Dear editor and reviewers,

Thank you very much for your great efforts, comments and suggestion! According to your comments and suggestion, we revised the manuscript carefully and thoroughly. Please see, below, our point-to-point response.

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

Sincerely,

Xiaolu Tang, Wenjie Zhang and Sicong Gao, on behalf of all co-authors.

Response to reviewer #1

This manuscript deals with estimation of global belowground autotrophic respiration (RA) in terrestrial ecosystems. I have some questions in this study. 1) Authors compared global RA by data-derived with that by Hashimoto et al.(2015). However, authors did not refer the data of Hashimoto et al. (2015) in the manuscript. How did authors get the data from Hashimoto? Please explain the difference between Random forest model and methods of Hashimoto et al. (2015).

Response: we apologize for the unclear statement of Hashimoto RH.

Hashimoto RA was publicly available at:

http://cse.ffpri.affrc.go.jp/shojih/data/index.html, therefore, we obtained the annual RA product for our study. Such information was added in text:

“In order to compare with the solely global RA product generated by Hashimoto et al. (2015), which was estimated by a climate-driven model using temperature and precipitation only and obtained from the public available dataset (http://cse.ffpri.affrc.go.jp/shojih/data/index.html)”.

1
Hashimoto et al. (2015) proposed a global RA based on the difference of heterotrophic respiration and total soil respiration, and total soil respiration was predicted by a climate-driven model using temperature and precipitation only and global soil respiration dataset. Therefore, Hashimoto RA did not consider other environmental control, such as soil carbon, on RA (Hashimoto et al., 2015).

To fill such knowledge gap, we applied a Random Forest algorithm to model global RA with field observations and 11 environmental variables in terms of different aspects of environmental controls on RA, and we obtained a much higher model efficiency (52%) compared to Hashimoto RA (32%). Furthermore, Random Forest algorithm have great potentials to address the non-linear correlation between RA and environmental variables, and remove auto-correlations among environmental variables.

2) Authors used PgC a-1 or gC m-2 a-1 for the unit of RA, but I guess that a-1 should be yr-1. Please correct all unit in the manuscript and figures.

Response: yes, it means Pg C per year. Corrected to “Pg C yr⁻¹” or “g C m⁻² yr⁻¹” throughout the manuscript and figures!

3) Authors discussed about importance of the dominant environmental factors for estimate spatio-temporal variation in RA. I think that it is important not only environmental factors for plant production but also plant biomass because root respiration would have positive correlation with plant biomass. Why did authors ignore the global pattern of plant biomass??

Response: thank you for the good comments. We agree with you that plant biomass, particularly root biomass, would have positive correlations with RA. However, selecting variables is constrained by the fact that a variable must be available at all sites and at the corresponding global product simultaneously. For instance, if a variable is measured accurately at sites, but with large uncertainties in the corresponding global product, it may be advantageous to exclude this variable from the analysis (Jung et al., 2011).

Although we tried to include global plant or root biomass as a driving variable, we
found such product was only available for a single year, or mean values of several years (Huang et al., 2017), or forests (Hengeveld et al., 2015), and there was a lack of time-series global biomass product covering all land covers. Given the fact that plant biomass was highly dynamic due to annual accumulation, using a global biomass for a given year or particularly ecosystem type to represent the biomass dynamics covering all terrestrial ecosystems would cause a great uncertainty to RA estimation. Therefore, the lack of global biomass product constrained the use of plant biomass as a driving variable for RA in this study. Instead, we used MODIS land cover as one of driving variables, which could indirectly reflect the biotic or biomass control on RA to some extent.

Finally, please considering my specific comments and get some English proofreading. In addition, please reconsider carefully about all figures, because I feel that some figures are not important in this manuscript. If authors resolve these questions, I think that this manuscript would be better for global data science.

Response: we answered each of your specific comment carefully, and we improved the English.

As you suggested, see specific comments below, Figure 6c was not important and removed.

Specific comments

Page 3, line 54, “which is almost 5 times of:” I cannot understand relationship between this sentence and preceding sentence. Page 3, line 56, “Therefore, an accurate estimate of:” I think that authors did not enough explain the reasons before the sentences. Please add more explanation.

Response: Since the two comments link with each other, we answer the two comments together.

We apologize for the unclear statement. We revised and added more explanation for it
as follows:

“RA could amount roughly up to 54 Pg C yr\(^{-1}\) (1 Pg = 10\(^{15}\) g, calculating RA as an approximate ratio of 0.5 of soil respiration, more details in Hanson et al., 2000) according to different estimates of global soil respiration (Bond-Lamberty, 2018), which is almost 5 times of the carbon release from human activities (Le Quéré et al., 2018). However, the contribution of RA to soil respiration varied greatly from 10% to 90% across biomes, climate zones and among years (Hanson et al., 2000), leading to the strong spatial and temporal variability in RA. Thus, whether RA varies with ecosystem types or climate zones remains an open question at the global scale (Ballantyne et al., 2017). Consequently, an accurate estimate of RA and its spatial-temporal dynamics are critical to understand the response of terrestrial ecosystems to global carbon cycling and climate change.”

Page 8, Figure 2: I cannot understand the meaning of the figure 2c and 2b. Why did authors indicate the standard deviation of temporal variation in RA??

Response: Fig. 2b is the mean value of Hashimoto RA over 1980-2012, while Fig. 2c represents the standard deviation of predicted RA in this study. The figure caption was revised as:

“Spatial patterns of annual mean and standard deviation of belowground autotrophic respiration (RA) from 1980 to 2012 for this study (a, c) and Hashimoto RA (b, d), respectively”

Due to the inter-annual variability of environmental controls on RA, RA varied annually. Although Fig. 6 describes the annual variability of total RA, the spatial pattern of annual variability of RA is lacking. To characterize the spatial pattern of annual variability of RA, the standard deviation of RA from 1980-2012 was employed. Such analysis was also conducted in other studies, e.g. Yao et al. (2018). Therefore, we used standard deviation to represent the temporal pattern of RA.

Page 11, Figure 6: what the difference of Fig.6a and Fig.6b? Please add more explanation. And, please make the same value of y-axis in both of Fig.6a and Fig.6b.
And I think that Fig.6c is not needed.

Response: Fig. 6a represents the annual variability of predicted RA in this study, while Fig. 6b represents the annual variability of Hashimoto RA. The same value of yaxis from 39 – 45 Pg C yr\(^{-1}\) was applied.

Fig. 6c was not important and removed.

We corrected the figure description more clearly:

“Figure 6 Annual variability of belowground autotrophic respiration (RA) for this study (a) and Hashimoto RA (b) from 1980 to 2012. The grey area represents 95% confidence interval.”

Page 12, Line 254 to 227, “All the biomes, except: : :, respectively”: please rewrite these sentences. Grammatical subject is RA, I think.

Response: we rewrite these sentences:

“RA showed a significantly increasing trend during 1980-2012 (p\(_s\) < 0.01) in most of the biomes, except temperate forest, savannas and wetland. RA in tropical forests, boreal forests and cropland increased by 0.0076±0.0015, 0.0047±0.0016, 0.0036±0.0014 Pg C yr\(^{-2}\), respectively.”

Page 12, Line 259, “a significant increasing trend of: : : ’”: is this “a significant increasing trend of total RA in temperate zones,: : :”??

Response: thank you for your careful revision. Yes, we mean “a significant increasing trend of total RA in temperate zones….”. We revised the text:

“there were significant increasing trends of total RA in temperate zones, temperate forest, savannas and wetland of Hashimoto RA, which were not observed in data-derived RA”.

Page13, Figure 9: I cannot understand the importance of this figure.

