

1 Title: Soil respiration database at different time scales in forest ecosystems across China

2

3 **Dear Editor and Referees:**

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

6 **Referee #1:**

7 **Comment:** Soil respiration is an important indicator for a wide range of applications, especially
8 those related to evaluating carbon cycle regionally or globally. The authors did a great job for
9 collecting a total of 10288 monthly and 634 annual soil respiration data from 568 publications. I
10 enjoyed reading this well-reasoned and well written study. In a certain extent, it is helpful for
11 building robustness of this dataset for comparison between that from reference and software.
12 However, it still requires substantive effort for the following reason:

13 1. Authors mentioned Yu et al. (2010) established a geostatistical model with 390 monthly data and
14 Jian et al. (2020) analyzed the spatial patterns and temporal trends with 1782 monthly data. The
15 authors need to justify the importance of their research in comparison with these researches. For
16 example, using any quantitative method to address advantage of their dataset. It just looks like a
17 supplement for the research mentioned above right now.

18 **Response:** The importance of the dataset was discussed in the section "**4.3 Improvements of the**
19 **dataset**" in Lines 260-290, mainly including the following four aspects: 1) the large increase of
20 samples in annual Rs (N=634) and monthly mean Rs (N=5003), 2) the concurrently measured Rs &
21 T₅ (N=6341) and Rs & T₁₀ (N=2878), which were extracted from the figures in the original papers
22 with the digital software (WEBPLOTDIGITIZER) and not supplied in the previous datasets, 3) the
23 consistency of the selected measurement methods (Li-6400, Li-8100, Li-8150 and gas
24 chromatography), 4) Bamboo forests included, which were seldom considered in the previous
25 datasets.

26 **Comment:** 2. What is the difference of soil respiration among different equipment and method?
27 How does these affect the robustness of the dataset?

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

33 **Comment:** 3. Authors need to add more information for 'cross-checked'. It will be helpful to add
34 one table or figure to address the different sources.

35 **Response:** There are some data from the same authors and different sources (e.g. master or Ph. D.
36 dissertation and journal article). Here, "cross-checked" means the data from different sources was
37 checked. To avoid to misunderstand, "Moreover, the data has been carefully cross-checked by the
38 authors and from different sources." **was revised to** "Moreover, the data from the same authors and
39 different sources (e.g. master or Ph. D. dissertation and journal article) has been carefully
40 cross-checked and supplemented." in Lines 106-108.

41 **Comment:** 4. In Table 1, it is confused that the number of latitude and longitude are more than
42 Study site. I only found 251, 122 and 180 different values for latitude, MAT and MAP, respectively.
43 It needs more clarification for this table.

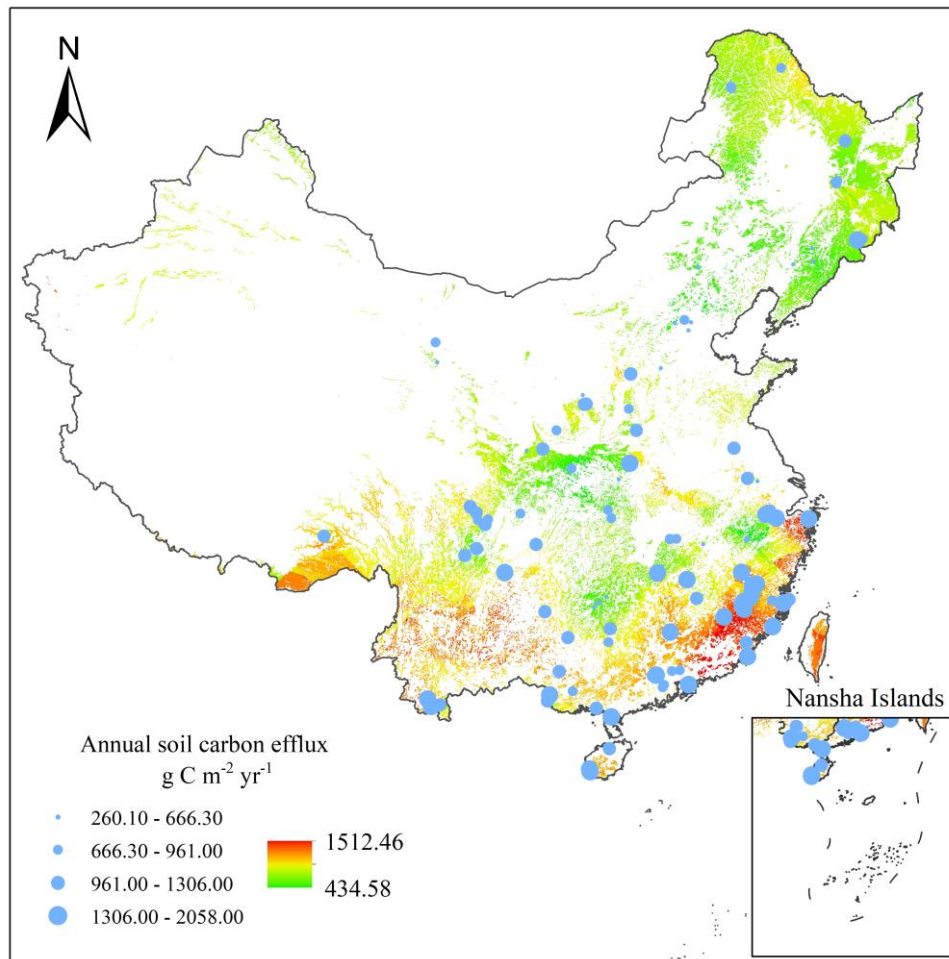
44 **Response:** Thanks for your reminder. In Table 1, we showed the number of the different study sites,
45 but the numbers of the latitude, longitude, altitude, MAT and MAP were all. Thus, to keep the
46 consistency, the numbers of the different latitude, longitude, altitude, MAT and MAP were listed in
47 Table 1, i.e. 208, 218, 329, 122 and 180, respectively.

