ECOLOGY

Spatial and temporal variations in global soil respiration and their relationships with climate and land cover

Ni Huang[1], Li Wang[1]*, Xiao-Peng Song[2], T. Andrew Black[3], Rachhpal S. Jassal[3], Ranga B. Myneni[4], Chaoyang Wu[5], Lei Wang[1], Wanjuan Song[1], Dabin Ji[1], Shanshan Yu[1], Zheng Niu[1,6]

Soil respiration (R_s) represents the largest flux of CO_2 from terrestrial ecosystems to the atmosphere, but its spatial and temporal changes as well as the driving forces are not well understood. We derived a product of annual global R_s from 2000 to 2014 at 1 km by 1 km spatial resolution using remote sensing data and biome-specific statistical models. Different from the existing view that climate change dominated changes in R_s, we showed that land-cover change played a more important role in regulating R_s changes in temperate and boreal regions during 2000–2014. Significant changes in R_s occurred more frequently in areas with significant changes in short vegetation cover (i.e., all vegetation shorter than 5 m in height) than in areas with significant climate change. These results contribute to our understanding of global R_s patterns and highlight the importance of land-cover change in driving global and regional R_s changes.

INTRODUCTION

Large uncertainties in the global carbon budget are associated with the terrestrial carbon cycle, and reducing these uncertainties requires an improved capacity to estimate carbon fluxes between the atmosphere and terrestrial ecosystems (*1*). Over two-thirds of terrestrial carbon is stored belowground, and a significant amount of the atmospheric CO_2 assimilated by plants is respired by roots and microbes in soils (*2*, *3*). Soil respiration (R_s), consisting of root (autotrophic) respiration (R_a) and microbial (heterotrophic) respiration (R_h), is therefore a key process in the terrestrial carbon cycle. However, R_s is difficult to estimate at the global scale because of the limited understanding of the complex interactions of physical, chemical, and biological processes and the resulting high spatiotemporal dynamics (*4*).

Compared to the large number of studies on plant productivity in terrestrial carbon cycling, there are few studies on global R_s dynamics (*4*, *5*). Global R_s can be estimated using top-down or bottom-up approaches. Top-down approaches typically estimate R_s as the residual of the carbon balance, and thus, the indirect estimates can naturally capture average soil carbon fluxes over large areas, although the results inevitably suffer from errors and uncertainties from any of the source datasets (*6*). By contrast, bottom-up methods, including process-based models and empirical statistical models, use direct observational data (*6*, *7*). Process-based models rely on knowledge of the interactions among different physical and ecological processes, but many models rarely simulate R_s explicitly and instead generate only an explicit output of R_h (*8*). Global R_h may provide some information on R_s, but the use of different model structures and parameters has resulted in large discrepancies among them (*3*, *9*). Compared to process-based models, statistical models are simpler and can provide data-oriented estimates of R_s using various biotic or abiotic factors as input, such as temperature, moisture, plant productivity, and soil properties (*2*, *3*, *9*–*17*). Previous studies using statistical models have produced mean annual global R_s estimates from 78 to 108 Pg C $year^{-1}$, and the error associated with each estimation was high (table S1). One study (*7*) noted that the lack of large-scale, observation-driven R_s estimates was a major problem in constraining regional-scale to global-scale C fluxes.

Given that multiple regional-scale studies (*18*, *19*) have shown that remote sensing data can be used to estimate R_s with high accuracy, we expect that it is also possible to estimate R_s beyond the regional scale. The openly available remote sensing data of high spatial resolution provide large-scale observations, and remote sensing is anticipated to play an increasingly important role in carbon cycle research in the future (*20*). For instance, current remote sensing products provide a range of key ecosystem variables that accurately reflect the spatiotemporal variations in surface temperature, moisture, and plant productivity of terrestrial ecosystems (*21*–*23*). Adachi *et al.* (*14*) provided insights into the estimation of global R_s by combining empirical models derived from field studies with remote sensing data. However, Adachi *et al.* (*14*) used one empirical model from a site-specific study for all ecosystem types, which is unlikely to be representative and may lead to large uncertainty in global R_s estimation.

R_s change can be affected by many factors, and climate factors (e.g., air temperature and precipitation) have been commonly investigated because of their direct or indirect effects on R_s metabolism (*9*–*15*). Land-cover change can also greatly affect R_s by changing vegetation structure, plant species composition, local microclimate, and soil properties (*24*, *25*). For instance, global greening and the associated vegetation structural change could affect R_s by altering biogeophysical processes (*26*, *27*). However, few studies have comprehensively considered both climate and land-cover change effects on the spatial and temporal variations in global R_s. As soil nutrient availability ultimately depends on photosynthate supply (*28*), we assume that the estimation of global R_s can be achieved by including three factors: temperature, moisture, and plant productivity.