Response: We appreciate your question. Figure 9 showed the relative importance of three main environmental drivers – MAT, MAP and SWR, by colors with RGB plot.
Due to different ecosystem types, or plant functional types or climate zones, the dominant factors may vary. As indicated by Fig. S7, 56% of land area was dominated by precipitation, while temperature and shortwave radiation dominated 19% and 25% of global land areas, which indicated an uneven control of environmental factors on RA. Therefore, Figure 9 showed the spatial variability of dominance of MAT, MAP and SWR on RA. It was found that the dominance of precipitation on RA was globally distributed, particularly dry or semi-arid areas, such as Northwest China, Southern Africa, Middle Australia and America, while temperature controlled RA mainly in in tropical Africa, Southern Amazon rainforests, Siberia and partly tundra, and shortwave radiation dominated high latitudinal areas, e.g. Eastern America and middle and Eastern Russian. Such analysis have been widely used in other studies, e.g. gross primary production (Yao et al., 2018), earth greening (Zhu et al., 2016), vegetation productivity (Seddon et al., 2016).

RGB synthesis (Fig. 9) was performed on stretched values of partial correlation coefficients, an effective way to illustrate the spatial distribution of dominant driving factors of RA (Yao et al., 2018), which could increase our understanding the mechanisms and spatial variability of environmental controls on RA at the global scale.

Page 14, Line 290 “For example, temperature was the: :Australia” is that the result of Hashimoto et al.(2015)?

Response: thank you for your careful revision again. Yes, we mean “Hashimoto RA”, and revised in text:

“temperature was the main dominant factor for most area of Australia for Hashimoto RA”.

References


Response to Reviewer #2

I have read "Global variability of belowground autotrophic respiration in terrestrial ecosystems". In the manuscript, the authors estimated global belowground autotrophic respiration from 1980-2012, analyzed the temporal trend, and explored the dominant factors for autotrophic variability. Global autotrophic respiration is a big carbon exchange between the atmosphere and terrestrial, but was rarely studies in the past years. Global temporal and spatial variability of autotrophic respiration is clearly a timely and interesting topic. Generally, this manuscript is well organized and easy to follow. The results and conclusions are reasonable. The production (Global belowground autotrophic respiration shared in the figShare) is a contribution to the community and potentially can serve as a benchmark for ecosystem models, it will be useful also make the analysis (include the codes) public available to make the analysis reproducible. But I think the authors have to better address the limitation, weakness, and uncertainty of this study. In my opinion, some major limitation including: 1) The sample size of RA: there are much less annual RA comparing with annual Rs (less than 10%), even though the authors extended the RA dataset by new papers from China Knowledge Resource Integrated (CNKI) Database, the total samples is only 449. And the majority of the samples are from the forest, samples from wetland and shrubland are extremely lacking (only 5 observations).

Response: we also attached the dataset and the R codes to generate the main results to figshare at https://doi.org/10.6084/m9.figshare.7636193.

Based on SRDB v4, including new observations from CNKI, we got a total of 4276 observations for soil respiration, however, there were 697 observations for RA. According to our selecting criteria: e.g. RA measurement lasting for one year; excluding measurements with Alkali absorption and soda lime; no site management, we got a RA dataset of 449 observations. Our dataset are mainly from forests, but a lack of observations in wetland and shrubland, which could be the limitation in this study.

We have discussed the limitation in “4.4 Advantages, limitations and uncertainties”
section as follows:

“Finally, uneven coverage of observations in the updated database would be another source of uncertainties. Although our dataset had a wide range of land cover, the observational sites mainly distributed in China, Europe and North America and were dominated by forests. There was a great lack of observations in areas, such as Africa, Austria and Russia, and biomes, such as tropical forest, shrubland, wetland and cropland. Consequently, RA observations caused bias of RF model toward the regions with more observations.”

2) How can you evaluate the quality of the RA data? Even though the authors conducted quality control on the RA data, but it does not guarantee the reliability of the RA data. We lack reliable methods to separate RA and RH, current ways (e.g., trend, gap, girdling, clip, and isotope) have their own problem. Further, usually RH is measured, and RA was calculated as the difference between RS and RH, which also bring uncertainties. All those issues were not addressed and discussed in the manuscript. If the data reliability cannot be guaranteed, the estimates, trend, and dominant factors should also be questioned. Despite the above problems, I still think this study tend to address an important topic and may inspire more research in the future.

Response: we evaluate the quality of RA from different aspects to guarantee the reliability of RA: (1) measuring approaches: Alkali absorption and soda lime were not included due to the potential underestimate of respiration rate with the increasing pressure inside chamber (Pumpanen et al., 2004); (2) data quality control by quality flag: Q01 (estimated from figure), Q02 (data from another study), Q03 (data estimated-other), Q04 (potentially useful future data), Q10 (potential problem with data), Q11 (suspected problem with data), Q12 (known problem with data), Q13 (duplicate?), Q14 (inconsistency). Therefore, RA or total soil respiration observations labelled by “Q10”, “Q11”, “Q12”, “Q13” and “Q14” were removed in this study. More details on data quality controls can be found in Bond-Lamberty and Thomson (2010a).

We agree with you that there was a lack reliable method to separate RA and RH, and
current ways (e.g., trend, gap, girdling, clip, and isotope) have their own problem.

We have discussed the data quality and limitation of unreliable method to separate RA and RH in **“4.4 Advantages, limitations and uncertainties”** as follows:

“First, although we conducted a data quality control in this study, a lack of reliable approach to separate RA and heterotrophic respiration may lead to an uncertainty of RA values. There are several approaches, e.g. trenching, stable or radioactive isotope, gridding (Bond-Lambery et al., 2004; Högberg et al., 2001; Hanson et al., 2000), however, each of these approaches has its own limitations. For example, trenching has been widely applied to partition RA and heterotrophic respiration due to easy operation and low cost, on the other hand, heterotrophic respiration may be increased due to the termination of water uptake by roots and the decomposition of remaining dead roots in trenching plots (Hanson et al., 2000; Tang et al., 2016). Commonly, RA was calculated from the difference between total soil respiration and heterotrophic respiration, thus the trenching approach might lead to an underestimation of RA. In our dataset, a total of 254 RA observations were estimated by trenching approach, while the rest RA observations were estimated by other separation approaches, e.g. isotope, radiocarbon, mass balance. Thus, inconsistent separation approaches could be another source of uncertainty of RA values.”

Specific comments Abstract

Line 22: (srdb v4) but later (line 97) you used (srdb version 4), be consistent.

Response: done!

Line 24: the unit for RA increasing trend should be Pg C a-2? Please see this paper: Ballantyne, A., Smith, W., Anderegg, W., Kauppi, P., Sarmiento, J., Tans, P., Shevliakova, E., et al. (2017). the warming hiatus due to reduced respiration. Nature Climate Change, 7(2), 148. [https://doi.org/10.1038/NCLIMATE3204 – 152].

Response: thank you for your kind recommendation, and we corrected the increasing unit to Pg a yr-2 or g C m-2 yr-2 throughout the text and figures.

https://doi.org/10.1038/NCLIMATE3204 – 152.

Response: thank you for your kind recommendation. Huang et al. (2012) mainly discussed the uneven changes of temperature, not RA.

Jian et al. (2018) found uneven changes of soil respiration in different areas, and Ballantyne et al. (2017) also proposed that belowground autotrophic respiration may be varied among ecosystem types. These references have been cited to support our conclusions, and revised in the text as follows:

“However, RA increment varied with climate zones and ecosystem types (Figs. S2 and S3), which was similar to previous findings (Ballantyne et al., 2017; Jian et al., 2018a), who found that total soil respiration or RA varied with climate zones or ecosystem types.”

Introduction

Line 48: It is not accurate to say RA is the second largest source of carbon fluxes from soil because we don’t know whether Ra is larger than Rh. And does the (Raich and Schlesinger 192) paper really say that? And in line 309 you said Rh account for 0.54-0.63, means RH > RA.

Response: we apologize for the improper statement. We mean soil respiration is the
second largest carbon flux. We revise the text:

“RA is one main component of soil respiration (Hanson et al., 2000), and soil respiration represents the second largest source of carbon fluxes from soil to the atmosphere (after gross primary production, GPP) in the global carbon cycle (Raich and Schlesinger, 1992).”


Response: thank you for your recommendation. We cited the global estimates of soil respiration summarized by Bond-Lamberty (2018) to support our study.