48 **Comment:** 5. What are the patterns for soil respiration along MAT and MAP?

49 **Response:** The patterns for annual soil respiration along MAT and MAP **were supplemented** in
50 Figure S3.

51 **Comment:** 6. I strongly recommend that the authors rasterize this dataset to about 10 km resolution.
52 You can exclude the Northwest and Southwest part of China which did not covered with much
53 forests. Then, it would be more compatible to do analysis with spatial climate data.

54 **Response:** Thank you for your suggestion. This is a good advice. We rasterized this dataset to 1 km
55 resolution in the following figure. Yet compared with the large forest area of China ($188 \times 10^4 \text{ km}^2$),
56 the small samples (N=634) and their unbalanced distributions would increase the uncertainties.
57 Data coverage is a challenge for larger-scale extrapolation, and the simple interpolation analysis
58 based on this data may be inaccurate. In order to make a more precisely map, in the future study, we
59 are going to use more reasonable methods like machine learning, and have climate and soil
60 variables included as predictors.



61
62

63 **Referee #2:**

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

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

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

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

77 **Response:** Indeed, soil moisture is an important factor. Many soil moisture datasets have been

78 developed (e.g., Chen et al., 2021; Guevara et al., 2021; Meng et al., 2021; Wang et al., 2021), we
79 are going to carry out the related analyses in the future study.

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

82 Guevara M, Taufer M, Vargas R. 2021. Gap-free global annual soil moisture: 15 km grids for
83 1991-2018. *Earth System Science Data*, 13: 1711-1735.

84 Meng X, Mao K, Meng F, et al. 2021. A fine-resolution soil moisture dataset for China in
85 2002-2018. *Earth System Science Data*, 13: 3239-3261.

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

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

90 **Response:** Yes, the extracted data with WEBPLOTDIGITIZER were verified in Section 2.3 in Lines
91 96-106.

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

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

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

102 **Comment: Data quality** The dataset is easily accessible via the given identifier.

103 I would normally expect a non-proprietary format for long-term storage/publication of data – e.g.
104 comma separated values (.csv) rather than Microsoft Excel (.xlsx), for purposes of longevity, and to
105 ensure the maximum number of users are able to open the dataset in freely available software.

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

108 **Comment:** I would have expected the sample information (Province, Study site, Lat/long etc.) to
109 persist for each data point, rather than there being rows of blank information. The assumption is that
110 samples below the first instance of each Province, Study site etc., are the same/related, however, if

111 you were to re-sort the spreadsheet, you would lose this associated information from the samples
112 with blanks - each row is not 'stand-alone' as it should be.

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

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

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

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

127 **Response:** Means of soil respiration rate, soil temperature at 5 cm depth and 10 cm depth in each
128 study were given in the columns of "Rs", "T₅" and "T₁₀", respectively. In order to calculate
129 conveniently, the missing values (i.e. NA) in the numeric value columns of "Rs", "T₅" and "T₁₀"
130 **were deleted** in the dataset.

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

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

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

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

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

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

142

143 **Referee #3:**

144 **Comment:** Based on a thorough review of 568 original research articles and other publications, Sun

145 et al. compiled a comprehensive soil respiration dataset that covers a wide range of climates,
146 elevations and forest ecosystems across China. The dataset comprises a total of 10288 monthly and
147 634 annual soil respiration measurements and for some sites also monthly soil temperature
148 measurements (at 5 and/or 10 cm depth). In addition, specific information (geographic location,
149 forest type, mean annual air temperature and precipitation, etc.) are provided for each site. In view
150 of the vast number of independent soil respiration studies from different regions, consistent datasets
151 that summarise the state of the art and present the available data in a common format are of great
152 benefit for the research community as they facilitate, for example, the analysis of spatial variations
153 and temporal trends in soil carbon emission. The manuscript is generally well-written and the entire
154 dataset is made publicly available through the open-access data repository PANGAEA. However, I
155 have some concerns and suggestions that I would like to see addressed before I can recommend the
156 manuscript for publication.