With the number of measurements of R_s rapidly increasing worldwide, future modeling efforts that are focused on global R_s estimation should take advantage of remote sensing data and data mining techniques to link observations at different scales. In this study, we have

[1]State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing Normal University, Beijing, China. [2]Department of Geosciences, Texas Tech University, Lubbock, TX, USA. [3]Faculty of Land and Food Systems, University of British Columbia, Vancouver, BC, Canada. [4]Department of Earth and Environment, Boston University, Boston, MA, USA. [5]The Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China. [6]University of Chinese Academy of Sciences, Beijing, China.
*Corresponding author. Email: wangli@radi.ac.cn

developed an annual global R_s dataset at 1 km by 1 km spatial resolution for the time period from 2000 to 2014. The product was generated by combining globally distributed in situ R_s measurements, satellite remote sensing data, and biome-specific statistical models, including parametric models and nonparametric machine learning models. We analyzed spatial and temporal variations in global R_s. We then investigated how the changing trends in R_s were related to changes in climate (temperature, precipitation, and drought) and land cover.

RESULTS

Deriving annual global R_s product from 2000 to 2014

On the basis of a 10-fold cross-validation method, we evaluated the accuracy of the four statistical models [i.e., multiple nonlinear regression (MNLR), random forest regression (RFR), support vector regression (SVR), and artificial neural network (ANN) models] by comparing RMSE (root mean square error) and R^2 (coefficient of determination) (table S2) using available measured R_s datasets as a reference. We determined the biome-scale model for the estimation of annual R_s of each biome by selecting the model with the highest cross-validation accuracy (table S3). Although the selected models were not the same across the 10 biomes, they were all machine learning algorithm models. The selected models explained 62 to 84% of the interannual and intersite variabilities in annual R_s with an RMSE ranging from 107 to 413 g C m^{-2} $year^{-1}$ (table S3). On the basis of the selected biome-scale models (tables S3 and S4), we produced the annual global R_s dataset at the global scale from 2000 to 2014.

Spatial and temporal patterns of global R_s

The regions with high R_s are located in the tropics, such as the Amazon Basin, Central Africa, and Southeast Asia. Low R_s values are widely distributed in the Northern Hemisphere high-latitude regions, western United States, Canada, Central Asia, northern Mongolia, northeast China, Argentina, and Australia (Fig. 1). Boreal, temperate, and tropical regions contributed 15, 24, and 61%, respectively, to the total mean annual global R_s. Our estimated mean annual global R_s from 2000 to 2014 is 72.6 Pg C $year^{-1}$ [95% confidence interval (CI) = 69.8-75.4 Pg C $year^{-1}$], which was lower than the estimated global R_s for recent years (table S1). The estimated annual global R_s showed fluctuations over time (fig. S2), with the lowest value (70.6 Pg C $year^{-1}$) occurring in 2000 and the highest value (74.5 Pg C $year^{-1}$) occurring in 2010.

Over the 15 years of the study period, the R_s trends varied at different spatial scales (Fig. 2A). R_s showed a significantly increasing trend globally (0.13 ± 0.02 Pg C $year^{-1}$, $P < 0.05$), in the boreal region (0.05 ± 0.01 Pg C $year^{-1}$, $P < 0.05$), and in the tropical region (0.11 ± 0.02 Pg C $year^{-1}$, $P < 0.05$) and a nonsignificantly decreasing trend in the temperate region (−0.02 ± 0.01 Pg C $year^{-1}$, $P = 0.42$). The tropics accounted for approximately 85% of the total increase in the global R_s. The largest R_s net change occurred in the tropical region (1.54 Pg C, 95% CI = 1.30-1.78 Pg C; +3.66% change relative to the global R_s in 2000), followed by those in the boreal region (0.65 Pg C, 95% CI = 0.49-0.81 Pg C; +6.80%) and temperate region (−0.27 Pg C, 95% CI = −0.43-−0.11 Pg C; −1.53%).

