Cover letter

Dear editor and reviewer,

Thank you very much for your great efforts, comments and suggestion! According to the third reviewer's comments, we revised our manuscript carefully and thoroughly, particularly on language issues from co-authors' efforts and few logic issues. Since our manuscript has been sent for a language proof before submission, we did not send our manuscript again, but combine co-authors' efforts on language. Please see, below, our point-to-point response.

Suggestion and comments from referees are marked in **<u>Black</u>**.

Responses to referee's comments are labelled in **<u>blue</u>**.

Cited changes made in the manuscript are marked in **<u>red</u>**.

Please do not hesitate to let us know if you have further questions and/or comments.

Sincerely,

Xiaolu Tang, Shaohui Fan and Wunian Yang, on behalf of all co-authors.

Comments to the Author:

Dear Authors,

One reviewer reviewed the revised version. The reviewer agreed that you made efforts to improve the manuscript but had some minor suggestions that you should follow during the final version. The reviewer also thought the paper has some linguistic problems. Please ask a native speaker to check the manuscript and correct these problems before the paper can be accepted.

Response: thank you very much for your efforts again. Since our manuscript has been sent for a language proof before submission, we did not send our manuscript again, but combine co-authors' efforts on language. Please see, below, our point-to-point response.

Suggestions for revision or reasons for rejection (will be published if the paper is accepted for final publication)

Tang et al presented to the carbon science community an interesting dataset of RH, which may be helpful for benchmarking terrestrial biogeochemical models. They also conducted a number of statistical comparisons with other datasets, including the Hashimoto RH, and the TRENDY model outputs, and found some very interesting and logically sound patterns. Overall, I agree with the previous reviewers that this study merits publication. And the authors have worked hard to address the concerns raised in the first 3 reviewers. That being said, however, I still think the paper has some linguistic problems (more than grammar and syntax), which need to be carefully attended before the paper can be accepted. I will give my details below.

Response: thank you very much for your comments and suggestion to improve the manuscript. As suggested, we revised our manuscript carefully again with a focus on language problems. Please see the point-to-point change and the revised manuscript with changes tracked.

P1, L21-22: "high variations and uncertainties of RH existing in global carbon cycling models require an urgent development of data-derived RH dataset". Although I understand the value of a data-derived RH dataset, I don't think the logic here is well established. For instance, I can equally say a more mechanistically based modeling approach is more urgently needed. What a data-derived RH dataset will do is to provide a different angle to quantify the RH besides the biogeochemical modeling approach.

Response: thank you for the good suggestion. We agree with the reviewer and our RH dataset could provide another different angle to quantify RH. We revised the manuscript as:

"However, high variations and uncertainties of RH existing in global carbon cycling models require RH estimates from different angles, e.g. a data-driven angle."

P2, L51-2: "Although it is widely recognized that warming enhances CO2 release from soils, the magnitude of release is uncertain due to variations in the temperature sensitivity of soil organic matter decomposition". This statement is generally right, but insufficient. Because when talking about climate responses (aka warming here), the uncertainty involves both temperature and moisture, with all other granularities under the hood. I believer, this sentence need be revised to better connect with the next sentence.

Response: thank you for the suggestion. As suggested, we revised the manuscript as:

"In addition, environmental drivers of RH, e.g. temperature and soil moisture, are still undergoing changes under climate warming and can affect RH individually or interactively."

P3, L66-67: "potential new numerical/algorithmic methods to better quantification and understanding of large-scale soil carbon flux". There are misuses of adverb, and verb here. Please correct.

Response: we are sorry. the verbs "quantification" and "understanding" were correct to "quantify" and "understand", respectively.

P3, L79: "empirically" should be "empirical".

Response: done.

P3, L89 and throughout the paper: "global vegetation models" is not the right jargon, "terrestrial biogeochemical models" or "land biogeochemical models" are more often used.

Response: done as suggested. We replace "global vegetation models" with "terrestrial biogeochemical models".

However, we kept Dynamic Global Vegetation Model from TRENDY model ensembles, because this is the office name and widely used in publications (Ballantyne et al., 2017; Piao et al., 2019; Sitch et al., 2008).

P6, L169: In what way "the other two proxies were controlled"?

Response: We meant "the other two proxies were controlled to remove their confound effects on RH", and we revised as:

"When analysing the partial correlations between RH and the proxy, the other two proxies were controlled to remove their confounding effects on RH."

P7, L218: "reflecting less stress from environmental limitations". This statement is unclear. The first

half of the sentence made a contrast, manifesting a gradient-like transition, but this statement does not describe this clearly.

Response: sorry for the unclear statement. We revised the sentence as:

"reflecting a higher resource limitation in high latitude areas and a less resource limitation in low latitude areas."

P8, L286: "than RH from satellite-driven estimates", seems missing a word like "more" or "higher".

Response: sorry, a "higher" was added.

P12, L360: "negatively feeding back to future climate change". I was not able to figure out the logic behind this inference.

Response: we remove the sentence.

P12, L374-375: The sentence should be better broken into two, because it is mumbling currently.

Response: we reduce the long sentence into two as below:

"Compared to linear regression analysis for predicting soil respiration (as no such global RH predictions were previously available for comparison) with model efficiencies of <50% (Bond-Lamberty and Thomson, 2010; Hashimoto et al., 2015; Hursh et al., 2017), RF algorithms achieved a higher model efficiency of 50% in this study. In addition to a feature selection according the importance value of each variable and avoiding overfitting (Bodesheim et al., 2018; Jian et al., 2018), RF could improve RH modelling accuracies and reduce uncertainties."

P12, L376-377: "This developed RH dataset provides several advantages to the estimation of global RH". The advantage should be established within a comparison, however, what are compared here is unclear. After all, you don't want to people to guess, and guessing is prone to misunderstanding.

Response: thank you. We meant advantages compared to other studies and revised the manuscript as:

"This developed RH dataset provides several advantages to the estimation of global RH compared to previous studies, e.g. Hashimoto et al. (2015) and Konings et al. (2019)."

Section 4.5: Overall the authors did discuss the limitations of this study. But I think if they formulate this alternatively by stating what use can be made with this dataset will be more helpful. Of course, it is up to the authors' decision.

Response: after discussing with co-authors before submission and the suggestion from one reviewer during the review process, the co-authors and reviewer suggested a section on the limitation of this study.

References:

Ballantyne, A., Smith, W., Anderegg, W., Kauppi, P., Sarmiento, J., Tans, P., Shevliakova, E., Pan, Y., Poulter, B., Anav, A., Friedlingstein, P., Houghton, R., and Running, S.: Accelerating net terrestrial carbon uptake during the warming hiatus due to reduced respiration, Nature Clim. Change, 7, 148-152, http://dx.doi.org/10.1038/nclimate3204, 2017.

Bodesheim, P., Jung, M., Gans, F., Mahecha, M. D., and Reichstein, M.: Upscaled diurnal cycles of land–atmosphere fluxes: a new global half-hourly data product, Earth Syst. Sci. Data, 10, 1327-1365, http://dx.doi.org/10.5194/essd-10-1327-2018, 2018.

Bond-Lamberty, B. and Thomson, A.: Temperature-associated increases in the global soil respiration record, Nature, 464, 579-582, http://dx.doi.org/10.1038/nature08930, 2010.

Hashimoto, S., Carvalhais, N., Ito, A., Migliavacca, M., Nishina, K., and Reichstein, M.: Global spatiotemporal distribution of soil respiration modeled using a global database, Biogeosciences, 12, 4121–4132, http://dx.doi.org/10.5194/bgd-12-4331-2015, 2015.

Hursh, A., Ballantyne, A., Cooper, L., Maneta, M., Kimball, J., and Watts, J.: The sensitivity of soil respiration to soil temperature, moisture, and carbon supply at the global scale, Glob. Chang. Biol., 23, 2090-2103, http://dx.doi.org/10.1111/gcb.13489, 2017.

Jian, J., Steele, M. K., Thomas, R. Q., Day, S. D., and Hodges, S. C.: Constraining estimates of global soil respiration by quantifying sources of variability, Glob. Chang. Biol., 24, 4143-4159, http://dx.doi.org/10.1111/gcb.14301, 2018.

Konings, A. G., Bloom, A. A., Liu, J., Parazoo, N. C., Schimel, D. S., and Bowman, K. W.: Global satellite-driven estimates of heterotrophic respiration, Biogeosciences, 16, 2269-2284, http://dx.doi.org/10.5194/bg-16-2269-2019, 2019.

Piao, S., Wang, X., Park, T., Chen, C., Lian, X., He, Y., Bjerke, J. W., Chen, A., Ciais, P., Tømmervik, H., Nemani, R. R., and Myneni, R. B.: Characteristics, drivers and feedbacks of global greening, Nature Reviews Earth & Environment, doi: 10.1038/s43017-019-0001-x, 2019. 10.1038/s43017-019-0001-x, 2019.

Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., Betts, R., Ciais, P., Cox, P., Friedlingstein, P., Jones, C. D., Prentice, I. C., and Woodward, F. I.: Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), Glob. Chang. Biol., 14, 2015-2039, 10.1111/j.1365-2486.2008.01626.x, 2008.

Spatial- and temporal-patterns of global soil heterotrophic respiration in terrestrial ecosystems

Xiaolu Tang[1, 2], Shaohui Fan[3], Manyi Du[4], Wenjie Zhang[5, 6], Sicong Gao[6], Shibing Liu[1], Guo Chen[1], Zhen Yu[7, 8], Wunian Yang[1]

[1]College of Earth Science, Chengdu University of Technology, Chengdu 610059, Sichuan, P.R. China

[2]State Environmental Protection Key Laboratory of Synergetic Control and Joint Remediation for Soil & Water Pollution, Chengdu University of Technology, Chengdu 610059, P. R. China

[3]Key laboratory of Bamboo and Rattan, International Centre for Bamboo and Rattan, Beijing 100102, P.R. China

[4]Experimental Center of Forestry in North China, Chinese Academy of Forestry, Beijing 102300, China

[5]State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Beijing 100101, China

[6]School of Life Science, University of Technology Sydney, NSW 2007, Australia

[7]School of Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, P.R China

[8]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011, USA

Correspondence to: Shaohui Fan (fansh@icbr.ac.cn) ; Wunian Yang (ywn@cdut.edu.cn)

Abstract. Soil heterotrophic respiration (RH) is one of the largest and most uncertain components of the terrestrial carbon cycle, directly reflecting carbon loss from soils to the atmosphere. However, high variations and uncertainties of RH existing in global carbon cycling models require RH estimates from different angles, e.g., a data-driven angle.an urgent development of data-derived RH dataset. To fill this knowledge gap, this study applied a Random Forest (RF) algorithm – a machine learning approach, to (1) develop a globally gridded RH dataset and (2) investigate its spatial- and temporal-patterns from 1980 to 2016 at the global scale by linking field observations from the Global Soil Respiration Database and global environmental drivers – temperature, precipitation, soil water content, etc. Finally, a globally gridded RH dataset was developed covering from 1980 to 2016 with a spatial resolution of half degree and a temporal resolution of one year. Globally, the average annual RH was 57.2±0.6 Pg C a^{-1} from 1980 to 2016, with a significantly increasing trend of 0.036±0.007 Pg C a^{-2}. However, the temporal trend of the carbon loss from RH varied with climate zones that RH showed a significant and increasing trend in boreal and temperate areas, in contrast, such trend was absent in tropical regions. Temperature-driven RH dominated 39% of global land

1

and was primarily distributed at high latitude areas. While the areas dominated by precipitation and soil water content were mainly semi-arid and tropical areas, accounting for 36% and 25% of the global land, respectively, suggesting variations in the dominance of environmental controls on the spatial patterns of RH. The developed globally gridded RH dataset will further aid in understanding of the mechanisms of global soil carbon dynamics, serving as a benchmark to constrain terrestrial biogeochemical ~~global vegetation~~ models. The dataset is publicly available at https://doi.org/10.6084/m9.figshare.8882567 (Tang et al., 2019a).