157 As I am not an expert in the field of soil respiration, my comments focus mainly on the overall
158 content of the manuscript and the structure of the dataset.

159 **Response:** Thank you very much to take the time to review and improve our manuscript and the
160 dataset.

161 **Comment: General comments**

162 Uniqueness of the dataset: A similar dataset as the one presented by Sun et al. has been published
163 by Jian et al. (2020) for different forest ecosystems across China, although it is less comprehensive.
164 I therefore encourage the authors to clearly state the added value of their dataset compared to
165 previous studies.

166 **Response:** Uniqueness of the dataset was discussed in the section "**4.3 Improvements of the**
167 **dataset**" in Lines 260-286, for example, the consistency of the measurement methods, the large
168 increase of monthly and annual Rs samples, soil respiration rates and concurrently measured soil
169 temperature extracted from the figures with the digital software (WEBPLOTDIGITIZER), the
170 extension of forest types (including Bamboo forests).

171 **Comment:** Period of the dataset: I acknowledge the effort of the authors to screen 568 publications
172 and compile all the data, but it would be highly desirable if the dataset could be extended until 2020
173 so that it would cover a 20-year period (2000-2020), in line with the last two decades of the latest
174 climate reference period (1991-2020). Such a dataset would facilitate the joint analysis of
175 spatio-temporal climate and soil respiration variations (in the context of climate change). Similar to
176 the title for the data repository (Sun et al. 2021), the period of the presented dataset should also be
177 included in the title and abstract of the manuscript. Moreover, it would be helpful to mention the

178 (average) length of the individual time series from the different sites somewhere in the text or to
179 provide a respective figure (e.g. histogram) in the supplements.

180 **Response:** Thanks for your suggestions, it took us three years (from 2019 to now) to establish the
181 dataset, it's an interesting and valuable work, we will go on updating the dataset in the future study.
182 Title "Soil respiration database at different time scales in forest ecosystems across China" **was**
183 **revised to** "A compiled soil respiration dataset at different time scales for forest ecosystems across
184 China from 2000 to 2018".

185 Frequency distribution histogram of the length of the individual time series from the different sites
186 was supplemented in Fig. S4. "Fig. S4 showed the length of the individual time series from the
187 different sites, the high frequencies were 12 months (38%), 6–7 months (20%) and 13–24 months
188 (15%)." **was added** in Lines 280-282.

189 **Comment:** Uncertainties: As mentioned in chapter 2.3, most of the soil temperature and respiration
190 data (82 %) were extracted with the WEBPLOTDIGITIZER. This in an interesting approach that
191 provides a workaround to compile scientific data that are not made publicly available in the original
192 studies. However, I was wondering if the authors of these studies have been contacted to request
193 access to the numeric data or was this not feasible due to the number of studies? Were the data from
194 the 568 studies (many of them non-peer-reviewed) checked manually or automatically to identify
195 potential errors or inconsistencies?

196 **Response:** The database included 568 literatures, it was not feasible to contact all of the authors.
197 The data in figures were extracted by Hongru Sun and checked manually by Bingrui Jia.
198 Additionally, the data from the same authors and different sources (e.g. master or Ph. D. dissertation
199 and journal article) has been carefully cross-checked.

200 **Comment:** The given R^2 values of 0.99 for the simple linear regressions (original mean soil
201 respiration data vs. digitised soil respiration data) seem promising, but how does it look like for the
202 monthly data? As a measure for uncertainty, the RMSD or MAE should be provided as well. I am
203 also missing a section in the manuscript that discusses (at least qualitatively) the potential
204 uncertainties originating from the different instruments and experimental setups at the different
205 study sites as well as from the varying time periods of the datasets used for characterising
206 differences between the four climate zones (cold-temperature, temperate, subtropical and tropical).
207 Lastly, are there forest types (e.g. mountain forests) that are potentially under-represented in the
208 dataset (due to a lack of respective studies)? Such a potential bias might affect the numbers
209 provided for the temperature sensitivity of soil respiration and for the annual soil carbon emission
210 originating from forest ecosystems in China. This needs to be to discussed at least briefly.

211 **Response:** Yes, these were monthly data in Fig. S1, because most of the monthly data were shown
212 with figures and the annual data were directly given in the original papers. Root Mean Square Error
213 (**RMSE**) (also called the root mean square deviation, **RMSD**) was supplemented in Fig. S1.

214 "Rs measurements were mainly from Li-8100 (47%) and Li-6400 (33%), secondary from gas
215 chromatography (18%), and Li-8150 only accounted for 2%." **was added** in Lines 261-262.

216 "The spanning years were 2003–2014 in cold-temperate zone, 2000–2018 in temperate zone,
217 2002–2017 in subtropical zone and 2003–2017 in tropical zone." **was added** in Lines 172-174.