At the global scale, 50% of our study area experienced an increasing trend in R_s, and the regions with a significant increasing trend (8.5%) were spatially aggregated and mainly located at the high latitudes (e.g., northern Canada, northern Russia, and northern Europe), northern Mongolia, Loess Plateau and northeastern Tibetan Plateau of China, India, Amazon Basin, and Congo Basin (Fig. 2, B and C). The annual R_s significantly decreased in less than 5% of the study regions; these regions were mainly scattered in southeastern Russia, midlatitude temperate regions (i.e., United States, Kazakhstan, and northeast China), southern Africa, and Argentina (Fig. 2B). At the regional scale, the area with the highest proportion of increasing R_s (60.7%, 11.1% with a significant increase) was the boreal region, and the area with the highest proportion of decreasing R_s (39.9%, 5.1% with a significant decrease) was the temperate region (Fig. 2C).

The relationships between R_s and its driving factors at global and regional scales

Global R_s showed a significant positive partial correlation with annual mean air temperature (TEM) from 2000 to 2014 ($R = 0.67$, $P < 0.05$; Fig. 3). A significant positive partial correlation between R_s

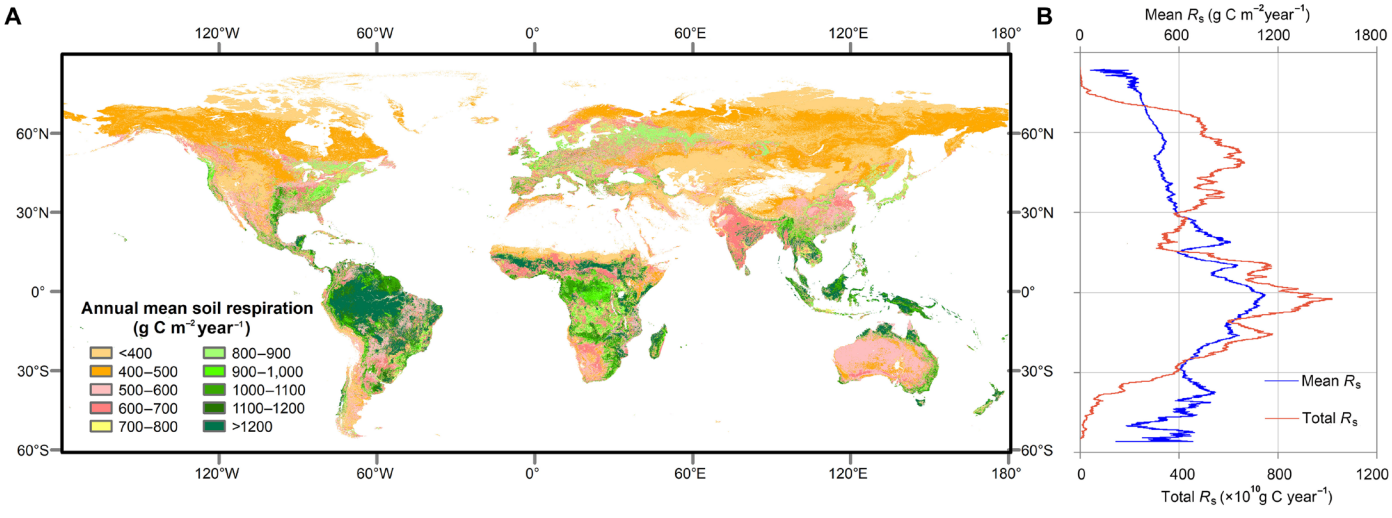

Fig. 1. Global distribution of mean annual R_s between 2000 and 2014. (**A**) Global map of mean annual R_s at 1 km by 1 km spatial resolution derived using satellite data. (**B**) Latitudinal distribution of mean annual R_s (blue line) and total annual R_s (orange line). All land grids along a latitudinal row in the global map were averaged to derive mean R_s and summed to derive total R_s.