1 Introduction

Global soils and surface litter store up to 2- or 3-fold the amount of carbon present in the atmosphere (Trumbore, 2009) and therefore, a small change in soil carbon content could have profound effects on atmospheric CO_2 and climate change (Köchy et al., 2015). Although global carbon flux from soil-to-atmosphere is increasing (Zhao et al., 2017), the degree to which future climate change will stimulate soil carbon loss via heterotrophic respiration (RH) remains highly uncertain (Bond-Lamberty et al., 2018; Friedlingstein et al., 2014; Trumbore and Czimczik, 2008), particularly for areas with a high temperature sensitivity, or rapid changes in precipitation frequency and intensity.

Soil RH represents the carbon loss from the decomposition of litter detritus and soil organic matter by microorganisms (Subke et al., 2006), accounting for one of the largest components of the terrestrial carbon cycle (Bond-Lamberty et al., 2016). However, RH's feedback to climate variability is poorly understood. RH could affect future climate change via the mineralization of long-stored soil carbon, offsetting net primary production (NPP) and even converting terrestrial ecosystems from a carbon sink to a carbon source (Tremblay et al., 2018). Conversion of the sink/source role depends on how strongly large-scale process affected by environmental drivers, e.g. temperature, precipitation and soil organic carbon content (Hursh et al., 2017; Sierra et al., 2015), or extreme conditions, such as fire, human disturbance and drought (Kurz et al., 2013; Metsaranta et al., 2011). Although it is widely recognized that warming enhances CO_2 release from soils, the magnitude of such release is uncertain due to variations in the temperature sensitivity of soil organic matter decomposition (Suseela et al., 2012). In addition, environmental drivers of RH, e.g. temperature and soil moisture, are still undergoing changes under climate warming and can affect RH individually or interactively. Therefore, reducing RH uncertainty and clarifying the response of RH to environmental factors are essential for future projections of the impact of climate change on the terrestrial carbon balance.

Due to the diurnal, seasonal and annual variability in RH, in addition to the difficulties~~ty~~ of large-scale measurements, regional and global RH estimations mainly depend on modelling approaches using regional or global variables, such as temperature, precipitation and carbon supply (Bond-Lamberty and Thomson, 2010b; Hashimoto et al., 2015; Hursh et al., 2017). Besides temperature and precipitation, soil variables, such as water, carbon and nitrogen contents, are also important factors in the regulation of RH and should be considered for accurate RH estimations (Hursh et al., 2017; Zhao et al., 2017), although these variables vary with biome and climate.

2

Observational studies have examined the responses of soil respiration to different climatic variables at different locations across the globe (Bond-Lamberty and Thomson, 2010a; Zhou et al., 2016). Hashimoto et al. (2015) and Bond-Lamberty and Thomson (2010b) predicted global soil respiration rates using climate-derived models driving by temperature and precipitation,

65 however, these models commonly explain less than 50% variations of soil respiration, requiring new techniques, potential new numerical/algorithmic methods to better quantif~~ication~~y and understand~~ing of~~ the large-scale soil carbon flux~~es~~ (Bond-Lamberty, 2018). To improve the modelling accuracy, more recent studies have used linear regression or machine learning approaches including more abiotic or biotic variables, such as soil carbon supply, soil properties and NPP (Hursh et al., 2017; Zhao et al., 2017) and observations collected from newly published measurements (Jian et al., 2018; Zhao et al., 2017). On the

70 other hand, including more variables in linear or non-linear regression models may cause overfitting and autocorrelation issues (Long and Freese, 2006). To overcome overfitting and autocorrelation, machine learning approaches, such as the Random Forest (RF, Breiman, 2001), have been applied to explore the hierarchical importance of environmental factors, such as temperature, soil water content (SWC), NPP and soil pH (Hursh et al., 2017). Machine learning techniques are highly effective ~~as~~ because they are fully data adaptive, and do not require initial assumptions on functional relationships and can function with

75 nonlinear dependencies (Bodesheim et al., 2018). Therefore, these approaches are beginning ~~used~~ using in earth science, particularly in carbon and water flux modelling (Jung et al., 2010; Jung et al., 2017; Yao et al., 2018b), and may provide more reliable estimates of soil respiration (Bond-Lamberty, 2018; Zhao et al., 2017). However, no study to date has assessed the global variability of RH using empirical~~ly~~ field observations to bridge the knowledge gap between local, regional and global scales.

80 The newly-emerged Dynamic Global Vegetation Model~~s~~ from the TRENDY model ensembles and Earth System Models have been widely used to investigate major physiological and ecological processes and ecosystem structures, providing a novel database and approach to examine and estimate RH at the global scale (Zhu et al., 2017), although RH improvements in Earth System Models are still required (Shao et al., 2013). TRENDY and Earth System Model simulations incorporating a RH component are commonly calibrated and validated by eddy covariance measurements, e.g. net ecosystem carbon exchange

85 ~~(Yang et al., 2013)~~, however, modelled RHs from these models have not yet been calibrated and validated using field ~~RH~~ observations. Therefore, these modelled RHs may be fundamentally different from observed values and no global observations exist to evaluate model effectiveness. Consequently, ~~the~~ a data-driven RH ~~database~~ dataset could improve our understanding of the underlying mechanisms of RH variability to ~~future~~ climate change at the global scale, and could serve as a benchmark to constrain terrestrial biogeochemical~~global vegetation~~ models.

90 Thus, we used a RF algorithm to estimate global RH based on updated RH observations from the Global Soil Respiration Dataset (SRDB, Bond-Lamberty and Thomson, 2010a) with the objectives of: (1) developing a globally gridded RH product (named data-derived RH); (2) detecting the ~~temporal and~~ spatial- and temporal- patterns of RH; (3) identifying the dominant driving factors for spatial- and temporal-variabilities of RHs; and (4) comparing the data-derived RH dataset with ~~data~~ RH generated by Dynamic Global Vegetation Models from the TRENDY ensembles.

2 Materials and methods

2.1 Soil heterotrophic respiration database development

The basis of the database developed here included observed global RH values from SRDB (Bond-Lamberty and Thomson, 2010a), which was freely obtained at https://github.com/bpbond/srdb. The database was further updated using observations collected from Chinese peer-review publications at the China Knowledge Resource Integrated Database (www.cnki.net) until March 2018. This study included the RH data for: (1) annual RH as directly reported in publications with at least one year continuous measurements; (2) the start- and end-year were extracted from SRDB, directly from publications or calculated by the "years of data" in the SRDB; (3) observations measured by alkali absorption or soda lime approaches were not included because of their potential underestimation of respiration flux with an increasing pressure in the measurement chamber (Pumpanen et al., 2004); (4) experiments with treatments, such as nitrogen manipulation, or fertilization, were excluded, and only RH measurements from the control treatment were included (Jian et al., 2018); (5) SRDB observations labelled as "potential problem" ("Q10"), "suspected problem" ("Q11"), "known problem" ("Q12"), "duplicate" ("Q13") and "inconsistency" ("Q14") were not included (Bond-Lamberty and Thomson, 2010a). Finally, the newly updated database~~set~~ included 504 RH observations in total. Although most of the observations were from China, America and Europe, this database covered all the major terrestrial biomes across the world (Fig. 1).

2.2 Climate and soil data

To investigate the global spatial-temporal RH patterns, global spatial-temporal grids of RH driving factors were required. A total of ~~9~~ nine global variables were included (Supplementary Table S1): ~~M~~monthly gridded data of temperature, precipitation, diurnal temperature range from Climatic Research Unit TS v.4.01 over 1901-2016 (https://crudata.uea.ac.uk, Harris et al., 2014); shortwave radiation (SWR, https://www.esrl.noaa.gov, Kalnay et al., 1996); gridded soil organic carbon content (Hengl et al., 2017) and nitrogen content from https://webmap.ornl.gov/ogc/index.jsp (Global Soil Data, 2000); monthly gridded nitrogen deposition dataset from the global Earth System Models of GISS-E2-R, CCSM-CAM3.5 and GFDL-AM3 from 1850 to 2000s (https://www.isimip.org, Lamarque et al., 2013); monthly Palmer Drought Severity Index (PDSI, https://www.esrl.noaa.gov/psd/, Dai et al., 2004) and (SWC, https://www.esrl.noaa.gov, van den Dool et al., 2003). Before further data analysis, monthly data were aggregated to an annual scale. These variables could stand for different environmental controls on RH. For example, temperature, precipitation and SWC are critical environmental controls on microbial activities for soil organic ~~carbon~~ matter decomposition (Jian et al., 2018; Suseela et al., 2012; Tremblay et al., 2018). Soil organic matter, soil carbon stock and soil nitrogen are important carbon and nitrogen substrates for microbes that are related to the decomposition of soil organic matter (Tremblay et al., 2018). The drought index (PDSI) and diurnal temperature range to represent water and temperature stress on RH (Berryman et al., 2015; Zhu and Cheng, 2011). Then t~~T~~he global environmental drivers for each given site were extracted by site longitudes and latitudes corresponding to annual RH observations. If the

environmental driver is not in a spatial resolution of 0.5º, we first re-sampled this environmental driver to a 0.5º resolution using a the bilinear interpolation.

2.3 RH from TRENDY models

130   In the last several decades, TRENDY models were developed to simulate key processes (e.g. photosynthesis, respiration, evapotranspiration, phenology and carbon allocation) that drive the dynamics of global terrestrial ecosystems (Piao et al., 2015). TRENDY models follow a common protocol and use the same climate-forcing data from National Centres for Environmental Prediction at a spatial resolution of 0.5°. For modelled products with different spatial resolutions, new errors will be produced when re-sampling to 0.5°. Therefore, to compare the dynamics in the data-derived RH dataset and TRENDY

135   RH dataset, we used model outputs from seven TRENDY models: (Community Land Model-4.5 (CLM4, Lawrence et al., 2011); Integrated Science Assessment Model (ISAM, Cao, 2005); Lund-Potsdam-Jena (LPJ, Sitch et al., 2003); Lund-Potsdam-Jena General Ecosystem Simulator (LPJ-GUESS, Smith et al., 2001); VEgetation-Global-Atmosphere-Soil (VEGAS, Zeng et al., 2005); and Vegetation Integrative Simulator for Trace gases (VISIT, Kato et al., 2013). Additionally, the RH dataset generated empirically by Hashimoto et al. (2015) was compared (termed as Hashimoto RH), which was publicly

140   available (http://cse.ffpri.affrc.go.jp/shojih/data/index.html) and estimated from a global relationship between RH and soil respiration (Bond-Lamberty et al., 2004), and the total soil respiration was predicted from a climate-driven model using precipitation and temperature estimated using empirical total soil respiration relationships using a climate driven model (termed as Hashimoto RH) based on the observations from SRDB. More details can be found in Hashimoto et al. (2015).

2.4 RF-based RH Modelling

145   RF is a machine learning approach that uses a large number of ensemble regression trees, but a random selection of predictive variables (Breiman, 2001). Two free parameter settings are required, which are the number of trees and candidate variables for each split. However, the RF model is not usually sensitive to the number of trees or variables. A RF regression can deal with a large number of features, assisting a feature selection based on the importance value of each variable and the avoidance of overfitting (Bodesheim et al., 2018; Jian et al., 2018). In the present study, a RF model was trained using 9 variables

150   (supplementary Table 1) in the "*caret*" package (version 6.0-80, accessed on May 27, 2018) in R (R Core Team, 2018), which was then implemented to predict RH for each grid at a spatial resolution of 0.5°. To characterize the performance of RF, a 10-fold cross-validation was applied, which means that the dataset was stratified into 10 parts and each part contained roughly equal number of samples. The target values for each of these 10 parts were predicted based on the training models using the remaining nine parts. Two model evaluation statistics were used, including modelling efficiency (R^2) and root mean square

155   error (RMSE, Tang et al., 2019b; Yao et al., 2018b).

2.5 Trend analysis

A Ttrend analysis of RH was estimated by the Theil-Sen linear regression and tested with the Mann-Kendall non-parametric test. The Theil-Sen estimator is a non-parametric slope estimator based on median values, and this approach was widely used

for time-series analysis, e.g. carbon fluxes (Dai et al., 2016), and vegetation greening and browning (Pan et al., 2018) . The
160 Mann-Kendall non-parametric test was employed to investigate the significant changes in RH trend at a. ~~The~~ significance level
~~was at~~of 0.05.