“RA could amount roughly up to 54 Pg C yr\(^{-1}\) (1 Pg = 10\(^{15}\) g, calculating RA as an approximate ratio of 0.5 of soil respiration, more details in Hanson et al., 2000) according to different estimates of global soil respiration (Bond-Lamberty, 2018), which is almost 5 times of the carbon release from human activities (Le Quéré et al., 2018).”

Line 62-63: a citation needs to support this statement.

Response: done! We revised the text as follows:

“the globally spatial and temporal pattern of RA has not been explored and still acts as a “black box” in global carbon cycling (Ballantyne et al., 2017)”.

Line 63-64: need a citation.

Response: Revised as follows in the text:

“This “black box” is not well constrained and validated, because most terrestrial ecosystem models and earth system models were commonly calibrated and validated against eddy covariance measurements of net ecosystem carbon exchange (Yang et al., 2013)”.

Line 85: “linear of non-linear models” change to “linear and non-linear models”.
Response: done!

Line 86: But in line 94, you said RF model can avoid overfitting. Zhao et al 2017 used ANN models; and Jian et al 2018 also include RF models. So you need to be concise to avoid inconsistent.

Response: Zhao et al 2017 was appropriate and removed in L94!

Line 95: Zhao et al. 2017 used ANN models, it is not appropriate to cite here.

Response: Zhao et al. 2017 was removed, while Bodesheim et al., 2018 and Jung et al. 2017 were cited here.

Line 96: It is better also include the GitHub commit number of SRDB.

Response: the doi number was added.

Line 105: other environmental factors is too broad, please to be more specific.

Response: revised! We specified the soil and vegetation factors.

“It will also advance our knowledge of the co-variation of RA with climate, soil and vegetation factors”

Material and methods

A big point in this study is you compared your results with that from Hashimoto (2015), you need to talk about how you get the RA data of Hashimoto (2015). You directly used their data or you reproduced their estimates. If you reproduced, how and whether you used the same climate data as Hashimoto?

Response: we apologize for the misleading of Hashimoto RA. Hashimoto RA is publicly available at http://cse.ffpri.affrc.go.jp/shojih/data/index.html, therefore, we obtained the annual Hashimoto RA product for our study. Such information was added in text:

“In order to compare with the solely global RA product generated by Hashimoto et al. (2015), which was estimated by a climate-driven model using temperature and
precipitation only and obtained from the public available dataset (http://cse.ffpri.affrc.go.jp/shojih/data/index.html)

Line 110-112: are those papers from CNKI all in Chinese? How many studies and how many more data records you got from that? Please clarify that.

Response: yes, those papers from CNKI are all in Chinese with English abstract. We added 68 more RA observations and revised in the text:

“Finally, this study included a total of 449 field observations (Fig. 1), including 68 observations from CNKI.”

Line 122: Australia, Russia, Africa, and South America.

Response: done!


Results: from 1960 to 1980, there are only 11 observations, which might bring uncertainties. Our study covered the period until 2012 for easily comparing with Hashimoto RA, which covered the period up to 2012.

Line 224: ‘-4 – 4’ change to ‘-4 to 4’.

Response: done!

Line 224-225: ‘East Russia and tropical and Eastern regions in Africa’ change to ‘East Russia, tropical, and Eastern regions in Africa’.

Response: done!

Line 264-265: Usually anomaly was the difference between temperature/precipitation of corresponding year to the mean of a period (e.g., 1980-2012 in this study). But this should not change the results, if previous studies calculate anomaly like yours, please provide a citation to support.

Response: thank you for your suggestion, and we followed the suggestion. The anomaly of temperature/precipitation of corresponding year to the mean of 1980-2012, and the
results did not change.

Line 270-273: why in temperate zone/savannas/wetland there is no correlation between RA and temperature anomaly? That is interesting, usually, in tropical and subtropical regions, Rs is less correlated with temperature (and should be also true for the temperature anomaly). I think it worth to analyze in more details and try to explain the mechanism or maybe just because of the uncertainty.

Response: the different responses of ecosystem types or climate zones to climatic variables may be related to regional heterogeneity and plant functional trait. For example, regional temperature significantly differed from global averages (Huang et al., 2012), with much faster change in high-latitude regions (Hartmann et al., 2014), and semi-arid dominated the trend and variability of global land CO$_2$ sink (Ahlström et al., 2015). Similar studies were also found in other studies, e.g. total soil respiration or RA (Ballantyne et al., 2017; Jian et al., 2018a). Therefore, the regionally uneven responses of RA to climatic variables were unlikely due to model uncertainty.

These results have been discussed in “4.1 Global RA” section, and we revised the text as:

“However, RA increment varied with climate zones and ecosystem types (Figs. S2 and S3), which was similar to previous findings (Ballantyne et al., 2017; Jian et al., 2018a), who found that total soil respiration or RA varied with climate zones or ecosystem types. These differences may be related to regional heterogeneity and plant functional trait. For example, regional temperature significantly differed from global averages (Huang et al., 2012), with much faster change in high-latitude regions (Hartmann et al., 2014), and semi-arid dominated the trend and variability of global land CO$_2$ sink (Ahlström et al., 2015). Therefore, the regionally uneven responses of RA to climatic variables highlights the urgent need to account for regional heterogeneity when studying the effects of climate change on ecosystem carbon dynamics in future.”

Line 310-311: See also Lamberty 2018 Earth’s Future paper. "New techniques and data for understanding the global soil respiration flux." Earth’s Future 6.9 (2018): 1176-
Response: thank you for the recommendation, and we cited the global soil respiration estimates from Bond-Lamberty (2018):

“Bond-Lamberty et al. (2018) proposed that the global average proportion of heterotrophic respiration ranged from 0.54 to 0.63 over 1990-2014 and global total soil respiration was 67 to 108 Pg C yr\(^{-1}\) using different approaches and datasets Bond-Lamberty (2018); (Bond-Lamberty and Thomson, 2010b; Hashimoto et al., 2015; Hursh et al., 2017; Jian et al., 2018b), thus global RA varied from 25 to 51 Pg C yr\(^{-1}\).”

Discussion

Dominant factors: all you talked were about driving factors of RA spatial variability, right? Did you also analyze the dominant factors of temporal variability? Limitation and uncertainty: see my previous overall comment. In addition, Jian et al. "Constraining estimates of global soil respiration by quantifying sources of variability." Global change biology 24.9 (2018): 4143-4159 talked about uncertainty related to time-scaling and Rs upscaling. How about RA upscaling and timescale?

Response: we analyzed the dominate factors at both spatial and temporal patterns. We used partial correlation analysis based on a timescale from 1980 to 2012 for each grid cell (see methodology section 2.5), and the correlation coefficient was applied to derive the dominant factor map (Fig. 9). However, we did not analyze the dominant factors for each given year.

We additionally discussed the potential variability of RA using different time scale variables in “4.4 Advantages, limitations and uncertainties”.

“Second, due to the limited observations of RA at a daily or monthly scale, this study only predicted RA at an annual scale. Although there was no direct study to compare the difference of RA upscaling from daily or monthly and annual scale, substantial difference of total soil respiration upscaling from daily or monthly and annual scales (Jian et al., 2018b) indirectly illustrated the potential difference of RA upscaling from
different timescales."

Author contributions

Line 445: ‘to the review the manuscript’ change to ‘to review the manuscript’.

Response: done.

References


Huang, J., Guan, X., and Ji, F.: Enhanced cold-season warming in semi-arid regions, Atmospheric Chemistry and Physics, 12, 5391-5398, http://dx.doi.org/10.5194/acp-12-5391-2012, 2012.