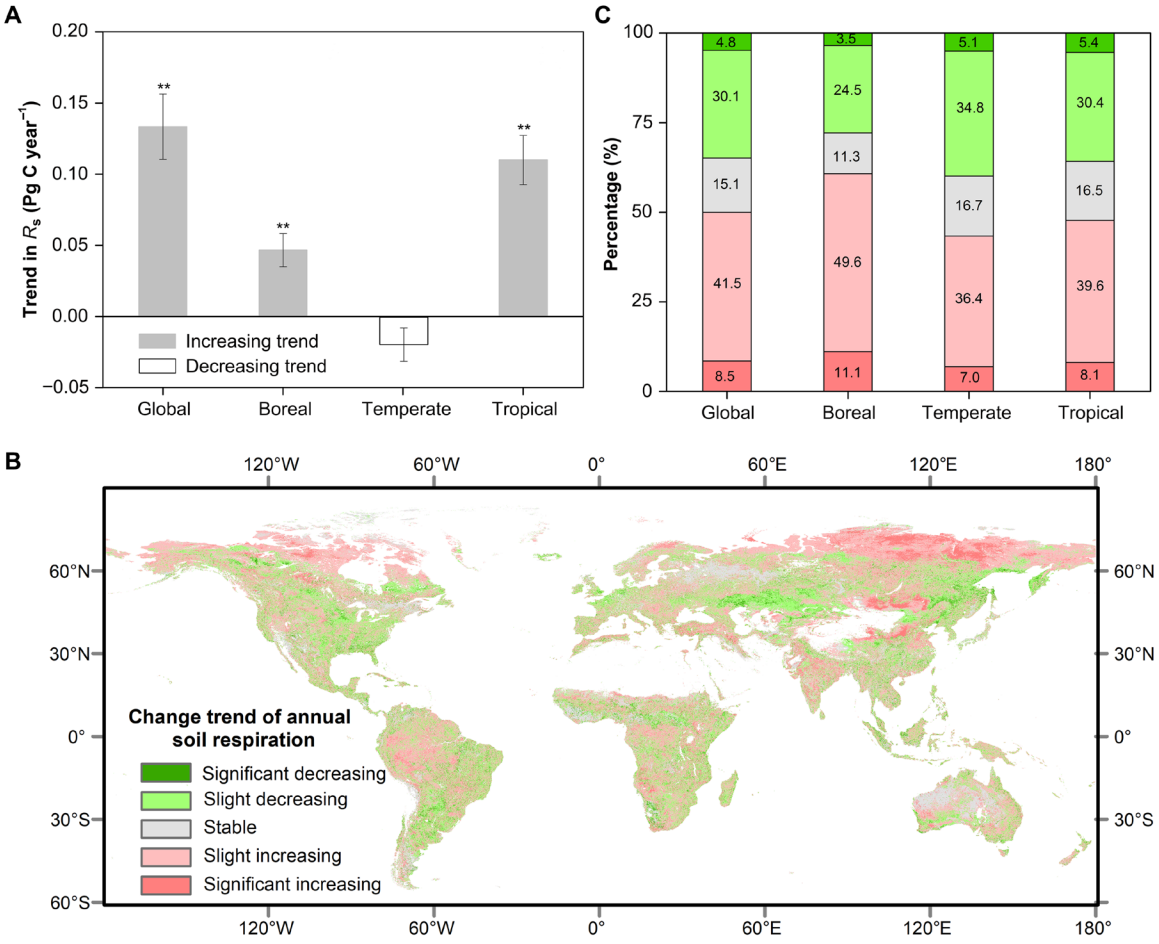

Fig. 2. Trends in annual R_s from 2000 to 2014. (**A**) Overall trends in annual R_s. To calculate the overall trend at the global scale, global average R_s for each year was first derived. Then, a two-sided Mann-Kendall test and a Theil-Sen median trend analysis were performed. Trends at the regional scales were derived following the same method as the global trend. Gray bars in the upward direction indicate an increasing trend, whereas the hollow bar in the downward direction indicates a decreasing trend, with two asterisks denoting significant trends ($P < 0.05$). Error bars represent the 95% CIs estimated via a 1000-bootstrap analysis. (**B**) Spatially explicit trends in annual R_s at 1 km by 1 km spatial resolution. Similar to overall trends in (A), per-pixel trend was characterized using a two-sided Mann-Kendall test and a Theil-Sen median trend analysis. Significant increasing and decreasing trends correspond to positive and negative Theil-Sen estimators, respectively, with significant Mann-Kendall test ($P < 0.05$). Slightly increasing and decreasing trends correspond to positive and negative Theil-Sen estimators, respectively, with nonsignificant Mann-Kendall test results ($P > 0.05$). The residual pixels belong to a stable class. (**C**) Normalized frequency distribution of trends derived using the map presented in (B). The color scheme of the histogram bars matches that of the map legend.

and TEM was also found in the tropical region over the study period ($R = 0.83$, $P < 0.05$; Fig. 3). However, for boreal and temperate regions, land-cover factors, especially the short vegetation (SV) cover, showed consistently significant positive partial correlation with R_s during 2000–2014 (Fig. 3). Through the statistical analyses of the per-pixel significant change trends in R_s and the six driving factors over the globe and three regions (Fig. 2B and figs. S4 and S5), we found that the highest proportion of the areas where R_s changed significantly was located in areas with significant change in TEM at the global scale and in the tropical region (Fig. 4). However, in temperate and boreal regions, the significantly changed R_s was more commonly located in areas with significant changes in SV cover than in areas with significant climate change (Fig. 4). Separating both effects (Fig. 5), climate change accounted for the majority of the observed R_s change at the global scale (56%) and in tropical regions (66%), but land-cover change contributed the most to the R_s change in temperate (58%) and boreal regions (55%).