2.6 Relationships between RH and temperature, precipitation and SWC

Although previous studies have used precipitation as a proxy for SWC (Bond-Lamberty and Thomson, 2010b; Chen et al.,
2010), this may result in variability in soil respiration estimates (Jassal et al., 2007; Zhang et al., 2006), because the relationship
165 between SWC and soil respiration ~~is~~ was much more complex than that between soil respiration and temperature or
precipitation (Jian et al., 2018). Therefore, ~~the~~ mean annual temperature (MAT), precipitation (MAP) and SWC were all
considered as potentially important proxies driving RH (Bond-Lamberty et al., 2016; Reichstein and Beer, 2008). Annual mean
RH was regressed against the three proxies. The relationships between the data-derived RH and
~~temperature~~MAT/~~precipitation~~MAP/SWC were assessed locally for each grid cell by calculating the correlations using partial
170 correlation analysis. When analysing the partial correlations between RH and the proxy, the other two proxies were controlled
to remove their confounding effects on RH. The correlation strengths of ~~temperature~~MAT, ~~precipitation~~MAP and SWC were
used to derive RGB combinations and indicate the drivers of RH.

2.7 The comparison map profile method

To detect the spatial similarity and difference patterns between the data-derived RH and TRENDY and Hashimoto ~~or~~
175 ~~reported~~ RH -values from 1981 to 2010, we utilized the comparison map profile (CMP) method (Gaucherel et al., 2008). This
method was based on the absolute distances (D) and cross-correlation coefficients (CC) across multiple scales, with D and CC
reflecting the similarity and the spatial structures of two compared images with the same sizes, respectively (Gaucherel et al.,
2008). The D value between moving windows (from 3×3 to 41×41 pixels in present study) of two compared images w~~er~~as
calculated by Equation (1):

180 $$D = abs(\bar{x} - \bar{y}) \tag{1}$$

Where, $\bar{x}$ and $\bar{y}$ represent mean values calculated over two moving windows. Finally, the mean D value was calculated as an
average of different moving windows.

The CC was calculated according to Equation (2):

$$CC = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{(x_{ij} - \bar{x}) \times (y_{ij} - \bar{y})}{\sigma_x \times \sigma_y} \tag{2}$$

185 $$\sigma_x^2 = \frac{1}{N^2 - 1} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_{ij} - \bar{x})^2 \tag{3}$$

Where, x_{ij} and y_{ij} are pixel values at i^{th} row and j^{th} column of the moving windows of two compared images, respectively; N
represents the total number of pixels covered by each of moving windows; σ_x and σ_y stand for standard deviations of two

6

moving windows. Low D values reflect a goodness between the compared images, while low CC values suggest a low similarity. Finally, the mean D and CC were calculated as the average from different moving windows.

190 All data analysis mentioned above were conducted in R (version 3.5.0, access on April 2018).

3 Results

3.1 Spatial patterns of RH

Based on the 10-fold cross-validation, ~~model efficiency (~~R^2~~)~~ and *RMSE*~~,~~ were 50% and 143 g C m^{-2} a^{-1} (Fig. S1), respectively. This indicates that the RF algorithm effectively captured the spatial- and temporal-variability of RH, therefore enabling

195 deriving of a global gridded RH dataset.

The ~~D~~data-derived RH dataset showed a strong spatial pattern globally (Fig. 2a). The largest RH fluxes occurred in tropical areas (e.g. Amazon tropical forests) at > 700 g C m^{-2} a^{-1}, followed by the subtropics, such as South China and America, and humid temperate areas, e.g. North America, Western and Central Europe, with an annual RH of 400-600 g C m^{-2} a^{-1}. Relatively low annual RH less than 200 g C m^{-2} a^{-1} was generally observed in areas with cold and dry climates, such as boreal areas,

200 characterized by low temperatures and short growing seasons, dry or semi-arid regions (e.g. Northwest China), where water availability limits ecosystem development. However, the most variable changes in RH over the time from 1980 to 2016 - using standard deviation and coefficient of variation (CV, the ratio of the standard deviation and the mean) as ~~proxies~~ a proxy (Fig. 2b an S3), were found in boreal regions with RH ~~more~~ higher than 70 g C m^{-2} a^{-1} or a CV > 0.7~~,~~. ~~w~~While the majority areas of RH variabilit~~ies~~y ~~exhibited smaller~~were lower than 30 g C m^{-2} a^{-1} or a CV < 0.3. Similarly, TRENDY ~~and previously~~

205 ~~reported~~ RH ~~values~~ showed similar patterns with the highest RH in warm and humid areas and the lowest RH in cold and dry regions (Fig. S2). However, differences existed in the absolute RH fluxes (Fig. S2). For example, CLM4 and VISIT models predicted RH to be ~~more~~ higher than 1400 g C m^{-2} a^{-1} within Amazon forest regions, while ISAM and LPJ-GUESS estimates were typically low at around 1000 g C m^{-2} a^{-1}. However, the data-derived RH dataset and Hashimoto RH showed the highest RH fluxes in tropical regions of about 800 g C m^{-2} a^{-1}.

210 To examine the similarity in the patterns ~~of~~ between the data-derived RH dataset and ~~those established by~~ TRENDY ~~models and~~ /Hashimoto RH, the CMP method was employed (Fig. 3). Larger D and lower CC values indicate less consistent magnitudes and a local gradient distribution between the two compared images. The ~~D~~data-derived RH dataset and Hashimoto RH differed greatly in East Canada and the Middle East with D values ~~above~~ higher 200 g C m^{-2} a^{-1} and CC values lower than -0.5. Interestingly, the most noticeable differences between the data-derived RH ~~values~~ and ~~TRENDY model~~ mean TRENDY

215 RH ~~values,~~ occurred in East Asia and the Middle East, where D was higher than 500 g C m^{-2} a^{-1}, while CC~~s~~ ~~was~~ were around -0.1. When assessing each TRENDY model individually (Figs. S4 and S5), the differences between the data-derived RH dataset and TRENDY RH were even larger. The most remarkable difference~~s~~ ~~was~~ were found for CLM4 and VISIT models in regions where D was above 800 g C m^{-2} a^{-1} with CC values of about -0.3 (East Asia and America).

7

Across the latitudinal gradients, zonal mean RH values increased from cold or dry areas (e.g. tundra, and desert or semi-arid
220 areas) to warm or humid areas (e.g. temperate and tropical areas, Fig. S6). The ~~D~~data-derived RH dataset varied from 60±12
at about 75ºN to 640±71 g C m^{-2} a^{-1} at the equator, reflecting a higher resource limitation in high latitude areas and a less
resource limitation in low latitude areas~~less stress from environmental limitations~~. In the dry tropical areas (10ºS -25ºS and
10ºN -25ºN) limited by water, the zonal mean RH decreased slightly. With the increase of water availability, RH showed a
second peak in the Northern and Southern Hemispheres around 20ºN and 40ºS, respectively. Nonetheless~~However~~, there was
225 a high level of variability between the data-derived RH and ~~those predicted by~~ TRENDY/Hashimoto RH in the equatorial
regions (Fig. S6), with predictions generally overestimating RH at the equator. Peak RH values in the equatorial region ranged
from 660±65g C m^{-2} a^{-1} for ORCHIDEE model~~in previously published estimations,~~ to above 1200±460 g C m^{-2} a^{-1} for CLM4
model, resulting in a considerably higher peak RH value for the model mean (950±300 g C m^{-2} a^{-1}).

3.2 Total RH

230 Over the last 37 years, the global RH has increased from 55.8 Pg C a^{-1} (1 Pg = 1×10^{15} g) in 1992 to 58.3 Pg C a^{-1} in 2010,
with an average of 57.2±0.6 Pg C a^{-1} and strong annual variabilit~~ies~~y (Fig. 4). Compared to the data-derived RH dataset,
TRENDY/Hashimoto RH was underestimated ~~global RH values~~ (Fig. 5~~6~~a), with the exception of the VISIT model. ISAM
predicted the lowest global RH of 34.8±0.4 Pg C a^{-1}, while the VISIT model produced the highest RH of 59.9±0.6 Pg C a^{-1}
(Fig. 5a). The model mean RH was 47.6±0.5 Pg C a^{-1}, underestimating RH by 9.6 Pg C a^{-1} (16%) in comparison to the data-
235 derived RH dataset. Due to this large divergence, the strength of correlation between the data-derived RH and
TRENDY/Hashimoto RH varied greatly from 0.06 to 0.72 (Fig. 5b). Boreal, temperate and tropical regions were the three
most important contributors for the global RH according to the Köppen – Geiger climate classification system (Peel et al.,
2007), contributing 76% of the total global RH. The mean RH of boreal, temperate and tropical areas were 10.8±0.3, 12.9±0.1
and 19.5±0.2 Pg C a^{-1}, accounting for 19%, 22% and 35% of the total global RH, respectively (Fig. S8).

240 **3.3 Trends in RH**

Globally, although there was a great inter-annual variability in RH, the total RH has significantly increased at ~~the~~ a rate of
0.036±0.007 Pg C a^{-2} from 1980 to 2016 (p = 0.000, Fig. 4). Comparison of the data-derived RH dataset and ~~that~~ TRENDY
RH ~~generated by TRENDY models~~ during the period of 1981 to 2010 was performed. The ~~D~~data-derived RH increased at
0.041±0.01 Pg C a^{-2} (Fig. S7), which was lower than that of TRENDY RH (0.057±0.009 Pg C a^{-2}) and Hashimoto RH
245 (0.057±0.009 Pg C a^{-2})~~exhibiting slower increasing trends of 0.05±0.007 and 0.057±0.009 Pg C a^{-2} of TRENDY/Hashimoto
RH, respectively~~. Additionally, temporal trends varied greatly among TRENDY models (Fig. S7), with the largest increasing
trend of 0.123±0.013 Pg C a^{-2} for LPJ-GUESS and the largest decreasing trend of -0.018±0.007 Pg C a^{-2} for ISAM.

Temporal trends varied ~~according to~~among climate zones. RH in boreal and temperate areas increase~~ing~~d by 0.020±0.004
and 0.007±0.002 Pg C a^{-2} from 1980 to 2016 (both ps =~~<~~ 0.~~000~~001, Fig. S8), respectively, while RH in tropical areas did not
250 show a significant temporal trend, although inter-annual variabilities were observed (p = 0.362, Fig. S8). TRENDY/Hashimoto

RHs showed significant increasing temporal trends in boreal, temperate and tropical regions, except ~~those estimated by~~ ISAM and ORCHIDEE models (Fig. S9-11). However, the increasing magnitude varied among different TRENDY models.

From 1980 to 2016, the global RH was expected to be driven by multiple environmental factors, such as temperature and precipitation. During this period, MAT and MAP levels increased significantly by 0.34±0.032°C and 6.69±2.399 mm per
255 decade, respectively ($p < 0.01$, Fig's. S12a and b). Therefore, the correlations between RH and ~~temperature~~MAT/~~precipitation~~ MAP were evaluated. Globally, RH was significantly correlated with ~~temperature~~ MAT ($R^2 = 0.56$, p ~~=~~< 0.001~~0~~) and ~~precipitation~~ MAP ($R^2 = 0.42$, p ~~=~~< 0.~~000~~001, Fig. 6) anomalies. On average, the global RH increased by 1.08±0.163 Pg C a^{-1} per 1°C increase in MAT and 0.23±0.046 Pg C a^{-1} per 10 mm increment in MAP.

3.4 Spatial pattern of RH trends

260 Spatially, the data-derived RH trends presented heterogeneous geographical patterns (Fig. 2c). Positively increasing trends ~~in~~ of RH were found for more than half of the global land areas (59%, calculated from cell areas; Fig. S13). Generally, the increasing rates of RH were lower than 3 g C m^{-2} a^{-2}, in contrast, the highest RH increase was above 6 g C m^{-2} a^{-2} in boreal regions, such as Russia, North Canada and the Tibetan Plateau. RH exhibited a decreasing trend in 41% of the global land area and most considerably in South Asia (Fig. 2c). Similar to the data-derived RH trends, RH trends estimated by the
265 TRENDY/Hashimoto RH trend also showed heterogeneous geographical patterns (Fig. S14). However, large discrepancies were found among TRENDY/Hashimoto RH (Fig. 2c and S14). Generally, the largest increase in RH trends occurred in boreal areas, except for outputs by LPJ-GUESS and LPJ models, which showed a decreasing trend for most boreal areas. There was a decreasing trend across most of tropics (e.g. Southeast Asia), with the exception of VEGAS model ~~results~~ (Fig. S14).

