

We thank all three reviewers for their thorough and thoughtful assessment of our manuscript, and for the constructive comments and suggestions they have provided. We found the feedback extremely helpful, and it has allowed us to clarify our aims, strengthen the presentation, and improve the manuscript considerably. We have carefully considered each point and revised the manuscript accordingly; our detailed, point-by-point responses are set out below, with reviewer comments shown in black and our replies beneath each in blue.

Responses to R1

This paper is the fruit of tedious, carefully executed work. The authors are to be applauded. The listed sources might in fact contribute significantly to improved products and to improved understanding of climate and atmospheric circulation in Southeast Asia.

The procedure is rigorous and the criteria and methods are described in detail. The paper is well structured and well written. I have some more technical comments (see below). I recommend the paper for publication after revisions. This is an excellent contribution to the literature and to the data rescue community.

Comments:

- Did you also consider graphical data (such as thermo-barographs etc.)?

We did not find graphical data while preparing this inventory, however, we may soon discover barographs, or hydrographs (river gauges). Will be added to inventory as and when discovered.

- The inventory seems to list surface data only, please state this explicitly. It is confusing because you show a balloon launch.

We did not find upper-air observations at the time of initial submission, but we have found now and that relates to balloon launch photo shown. The inventory is updated with new items discovered since the initial submission. The flexible structure of the inventory allows for such later additions.

- I am also confused about Table 1, which also lists transcribed data, and the main file suggests this is a large portion if not the majority. From the text I understood the inventory to list un-transcribed records. Could you perhaps clarify?

Thanks for suggesting. This line is added at L219: The inventory contains both un-transcribed data-sources to help locate and plan for data-rescue missions, and transcribed data-sources to provide provenance and avoid duplicated efforts. Hope this clarifies.

- Your tables mark corrections (e.g., units), but do you also mark records that have been homogenised or merged?

Yes, we include fields such as pressure reduced (corrected) to sea-level (in Secondary tables), and a flag if the data-source is a statistical summary (in primary table). As historical data-sources usually do not contain homogenised or merged observations, this flag was not included in the inventory.

- It would be interesting to show a figure of the amount of information per different variables, if this is possible (I know this might be more difficult)

We could generate such figure but that would be misleading as unless a thorough data-rescue is executed accurate number of each type of observations cannot be known. Data-rescue of that scale is beyond our aims for this work. This inventory enables the data-rescue community to make informed decisions as to where to focus their efforts.

- The station table lists the resolution, and this is exclusively monthly. Are there no daily or subdaily station data?

Thanks for pointing it out. The inventory does contain sources with daily and sub-daily observations, the inventory file uploaded did not have it. Apologies now rectified.

- I realise that there are also many other columns whose category is always the same. In the main file completeness is always "continuous", quantitative is always TRUE, legibility is always "high" etc. for many others. The same holds for the station file: Is pressure really always reduced to sea level?

We are aware that some fields have 'default' values because the sample of images/pages of a data-source examined are usually continuous. Data-sources are only included if they are quantitative. Because most if not all stations are taken at a weather station or ship, the pressure is recorded as reduced to sea-level. We don't require so many flags now but as the diversity of data-sources increases, these flags help to record diverse meta-data.

- For all sources, copyright is listed as "public". However, I still think it is a rather delicate issue. I know that legally, facts or raw data cannot be owned nor copyrighted. However, there is still some kind of ownership in the data, and this ownership may not be with the person transcribing the data. Do you have any advice for the reader on how to proceed? Given the colonial history of many of the observations, this is might be relevant.

That's true. We have deliberately not included data-sources whose owners do not share images or data openly for research. For example, BKMG doesn't allow access to its archives either scanned images or raw data files. Many of the sources listed in the inventory do not make scanned images available online or have been since removed due to operational changes (e.g. NOAA). In the same way copyright policy might have changed as well since we obtained those images.

Even though we have scanned images of most of data-sources listed in the inventory, but the ultimate owners are the institutions where these data sources are archived. Our suggestion would be, even if one gets hands on the images, to notify data-source's custodians of the impending data-rescue operation and planned destination of rescued data. It is a good data stewardship practice.

- Avoid "Govt", "Dept" etc.

Changed. Thanks for the suggestions