DISCUSSION

Comparison to previous studies on global R_s

Differences in estimated mean annual global R_s observed between recent studies and the present work (table S1) can be explained largely by differences in the spatial extent of study areas. Our study does not consider R_s from the permanent bare ground (BG) land-cover types (LCTs) because there are very few measurements of the annual R_s from bare lands. On the basis of the Moderate Resolution Imaging Spectroradiometer (MODIS) LCT data, global BG lands have a mean area of 20.6×10^6 km^2 between 2000 and 2014. If the annual R_s of BG lands were 421.3 ± 167.9 g C m^{-2} year^{-1} (*2*), then mean annual R_s of global BG lands would range from 5.2 to 12.1 Pg C year^{-1}. This result explains, to some extent, the underestimation of our global annual R_s and demonstrated that the estimated annual global R_s is likely in line with previous estimates.

The increasing global R_s over the period 2000–2014 was consistent with an increasing trend in global R_h in the periods 1998–2012 (*29*)

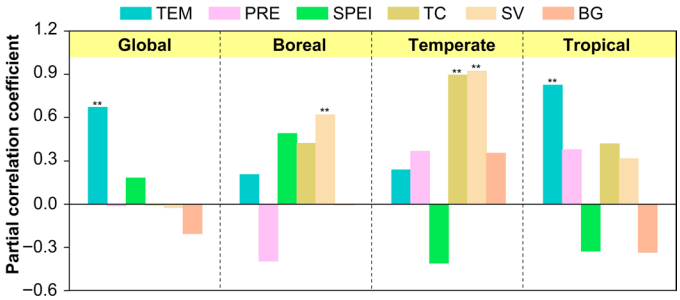

Fig. 3. Relationships between spatially averaged annual R_s and six driving factors from 2000 to 2014. The plot shows partial correlation coefficient between spatially averaged annual R_s and each of the six driving factors at global and regional scales. The six driving factors are annual mean air temperature (TEM), annual precipitation (PRE), annual mean standardized precipitation-evapotranspiration (ET) index (SPEI), tree canopy (TC) cover, short vegetation (SV) cover, and bare ground (BG) cover. All variables (i.e., 15-year spatially averaged annual R_s, TEM, PRE, SPEI, TC cover, SV cover, and BG cover at the global and regional scales) were detrended before conducting partial correlation analysis. Two asterisks indicate that the partial correlation is significant at the 0.05 level (two-tailed).

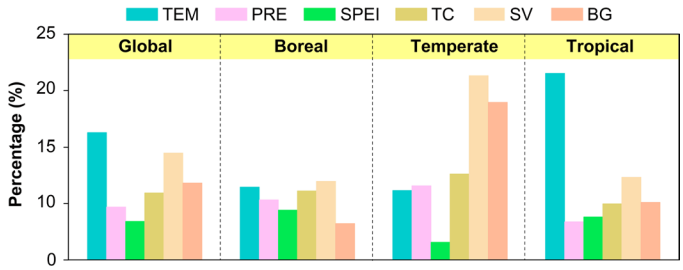

Fig. 4. Proportion of colocated annual R_s change and each of the six driving factors at global and regional scales from 2000 to 2014. The plot represents the percentage of the areas with both significant climate (or land-cover) change and significant annual R_s change to the areas with significant annual R_s change. Climate factors include TEM, PRE, and SPEI. Land-cover factors include TC cover, SV cover, and BG cover. This plot indicates that the significantly changed R_s in temperate and boreal regions was more commonly located in areas with significant changes in SV cover than in areas with significant climate change.

and 2000–2015 (*30*). Tropical regions contributing the largest proportion to global R_s change is consistent with earlier studies (*9*, *11*). R_s in the temperate region showed a nonsignificant decreasing trend between 2000 and 2014 (Fig. 2A), which was different from the temperature-associated significant increases in temperate R_s found by previous studies (*9*, *11*). The differences in the time period considered for these estimates (table S1) could partially explain this result. The relatively greater R_s increase in the boreal region may be explained by the larger magnitude of temperature increase in this region (fig. S3A) because the cold regions exhibited a higher R_s temperature sensitivity and larger carbon stock than those of warm regions (*31*–*33*).