Global variability of belowground autotrophic respiration in terrestrial ecosystems

Xiaolu Tang¹,², Shaohui Fan³, Wenjie Zhang⁴,⁵,*, Sicong Gao⁵,*, Guo Chen¹, Leilei Shi⁶

¹College of Earth Science, Chengdu University of Technology, Chengdu, P. R. China
²State Environmental Protection Key Laboratory of Synergetic Control and Joint Remediation for Soil & Water Pollution, Chengdu University of Technology, Chengdu, P. R. China
³Key laboratory of Bamboo and Rattan, International Centre for Bamboo and Rattan, Beijing, P. R. China
⁴State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Beijing, P. R. China
⁵School of Life Science, University of Technology Sydney, NSW, Australia
⁶Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions, College of Environment and Planning, Henan University, Jinming Avenue, Kaifeng, P. R. China

Correspondence to: Wenjie.Zhang@uts.edu.au; Sicong.Gao@student.uts.edu.au
Abstract

Belowground autotrophic respiration (RA) is one of the largest, but highly uncertain carbon flux components in terrestrial ecosystems. However, RA has not been explored globally before and still acted as a “black box” in global carbon cycling currently. Such progress and uncertainty motivate a development of global RA dataset and understand its spatial and temporal pattern, causes and responses to future climate change. This study We used Random Forest to study investigate RA’s spatial and temporal pattern at the global scale by linking the updated field observations from Global Soil Respiration Database (v4) with global grid temperature, precipitation and other environmental variables. Globally, mean RA was 43.8±0.4 Pg C ayrt−1 with a temporally increasing trend of 0.025±0.006 Pg C ayrt−2 from 1980 to 2012. Such increment trend was widely spread with 58% global land areas. For each 1 °C increase in annual mean temperature, global RA increased by 0.85±0.13 Pg C ayrt−2, and it was 0.17±0.03 Pg C ayrt−2 for 10 mm increase in annual mean precipitation, indicating a positive feedback of RA to future climate change. At a global scale, Precipitation was the main dominant climatic drivers of the spatial pattern of controlling RA, accounting for 56% of global land areas with widely spread globally, particularly in dry or semi-arid areas, followed by shortwave radiation (25%) and temperature (19%). Different temporal patterns for varying climate zones and biomes indicated uneven response of RA to future climate change, challenging the perspective that the parameters of global carbon stimulation independent on climate zones and biomes. The developed RA database, the missing carbon flux component that is not constrained and validated in terrestrial ecosystem models and earth system models, will provide insights into understanding mechanisms underlying the spatial and temporal variability of belowground carbon dynamics. The developed RA database also has great potentials to serve as a benchmark for future data-model comparisons. The RA product is freely available at

https://doi.org/10.6084/m9.figshare.7636193.
1 Introduction

Belowground autotrophic respiration (RA) mainly originated from plant roots, mycorrhizae, and other micro-organisms in the rhizosphere directly relying on labile carbon component leaked from roots (Hanson et al., 2000; Tang et al., 2016; Wang and Yang, 2007). Thus, RA reflects the photosynthesis derived carbon respired back to the atmosphere by roots and regulates the net photosynthetic production allocation to belowground tissues (Högberg et al., 2002). RA is also one main component of soil respiration (Hanson et al., 2000), which and soil respiration represents the second largest source of carbon fluxes from soil to the atmosphere (after gross primary production, GPP) in the global carbon cycle (Raich and Schlesinger, 1992). Globally, RA could amount roughly up to 54 Pg C yr\(^{-1}\) (1 Pg = 10\(^{15}\) g, calculating RA as an approximate ratio of 0.5 of soil respiration, more details in Hanson et al., 2000) according to different estimates of global soil respiration- (Bond-Lamberty, 2018), which is almost 5 times of the carbon release from human activities (Le Quéré et al., 2018). RA contributed about 50% of total soil respiration on average. However, the contribution of RA to soil respiration varied greatly from 10% to 90% across biomes, climate zones and among years (Hanson et al., 2000), leading to the strong spatial and temporal variability in biomes and climate regions. Thus, whether RA varies with ecosystem types or climate zones remains an open question at the global scale (Ballantyne et al., 2017). Thus global mean RA would amount from roughly 42 to 54 Pg C a\(^{-1}\) (1 Pg = 10\(^{15}\) g) according to recent estimates of soil respiration (Bond-Lamberty and Thomson, 2010b; Hursh et al., 2017), which is almost 5 times of the carbon release from fossil fuel combustion from human activities (Le Quéré et al., 2018). Therefore, an accurate estimate of RA and its spatial-temporally dynamics is critical to understand the response of terrestrial ecosystems to global carbon cycling and climate change.

Due to the difficulties of separation and direct measurement of RA at varying spatial scales and its diurnal, seasonal and annual variabilities, RA becomes one of the largest but highly uncertain carbon flux components in terrestrial ecosystems. Although individual site measurements of RA have been conducted across ecosystem types and biomes, knowledge gap still remains even though with a large number of field measurements (Hashimoto et al., 2015). Consequently, the globally spatial and temporal pattern of RA has not been explored and still acts as a “black box” in global carbon cycling (Ballantyne et al., 2017). This “black box” is not well constrained and validated, because most terrestrial ecosystem models and earth system models were commonly calibrated and validated against eddy covariance measurements of
net ecosystem carbon exchange (Yang et al., 2013). Such progress and uncertainty motivate a development of global RA dataset from observations and understand its spatial and temporal pattern, causes and responses to future climate change. Despite of the general agreement that global soil respiration increased during last several decades (Bond-Lamberty et al., 2018; Bond-Lamberty and Thomson, 2010b; Zhao et al., 2017), how global RA responding to climate change is far from certain because of different temperature sensitivities of RA across terrestrial ecosystems (Liu et al., 2016; Wang et al., 2014). Therefore, reducing RA uncertainty and clarifying its response to climate change, particularly to temperature and precipitation, is essential for global carbon allocation and future projection of the impact of climate change's effects on global terrestrial carbon cycle.

Although several studies have globally estimated soil respiration and its response to climate variables (Bond-Lamberty and Thomson, 2010b; Hursh et al., 2017; Zhao et al., 2017), such efforts have not been conducted for global RA directly. Although Hashimoto et al. (2015) indirectly derived RA via the difference between total soil respiration and heterotrophic respiration, it probably led to uncertainty due to merely using include the temperature and precipitation as the only model drivers and a low model efficiency (32%). Besides temperature and precipitation, other variables, e.g. soil water, carbon and nitrogen content, are additionally critical factors regulating RA, and those factors generally varied with biomes and climate zones. Consequently, Hashimoto et al. (2015) may not reflect the key processes affecting RA, such as soil nutrient limitation.

On the other hand, the climate-derived models usually explain < 50% variability of soil respiration (Bond-Lamberty and Thomson, 2010b; Hashimoto et al., 2015; Hursh et al., 2017), which might be another uncertainty source. Recent studies have included more variables and field observations to promote improve the prediction ability of the linear and non-linear models (Jian et al., 2018b; Zhao et al., 2017). Yet, however, it may propagate error because of the overfitting and autocorrelation among these variables (Long and Scott, 2006). Random Forest (RF, Breiman, 2001), a machine learning approach, could overcome these issues based on the hierarchical structure, and be insensitive to outliers and noise compared to single classifiers (Breiman, 2001; Tian et al., 2017). RF uses a large number of ensemble regression trees but a random selection of predictive variables (Breiman, 2001). RF only requires two free parameter settings: the number of variables sampled as candidates for each split and the number of trees. The performance of the RF model usually is not sensitive to the number of trees and number of
variables. Moreover, RF regression can deal with a large number of features, and which could help feature selection based on the variable importance and can avoid overfitting (Jian et al., 2018b). RF has been widely used for carbon fluxes modelling in recent years (Bodesheim et al., 2018; Jung et al., 2017; Jung et al., 2017; Li et al., 2017; Zhao et al., 2017).

Therefore, this study we firstly applied RF algorithm to retrieve global RA based on the updated RA field observations from the most updated global soil respiration dataset (SRDB v4, doi: 10.5194/bg-7-1915-2010, Bond-Lamberty and Thomson, 2010a) with the linkage of other global variables (see “materials and methods” part) in this study for the first time, aiming to: (1) develop a global RA product using field observations across the globe (named data-derived RA); (2) estimate RA’s spatial and temporal pattern; (3) identify the main driving factors of RA’s spatial and temporal variabilities; (4) compare with the previous RA estimates from Hashimoto et al. (2015). The outcome of this study will advance our understanding of global RA and its spatial and temporal variabilities. The proposed RA product is expected to serve as a benchmark for global vegetation models and its role in global carbon cycling. It will also advance our knowledge of the co-variation of RA with climate, and soil and vegetation other environmental factors, further link the empirical observations temporally and spatially to bridge the knowledge gap among local, regional and global scales.