218 "It's worth noting that the Rs studies were fewer in the regions of latitude larger than 48 ° (~2%) or
219 elevation higher than 3000 m (~4%). The potentially under-represented forest types might affect the
220 evaluation of temperature sensitivity of soil respiration and annual soil carbon emission at the
221 regional and national scale." **was added** in Lines 287-290.

222 **Comment:** I would suggest the following modifications for the dataset:

223 Use a non-proprietary data format (e.g. CSV file) so that the dataset can be easily read by any
224 software.

225 **Response:** The format of the dataset was changed to a non-proprietary data format (i.e. CSV file) in
226 the repository in Pangaea.

227 **Comment:** Add a metadata file or readme file that contains all necessary information (e.g. those
228 from Table 1 in the manuscript) so that the dataset can theoretically be used independently of the
229 data paper. Nevertheless, add a reference to the data paper in the metadata/readme file and on the
230 landing page of the repository.

231 **Response:** A readme file was supplemented in the dataset in Pangaea. There was a reference on the
232 landing page of the repository in Pangaea.

233 **Comment:** Create a GeoPackage (.gpkg) or Shapefile (.shp) that contains the metadata (coordinates,
234 elevation, study site name, forest type, etc.) for each study site. Include it in the repository so that it
235 can be easily imported in a GIS by potential users for spatial data visualisation and analysis.

236 **Response:** Thank you for your suggestions. We created an ESRI Map Package (.mpk), the metadata
237 was included in the package and was added in the repository in Pangaea. Potential users for spatial
238 data visualisation and analysis can use this ESRI Map Package.

239 **Comment:** Add the units (of each column) either in the header or in the metadata/readme file.

240 **Response:** The units were added in the readme file in the repository in Pangaea.

241 **Comment:** Column "Month": Use the international date format (ISO 8601) or another common
242 date format that can be easier interpreted by a machine (e.g. 2013-07 instead of Jul.,2013). Replace
243 "Month" by "Date".

244 **Response:** Column “Month” **was revised to** a common date format (e.g. 2013-07). "Month" **was**
245 **revised to** "Date".

246 **Comment:** No need for column “Period” as the necessary information are already provided in
247 column “Month” (“Date”).

248 **Response:** Column “Period” means the period of annual soil carbon efflux and is necessary,
249 because some studies only supplied annual data, but not monthly data or only part of monthly data.

250 **Comment:** Column “Time”: which time is meant here? No time zone provided. It is unclear to
251 which data the time refers. Split into two columns as well: e.g. “Start” and “End”.

252 **Response:** Soil respiration is usually measured a few days per month, here, “Time” means the
253 observation time per day. "Observation time of Rs" **was revised to** "Observation time of Rs per day
254 (Beijing time)" in Table 1. There were 3 observation times (daily, monthly, yearly) in the dataset,
255 splitting into two columns would be more confused.

256 **Comment:** Remove the tilde in “Age” as this special character is difficult to handle during
257 automatic processing. Alternatively include a column before or after and use a flag (1, 0) to indicate
258 whether the “Age” is measured/precise (e.g. 1) or estimated/approximate (e.g. 0).

259 **Response:** Age of a natural forest is generally estimated from historical records or dominant tree
260 rings, and the age of a planted forest is defined since planting. Thus, the tilde “~” in column “Age”
261 was deleted, and "estimated from historical records or dominant tree rings in natural forest, defined
262 since planting in planted forest" **was added** in Table 1.

263 **Comment:** Column “Forest type”: would it be possible to use an integer or acronym for each forest
264 type in the database and provide the full name in the metadata/readme file?

265 **Response:** To give the detailed forest community of the study site, forest type was characterized by
266 the dominant tree species, or the ecological similarities (e.g. life form and biotope). The number of
267 forest type was 180, it was not suitable to substitute with an integer or acronym.

268 **Comment:** Columns “Rs”, “T5”, “T10”: Better remove “NA” (leave cells empty) and create
269 another column before or after indicating with a flag whether data are available (e.g. 1) or lacking
270 (e.g. 0). The same for the other columns where NA values exist.

271 **Response:** The missing values (i.e. NA) in the numeric value columns of "Rs", "T₅" and "T₁₀" **were**
272 **deleted** in the dataset. A total of 17 columns included "NA" in the dataset. If we created the
273 additional 17 columns, the dataset would become complicated.

274 **Comment:** Column “Annual Rs”: Does this column indeed provides annual averages or rather the
275 mean over the study period? I think it can be deleted as the mean can be easily calculated from the
276 monthly data provided.

277 **Response:** Column “Annual Rs” is the annual soil carbon efflux ($\text{g C m}^{-2} \text{ year}^{-1}$), not the mean over
278 the study period.

279 **Comment:** Column “Altitude”: Replace “Altitude” by “Elevation”.

280 **Response:** Column "Altitude" **was replaced by** "Elevation".