Global R_s in relation to climate and land-cover changes

On a global scale, R_s correlated significantly with TEM (Fig. 3). This result was consistent with the findings of previous studies (*9*, *11*), which exhibited an increasing temporal trend for the annual global R_s, primarily driven by air temperature anomalies. Temperature also played an important role in explaining the interannual variation in R_s in the tropical region (Fig. 3). The positive response of R_s to temperature in the tropical region supported the findings of Fernández-Martínez *et al.* (*33*) that increasing temperature reduced the carbon sink capacity of the tropical region because of a greater stimulation of ecosystem respiration than photosynthesis at higher temperatures (*34*). However, vegetation change in SV showed stronger link with R_s than climate change in both boreal and temperate regions (Fig. 3). Furthermore, the globally dominant, coupled land-cover changes are the changes in tree canopy (TC) cover colocated with changes in SV and changes in SV colocated with BG (fig. S4) (*35*). The close link between TC and SV cover (table S5) explained why the partial correlation coefficient between R_s and SV cover had a similar magnitude as that between R_s and TC cover at the regional scale, especially in the temperate region (Fig. 3).

Overall, although climate change dominated global R_s change, it did not have consistent influences on the change in R_s for the period of 2000–2014 at the regional scale (Fig. 5). Land-cover change, especially changes in SV cover, exerted more effects on R_s in both boreal and temperate regions than in tropical regions (Figs. 3 and 4). However, at the pixel scale, the specific R_s drivers are diverse and interactive (*36*), as discussed in detail below.

Distinct effects of climate change on R_s change

In most tropical regions, including the Amazon Basin, Congo Basin, India, Myanmar, and large areas of Australia, the widespread increase in R_s might be attributed to increased temperature (Fig. 2B and fig. S5A), which was further verified by the positive partial correlation between annual R_s and TEM at the pixel scales (Fig. 6A). However, increased R_s in central and eastern China (Fig. 2B) may be due to increased temperature and precipitation (fig. S5, A and B). A slightly increased annual mean standardized precipitation-evapotranspiration (ET) index (SPEI) (fig. S3C) may explain the spatial hot spots of increased R_s in the western United States where limiting soil moisture has been constraining R_s (*37*).

Some previous studies (*11*, *15*) have shown that warmer air temperatures are associated with lower R_s values in the boreal region. However, our results found that annual R_s showed a strong positive partial correlation with TEM throughout most of the northern high latitudes (over 55°N) and the Tibetan Plateau (Fig. 6A). A pixel-based partial correlation analysis revealed that 65.6% of the boreal region showed a positive partial correlation between annual R_s and TEM, and 17.8% was statistically significant at the 95% confidence level ($P < 0.05$; Fig. 6A). This result supports earlier findings that air temperature is one of the dominant factors constraining R_s in high-latitude and high-altitude carbon cycling (*3*, *33*).

Complex regional interactive effects of climate and land-cover changes on R_s

Hot spots of increasing R_s, found spatially coincident with increases in SV cover, SPEI, and annual precipitation (PRE) and decreases in BG cover, were observed in northern Canada, northern Mongolia, Loess Plateau of China, India, and eastern Australia (Fig. 2B and figs. S4, B and C, and S5, B and C). The widespread distribution of positive partial correlation between R_s and each of the four variables (i.e., SV cover, SPEI, PRE, and BG cover) also supported this finding (Fig. 6, B to F, respectively). This suggests that an increase in SV cover from BG cover coupled with increased moisture jointly drive the increase in R_s through increasing carbon inputs to the soil (*28*).

Most of the increasing R_s also showed a significant positive partial correlation with interannual variations in SPEI, TC, and SV

cover in the boreal region, such as central Russia and northwest Canada (Fig. 6, C to E, respectively). These two areas experienced a widespread increase in SPEI and TC cover and decrease in SV cover (figs. S4, A and B, and S5C), which were probably related to woody expansion (*27*) and permafrost thawing (*38*) induced by climate warming. This result indicates the important effect of moisture on R_s (*16*) and that transformation of SV into woody vegetation may increase R_s in the boreal region (*15*). However, central Quebec, Canada also experienced widespread SV loss and TC gain, but R_s in this area showed a decreasing trend (Fig. 2B and fig. S4, A and B), possibly due to the effect of drought stress (*39*). The alteration of forests and woodlands by deforestation into cultivation is pervasive across tropical regions (*40*), which could increase soil temperature and thus increase R_s (*41*).