2 Material and methods

2.1 RA database development

First, RA database was developed based on observations across the globe from SRDB (Bond-Lamberty and Thomson, 2010a), which is publicly available at https://github.com/bpbond/srdb. Then we further updated the database using observations collected from China Knowledge Resource Integrated Database (www.cnki.net) up to March 2018, which followed the identical criteria applied in SRDB development. To control the data quality, annual RA observations were filtered that: (1) annual RA was directly reported in publications indicated by “years of data” of SRDB; (2) the start and end years were recorded in literatures or expanded from “years of data” of SRDB; (3) soil respiration measurements with Alkali absorption and soda lime were not included due to the potential underestimate of respiration rate with the increasing pressure inside chamber (Pumpanen et al., 2004); (4) observations with treatments of nitrogen addition, air/soil warming, and rain/litter exclusion were not included, except cropland; (5)
potential problems observations (labelled by “Q10”, “Q11”, “Q12”, “Q13” and “Q14”) were excluded. Finally, this study included a total of 449 field observations (Fig. 1), including 68 observations from CNKI. RA observations were absolutely dominated by forest ecosystem (379 observations) with globally unevenly distributed, mainly from China, America and Europe. There was a great lack of RA observations in Australia, Russia, and Africa, and South America.

![Observational sites used in this study](image)

**Figure 1** Observational sites used in this study

### 2.2 Vegetation, climate and soil data

Global land cover with a half degree resolution was obtained from MODIS land cover ([https://gfc.umd.edu/data/lc/](https://gfc.umd.edu/data/lc/)). Monthly grid data of temperature, precipitation, diurnal temperature range and potential evapotranspiration at 0.5° resolution were obtained from CRU TS Version 4.01 from 1901 to 2016 ([https://crudata.uea.ac.uk](https://crudata.uea.ac.uk)) (Harris et al., 2014). Monthly shortwave radiation (SWR), Palmer Drought Severity Index (PDSI) and soil water content at 0.5° resolution were from NOAA/ESRL Physical Sciences Division ([https://www.esrl.noaa.gov](https://www.esrl.noaa.gov)) (Kalnay et al., 1996). Soil organic carbon content with a resolution of 250 m was downloaded from soil grid data ([https://soilgrids.org](https://soilgrids.org)) (Hengl et al., 2017), and soil nitrogen content was from Spatial Data Access Tool ([https://webmap.ornl.gov/ogc/index.jsp](https://webmap.ornl.gov/ogc/index.jsp)), while monthly nitrogen deposition data with a resolution of 0.5° were downloaded from the Earth System Models of GISS-E2-R, CCSM-CAM3.5 and GFDL-AM3, covering since 1850s ([https://www.isimip.org](https://www.isimip.org)). The monthly global variables were first aggregated to year scale and then resampled to a 0.5° resolution using bilinear interpolation for those variables without a 0.5° resolution. These variables could represent different aspects controlling RA variability. For instance, temperature,
precipitation and soil water content are most important variables controlling plant photosynthesis, which is the primary carbon source of RA (Högberg et al., 2002; Högberg et al., 2001). Finally, global variables of each given site extracted by coordinates corresponding with annual RA estimates from the SRDB.

2.3 Random Forest-based RA Modelling

In this study, a RF model was trained with the 11 variables listed in Table S1 by caret by linking RandomForest package in R 3.4.4 (Kabacoff, 2015), then the trained model was implemented to estimate grid RA at 0.5° × 0.5° resolution over 1980-2012. The performance of RF was assessed by a 10-fold cross-validation (CV). A 10-fold CV suggested that the whole dataset was subdivided into 10 parts with approximately an equal number of samples. The target values for each of these 10 parts were predicted on the training using the remaining nine parts. Two statistics were employed in model assessment: modelling efficiency ($R^2$) and root mean square error ($RMSE$) (Yao et al., 2018). The 10-fold CV result showed that RF performed well and could capture the spatial- and temporal-pattern of RA (Fig. S1 in supplementary materials).

2.4 Temporal trend analysis

This study applied Theil-Sen linear regression to estimate temporal trend analysis of RA and its driving variables for each grid cell. The Theil-Sen estimator is a median-based non-parametric slope estimator, which has been widely used for spatial analysis of time series carbon flux analysis (Forkel et al., 2016; Zhang et al., 2017). Mann-kendall non-parametric test was applied for the significant change trend in RA and its driving factors for each grid cell ($p < 0.05$).

2.5 Relationships between RA and climate variables

In this study, mean annual temperature, mean annual precipitation and mean annual shortwave radiation were considered as the most important proxies driving RA. The relationships between RA and temperature, precipitation and shortwave radiation were analyzed by partial correlation for each grid cell. The absolute value of the correlation coefficient of these three variables was used in RGB combination to indicate the dominant factors of RA.

2.6 The comparison map profile method

In order to compare with the solely global RA product generated by Hashimoto et al. (2015), which was
estimated by a climate-driven model using temperature and precipitation only and obtained from the public available dataset (http://cse.ffpri.affrc.go.jp/shojih/data/index.html), this study applied the comparison map profile (CMP) method. CMP was developed based on absolute distance (D) and cross-correlation coefficient (CC) through multiple scales (Gaucherel et al., 2008). D and CC reflect the similarity of data values and spatial structure of two images with the same size, respectively (Gaucherel et al., 2008). Low D and higher CC reflects goodness between the compared images, and vice versa. The D among moving windows of two compared images was calculated by equation (1) (Gaucherel et al., 2008):

\[ D = \text{abs}(\bar{x} - \bar{y}) \]  

(1)

\( \bar{x} \) and \( \bar{y} \) are averages calculated over two moving windows. This study used 3×3 to 41×41 pixels. Finally, the mean D was averaged for different scales.

The CC was calculated by equation (2) (Gaucherel et al., 2008):

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

(2)

\[ \sigma_x^2 = \frac{1}{N^2 - 1} \sum_{i=1}^{N} \sum_{j=1}^{N} (x_{ij} - \bar{x}) \]  

(3)

Where \( x_{ij} \) and \( y_{ij} \) are the pixel values at row \( i \) and column \( j \) of two moving windows of the two compared images, respectively. \( N \) represents the number of pixels for each moving window, while \( \sigma_x \) and \( \sigma_y \) are the standard deviation calculated from the two moving windows. Finally, like D calculations, CC was calculated as the mean of different scales.

3 Results

3.1 Spatial patterns of RA
Figure 2 Spatial patterns of annual mean (a, b) and standard deviation (c, d) of belowground autotrophic respiration (RA) from 1980 to 2012 for this study (a, c) and Hashimoto RA (b, d) during 1980-2012, respectively. The standard deviation was applied to characterize the inter-annual variability following Yao et al. (2018).

The data-derived RA in this study presented a great globally spatial variability during 1980-2012 (Fig. 2a and Fig. 3). Largest RA fluxes commenced from tropical regions, particularly in Amazon tropical and Southeast areas, where generally have a high RA > 700 g C m$^{-2}$ yr$^{-1}$. Following the tropical areas, subtropics, e.g. South China, East America, and humid temperate areas, e.g. North America, West and Middle Europe, had typical moderate RA fluxes of 400-600 g C m$^{-2}$ yr$^{-1}$. By contrast, the relative low RA fluxes occurred in the areas with sparse vegetation cover, cold and dry climate, e.g. boreal and tundra, which had low temperature and short growing season. Besides, dry or semi-arid areas, e.g. Northwest China and Middle East, also had typical low RA fluxes below 200 g C m$^{-2}$ yr$^{-1}$, where were often limited by water availability.
The most significant RA inter-annual variability (expressed by standard deviation, Fig. 2c) was found in topical or subtropical regions with above 80 g C m\(^{-2}\) yr\(^{-1}\), while most areas remained less variable with less than 40 g C m\(^{-2}\) yr\(^{-1}\). Latitudinally, zonal mean RA increased from cold and dry biomes (Tundra and semi-arid) to warm and humid biomes (temperate and tropical forests, Fig. 3), reflecting from more to less environmental limitations. RA varied from 112\(\pm\)21 g C m\(^{-2}\) yr\(^{-1}\) at about 70°N to 552\(\pm\)101 g C m\(^{-2}\) yr\(^{-1}\) at equator. Within in 10°S -25°S and 15°N -20°N, due to the limitation of water, zonal mean RA experienced a slight decrease. Therefore, with the increase of water availability, RA led to a second peak in around 20°N and 40°S, respectively.