3.5 Dominant factors in RH annual variability

270 ~~Annual mean temperature and precipitation~~MAT and MAP were the most important factors dominating RH in 39% and 36% of global land area~~s~~, respectively (Fig. S15). While SWC dominated the remaining 25% of global land area~~s~~. Spatially, the dominant drivers controlling RH varied greatly across the globe (Fig. 2d), with the area dominated by temperature mainly distributed in boreal areas above 50°N. This was also observed in the relatively high and positive partial correlation coefficient between temperature and RH (Fig. S15a). While precipitation dominated temperate areas between 25°N and 50°N (such as
275 North China, the Middle East and America), where a wide distribution of desert or semi-arid regions occur, SWC dominated in tropic areas, such the Amazon, India and Africa. Similarly, water availability (SWC and precipitation) were also main driving factors for RH in Australia.

Spatial patterns in environmental controls on TRENDY/Hashimoto RH varied greatly compared to the data-derived RH dataset or among TRENDY models (Figs. 2d and S15-17). Water availability (including precipitation and SWC) appeared to
280 be more important than temperature. The percentage of the areas dominated by temperature (mainly distributed in boreal areas, except for in ISAM model outputs)~~,~~ wa~~s~~ere less than the areas dominated by precipitation and SWC (globally distributed) (Fig.

S17). In terms of the ~~modelled~~ mean TRENDY RH, precipitation dominated most of global land areas (43%), followed by SWC (36%) and temperature (21%) (Fig. S15).

4 Discussion

4.1 Annual RH

4.1.1 Comparison with Hashimoto RH

Despite the increasing efforts to quantify the global carbon cycle, large uncertainties still remain in the spatial- and temporal-patterns in RH. To the best of our knowledge, this is the first study to apply the RF approach to predict ~~of~~ the ~~temporal~~spatial- and ~~spatial~~temporal-patterns of global RH using field observations. Globally, the mean RH amounted to 57.2 ± 0.6 Pg C a^{-1} from 1980 to 2016, 13.6 Pg C a^{-1} higher than RH from a satellite-driven estimate~~s~~ (Konings et al., 2019)~~-~~, and 6.4 Pg C a^{-1} higher than Hashimoto RH (Hashimoto et al., 2015). The differences between the data-driven RH ~~in this study~~ and Hashimoto may be due to several reasons, Firstly, the two RH products covered different land areas, with the data-derived RH dataset ~~in the present study~~ covering a higher land area. If the data-derived RH dataset was masked by Hashimoto RH over 1981-2010, the total RH was 51.8 ± 0.6 Pg C a^{-1}, close to that of Hashimoto RH with 51.1 ± 0.7 Pg C a^{-1} (Fig. S18), respectively. However, the ~~temporal~~spatial- and ~~spatial~~temporal-patterns varied greatly (Figs. 3 and 5).

Secondly, the two RH products used different variables and algorithms for RH predictions. RH was not only affected by temperature and precipitation, but also by carbon substrates, soil nutrient levels and other variables (Hursh et al., 2017). Besides temperature and precipitation, we also included SWC, soil nitrogen and carbon contents as indicators for environmental and nutrient constraints on RH. Conversely, Hashimoto RH was estimated from a climate-driven model including only temperature and precipitation as the driving variables (Hashimoto et al., 2015). This simple model can partly explain the reasons that Hashimoto RH could not capture the significant decrease in RH in 1982 and 1991 due to El Chichón and Pinatubo eruptions, respectively (Zhu et al., 2016), while the data-derived RH dataset and TRENDY RH successfully captured such effects.

Thirdly, the linear model between total soil respiration and RH was developed based on forest ecosystems (Bond-Lamberty et al., 2004; Hashimoto et al., 2015), which could be another uncertainty when applying this linear model to other ecosystems, e.g. croplands and grasslands.

4.1.2 Comparison with TRENDY models

As data-derived RH dataset often serve as a benchmark for terrestrial biogeochemical~~global vegetation~~ models, the data-derived RH dataset~~empirically derived global patterns of annual RH~~ ~~were~~ was compared with TRENDY model~~s~~ from 1980 to 2010 ~~results~~. Although the data-derived RH dataset lied within the model range (34.8 ± 0.4 Pg C a^{-1} for ISAM to 59.9 ± 0.6 Pg C a^{-1} for VISIT, Fig. 5a), the ~~model~~ mean TRENDY RH was underestimated ~~RH~~ by 16% compared to the data-derived RH dataset. Due to the different temporal trends among TRENDY model~~s~~ ~~outputs~~ and their low spatial correlation~~s~~ to the ~~of~~ data-derived RH dataset ~~and TRENDY RH~~ (correlation efficiencies ranging from 0.06 to 0.72, Fig. 6b), TRENDY RH clearly have

different sensitivities to climate variations. Additionally, the difference in RH magnitude and spatial pattern varied considerably, as shown by analysis of absolute distances and cross-correlations. This effect was mostly notable in tropical areas in VISIT and CLM4 models outputs (Figs. S4 and S5). This phenomenon may be associated to several reasons. Firstly, plant functional types differed among TRENDY models. For example, the VEGAS model included four plant functional types (Zeng et al., 2005), while the LPJ model defined ten plant functional types (Sitch et al., 2003).

Secondly, for each set of equations, constant vegetation parameters (e.g. photosynthetic capacity) were applied across time and space for most TRENDY models, which may induce an RH bias. Model parameters using short-term observations do not account for the inter-annual variability of climatic and soil conditions, generating a simplistic representation of RH due to the inability to capture the response of RH to new environmental controls in short-term observations.

Thirdly, models that do not consider nitrogen constraint could overestimate the increasing trend of RH, because nitrogen limitation was globally observed (LeBauer and Treseder, 2008). This could explain why the CLM4 model with a nitrogen control constraint produced a much smaller increasing trend compared to other TRENDY models, with the exception of ISAM (Fig. S7). Therefore, including soil nitrogen as a driving variable in modelling RH in this dataset had the advantage to detect the nitrogen constrain on RH.

Fourthly, the lacking of the representation of human activities and agricultural management (e.g. fertilization and irrigation) may underestimate RH, because fertilization and irrigation were important practices to increase RH (Chen et al., 2018; Zhou et al., 2016). This could explain why five of seven TRENDY models could not explain the significant increasing change of RH in middle China (Fig. S14), which experienced an intensive use of fertilization for food security in recent decades.

Finally, uncertainties and differences in model structures could also lead to inconclusive RH estimations. Although the same climatic data, e.g. temperature and precipitation, were used for TRENDY models to reduce the uncertainty causing by various meteorological forcings, systematic errors may be caused by applying a particular forcing and the errors might be propagated to model outputs (Anav et al., 2015). Therefore, TRENDY models should be improved by incorporating more processes such as nutrient constrains and an assessment of the model response to environmental variability (Keenan et al., 2012; Wang et al., 2014; Yao et al., 2018b).

4.2 Linkage to global carbon balance

If assuming the global ratio of RH/total soil respiration ranged from 0.56 ((Hashimoto et al., 2015)) to 0.63 ((Bond-Lamberty et al., 2018)), annual soil respiration varied from 90.8 to 102.1 Pg C a^{-1}, within the reported values of 83 to 108 Pg C a^{-1} based on recent studies (Bond-Lamberty and Thomson, 2010b; Hursh et al., 2017). This indirectly highlights the reliability of the use of RF for global RH prediction. Moreover, these findings also have important indications to carbon balance estimations. According to a recent NPP estimate from observations and IPCC report data (IPCC, 2013; Li et al., 2017), the global NPP ranged from 61.5 to 60 Pg C a^{-1}, respectively. The residual between RH and NPP (net ecosystem production) was 2.8-4.3 Pg C a^{-1}, which is similar to the global estimates of net ecosystem production from the International Geosphere-Biosphere

345 Programme, which ranged from 1.9 to 4.1 Pg C a^{-1} from 1959 to 2016 (Le Quéré et al., 2013; Le Quéré et al., 2016; Le Quéré et al., 2014). With a 1°C increase in global MAT, RH will increase by 1.08 Pg C a^{-1} globally, and ~~it~~ such increase is 0.23 Pg C a^{-1} for 10 mm increment in global MAP. These findings indicate that carbon flux~~es~~ from the decomposition of soil organic matter and litter (RH) maybe positively feedback to ~~future~~ global climate change - typically characterized by the increasing temperature and ~~alterations~~ the changes in precipitation (IPCC, 2013).

350 **4.3 Dominant factors in RH**

Dominant factors driving RH varied spatially. As temperature and energy were the most limited climatic factor~~s~~ in high latitude areas, temperature was a dominant factor for RH in high latitudinal regions above 50°N (Fig. 2d), with low temperatures leading to low RH (Fig. 2a). Similarly, due to the limited amount of precipitation, RH in semi-arid areas was mainly controlled by precipitation, which ~~is~~was consistent with reported both field observations (Bai et al., 2008) and

355 modelling studies (Gerten et al., 2008). SWC control of RH in tropical areas could be explained by the mechanisms of RH. Excessively high SWC can reduce the diffusion of oxygen, while excessively low SWCs could limit water and soluble substrate availabilit~~y~~ies, preventing microbial activit~~ies~~y (Luo and Zhou, 2006; Xu et al., 2004). Suseela et al. (2012) proposed that RH fluxes declined sharply when volumetric soil moisture reduced below ~15% or exceeded ~26%, which support~~s~~ed the findings of the present study. However, it should be noted that dominant environmental controls on spatial carbon flux gradients might

360 vary among different years (Reichstein et al., 2007), such as with climatic extremes.

4.4 Temporal variability of tropical, temperate and boreal areas

~~Temporally,~~ RH in tropical areas did not exhibit a significantly temporal pattern between 1981 and 2010 (Fig. S8, $p = 0.362$), indicating that in our model, climate change did not affect RH fluxes in these areas~~, negatively feeding back to future climate change~~. However, RH in boreal and temperature areas experienced significant increasing trends of 0.020±0.004 and

365 0.007±0.002 Pg C a^{-2}~~1~~, respectively (Fig. S7), suggesting that a positive feedback may occur with ~~future~~ climate change. Tremblay et al. (2018) proposed that the increased RH was mainly related to the increasing temperature~~s~~ in boreal forest soils, which support~~s~~ed the findings of the present study. It should be noted that both the data-derived RH dataset and Hashimoto/TRENDY RH in boreal areas showed a temporally increasing trend from 1981 to 2010, although the magnitude of increase differed (Fig. S9). Furthermore, despite the ISAM model showing a decreasing trend for temperate and tropic regions,

370 the ISAM model had an increasing trend in RH from 1981 to 2010 in boreal areas (Fig. S9). These results indicate that boreal regions are becoming increasingly important in global carbon cycling and that the increasing trend may continue due to ~~the~~ a large amount of carbon stored in soil~~s~~. Therefore, climate change may fundamentally alter carbon cycling in boreal areas through changes in the decomposition rate of soil organic matter (Crowther et al., 2016; Hashimoto et al., 2015; Schuur et al., 2015). Furthermore, the response of RH to climate variability varied with climate zone, indicating that different carbon loss

375 rates from RH will occur in different regions to ~~future~~ climate change.