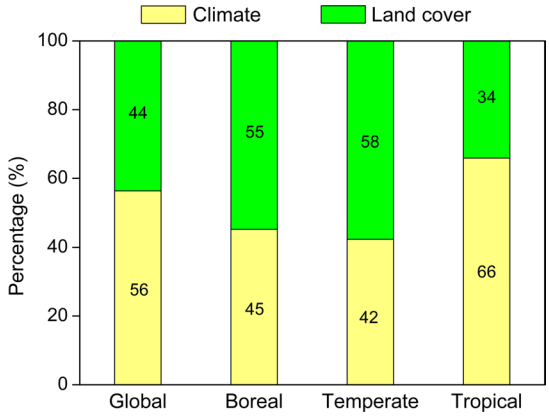

Fig. 5. Contributions of climate change and land-cover change to changes in annual R_s from 2000 to 2014. The plot indicates that climate change dominates global R_s change, but it does not have consistent influence on R_s change at the regional scale.

Decreasing R_s have been observed to be widely distributed in water-limited areas (*42*, *43*), including the central United States, western Europe, northeast China, Kazakhstan, Argentina, east Brazil, east Africa, south Africa, and western Australia (Fig. 2B). These drought-prone areas spatially matched the areas with decreasing trends in SV cover (fig. S4B), with most of these areas corresponding to the areas with decreasing trends in PRE and SPEI (fig. S5, B and C) and some of these areas (i.e., Kazakhstan, Argentina, east Brazil, east Africa, and south Africa) corresponding to widespread increases in BG cover (fig. S4C). The widely distributed positive partial correlations between R_s and SV (or PRE or SPEI; Fig. 6, B, C, and E) in these areas indicated that SV loss and decreased water availability jointly determined the decrease in R_s in these water-limited regions. This phenomenon was particularly pronounced in the temperate region