Compared to data-derived RA, Hashimoto RA presented a similar spatial pattern, with highest RA fluxes in tropical regions characterized by warm and humid climate, followed by subtropical regions, and lowest RA in boreal areas featured by in cold and dry climate (Fig. 2b). The most significant change occurred in tropical areas and middle Australia. However, it is worth noted that some clear differences between
data-derived and Hashimoto RA existed (Fig. 4): specifically, there was a remarkable difference of above 300 g C m\(^{-2}\) a\(^{-1}\) for South Amazon and larger than 200 g C m\(^{-2}\) a\(^{-1}\) for subtropical China. Although most areas between data-derived RA and Hashimoto RA expressed high and positive correlations, some areas, such as Middle East, West Russia and East America and North Japan, showed negative correlations.

**Figure 4** Comparison of data-derived RA with Hashimoto RA based on absolute distance (a) and cross-correlation (b)

### 3.2 Spatial pattern of RA trend

**Figure 5** Spatial patterns of the temporal trend for belowground autotrophic respiration (RA) for this study and Hashimoto RA during 1980-2012

The trend of data-derived RA showed heterogeneous patterns in spatial (Fig. 5). A total of 58% of global areas experienced an increasing trend during 1980-2012 (calculating from cell areas), and 33% of these
areas showed a significant change \( (p < 0.05) \). Generally, the change trend for the majority areas was within \(-4 - 4 \) g C m\(^{-2}\) \text{yr}^{-1},\) while the most striking increasing change occurred in East Russia, and tropical, and Eastern regions in Africa with an increasing trend of above 5 g C m\(^{-2}\) \text{yr}^{-1}. Similarly, 77% of global areas of Hashimoto RA had an increasing trend, 46% of which were statistically significant \( (p < 0.05) \).

### 3.4 | Total RA and its temporal trend

#### Figure 6

Inter-annual variability (a and b) and total amount (c) of belowground autotrophic respiration (RA) for this study (a) and Hashimoto RA (b) during from 1980 to 2012. The grey area represents 95% confidence interval. The error bars mean standard deviation.

Mean global RA was 43.8±0.4 Pg C \text{yr}^{-1} during 1980-2012 (Figure 6c), varying from 42.9 Pg C \text{yr}^{-1} in 1992 to 44.9 Pg C \text{yr}^{-1} in 2010, with a significant trend of 0.025±0.006 Pg C \text{yr}^{-2}\text{per year} despite of high
annual variabilities (0.06% \text{a}^{-1}, p < 0.001, \text{Fig. } 6a). Similarly, a rising trend was also observed for Hashimoto RA, however, its annual increasing trend (0.073±0.009 Pg C \text{a}^{-1}, p < 0.001, \text{Fig. } 6b) was higher than that of data-derived RA. Annual mean of Hashimoto RA was 40.5±0.9 Pg C \text{a}^{-1} (\text{Fig. } 6c).

\textbf{Figure 7} Total amount of belowground autotrophic respiration (RA) for this study and Hashimoto RA for three climate zones and eight biomes during 1980-2012. Three climate zones defined as boreal, temperature and tropical regions according to Peel et al. (2007), while eight biomes include boreal forest, cropland, grassland, savannas, shrubland, temperate forest, tropical forest and wetland. The error bars indicated standard deviation.
RA and its trend was also evaluated for three climate zones (boreal, temporal and tropical areas based on Köppen-Geiger climate classification) and eight major biomes (boreal forest, cropland, grassland, savannas, shrubland, temperate forest, tropical forest and wetland, Fig. 7). Tropics had highest RA of 15.6±0.2 Pg C \( \text{a} \text{yr}^{-1} \), followed by temperate regions with 9.3±0.1 Pg C \( \text{a} \text{yr}^{-1} \), and boreal areas represented the lowest RA of 6.7±0.1 Pg C \( \text{a} \text{yr}^{-1} \). These three climate zones were main contributors of global RA, accounting for 72%. Temporally, considerable RA inter-annual variability of these three climate zones existed (Fig. S2). Specifically, RA in tropical and boreal zones showed a significantly increasing trend from 1980 to 2012, with an increasing rate of 0.013±0.003 and 0.008±0.002 Pg C \( \text{a} \text{yr}^{-2} \), respectively. However, RA in temperate zones presented a slightly decreasing trend of -0.003±0.001 Pg C \( \text{a} \text{yr}^{-2} \) (\( p = 0.048 \)) although strong variability was observed.

In terms of biomes, tropical forest had the highest RA, followed by the widely distributed cropland and savannas (Fig. 7), while wetland had the lowest RA due to its limited land cover. All the biomes, except temperate forest, savannas and wetland, RA showed a significantly increasing trend during 1980-2012 (\( p < 0.015 \)) in majority biomes, except temperate forest, savannas and wetland. RA in tropical forests, boreal forests and cropland had the highest increasing trend increased by of –0.0076±0.0015, 0.0047±0.0016, 0.0036±0.0014 Pg C \( \text{a} \text{yr}^{-2} \), respectively. Compared to data-derived RA, Hashimoto RA for the three climate zones and eight biomes generally produced similar change patterns, although the magnitude difference existed (Figs. 7, S2 and S3). However, there were significant increasing trends of RA in temperate zones, temperate forest, savannas and wetland for of derived from Hashimoto RA, which were not observed in data-derived RA.
RA was significantly correlated with temperature anomaly ($R^2 = 0.59$, $p < 0.001$) and precipitation anomaly ($R^2 = 0.50$, $p < 0.001$, Fig. 8). On average, RA increased by $0.85\pm0.13$ Pg C yr$^{-1}$ for 1°C increment in mean annual temperature, and $0.17\pm0.03$ Pg C yr$^{-1}$ for 10 mm increase in mean annual precipitation. However, different biomes and climate zones showed uneven responses to the temperature and precipitation change (Fig. S4 and S5). For example, no significant correlations were found between RA in temperate zone/savannas/wetland and temperature anomaly, while other climate zones and biomes were significantly correlated with temperature/precipitation anomaly.

4. Dominant factors for RA variability

Figure 9 Dominant factors for belowground autotrophic respiration (RA) for this study and Hashimoto
RA, MAT = mean annual temperature, MAP = mean annual precipitation; SWR = shortwave radiation.

The dominant environmental factor was examined with partial regression coefficients when regressing RA against annual mean temperature, annual mean precipitation and shortwave radiation. Latitudinally, higher mean annual temperature, precipitation and shortwave radiation were associated with higher RA in the major latitudinal gradients (positive partial correlations, Fig. S6). Spatially, the dominant environmental factor varied greatly globally (Fig. 9). Precipitation was the most important dominant factor for the spatial pattern of RA among the three environmental controls, covering about 56% of global land areas (Fig. S7), which was widely distributed globally, particularly in dry or semi-arid areas, such as Northwest China, Southern Africa, Middle Australia and America. Temperature dominated about 19% of global land areas, which mainly occurred in tropical Africa, Southern Amazon rainforests, Siberia and partly tundra. The rest land area (25%) was dominated by shortwave radiation, primarily covering boreal areas above 50°N, Eastern America and middle and Eastern Russian. Similarly, precipitation was also the most important dominant factor for Hashimoto RA, dominating about 77% land area, while temperature and shortwave radiation dominated 13% and 10% land area. However, their spatial patterns varied greatly compared to data-derived RA. For example, temperature was the main dominant factor for most area of Australia for Hashimoto RA, while data-derived RA indicated that precipitation and shortwave radiation dominated such areas (Fig. 9).