281 **Comment:** Although the redundancy may increase the file size of the dataset considerably, I would
282 recommend to copy the metadata (geographic information etc.) into each line (not only in the first
283 row from each site). Otherwise, complications may arise when the dataset is reformatted or
284 analysed. Alternatively, the table could be split into two related datasets. One would include the soil
285 respiration and temperature data and the other one the metadata for each site. An additional ID
286 could be provided for each study site to link the two datasets...

287 **Response:** With your suggestions, the table was split into two related datasets, one includes the soil
288 respiration and temperature data and the other one the metadata for each study.

289 **Comment: Specific comments**

290 Title: Mention the timeframe of the dataset (“2000-2018” or “2000-2020”) and replace “database”
291 by “dataset”. A database describes a collection of multiple datasets that are generally stored and
292 accessed electronically from a computer system... Title suggestion: “A compiled monthly soil
293 respiration dataset for [various] forest ecosystems across China from 2000 to 2018”

294 **Response:** Title "Soil respiration database at different time scales in forest ecosystems across
295 China" **was revised to** "A compiled soil respiration dataset at different time scales for forest
296 ecosystems across China from 2000 to 2018".

297 **Comment:** Line 64-73: Maybe just quote the database here and add the URL (with the access date)
298 in the reference list.

299 **Response:** The Five databases were searched the related soil respiration studies to compile our
300 dataset, it was more suitable to directly introduce them in the Section "**2.1 Data sources**".

301 **Comment:** Line 75-83: Indicate the period that has been considered. From 2000 until 2018 I think.

302 **Response:** "The observation years were from 2000 until 2018." **was added** in Lines 94-95.

303 **Comment:** Line 85: Do the 568 publications represent 568 study sites or are some data from the
304 same site? Please state the number of considered sites and their geographic and elevational
305 distribution somewhere in the text.

306 **Response:** Some data from multi-sources in the same site, forest type and author were merged in the
307 dataset. "The dataset covers 28 provinces in China ($18.61\text{--}52.86^\circ\text{N}$, $84.91\text{--}129.08^\circ\text{E}$)" **was revised to**
308 "There were 155 study sites from 28 provinces in China ($18.61\text{--}52.86^\circ\text{N}$, $84.91\text{--}129.08^\circ\text{E}$, $7\text{--}4200$
309 m)" in Lines 91-92.

310 **Comment:** Line 86-91: Why have the other provided variables (e.g. mean annual temperature and
311 precipitation as well as elevation) not been included in the analysis? I am aware that an in-depth
312 analysis is beyond the scope of this data paper, but some additional plots (soil respiration vs.
313 elevation, or soil respiration along selected temperature or precipitation transects) would emphasise
314 and showcase potential of this dataset.

315 **Response:** Thanks for your suggestions. The plots of annual soil carbon effluxes with mean annual
316 temperature and precipitation was added in Figure S3 in supplementary material. "The annual soil
317 carbon effluxes increased with the increasing of mean annual temperature and precipitation at the
318 national scale (Fig. S3)." **was added** in Lines 176-178.

319 **Comment:** Line 100: In addition to R^2 , provide the RMSD or the MAE as a measure for
320 uncertainty.

321 **Response:** "The Root Mean Square Errors (RMSE) of R_s , T_5 and T_{10} were $0.09 \mu\text{mol m}^{-2} \text{s}^{-1}$,
322 $0.35 \text{ }^\circ\text{C}$ and $0.44 \text{ }^\circ\text{C}$, respectively" **was added** in Lines 104-105.

323 **Comment:** Line 105-107: Please clarify whether this sentence describes the procedure in the
324 original studies or how you have modified and analysed the data.

325 **Response:** This sentence describes the procedure in the original studies, we didn't modify and analyze
326 the data. To avoid to misunderstand, "The typical days were usually selected to calculate mean monthly
327 R_s " **was deleted**.

328 **Comment:** Line 115: I am missing a brief section in the methods about the (statistical) analysis that
329 had been performed to present the data.

330 **Response:** The section "2.5 Statistical analysis" **was added** in Lines 122-132.

331 **Comment:** Line 119: Provide in addition to the total number of paired measurements (6341 and
332 2878) also the percentage (in parenthesis) with respect to the total number of considered data.

333 **Response:** "There were 6341 and 2878 samples of paired R_s & T_5 and R_s & T_{10} " **was revised to**
334 "The samples of the paired R_s & T_5 and R_s & T_{10} were 6341 (69%) and 2878 (31%)" in Lines
335 135-136.

336 **Comment:** Line 123-124: Note that R^2 is invalid/inappropriate for non-linear regressions! R^2
337 cannot differentiate between "good" and "bad" non-linear models. The standard error of the
338 regression could be used instead.

339 **Response:** "the coefficients of determination for tropical ecosystems ($R^2=0.225-0.291$) were
340 smaller than those in other three zones ($R^2=0.516-0.934$)" **was revised to** "RMSEs in
341 cold-temperate and temperate zones ($1.52-1.67 \mu\text{mol m}^{-2} \text{s}^{-1}$) were larger than those in subtropical
342 and tropical zones ($1.04-1.32 \mu\text{mol m}^{-2} \text{s}^{-1}$), except the smallest RMSE from T_{10} in cold-temperate
343 zone ($0.42 \mu\text{mol m}^{-2} \text{s}^{-1}$)." in Lines 140-143.