4.5 Advantages, limitations and uncertainties

Based on the updated SRDB data~~base~~set, we used a RF algorithm to predict the ~~temporal and~~ spatial- and temporal- patterns of RH at the global scale and its response to environmental variables, and ~~empirically derived global patterns of annual RH~~the data-derived RH could serve as a benchmark for terrestrial biogeochemical~~global vegetation~~ models and reduce RH uncertainties. This ~~developed~~ data-derived RH dataset provide~~d~~s several advantages to the estimation of global RH compared to previous studies, e.g. Hashimoto et al. (2015) and Konings et al. (2019). Firstly, we compiled up-to-date field observations from SRDB and Chinese peer-review literatures up to March 2018, including 504 observations in total covering the majority global terrestrial ecosystems and climate zones (Fig. 1). Secondly, total RH and its inter-annual variability were assessed for boreal, temperate and tropical zones ~~— three~~ three main global climate zones. Analysis from the data-derived RH dataset further conclude~~d~~s that RH in different climate zones responded differently to global climate change. Thirdly, we applied the RF to predict and map~~ped~~ RH at the global scale using climate and soil predictors. Compared to the linear regression analysis for predicting soil respiration (as no such global RH predictions were previously available for comparison)~~, which has reported with~~ model efficiencies of <50% (Bond-Lamberty and Thomson, 2010b; Hashimoto et al., 2015; Hursh et al., 2017), the RF algorithm~~s~~ achieve~~d~~ a higher model efficiency of 50%. In addition to~~, allowing~~ a feature selection according the importance value of each variable and avoiding overfitting (Bodesheim et al., 2018; Jian et al., 2018), the RF algorithm -improv~~ing~~ed RH modelling accuracies and reduc~~ing~~ed uncertainties. ~~Additionally~~Meanwhile, the data-derived RH dataset was cross-validated globally by a 10-fold cross-validation (see "materials and methods" section), which could improve its reliability and feasibility compared to TRENDY based RH that ~~are~~ were not validated and calibrated by field observations, bridging the knowledge gap between local, regional and global scales temporally and spatially with a large number of empirical field measurements.

However, although ~~empirically derived global patterns of annual~~the data-derived RH dataset could be used as a benchmark for the verification of global carbon cycle modelling, bridging the knowledge-gaps between local, regional and global scales, few uncertainties and limitations still remained. Firstly, the RF algorithm~~s~~ construct~~s~~ed a model based on the training dataset and ~~is~~ was typically data limited in terms of quantity, quality and representativeness. Uneven data distribution has been a known issue in many ecological studies across the world, e.g. Bond-Lamberty and Thomson (2010b), Jung et al. (2011), Xu and Shang (2016) and Yao et al. (2018a). The RH observations were mainly from China, Europe and North America, while there were a lack of RH observations in Russia, Afric, Australia and Southwestern Asia in our study. Thus ~~U~~the uneven coverage of the observations ~~is~~was an important source of uncertainty to develop the data-derived RH dataset, which may cause a bias of the RF model toward~~s~~ the areas with more observations. However, our dataset covered a large climatic and edaphic gradient covering the major land covers and climate zones. Therefore, in future studies, increasing field observations in unsampled areas should greatly improve our ability to evaluate spatial- and temporal- patterns of RH at the global scale and model global carbon cycle to ~~future~~ climate change.

Secondly, the misrepresentation of human activities, particularly regarding to land management and land use change, could result in uncertainties in RH (Bond-Lamberty et al., 2016; Tang et al., 2016). These human activities include both site-level *in situ* information and the corresponding global grids. Otherwise, such information must not be included as corresponding site

410 information or ~~respective~~ globally gridded datasets are missing or insufficient. Although soil organic carbon stock, soil nitrogen content, SWC and shortwave radiation were selected as inputs for the development of the RF model, which could partly capture land use change, the impacts of land use change on the inter-annual variability of RH have not been fully qualified in the present study. Therefore, further efforts are required to characterize and quantify the effects of land use change~~s~~ on the global RH.

415 Thirdly, the ~~developed~~ date-derived RH dataset was derived at an annual timescale, which may cause an additional uncertainty regarding to the inter-annual variability of RH. Therefore, the need for a larger number of global observations and to develop finer-scale temporal dynamics need further exploration, in combination with remote-sensing measurements and field observations, which may provide new insights into terrestrial ecosystem carbon dynamics at the global scale. Besides, without consideration of the temporal changes of soil organic carbon content from 1980 to 2016 might be another uncertainty 420 because the increase of productivity driven by CO_2 fertilization would increase litter input into soils. However, there is a lack of soil organic carbon content that considering its temporal changes based on observations, which has constrained the further analysis of the effects of the temporal changes of soil organic carbon content on RH.

Finally, we developed a global RH at a half-degree spatial resolution, which included a scale mismatch between the observations and global gridded variables. This could be a great challenge for spatial modelling and using global gridded 425 variables with a finer resolution is encouraged to overcome this limitation (Xu and Shang, 2016). On the other hand, the study sites were globally distributed and there was a large climatic and edaphic gradient covering the major land covers and biomes, which should reflect a larger variability than the site-to-grid mismatch.

5 Data availability

The developed globally gridded RH database, the field RH observation dataset and R codes to produce the main results are 430 publicly free for scientific purpose, which can be downloaded at https://doi.org/10.6084/m9.figshare.8882567 (Tang et al., 2019a).

6 Conclusions

Data-derived global RH dataset may be used as a benchmark for ~~global vegetation~~ terrestrial biogeochemical models, however, no such study has yet been conducted to assess the global variability in RH using a large dataset of empirical 435 measurements to bridge the knowledge gap between local, regional and global scales. To fill this knowledge gap, we developed a globally gridded RH dataset, which was $0.5^{o} \times 0.5^{o}$ from 1980 to 2016 with an annually temporal resolution, using a RF algorithm by linking field observations and global variables. Robust conclusions include: (1) Annual mean RH was 57.2±0.6 Pg C a^{-1} between 1980 and 2016, with an increasing trend of 0.036±0.007 Pg C a^{-2}, indicating an increase in carbon loss from soils to the atmosphere ~~with future climate change~~; (2) Significant temporal trends were observed in the RH in boreal and 440 temperate areas, although none were found in tropical regions. This indicates that the temporal trend in RH varied with climate zones, highlighting their different sensitivities to ~~future~~ climate change; (3) The magnitude and dominant factors ~~in~~ of the data-

derived RH TRENDY RHresults generated by data-derived and TRENDY models varied greatly, indicating that future efforts should focus on improving the representation of RH in ecosystem models and the ecosystemand its response to environmental variability in terrestrial biogeochemical models; (4) More field observations are required in areas with limited observational datasets, with the integration of smaller-scale temporal dynamics (rather than annual timescales) potentially providing new insights into terrestrial ecosystem carbon dynamics at the global scale; (5) The develop globally gridded data-derived RH dataset could serve as a benchmark to constrain the terrestrial biogeochemicalglobal vegetation models, further contributing to improve our understanding of the mechanisms of global soil carbon dynamics.

Author contributions. XT, SF and WY design the study; XT, WZ, SG, MD and ZY contributed to data analysis, including improving R code; SL and GC provided constructive comments. All authors contributed to review the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

References

Anav, A., Friedlingstein, P., Beer, C., Ciais, P., Harper, A., Jones, C., Murray-Tortarolo, G., Papale, D., Parazoo, N. C., Peylin,

470   P., Piao, S., Sitch, S., Viovy, N., Wiltshire, A., and Zhao, M.: Spatiotemporal patterns of terrestrial gross primary production: A review, Rev. Geophys., 53, 785-818, http://dx.doi.org/10.1002/2015rg000483, 2015.

Bai, Y. F., Wu, J. G., Xing, Q., Pan, Q. M., Huang, J. H., Yang, D. L., and Han, X. G.: Primary production and rain use efficiency across a precipitation gradient on the Mongolia plateau, Ecology, 89, 2140-2153, https://doi.org/10.1890/07-0992.1, 2008.

475   Berryman, E. M., Barnard, H. R., Adams, H. R., Burns, M. A., Gallo, E., and Brooks, P. D.: Complex terrain alters temperature and moisture limitations of forest soil respiration across a semiarid to subalpine gradient, J. Geophys. Res. Biogeosci., 120, 707-723, http://dx.doi.org/10.1002/2014jg002802, 2015.

Bodesheim, P., Jung, M., Gans, F., Mahecha, M. D., and Reichstein, M.: Upscaled diurnal cycles of land–atmosphere fluxes: a new global half-hourly data product, Earth Syst. Sci. Data, 10, 1327-1365, http://dx.doi.org/10.5194/essd-10-1327-2018,

480   2018.

Bond-Lamberty, B.: New Techniques and Data for Understanding the Global Soil Respiration Flux, Earth's Future, 6, 1176-1180, http://dx.doi.org/10.1029/2018ef000866, 2018.

Bond-Lamberty, B., Bailey, V. L., Chen, M., Gough, C. M., and Vargas, R.: Globally rising soil heterotrophic respiration over recent decades, Nature, 560, 80-83, http://dx.doi.org/10.1038/s41586-018-0358-x, 2018.

485   Bond-Lamberty, B. and Thomson, A.: A global database of soil respiration data, Biogeosciences, 7, 1915-1926, http://dx.doi.org/10.5194/bg-7-1915-2010, 2010a.

Bond-Lamberty, B. and Thomson, A.: Temperature-associated increases in the global soil respiration record, Nature, 464, 579-582, http://dx.doi.org/10.1038/nature08930, 2010b.

Bond-Lamberty, B., Wang, C., and Gower, S. T.: A global relationship between the heterotrophic and autotrophic components

490   of soil respiration?, Glob. Chang. Biol., 10, 1756-1766, http://dx.doi.org/10.1111/j.1365-2486.2004.00816.x, 2004.

Bond-Lamberty, B., Epron, D., Harden, J., Harmon, M. E., Hoffman, F., Kumar, J., McGuire, A. D., and Vargas, R.: Estimating heterotrophic respiration at large scales: challenges, approaches, and next steps, Ecosphere, 7, e01380, http://dx.doi.org/10.1002/ecs2.1380, 2016.

Breiman, L.: Random forests, Mach. Learn., 45, 5-32, http://dx.doi.org/10.1023/A:1010933404324, 2001.

495   Cao, L.: An Earth system model of intermediate complexity: Simulation of the role of ocean mixing parameterizations and climate change in estimated uptake for natural and bomb radiocarbon and anthropogenic CO_2, J. Geophys. Res., 110, http://dx.doi.org/10.1029/2005jc002919, 2005.

Chen, S., Huang, Y., Zou, J., Shen, Q., Hu, Z., Qin, Y., Chen, H., and Pan, G.: Modeling interannual variability of global soil respiration from climate and soil properties, Agric. For. Meteorol., 150, 590-605,

500   https://doi.org/10.1016/j.agrformet.2010.02.004, 2010.

Chen, Z., Xu, Y., He, Y., Zhou, X., Fan, J., Yu, H., and Ding, W.: Nitrogen fertilization stimulated soil heterotrophic but not autotrophic respiration in cropland soils: A greater role of organic over inorganic fertilizer, Soil Biol. Biochem., 116, 253-264, http://dx.doi.org/10.1016/j.soilbio.2017.10.029, 2018.

505 Crowther, T. W., Todd-Brown, K. E., Rowe, C. W., Wieder, W. R., Carey, J. C., Machmuller, M. B., Snoek, B. L., Fang, S., Zhou, G., Allison, S. D., Blair, J. M., Bridgham, S. D., Burton, A. J., Carrillo, Y., Reich, P. B., Clark, J. S., Classen, A. T., Dijkstra, F. A., Elberling, B., Emmett, B. A., Estiarte, M., Frey, S. D., Guo, J., Harte, J., Jiang, L., Johnson, B. R., Kroel-Dulay, G., Larsen, K. S., Laudon, H., Lavallee, J. M., Luo, Y., Lupascu, M., Ma, L. N., Marhan, S., Michelsen, A., Mohan, J., Niu, S., Pendall, E., Penuelas, J., Pfeifer-Meister, L., Poll, C., Reinsch, S., Reynolds, L. L., Schmidt, I. K., Sistla, S., Sokol, N. W., Templer, P. H., Treseder, K. K., Welker, J. M., and Bradford, M. A.: Quantifying global soil carbon losses in response to
510 warming, Nature, 540, 104-108, http://dx.doi.org/10.1038/nature20150, 2016.

Dai, A., Trenberth, K. E., and Qian, T.: A Global Dataset of Palmer Drought Severity Index for 1870–2002: Relationship with Soil Moisture and Effects of Surface Warming, J. Hydrometeorol., 5, 1117-1130, http://dx.doi.org/10.1175/JHM-386.1, 2004.