4 Discussion

4.1 Global RA

Despite of great efforts to quantify global soil carbon fluxes and their spatial and temporal patterns (Bond-Lamberty and Thomson, 2010b; Hursh et al., 2017; Jian et al., 2018b), to our knowledge, no attempt tried to assess RA using machine learning approach by linking a large number of empirical measurements, and RA’s spatial and temporal patterns remain large uncertainties. Such uncertainties justify a development of global RA product derived from observations to understand its spatial and temporal patterns, causes and responses to future climate change. Based on the most updated observations from SRDB released by the end of 2018, this study for the first time applied RF algorithm to estimate the temporal and spatial variability of global RA and its response to environmental variables, which indeed can contribute to reduce RA uncertainties.
Globally, mean annual RA amounted to 43.8±0.4 Pg C \(\text{year}^{-1}\) from 1980 to 2012 (Fig. 6). It was slightly higher than Hashimoto RA (40.5±0.9 Pg C \(\text{year}^{-1}\)), and there was great divergence of spatial and temporal patterns (see discussion part in “Comparison with Hashimoto RA”). Due to no direct estimate on global RA, this study compared other RA estimates using total soil respiration multiplied by the proportion of RA or heterotrophic respiration. Bond-Lamberty et al. (2018) proposed that the global average proportion of heterotrophic respiration ranged from 0.54 to 0.63 over 1990-2014 and global total soil respiration was 83-67 to 108 Pg C \(\text{year}^{-1}\) from using different approaches and datasets according to recent predictions by Bond-Lamberty (2018); (Bond-Lamberty and Thomson, 2010b; Hashimoto et al., 2015; Hursh et al., 2017; Jian et al., 2018b) (Bond-Lamberty and Thomson, 2010b; Hursh et al., 2017), thus global RA varied from 34-25 to 51 Pg C \(\text{year}^{-1}\). RA estimate in this study fell in this range. Similarly, during 1980-2012, RA increased by 0.025±0.006 Pg C \(\text{year}^{-2}\), which may be related to the increasing photosynthesis due to global warming and CO₂ fertilization effects, which could increase carbon availability in plant-derived substrate inputs into the soil (e.g. root exudates and biomass) for both root metabolism (Piñeiro et al., 2017; Zhou et al., 2016). Such annual increase accounted for about 25% of global soil respiration increase (0.09 and 0.1 Pg C \(\text{year}^{-2}\)) (Bond-Lamberty and Thomson, 2010b; Hashimoto et al., 2015), suggesting that about one quarter of the total soil respiration increment due to climate change came from RA. With 1 °C increase in global mean temperature, RA will increase by 0.85±0.13 Pg C \(\text{year}^{-2}\) and 0.17±0.03 Pg C \(\text{year}^{-2}\) for 10 mm increase in precipitation, which indicated that carbon fluxes from RA might positively feedback to future global climate change, which was typically characterized by increasing temperature and changes in precipitation (IPCC, 2013). However, RA increment varied with climate zones and ecosystem types (Figs. S2 and S3), which was similar to previous findings (Ballantyne et al., 2017; Jian et al., 2018a), who found that total soil respiration or RA varied with climate zones or ecosystem types. These differences may be related to regional heterogeneity and plant functional trait. For example, regional temperature significantly differed from global averages (Huang et al., 2012), with much faster change in high-latitude regions (Hartmann et al., 2014), and semi-arid dominated the trend and variability of global land CO₂ sink (Ahlström et al., 2015). Therefore, the regionally uneven responses of RA to climatic variables highlights the urgent need to account for regional heterogeneity when studying the effects of climate change on ecosystem carbon dynamics in future.

RA estimate in this study also has important indications of carbon allocation from photosynthesis. The
Immediate carbon substrates for RA were primarily derived from recent photosynthesis (Högberg et al., 2001; Subke et al., 2011). Strong correlation between photosynthesis and RA demonstrated the evidence for their close coupling relationships (Chen et al., 2014; Kuzyakov and Gavrichkova, 2010). Globally, GPP was about 125 Pg C yr⁻¹ during last few decades (Bodesheim et al., 2018; Zhang et al., 2017). Thus, roots respired more than one third carbon from GPP, suggesting that except the carbon used for constructing belowground tissues, a large proportion of carbon will be returned back to atmosphere respired by roots. However, it should be noted that through root respiration, soil nutrients for vegetation growth will be required, which may affect the RA flux.

4.2 Dominant factors

Spatially, the driving factors for RA varied greatly. Temperature and shortwave radiation were the main driving factors for high latitudinal areas above 50°N (Figure 9a). This result was not surprising because RA was positively correlated with temperature or photosynthesis (indirectly reflecting the solar radiation) (Chen et al., 2014; Tang et al., 2016), and high latitudinal regions was always limited by temperature or energy, leading to low RA as well (Fig. 3a).

Globally, precipitation was the most important factor, covering about 56% of land area (Fig. 9a and S7). Precipitation was always considered as a proxy for soil water content (Hursh et al., 2017; Yao et al., 2018), and such wide dominance of precipitation on RA was related to the mechanisms of the soil water availability driving RA. First, soil water exists in form of ice when temperature is below zero, that plant and soil microbes could not directly use for growth or respiration. This could be observed in some boreal areas where precipitation was the dominant factor of RA (Fig. 9a). Second, too high or too low soil water content (e.g. flooding and drought) could limit the mobility of substrates and carbon input to belowground, which could affect RA. Yan et al. (2014) found that soil respiration decreased once soil water content was below a lower (14.8 %) or above an upper (26.2%) threshold in a poplar plantation. Similarly, Gomez-Casanovas et al. (2012) also found that RA decreased when soil water content was above 30%. These results seemed to support the finding in this study. Third, the relationship between soil water content and RA or total soil respiration is more complex than the relationship between temperature and soil respiration. Numerous formula, such as linear (Tang et al., 2016), polynomial (Moyano et al., 2012), logarithmic (Schaefer et al., 2009), quadratic (Hursh et al., 2017) models have been widely applied to describe the relationship between soil water content and soil respiration. The
multifarious relationships between soil water content and RA may occur because soil water content affect RA in multiple ways. Meanwhile, seasonal variability of precipitation and soil water content is often correlated with temperature (Feng and Liu, 2015), making the relationship between soil water content and RA more complex.

Similarly, the dominance of precipitation in Hashimoto was also widely observed (Fig. 8), dominating 77% of land area (Fig. S7). Although this percentage was 17% higher than data-derived RA, both results demonstrated that global RA of the majority land cover was dominated by precipitation. However, it is noticeable that dominant environmental factor controlling spatial carbon fluxes gradient may differ among different years (Reichstein et al., 2007), e.g. climate extreme and disturbance.

4.3 Comparison with Hashimoto RA

Globally, total data-derived RA was slightly higher than Hashimoto RA, however, great divergence was observed both spatially and temporally (Fig. 6), particularly in tropical regions, where data-derived RA was much lower than Hashimoto RA (Fig. 3). These differences could be attributed to several reasons. First, two RA products had different land cover areas, especially in desert areas in North Africa, where existed very sparse or no vegetation. If data-derived RA was masked by Hashimoto RA, global RA was 39.6±0.4 Pg C yr⁻¹, which was pretty close to Hashimoto RA (Fig. S8). Second, different predictors and algorithms were applied for data-derived RA and Hashimoto RA prediction. Besides temperature and precipitation, RA was also affected by soil nutrient, carbon substrate supply, belowground carbon allocation, site disturbance and other variables (Chen et al., 2014; Hashimoto et al., 2015; Tang et al., 2016; Zhou et al., 2016). Hashimoto RA was calculated from the difference between total soil respiration and heterotrophic respiration, which were predicted by a simple climate-driven model using temperature and precipitation only (Hashimoto et al., 2015). Thus Hashimoto RA could not reflect its soil nutrient and other environmental constrains. To overcome such limitations, beside temperature and precipitation, this study included soil water content, soil nitrogen and soil organic carbon as proxies for environmental and nutrient constraints of RA and considered the interactions among these variables using RF. This study indeed improved the model efficiency to 0.52 for RA prediction (Fig. S1), which was 0.32 for Hashimoto soil respiration (Hashimoto et al., 2015). The simple climate-model for Hashimoto soil respiration could be its advantages and limitations (Hashimoto et al., 2015). Third, the empirical model (the relationship between total soil respiration and heterotrophic respiration) deriving Hashimoto RA originated from...
forest ecosystems (Bond-Lamberty et al., 2004; Hashimoto et al., 2015), which may bring uncertainties to other ecosystems. For example, the difference between data-derived RA and Hashimoto RA varied up to 350 g C yr$^{-1}$ in South, North Amazon areas and Madagascar, where the savannas widely distributed (Fig. 4), thus Hashimoto RA might not capture the spatial and temporal pattern of RA for non-forest ecosystems. Including more environmental variables and improving algorithm could be a good option to reduce the uncertainty in modelling RA.