344 **Comment:** Line 125-132: Uncertainties need to be provided for these values!

345 **Response:** These Q_{10} values were calculated with the exponential equations between all soil
346 respiration and soil temperature data in each climate zone, but not the means. Additionally, RMSEs
347 of the exponential equations between R_s and soil temperature were given in Figure 2 and Figure S2.

348 **Comment:** Line 134-150: One big problem I see here is that time series spanning different years
349 were used to determine seasonal and geographic differences in soil respiration. I am aware that this
350 is unavoidable when using a compiled dataset, but the uncertainties originating from this issue
351 should be at least discussed qualitatively. Which period do the data cover that were used to compute
352 these values (2000-2018?). Add this information.

353 **Response:** "which derived from the similar years in cold-temperate (2003–2016), temperate
354 (2002–2018), subtropical (2000–2017) and tropical zones (2003–2015)." **was added** in Lines
355 151-153.

356 **Comment:** Line 149: What do you mean with “winter” in the (sub)tropics? This term does not fit in
357 this context.

358 **Response:** "winter" **was revised to** "November–April" in Lines 168-169.

359 **Comment:** Line 154-176: Are the data precise enough to state two decimal places? Confidence in
360 these numbers would be increased if uncertainties were provided.

361 **Response:** Yes, the annual soil carbon effluxes were precise enough to state two decimal places.
362 With your suggestions, the standard errors were added in Lines 176-202.

363 **Comment:** Line 154-155: Are these mean annual values averaged across all study sites? What’s the
364 considered time period?

365 **Response:** "Mean annual soil carbon emission was $851.88 \text{ g C m}^{-2} \text{ yr}^{-1}$ in China’s forest ecosystems,
366 ranging from $260.10 \text{ g C m}^{-2} \text{ yr}^{-1}$ to $2058.00 \text{ g C m}^{-2} \text{ yr}^{-1}$ " **was revised to** "The annual soil carbon
367 effluxes ranged from $260.10 \text{ g C m}^{-2} \text{ yr}^{-1}$ to $2058.00 \text{ g C m}^{-2} \text{ yr}^{-1}$ in China’s forest ecosystems, and
368 the mean was $851.88 \pm 12.75 \text{ g C m}^{-2} \text{ yr}^{-1}$." in Lines 174-176. The observation period was not
369 considered.

370 **Comment:** Line 160: These acronyms were not defined before. Please specify.

371 **Response:** Full names of these acronyms **were added** in Lines 183-185.

372 **Comment:** Line 165-176: Too many numbers in these paragraphs. I would recommend to provide a
373 comparative figure instead.

374 **Response:** The comparative figure had been provided in Figure 4, and the comparisons of annual
375 soil carbon emissions among different forest types and climate zones were summarized from Figure
376 4 in these paragraphs.

377 **Comment:** Line 177: A general discussion on uncertainties is lacking!

378 **Response:** "Form Fig. 4 we could also found that the standard errors in tropical and temperate
379 zones ($\sim 16 \text{ g C m}^{-2} \text{ yr}^{-1}$) were smaller those in cold-temperate and tropical zones ($\sim 65 \text{ g C m}^{-2}$
380 yr^{-1})." **was added** in Lines 254-256.

381 **Comment:** Line 179-196: A brief comparison with values from other regions outside China could
382 be added.

383 **Response:** Our results were compared with global forest ecosystems in Line 210.

384 **Comment:** Line 193: See previous comments regarding the use of R^2 and non-linear models.

385 **Response:** "The correlations between R_s and soil temperature were lowest in tropical zone
386 ($R^2=0.225-0.291$, Fig. 2d)" **was revised to** "Soil temperature at the depth of 5 cm and 10 cm could
387 only explain 29% and 23% of the R_s variations and RMSEs were $1.09 \mu\text{mol m}^{-2} \text{ s}^{-1}$ and $1.13 \mu\text{mol}$
388 $\text{m}^{-2} \text{ s}^{-1}$ in tropical zone, respectively (Fig. 2d)" in Lines 218-220.

389 **Comment:** Line 207-2018: It is difficult to compare any total numbers of R_s (from different sites or
390 studies) if the respective measurement periods, for which these numbers have been computed, are
391 not stated.

392 **Response:** The annual soil respiration data in these large-scale syntheses were also from the
393 publications. Due to the limit of annual soil respiration data, it is hard to compare the annual means
394 in the same measurement periods. In this paragraph, we compared the mean annual soil carbon
395 efflux with the previous studies, not the total soil carbon efflux.

396 **Comment:** Line 214: I would suggest to write the full names and regret from using acronyms if the
397 terms are only used a few times throughout the manuscript as in this case. This enhances the
398 readability.