Dai, E., Huang, Y., Wu, Z., and Zhao, D.: Analysis of spatio-temporal features of a carbon source/sink and its relationship to climatic factors in the Inner Mongolia grassland ecosystem, Journal of Geographical Sciences, 26, 297-312,
515 http://dx.doi.org/10.1007/s11442-016-1269-0, 2016.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks, J. Clim., 27, 511-526, https://doi.org/10.1175/jcli-d-12-00579.1, 2014.

Gaucherel, C., Alleaume, S., and Hely, C.: The comparison map profile method: a strategy for multiscale comparison of
520 quantitative and qualitative images, IEEE Trans. Geosci. Remote Sens., 46, 2708-2719, http://dx.doi.org/10.1109/TGRS.2008.919379, 2008.

Gerten, D., Luo, Y., Le Maire, G., Parton, W. J., Keough, C., Weng, E., Beier, C., Ciais, P., Cramer, W., Dukes, J. S., Hanson, P. J., Knapp, A. A. K., Linder, S., Nepstad, D. A. N., Rustad, L., and Sowerby, A.: Modelled effects of precipitation on ecosystem carbon and water dynamics in different climatic zones, Glob. Chang. Biol., 14, 2365-2379,
525 https://doi.org/10.1111/j.1365-2486.2008.01651.x, 2008.

Global Soil Data, T.: Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS). ORNL Distributed Active Archive Center, http://dx.doi.org/10.3334/ornldaac/569, 2000.

Harris, I., Jones, P., Osborn, T., and Lister, D.: Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset, Int. J. Climatol., 34, 623-642, http://dx.doi.org/10.1002/joc.3711, 2014.

530 Hashimoto, S., Carvalhais, N., Ito, A., Migliavacca, M., Nishina, K., and Reichstein, M.: Global spatiotemporal distribution of soil respiration modeled using a global database, Biogeosciences, 12, 4121–4132, http://dx.doi.org/10.5194/bgd-12-4331-2015, 2015.

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotic, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G.,

535   Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLoS One, 12, e0169748, http://dx.doi.org/10.1371/journal.pone.0169748, 2017.

Hursh, A., Ballantyne, A., Cooper, L., Maneta, M., Kimball, J., and Watts, J.: The sensitivity of soil respiration to soil temperature, moisture, and carbon supply at the global scale, Glob. Chang. Biol., 23, 2090-2103, http://dx.doi.org/10.1111/gcb.13489, 2017.

540   IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Jassal, R. S., Black, T. A., Cai, T., Morgenstern, K., Li, Z., Gaumont-Guay, D., and Nesic, Z.: Components of ecosystem respiration and an estimate of net primary productivity of an intermediate-aged Douglas-fir stand, Agric. For. Meteorol., 144,
545   44-57, https://doi.org/10.1016/j.agrformet.2007.01.011, 2007.

Jian, J., Steele, M. K., Thomas, R. Q., Day, S. D., and Hodges, S. C.: Constraining estimates of global soil respiration by quantifying sources of variability, Glob. Chang. Biol., 24, 4143-4159, http://dx.doi.org/10.1111/gcb.14301, 2018.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu,
550   Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent decline in the global land evapotranspiration trend due to limited moisture supply, Nature, 467, 951-954, https://doi.org/10.1038/nature09396, 2010.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J. Q., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani,
555   L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res. Biogeosci., 116, G00J07, http://dx.doi.org/10.1029/2010jg001566, 2011.

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlstrom, A., Arneth, A., Camps-Valls, G., Ciais, P., Friedlingstein, P., Gans, F., Ichii, K., Jain, A. K., Kato, E., Papale, D., Poulter, B., Raduly, B., Rodenbeck, C., Tramontana,
560   G., Viovy, N., Wang, Y. P., Weber, U., Zaehle, S., and Zeng, N.: Compensatory water effects link yearly global land CO_2 sink changes to temperature, Nature, 541, 516-520, http://dx.doi.org/10.1038/nature20780, 2017.

Köchy, M., Hiederer, R., and Freibauer, A.: Global distribution of soil organic carbon – Part 1: Masses and frequency distributions of SOC stocks for the tropics, permafrost regions, wetlands, and the world, Soil, 1, 351-365, http://dx.doi.org/10.5194/soil-1-351-2015, 2015.

565   Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, Bull. Am. Meteorol. Soc., 77, 437-471, http://dx.doi.org/10.1175/1520-0477(1996)077<0437:Tnyrp>2.0.Co;2, 1996.
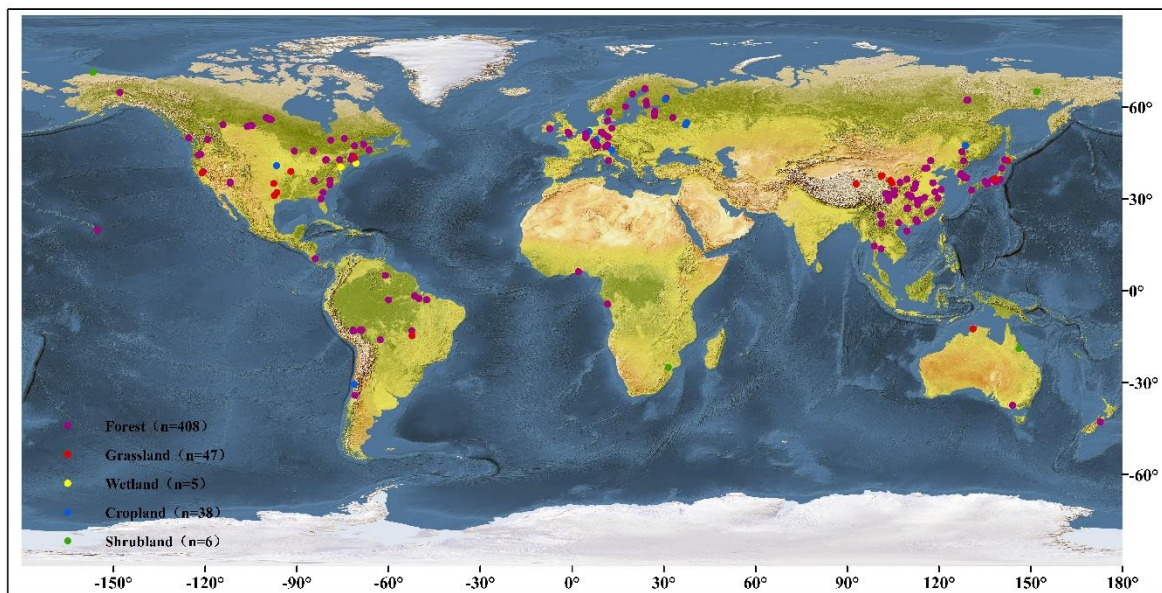
Kato, E., Kinoshita, T., Ito, A., Kawamiya, M., and Yamagata, Y.: Evaluation of spatially explicit emission scenario of land-use change and biomass burning using a process-based biogeochemical model, J. Land Use Sci., 8, 104-122, http://dx.doi.org/10.1080/1747423X.2011.628705, 2013.

Keenan, T., Baker, I., Barr, A., Ciais, P., Davis, K., Dietze, M., Dragoni, D., Gough, C. M., Grant, R., and Hollinger, D.: Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO_2 exchange, Glob. Chang. Biol., 18, 1971-1987, https://doi.org/10.1111/j.1365-2486.2012.02678.x, 2012.

Konings, A. G., Bloom, A. A., Liu, J., Parazoo, N. C., Schimel, D. S., and Bowman, K. W.: Global satellite-driven estimates of heterotrophic respiration, Biogeosciences, 16, 2269-2284, http://dx.doi.org/10.5194/bg-16-2269-2019, 2019.

Kurz, W. A., Shaw, C. H., Boisvenue, C., Stinson, G., Metsaranta, J., Leckie, D., Dyk, A., Smyth, C., and Neilson, E. T.: Carbon in Canada's boreal forest — A synthesis, Environ. Rev., 21, 260-292, https://doi.org/10.1139/er-2013-0041, 2013.

Lamarque, J. F., Dentener, F., McConnell, J., Ro, C. U., Shaw, M., Vet, R., Bergmann, D., Cameron-Smith, P., Dalsoren, S., Doherty, R., Faluvegi, G., Ghan, S. J., Josse, B., Lee, Y. H., MacKenzie, I. A., Plummer, D., Shindell, D. T., Skeie, R. B., Stevenson, D. S., Strode, S., Zeng, G., Curran, M., Dahl-Jensen, D., Das, S., Fritzsche, D., and Nolan, M.: Multi-model mean nitrogen and sulfur deposition from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP): evaluation of historical and projected future changes, Atmos. Chem. Phys., 13, 7997-8018, http://dx.doi.org/10.5194/acp-13-7997-2013, 2013.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X. B., Yang, Z. L., Levis, S., Sakaguchi, K., Bonan, G. B., and Slater, A. G.: Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model, J. Adv. Model Earth Syst., 3, M03001, http://dx.doi.org/10.1029/2011ms000045, 2011.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, Earth Syst. Sci. Data, 5, 165-185, https://doi.org/10.5194/essd-5-165-2013, 2013.

Le Quéré, C., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Peters, G. P., Manning, A. C., Boden, T. A., Tans, P. P., Houghton, R. A., Keeling, R. F., Alin, S., Andrews, O. D., Anthoni, P., Barbero, L., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Currie, K., Delire, C., Doney, S. C., Friedlingstein, P., Gkritzalis, T., Harris, I., Hauck, J., Haverd, V., Hoppema, M., Klein Goldewijk, K., Jain, A. K., Kato, E., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Melton, J. R., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-i., amp, apos, Brien, K., Olsen, A., Omar, A. M., Ono, T., Pierrot, D., Poulter, B., Rödenbeck, C., Salisbury, J., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D., Sutton, A. J., Takahashi, T., Tian, H., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2016, Earth Syst. Sci. Data, 8, 605-649, http://dx.doi.org/10.5194/essd-8-605-2016, 2016.
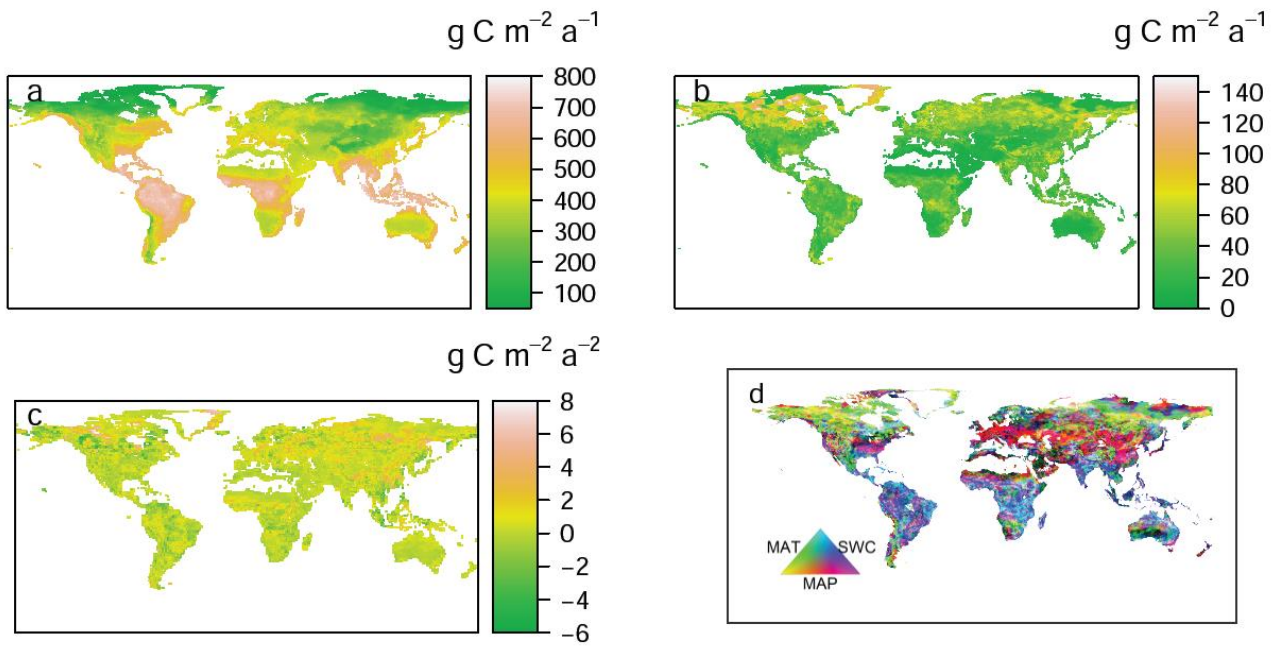
Le Quéré, C., Peters, G. P., Andres, R. J., Andrew, R. M., Boden, T. A., Ciais, P., Friedlingstein, P., Houghton, R. A., Marland, G., Moriarty, R., Sitch, S., Tans, P., Arneth, A., Arvanitis, A., Bakker, D. C. E., Bopp, L., Canadell, J. G., Chini, L. P., Doney,

605   S. C., Harper, A., Harris, I., House, J. I., Jain, A. K., Jones, S. D., Kato, E., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Koven, C., Lefèvre, N., Maignan, F., Omar, A., Ono, T., Park, G. H., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck, C., Saito, S., Schwinger, J., Segschneider, J., Stocker, B. D., Takahashi, T., Tilbrook, B., van Heuven, S., Viovy, N., Wanninkhof, R., Wiltshire, A., and Zaehle, S.: Global carbon budget 2013, Earth Syst. Sci. Data, 6, 235-263, https://doi.org/10.5194/essd-6-235-2014, 2014.

