data 6, 74 (2019).

O, S. and R. Orth, Global soil moisture data derived through machine learning trained with in-situ measurements, Sci. Data 8, 170 (2021).