

Earth Syst. Sci. Data Discuss., author comment AC2  
<https://doi.org/10.5194/essd-2021-454-AC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC2

Hongru Sun et al.

---

Author comment on "A compiled soil respiration dataset at different time scales for forest ecosystems across China from 2000 to 2018" by Hongru Sun et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-454-AC2>, 2022

---

### Dear Prof. Referee #2:

Thank you very much for your kind consideration and help to our manuscript! According to your suggestions, we revised our manuscript. All the modifications were listed as follows.

**Comment: Overall:** The dataset is a new compilation of soil respiration rates across China, available from a range of literature. The compilation has an impressive number of values collated. Whilst, by nature, the data is not unique on an individual basis due to being extracted from other sources, the compilation as a whole is a unique resource.

The dataset seems like a potentially useful compilation of values in terms of investigating climate change in China and globally. The values cover a broad range of climatic zones. The work attempts to standardise measurements in terms of different temporal scales, and explains the methods for doing this, which is commendable.

**Response:** Thank you very much for your kindly comments and encouragements. We are trying to establish a comprehensive and standardized forest soil respiration database across China, which could be useful in the related studies on carbon cycle and climate change in China and globally.

**Comment:** The discussion uses the data to link soil temperature to soil respiration, and thus climate change, and suggests further work could be carried out in relation to soil moisture.

**Response:** Indeed, soil moisture is an important factor. Many soil moisture datasets have been developed (e.g., Chen et al., 2021; Guevara et al., 2021; Meng et al., 2021; Wang et al., 2021), we are going to carry out the related analyses in the future study.

Chen Y, Feng X, Fu B. 2021. An improved global remote-sensing-based surface soil moisture (RSSM) dataset covering 2003-2018. *Earth System Science Data*, 13: 1-31.

Guevara M, Taufer M, Vargas R. 2021. Gap-free global annual soil moisture: 15 km grids for 1991-2018. *Earth System Science Data*, 13: 1711-1735.

Meng X, Mao K, Meng F, et al. 2021. A fine-resolution soil moisture dataset for China in 2002-2018. *Earth System Science Data*, 13: 3239-3261.

Wang Y, Mao J, Jin M, et al. 2021. Development of observation-based global multi-layer soil moisture products for 1970 to 2016. *Earth System Science Data*, 13: 4385-4405.

**Comment:** The WEBPLOTDIGITIZER method to extract values from figures seems interesting and an assurance of the quality of these data is given in Section 2.3.

**Response:** Yes, the extracted data with WEBPLOTDIGITIZER were verified in Section 2.3 in Lines 96-106.

**Comment:** Attempts are evident to show data consistency in collection in Section 2.2 in terms of only choosing Rs values measured from undisturbed ground, and in terms of the instruments used. It could perhaps be explained a little more as to the potential differences that could arise by using different equipment, and how this might affect the dataset.

The article overall is succinct, well-structured and clear.

**Response:** "The common measurement methods were selected, including Li-6400, Li-8100, Li-8150 and gas chromatography, which had been proved to be consistent" **was revised to** "Rs measurements were mainly from Li-8100 (47%) and Li-6400 (33%), secondary from gas chromatography (18%), and Li-8150 only accounted for 2%. The differences of the four common measurement methods had been proved to be small (~10%)" in Lines 261-263.

**Comment: Data quality** The dataset is easily accessible via the given identifier. I would normally expect a non-proprietary format for long-term storage/publication of data – e.g. comma separated values (.csv) rather than Microsoft Excel (.xlsx), for purposes of longevity, and to ensure the maximum number of users are able to open the dataset in freely available software.

**Response:** With your suggestions, the format of the dataset was changed to a non-proprietary data format (.csv) in the repository in Pangaea.

**Comment:** I would have expected the sample information (Province, Study site, Lat/long etc.) to persist for each data point, rather than there being rows of blank information. The assumption is that samples below the first instance of each Province, Study site etc., are the same/related, however, if you were to re-sort the spreadsheet, you would lose this associated information from the samples with blanks - each row is not 'stand-alone' as it should be.

This also means that it is not clear as to the difference between samples – for example, there are two data points marked with "Aug.,2013" – but what is the difference between the two? There is nothing on the individual rows to explain or describe.

**Response:** With your suggestions, the table was split into two related datasets with the same ID in the repository in Pangaea, one includes the soil respiration and temperature data and the other one the metadata of the sample information for each study.

Thanks for your reminder. Soil respiration was usually measured a few days per month. Means per month were only given in most studies, but a few values per month were all given in some studies. "monthly means or a few values per month" **was supplemented** in Table 1.

**Comment:** Again related to this, there is no way to automatically calculate the means in order to check their accuracy, because there is no field value by which to group the values to create the mean. The mix of data types in one column also precludes this – e.g. a numeric value column, with "NA" (string/character format) for missing values. I would have expected a numeric code to denote "NA", or a separate column containing the "NA".

**Response:** Means of soil respiration rate, soil temperature at 5 cm depth and 10 cm depth in each study were given in the columns of "Rs", "T<sub>5</sub>" and "T<sub>10</sub>", respectively. In order to calculate conveniently, the missing values (i.e. NA) in the numeric value columns of "Rs", "T<sub>5</sub>" and "T<sub>10</sub>" **were deleted** in the dataset.

**Comment:** Error estimates are not given in the dataset, although it is not clear as to whether this would be appropriate, based on the data extracted from the sources. Overall errors are presented in Figures 3 & 4.

**Response:** Root Mean Square Error (RMSE) was supplemented in Fig. 2, Fig. S1, Fig. S2, Fig. S3. "standard error" **was added** in Lines 176-202.

**Comment:** Whilst metadata is available in the article under review (e.g. table 1), I would have expected a metadata document (containing field level metadata) to accompany the data download, in addition to the summary given on the Pangaea landing page, which is not very detailed.

**Response:** A readme file was supplemented in the dataset in Pangaea.

**Comment: Specific Issues** Line 57 – bugedt -> budget

**Response:** "bugedt" **was revised to** "budget" in Line 57.

Thanks again for the reviewer and the editor for your kind consideration and help!

Best regards

Sincerely yours,

Hongru Sun, Zhenzhu Xu, Bingrui Jia