610   LeBauer, D. S. and Treseder, K. K.: Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed, Ecology, 89, 371-379, http://dx.doi.org/10.1890/06-2057.1, 2008.

Li, P., Peng, C., Wang, M., Li, W., Zhao, P., Wang, K., Yang, Y., and Zhu, Q.: Quantification of the response of global terrestrial net primary production to multifactor global change, Ecol. Indic., 76, 245-255, http://dx.doi.org/10.1016/j.ecolind.2017.01.021, 2017.

615   Long, S., J. and Freese, J.: Regression models for categorical dependent variables using Stata. College Station, TX: Stata Press, Texas2006.

Luo, Y. and Zhou, X.: Soil respiration and the environment, Academic press, San Diego, California, 2006.

Metsaranta, J. M., Dymond, C. C., Kurz, W. A., and Spittlehouse, D. L.: Uncertainty of 21st century growing stocks and GHG balance of forests in British Columbia, Canada resulting from potential climate change impacts on ecosystem processes, For.

620   Ecol. Manag., 262, 827-837, https://doi.org/10.1016/j.foreco.2011.05.016, 2011.

Pan, N., Feng, X., Fu, B., Wang, S., Ji, F., and Pan, S.: Increasing global vegetation browning hidden in overall vegetation greening: Insights from time-varying trends, Remote Sens. Environ., 214, 59-72, http://dx.doi.org/10.1016/j.rse.2018.05.018, 2018.

Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Koppen-Geiger climate classification, Hydrol.

625   Earth Syst. Sci., 11, 1633-1644, http://dx.doi.org/10.5194/hess-11-1633-2007, 2007.

Piao, S., Yin, G., Tan, J., Cheng, L., Huang, M., Li, Y., Liu, R., Mao, J., Myneni, R. B., Peng, S., Poulter, B., Shi, X., Xiao, Z., Zeng, N., Zeng, Z., and Wang, Y.: Detection and attribution of vegetation greening trend in China over the last 30 years, Glob. Chang. Biol., 21, 1601-1609, http://dx.doi.org/10.1111/gcb.12795, 2015.

Pumpanen, J., Kolari, P., Ilvesniemi, H., Minkkinen, K., Vesala, T., Niinist, S., Lohila, A., Larmola, T., Morero, M., Pihlatie,

630   M., Janssens, I., Yuste, J. C., Grünzweig, J. M., Reth, S., Subke, J.-A., Savage, K., Kutsch, W., Østreng, G., Ziegler, W., Anthoni, P., Lindroth, A., and Hari, P.: Comparison of different chamber techniques for measuring soil CO_2 efflux, Agric. For. Meteorol., 123, 159-176, http://dx.doi.org/10.1016/j.agrformet.2003.12.001, 2004.

R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. last access: accessed on 11 March, 2019 Access 2018.

635   Reichstein, M. and Beer, C.: Soil respiration across scales: The importance of a model–data integration framework for data interpretation, J. Plant Nutr. Soil Sci., 171, 344-354, http://dx.doi.org/10.1002/jpln.200700075, 2008.

Reichstein, M., Papale, D., Valentini, R., Aubinet, M., Bernhofer, C., Knohl, A., Laurila, T., Lindroth, A., Moors, E., Pilegaard, K., and Seufert, G.: Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites, Geophys. Res. Lett., 34, L01402, http://dx.doi.org/10.1029/2006GL027880, 2007.

640 Schuur, E. A., McGuire, A. D., Schadel, C., Grosse, G., Harden, J. W., Hayes, D. J., Hugelius, G., Koven, C. D., Kuhry, P., Lawrence, D. M., Natali, S. M., Olefeldt, D., Romanovsky, V. E., Schaefer, K., Turetsky, M. R., Treat, C. C., and Vonk, J. E.: Climate change and the permafrost carbon feedback, Nature, 520, 171-179, http://dx.doi.org/10.1038/nature14338, 2015.

Shao, P., Zeng, X., Moore, D. J. P., and Zeng, X.: Soil microbial respiration from observations and Earth System Models, Environmental Research Letters, 8, 034034, http://dx.doi.org/10.1088/1748-9326/8/3/034034, 2013.

645 Sierra, C. A., Trumbore, S. E., Davidson, E. A., Vicca, S., and Janssens, I.: Sensitivity of decomposition rates of soil organic matter with respect to simultaneous changes in temperature and moisture, J. Adv. Model Earth Syst., 7, 335-356, http://dx.doi.org/10.1002/2014ms000358, 2015.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ

650 dynamic global vegetation model, Glob. Chang. Biol., 9, 161-185, http://dx.doi.org/10.1046/j.1365-2486.2003.00569.x, 2003.

Smith, B., Prentice, I. C., and Sykes, M. T.: Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, Global Ecol. Biogeogr., 10, 621-637, http://dx.doi.org/10.1046/j.1466-822X.2001.t01-1-00256.x, 2001.

Subke, J.-A., Inglima, I., and Francesca Cotrufo, M.: Trends and methodological impacts in soil CO_2 efflux partitioning: A

655 metaanalytical review, Glob. Chang. Biol., 12, 921-943, http://dx.doi.org/10.1111/j.1365-2486.2006.01117.x, 2006.

Suseela, V., Conant, R. T., Wallenstein, M. D., and Dukes, J. S.: Effects of soil moisture on the temperature sensitivity of heterotrophic respiration vary seasonally in an old-field climate change experiment, Glob. Chang. Biol., 18, 336-348, https://doi.org/10.1111/j.1365-2486.2011.02516.x, 2012.

Tang, X., Fan, S., Du, M., Zhang, W., Gao, S., Liu, S., Chen, G., Yu, Z., Yao, Y., and Yang, W.: Spatial- and temporal-patterns

660 of global soil heterotrophic respiration in terrestrial ecosystemshttps://doi.org/10.6084/m9.figshare.8882567, 2019a.

Tang, X., Fan, S., Qi, L., Guan, F., Du, M., and Zhang, H.: Soil respiration and net ecosystem production in relation to intensive management in Moso bamboo forests, Catena, 137, 219-228, http://dx.doi.org/10.1016/j.catena.2015.09.008, 2016.

Tang, X., Fan, S., Zhang, W., Gao, S., Chen, G., and Shi, L.: Global variability of belowground autotrophic respiration in terrestrial ecosystems, Earth Syst. Sci. Data Discuss., 2019, 1-25, http://dx.doi.org/10.5194/essd-2019-18, 2019b.

665 Tremblay, S. L., D'Orangeville, L., Lambert, M.-C., and Houle, D.: Transplanting boreal soils to a warmer region increases soil heterotrophic respiration as well as its temperature sensitivity, Soil Biol. Biochem., 116, 203-212, https://doi.org/10.1016/j.soilbio.2017.10.018, 2018.

Trumbore, S.: Radiocarbon and Soil Carbon Dynamics, Annu. Rev. Earth Planet. Sci., 37, 47-66, https://doi.org/10.1146/annurev.earth.36.031207.124300, 2009.

670 Trumbore, S. E. and Czimczik, C. I.: An uncertain future for soil carbon, Science, 321, 1455-1456, https://doi.org/10.1126/science.1160232, 2008.

van den Dool, H., Huang, J., and Fan, Y.: Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981–2001, J. Geophys. Res. Atmos., 108, 8617, http://dx.doi.org/10.1029/2002jd003114, 2003.

Wang, X., Piao, S., Ciais, P., Friedlingstein, P., Myneni, R. B., Cox, P., Heimann, M., Miller, J., Peng, S., Wang, T., Yang, H.,

675 and Chen, A.: A two-fold increase of carbon cycle sensitivity to tropical temperature variations, Nature, 506, 212-215, http://dx.doi.org/10.1038/nature12915, 2014.

Xu, L., Baldocchi, D. D., and Tang, J.: How soil moisture, rain pulses, and growth alter the response of ecosystem respiration to temperature, Global Biogeochem. Cycles, 18, GB4002, https://doi.org/10.1029/2004gb002281, 2004.

Xu, M. and Shang, H.: Contribution of soil respiration to the global carbon equation, J. Plant Physiol., 203, 16-28,

680 https://doi.org/10.1016/j.jplph.2016.08.007, 2016.

Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Xu, B., Shi, J., and Dickinson, R.: The role of satellite remote sensing in climate change studies, Nat. Clim. Chang., 3, 875-883, https://doi.org/10.1038/nclimate1908, 2013.

Yao, Y., Piao, S., and Wang, T.: Future biomass carbon sequestration capacity of Chinese forests, Science Bulletin, 63, 1108-1117, http://dx.doi.org/10.1016/j.scib.2018.07.015, 2018a.

685 Yao, Y., Wang, X., Li, Y., Wang, T., Shen, M., Du, M., He, H., Li, Y., Luo, W., Ma, M., Ma, Y., Tang, Y., Wang, H., Zhang, X., Zhang, Y., Zhao, L., Zhou, G., and Piao, S.: Spatiotemporal pattern of gross primary productivity and its covariation with climate in China over the last thirty years, Glob. Chang. Biol., 24, 184-196, http://dx.doi.org/10.1111/gcb.13830, 2018b.

Zeng, N., Mariotti, A., and Wetzel, P.: Terrestrial mechanisms of interannual CO_2 variability, Global Biogeochem. Cycles, 19, GB1016, https://doi.org/10.1029/2004GB002273, 2005.

690 Zhang, Y., Sha, L., Yu, G., Song, Q., Tang, J., Yang, X., Wang, Y., Zheng, Z., Zhao, S., Yang, Z., and Sun, X.: Annual variation of carbon flux and impact factors in the tropical seasonal rain forest of xishuangbanna, SW China, Sci China Ser D Earth Sci, 49, 150-162, https://doi.org/10.1007/s11430-006-8150-4, 2006.

Zhao, Z., Peng, C., Yang, Q., Meng, F.-R., Song, X., Chen, S., Epule, T. E., Li, P., and Zhu, Q.: Model prediction of biome-specific global soil respiration from 1960 to 2012, Earth's Future, 5, 715-729, http://dx.doi.org/10.1002/2016EF000480, 2017.

695 Zhou, X., Zhou, L., Nie, Y., Fu, Y., Du, Z., Shao, J., Zheng, Z., and Wang, X.: Similar responses of soil carbon storage to drought and irrigation in terrestrial ecosystems but with contrasting mechanisms: A meta-analysis, Agric., Ecosyst. Environ., 228, 70-81, http://dx.doi.org/10.1016/j.agee.2016.04.030, 2016.

Zhu, B. and Cheng, W.: Constant and diurnally-varying temperature regimes lead to different temperature sensitivities of soil organic carbon decomposition, Soil Biol. Biochem., 43, 866-869, https://doi.org/10.1016/j.soilbio.2010.12.021, 2011.