399 **Response:** The acronyms of forest types were replaced by the full names in Lines 241-246.

400 **Comment:** Line 237: This number conflicts with the total number of R_s data ($=10288$) stated at the
401 beginning of the manuscript, doesn't it?

402 **Response:** A mean or a few values per month were shown in the original papers. The samples of
403 10288 included monthly means and a few values per month. "monthly means or a few values per
404 month" **was added** in Table 1.

405 **Comment:** Line 258: This sentence is a bit misleading as no in-situ measurement have been
406 performed. Specify that a comprehensive literature review has been conducted to generate the
407 dataset.

408 **Response:** "we collected in situ R_s measurements" **was revised to** "we reviewed the R_s -related
409 literatures and collected in situ R_s measurements" in Lines 297-298.

410 **Comment:** Line 251: Any kind of outlook is missing. Does this new compilation for example
411 indicates that there are particular regions or forest ecosystems that are under-represented with
412 respect to soil respiration and temperature studies and deserve more attention?

413 **Response:** "It's worth noting that the Rs studies were fewer in the regions of latitude larger than 48°
414 (~2%) or elevation higher than 3000 m (~4%). The potentially under-represented forest types might
415 affect the evaluation of temperature sensitivity of soil respiration and annual soil carbon emission at
416 the regional and national scale." **was added** in Lines 287-290.

417 **Comment:** Figure 1: I would suggest to include more information in this map. A digital elevation
418 model, hillshade, orthophoto or land surface cover classification could be displayed as a
419 background map. Different colours or symbols could be used for the study sites to indicate for
420 example the length of the time series (e.g. 1 year, 5 years, 10 years, 20 years), or the number of
421 available variables at each site (i.e. Rs, T5, T10). The overview map is too small and has no added
422 value. Better increase or remove it. It would also be helpful to indicate the considered climates
423 and/or forest types.

424 **Response:** Forest types were displayed as a background map, and the average length of the
425 individual time series from the different sites were added in Figure 1.

426 **Comment:** Figure 2: Note that R^2 is invalid/inappropriate for non-linear regressions (see previous
427 comment). Have monthly or annual data been used for the calculation? From which period do the
428 data originate?

429 **Response:** Root Mean Square Error (RMSE) was added in Figure 2. Monthly soil respiration rates
430 and soil temperature from 2001 to 2018 were used for the calculation.

431 **Comment:** Figure 4: How and for which period were the mean annual fluxes calculated? I assume
432 all annual data from different years and sites associated with one forest type were spatially and
433 temporally averaged. Is this correct?

434 **Response:** Yes, all annual data of the same forest type in each climate zone were averaged.

435 **Comment:** Figure S1: Does the mean reflects the entire study period at each site? Do the
436 correlations look similar if monthly data (collected vs. digitised) were compared?

437 **Response:** Yes, the mean reflects the entire observation period at each study. The correlations
438 (collected vs. digitized) were excellent and R^2 were all larger than 0.99.

439 **Comment:** Figure S2: I have the impression that there are other non-linear functions that describe
440 the relationship between the soil respiration rate and soil temperature better than the applied ones.

441 **Response:** Figure S2 showed the relationships of soil respiration rates with soil temperature at 5 cm
442 depth and 10 cm depth across China, Figure 2 showed the respective relationships in four climate
443 zones.

444

445 **Referee #4:**

446 **Comment:** Understanding soil respiration in large forest such as in China is of utmost importance,
447 and compiling a database with available data regarding this topic highly relevant for the scientific
448 community and stakeholders. Thus, the topic is worth of publication. However, clarifications should
449 be provided before publication.

450 In the Introduction, the aims of the paper must be clarified. It seems this will be a review paper but
451 latter the reader find there is data analysis also.

452 **Response:** Thanks a lot for your comments. "and analyze temperature sensitivity (Q_{10}), monthly
453 and annual Rs in cold-temperate, temperate, subtropical and tropical zones." **was added** in Lines
454 61-63.

455 **Comment:** In section 2.1, why was the search focused on publications performed 2018?

456 **Response:** We started from 2019 to collect the related literatures published up to December 31, 2018,
457 it took three years (2019-2021) to established the soil respiration dataset.

458 **Comment:** L83: it is not clear what you meant. Should this be provided in section 2.3?

459 **Response:** "Moreover, the data has been carefully cross-checked by the authors and from different
460 sources." **was revised to** "Moreover, the data from the same authors and different sources (e.g.
461 master or Ph. D. dissertation and journal article) has been carefully cross-checked and
462 supplemented.", and moved to section 2.3 in Lines 106-108.

463 **Comment:** In section 2.3 it is not clear why you used WEBPLOTDIGITIZER. It seems you used it
464 to extrapolate published data. Please, better explain. How much of your database is based on
465 published data and your estimations? Why the selection of 5cm and 10 cm soil depth? Was it linked
466 with data availability?