### 4.4 Advantages, limitations and uncertainties

Generally, this study had four main advantages to estimate global RA: first, this study, to our knowledge, was the first attempt to model RA using a large number of empirical field observations, and estimate spatial and temporal patterns globally. While most previous studies mainly focused on global total soil respiration, which was not partitioned into RA and heterotrophic respiration globally (Hursh et al., 2017; Jian et al., 2018b; Zhao et al., 2017). Second, this study used an up-to-date field observational database developed from SRDB up to the end 2018. This new updated database included a total of 449 field observations (Fig. 1). These observations had a wide coverage range of global terrestrial ecosystems and represented all major biomes and climate zones. Third, the global terrestrial ecosystems were separated into eight biomes, including boreal forest, cropland, grassland, savannas, shrubland, temperate forest, tropical forest and wetland. The total RA and its inter-annual variability were evaluated for each of the eight biomes (Fig. S3 and S4). Besides, total RA and its inter-annual variability was also assessed for three climate zones - boreal, temperate and tropical zones (Figs. S2 and S5), according to the Köppen-Geiger climate classification system (Peel et al., 2007). These were important climate zones, contributing 72% of global RA. Different temporal change trends across biomes and climate zones also further indicated an uneven response of RA to climate change. Fourth, this study used a RF algorithm to model and map global RA with the linkage of climate and soil predictors. The results showed that RF could accurately estimate the relationships between annual RA and predictors (Fig. S1). Compared to linear regressions for soil respiration prediction (because no global RA prediction before this study) with a model efficiency less than 35% (Bond-Lamberty and Thomson, 2010b; Hashimoto et al., 2015; Hursh et al., 2017), RF algorithm achieved a much higher model efficiency to 52%, which indeed improved the RA modelling and reduced the uncertainties.

Although data-derived global RA could serve as a benchmark for global carbon cycling modelling, and
this study had filled the data-gaps of global RA, limitations and uncertainties still remained in few aspects. First, although we conducted a data quality control in this study, a lack of reliable approach to separate RA and heterotrophic respiration may lead to an uncertainty of RA values. There are several approaches, e.g. trenching, stable or radioactive isotope, gridding (Bond-Lamberty et al., 2004; Högberg et al., 2001; Hanson et al., 2000), however, each of these approaches has its own limitations. For example, trenching has been widely applied to partition RA and heterotrophic respiration due to easy operation and low cost, on the other hand, heterotrophic respiration may be increased due to the termination of water uptake by roots and the decomposition of remaining dead roots in trenching plots (Hanson et al., 2000; Tang et al., 2016). Commonly, RA was calculated from the difference between total soil respiration and heterotrophic respiration, thus the trenching approach might lead to an underestimation of RA. In our dataset, a total of 254 RA observations were estimated by trenching approach, while the rest RA observations were estimated by other separation approaches, e.g. isotope, radiocarbon, mass balance. Thus, inconsistent separation approaches could be another source of uncertainty of RA values.

Second, due to the limited observations of RA at a daily or monthly scale, this study only predicted RA at an annual scale. Although there was no direct study to compare the difference of RA upscaling from daily or monthly and annual scale, substantial difference of soil respiration upscaling from daily or monthly and annual scales (Jian et al., 2018b) indirectly illustrated the potential difference of RA upscaling from different timescales.

Third, the effects of rising atmospheric CO₂ on root growth was not explicitly represented in this study, although CO₂ fertilization effects could partly be represented in the increase temperature. While the magnitude of CO₂ fertilization effects on photosynthesis is still uncertain (Gray et al., 2016), RF or other machine learning approaches are encouraged to quantify the uncertainties due to CO₂ fertilization.

Second, Fourth, this study did not consider the effects of human activities and historical changes in biomes on RA. However, important changes may occur in tropical forest, grassland and cropland during last several decades due to human activities (Hansen et al., 2013; Klein Goldewijk et al., 2011). Thus, changes in biomes should be included in future global RA and carbon cycling modelling. However, the lack of such data is the main constrain of detecting the effects of biome change on RA.

Finally, uneven coverage of observations in the updated database would be another source of
uncertainties. Although our dataset had a wide range of land cover, the observational sites mainly
distributed in China, Europe and North America and were dominated by forests. There was a great lack
of observations in areas, such as Africa, Austria and Russia, and biomes, such as tropical forest, shrubland,
wetland and cropland. Consequently, RA observations caused bias of RF model toward the regions with
more observations. Therefore, including more observations in these areas and biomes should largely
increase our capability to assess the spatial and temporal patterns of global RA and contribute to improve
the global carbon cycling modelling to future climate change.

5 Data availability

The datasets are freely downloadable from https://doi.org/10.6084/m9.figshare.7636193 (Tang et al.,
2019).

6 Conclusions

Although data-derived global RA may serve as a benchmark for ecosystem models, no such study has
assessed global variability in RA with a large number of empirical observations that can help bridge the
knowledge gap between local, regional, and global scales. This study has filled this knowledge gap by
linkage of field observations and global variables using RF algorithm, providing an annual global RA
product at a of 0.5° × 0.5° resolution from 1980 to 2012. Currently, robust findings include: (1) Annual
mean RA was 43.8±0.4 Pg C yr⁻¹ with a temporally increasing trend of 0.025±0.006 Pg C yr⁻² over
1980-2012, indicating an increasing carbon return from the roots to the atmosphere; (2) uneven temporal
and spatial variabilities in varying climate zones and biomes indicated their uneven temperature
sensitivities to future climate change, challenging the perspective that the parameters of global carbon
stimulation independent on climate zones and biomes; (3) precipitation dominated the spatial variabilities
of RA. However, further improvements in the approach should overcome shortcomings from reduced
data availability and the mismatch in spatial resolution between covariates and in situ RA.

Author contributions. XT, SF, WZ and SG designed the research and collected the data, XT, WZ and
SG contributed to the data processing and analysis. XT, WZ, SG, GC and LS wrote the manuscript, and
all authors contributed to the review the manuscript.

Competing interests. The authors declare that they have no conflict of interest.
Acknowledgements. This study was supported by Fundamental Research Funds of International Centre for Bamboo and Rattan (1632017003 and 1632018003); the National Natural Science Foundation of China (31800365); Innovation funding of Remote Sensing Science and Technology of Chengdu University of Technology (KYTD201501); Starting Funding of Chengdu University of Technology (10912-2018KYQD-06910), Foundation for University Key Teacher of Chengdu University of Technology (10912-2019JX-06910) and Open Funding from Key Laboratory of Geoscience Spatial Information Technology of Ministry of Land and Resources (Chengdu University of Technology). The authors express their great thanks to the contributors of Hashimoto soil respiration dataset, the contributors of collecting data from publication in SRDB. Great thanks to Dr. Shoji Hashimoto for his valuable comments to improve the manuscript.

References


Huang, J., Guan, X., and Ji, F.: Enhanced cold-season warming in semi-arid regions, Atmospheric Chemistry and Physics, 12, 5391-5398, http://dx.doi.org/10.5194/acp-12-5391-2012, 2012.


IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth


Yan, M. F., Zhou, G. S., and Zhang, X. S.: Effects of irrigation on the soil CO₂ efflux from different