700 Zhu, Z., Piao, S., Myneni, R. B., Huang, M., Zeng, Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneth, A., Cao, C., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y., Liu, R., Mao, J., Pan, Y., Peng, S., Peñuelas, J., Poulter, B., Pugh, T. A. M., Stocker, B. D., Viovy, N., Wang, X., Wang, Y., Xiao, Z., Yang, H., Zaehle, S., and Zeng, N.: Greening of the Earth and its drivers, Nat. Clim. Chang., 6, 791-795, http://dx.doi.org/10.1038/nclimate3004, 2016.

Zhu, Z., Piao, S., Xu, Y., Bastos, A., Ciais, P., and Peng, S.: The effects of teleconnections on carbon fluxes of global terrestrial
705   ecosystems, Geophys. Res. Lett., 44, 3209-3218, http://dx.doi.org/10.1002/2016GL071743, 2017.


**Figure 1.** Distributions of the study sites for RH observations.

710

Figure 2. Spatial patterns of (a) mean data-derived RH (b) standard deviations, (c) temporal trends of annual heterotrophic respiration (RH) from 1980 to 2016, while (d) dominant environmental drivers for the inter-annual variability of global RH. MAT = mean annual temperature; MAP = mean annual precipitation; SWC = soil water content.
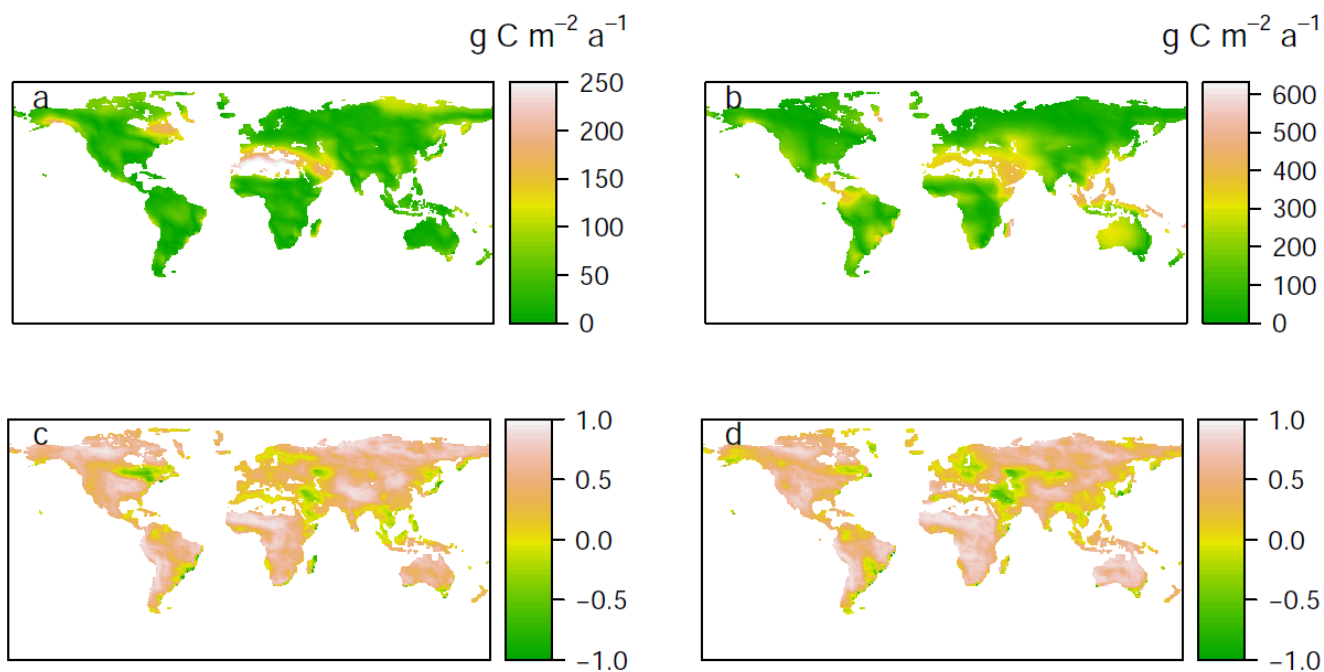
715

Figure 3. Comparing data-derived RH dataset with Hashimoto RH (a, c) and mean RH of TRENDY models (b, d) based on absolute distances (g C m^{-2} a^{-1}, a, b) and cross-correlations (c, d). The absolute distances and cross-correlations were calculated using comparison map profile method (Gaucherel et al., 2008).
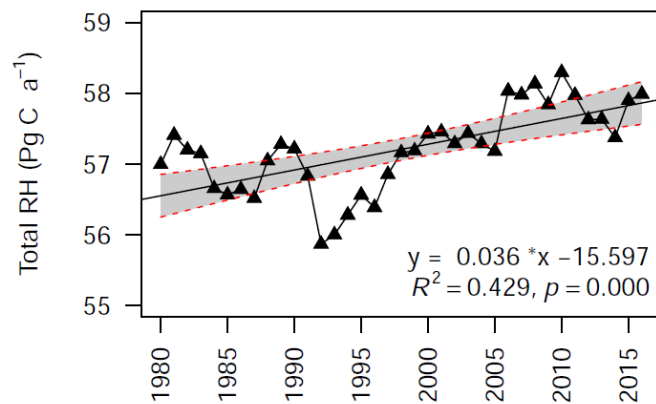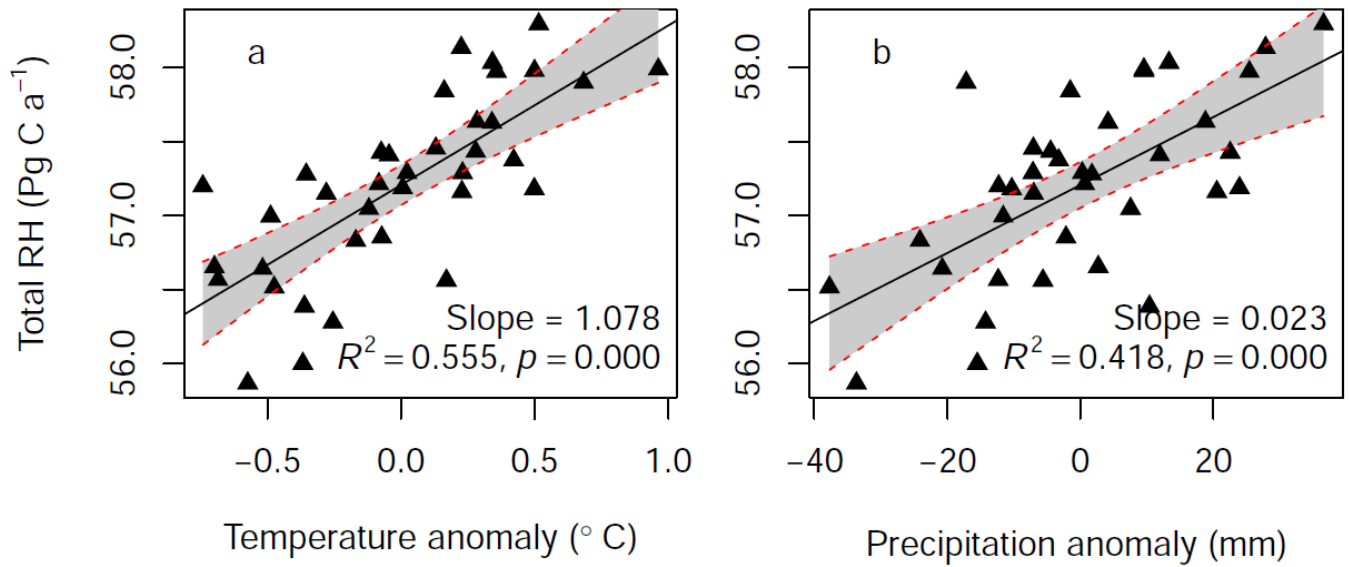
Figure 4. Inter-annual changes in global heterotrophic respiration (RH) from 1980 to 2016. The grey area indicates 95% confidence intervals. For the linear regression model, $R^2 = 0.429$ and $p < 0.01$.

720
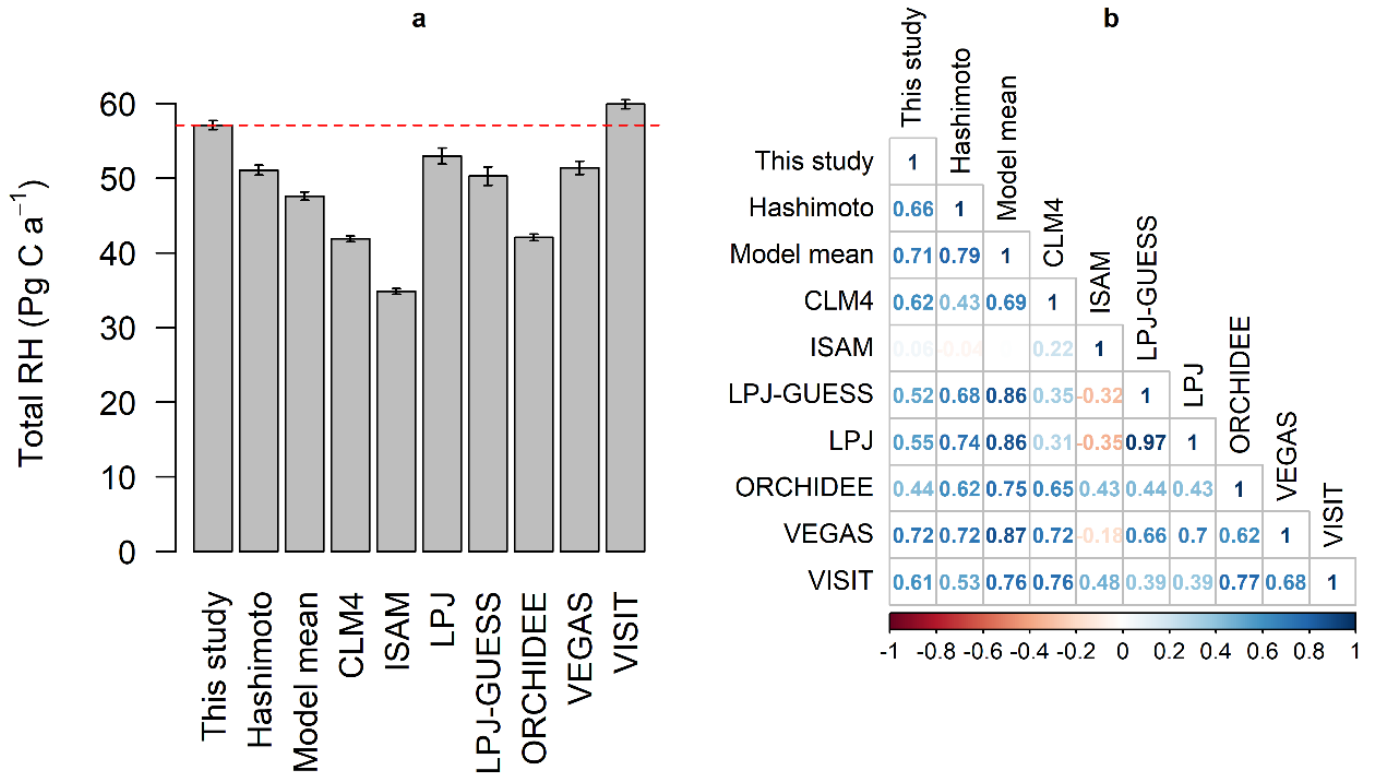
Figure 5. (a) Total global heterotrophic respiration (RH, mean ± standard deviation of annual RH from 1981 to 2010) fluxes and (b) the correlation coefficient analysed by Pearson correlation between data-derived RH and TRENDY /Hashimoto RH. The red dashed line (a) represents the average of data-derived RH from 1981 to 2010.

Figure 6. The relationships between heterotrophic respiration (RH) and mean annual temperature (a) or precipitation anomalies (b). The change was calculated as the difference of each given year to the average over 1980 to 2016. Grey areas indicate 95% confidence intervals.