467 **Response:** About 82% of the Rs data were extracted from monthly figures with
468 WEBPLOTDIGITIZER, others (e.g., minimum, maximum) were directly given in the original papers
469 (see in Lines 99-101.). Yes, most studies measured soil temperature at 5 cm depth and/or 10 cm depth.
470 "Soil temperature as a main influencing factor, was usually concurrently measured with Rs. Monthly
471 dynamics of Rs and soil temperature at 5 cm depth (T_5) and/or 10 cm depth (T_{10}) were shown with
472 figures in many literatures." **was supplemented** in Lines 97-99.

473 **Comment:** L103: I guess there may be other equipment models to perform these measurements, so
474 I suggest to use the name of the equipment (e.g. infrared gas analysers)

475 **Response:** "Long-term continuous Rs could be monitored with Li-8100 or Li-8150" **was revised to**
476 "Long-term continuous Rs could be monitored with infrared gas analyzers (e.g., Li-8100, Li-8150)" in
477 Lines 110-111.

478 **Comment:** L105: "The typical days" – does it mean that you did not compile all the available data?
479 (the reader needs to know the aim of the study)

480 **Response:** No, all the available data were collected and compiled in the database. Soil respiration was
481 usually measured in the typical days with no rain in the original papers. To avoid to misunderstand, "
482 The typical days were usually selected to calculate mean monthly Rs" **was deleted**.

483 **Comment:** The methodology used for data analysis must be clarified. Did you consider different
484 climatic zones in the analysis? If so, which ones? The methodology used to calculate temperature
485 sensitivity must be explained in section 2.

486 **Response:** "Monthly and annual Rs were averaged arithmetically in cold-temperate, temperate,
487 subtropical and tropical zones." **was added** in Lines 123-124.

488 "Temperature sensitivity (Q_{10}) is defined as the factor by which Rs is multiplied when temperature
489 increases by 10 °C (Davidson and Janssens, 2006; Lloyd and Taylor, 1994), which is usually
490 calculated with the van't Hoff equation ($R_s = a e^{\beta T}$ & $Q_{10} = e^{10\beta}$), where Rs is soil respiration rate
491 ($\mu\text{mol m}^{-2} \text{s}^{-1}$), T is temperature (°C)." **was added** in Lines 127-130.

492 **Comment:** L114: how did you estimated this value? Based on a weighted average between
493 emissions and the extent of each climate zone? Please, explain in the methodology.

494 **Response:** Monthly Rs and annual soil carbon efflux were collected from the original papers, not the
495 estimated values. Monthly Rs and annual soil carbon efflux with other units in some studies were
496 converted to the common unit of $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ and $\text{g C m}^{-2} \text{ year}^{-1}$, respectively. "Monthly and
497 annual Rs were averaged arithmetically in cold-temperate, temperate, subtropical and tropical
498 zones." **was added** in Lines 123-124.

499 **Comment:** L158: how did you assess the significance of the differences? Please, explain it in the
500 methodology

501 **Response:** With your suggestions, we have added the description of the method in the section "**2.5**
502 **Statistical analysis**", "Independent-Samples T Tests (2 groups) and One-Way ANOVA (≥ 3 groups)
503 at the $P = 0.05$ significance level were used to test the differences among different forest types in
504 the same climate zone and among the same forest type in different climate zones." **was added** in
505 Lines 124-127.

506 **Comment:** L160: what is the meaning of EBF, ENF and DNF?

507 **Response:** "EBF, ENF and DNF" **was revised to** "evergreen broadleaf forest (EBF), evergreen
508 needleleaf forest (ENF) and deciduous needleleaf forest (DNF)" in Lines 183-184.

509 **Comment:** L162: Meaning of MF?

510 **Response:** "MF" **was revised to** "Broadleaf and needleleaf mixed forest" in Lines 187-188.

511 **Comment:** L165: so the analysis was also performed based on ecosystem type? Please, explain it in
512 the methodology

513 **Response:** "Independent-Samples T Tests (2 groups) and One-Way ANOVA (≥ 3 groups) at the P =
514 0.05 significance level were used to test the differences among different forest types in the same
515 climate zone and among the same forest type in different climate zones." **was supplemented** in
516 Lines 124-127.

517 **Comment:** Section 5: you don't need a separate section to provide this information. You can
518 include it in the Conclusions

519 **Response:** "Data availability" need to be given as a separate section in Earth system science data.

520 **Comment:** Fig.1: it will be interesting to show also the climate zones and if possible the main
521 ecosystem types

522 **Response:** Forest types were displayed as a background map in Figure 1, including evergreen
523 broadleaf forest, deciduous broadleaf forest, evergreen needleleaf forest, deciduous needleleaf
524 forest, broadleaf and needleleaf mixed forest and Bamboo forest.

525

526

527

528

529 Thanks again for the reviewers and the editor for your kind consideration and help!

530

531 Best regards

532

533 Sincerely yours,

534

535 Hongru Sun, Zhenzhu Xu, Bingrui Jia