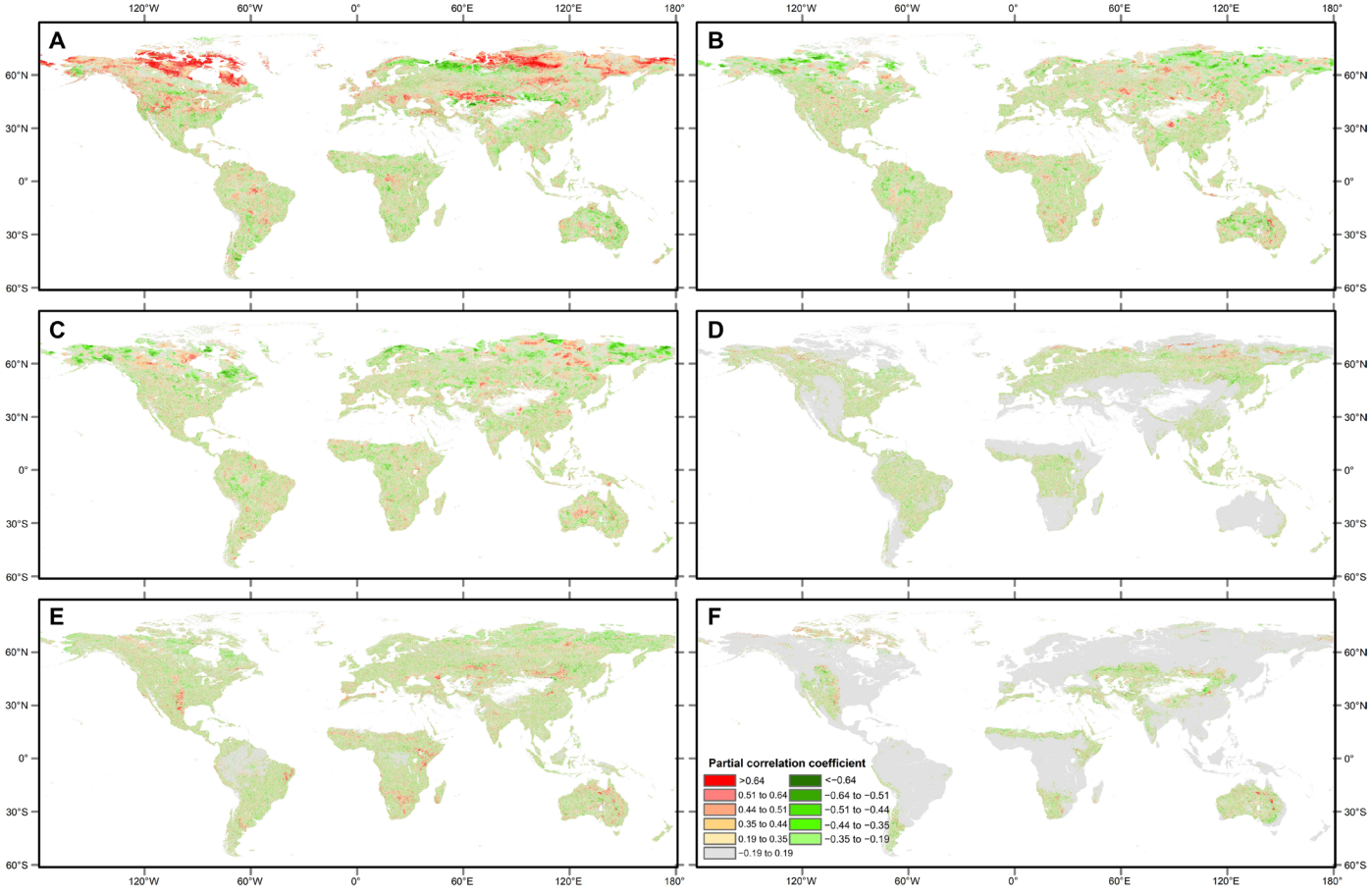

Fig. 6. Spatial patterns of partial correlation coefficient between annual R_s and its six driving factors from 2000 to 2014. Partial correlation coefficient (*R*) between detrended annual R_s and detrended driving factors are shown in (**A**) R_s and TEM, (**B**) R_s and PRE, (**C**) R_s and SPEI, (**D**) R_s and TC cover, (**E**) R_s and SV cover, and (**F**) R_s and BG cover. $R = \pm 0.64$, $R = \pm 0.51$, $R = \pm 0.44$, $R = \pm 0.35$, and $R = \pm 0.19$ correspond to the 0.01, 0.05, 0.1, 0.2, and 0.5 significance levels, respectively. To reduce the effects from the data acquisition error in the land-cover data, only the per-pixel percent cover of TC, SV, and BG greater than 25% (*61*) was used to conduct the partial correlation analysis.

(Fig. 6, B, C, and E), which could help explain the decreased R_s trend in this region (Fig. 2A). The mechanism behind these changes may be attributed to the response of R_s to drought. Drought can lead to a significant decline in vegetation productivity by reducing water availability for plant growth (*44*), inhibiting soil R_h and extracellular enzymes diffusion (*45*), and thus inducing a substantial reduction in R_s (*16*).

Limitations and future work

This analysis has a number of limitations. First, a lack of representativeness of in situ observations of annual R_s is the largest source of uncertainty in the global R_s estimation using bottom-up methods (*7*). Our study is based on an integrated global dataset of R_s, which had more records from temperate regions but a distinct lack of field data from boreal and tropical regions (fig. S1). Without proper correction of the effects of spatial sampling, the heterogeneous distribution of the measurements may bias global R_s estimates toward the rates of the most sampled biomes. Thus, our method carries uncertainties, particularly in the globally undersampled regions. Future efforts to improve our method for R_s estimation would likely benefit most from reduced uncertainty in sampling error. In this study, we established the optimal statistical model for each biome by a combination of remote sensing data and in situ data and then used remote sensing data to estimate global R_s. This approach can efficiently reduce spatial sampling error because use of space-based remote sensing techniques can reduce sampling bias in boreal and tropical regions by providing dense sampling in space and time to characterize the heterogeneity of specific ecosystem properties and parameters [e.g., (*46*)].

Second, these selected biome-specific models explained 62 to 84% of the spatiotemporal variability in the global R_s (table S3), highlighting the limitations of the input datasets and the resulting empirical models in accurately predicting global R_s spatiotemporal patterns. Despite the uncertainties in the remote sensing datasets used for global R_s estimation (table S3), our approach has the potential to provide better estimates of global R_s with improvements to remote sensing–based estimates of the input variables [e.g., gross primary production, land surface temperature (LST), and vegetation index (VI)] (*20*). To better understand spatial and temporal variability of global R_s, more information about the uncertainty in the remote sensing inputs (table S3) and driving datasets (i.e., climate and land-cover data) is needed.

Furthermore, multiple uncertainties remain in modeling global R_s. These uncertainties may be due to the omission of potentially important factors, such as transition periods from a land cover to another (*24*, *25*), nitrogen deposition (*47*), increasing atmospheric CO_2 (*48*), soil organic carbon (*49*), and ecological disturbances (*50*). Last, data at subannual time scales (e.g., monthly and daily data) are important for upscaling global R_s (*17*). Estimation of global R_s at a temporal resolution finer than the current annual time scale may derive different conclusions and deserves further exploration.

CONCLUSIONS

Identifying the factors that affect the contribution of the soil surface CO_2 efflux to the atmosphere will help increase confidence in future projections of the terrestrial carbon cycle in response to global climate change. Until recently, the analysis on driving forces of R_s change was largely limited to climate factors (*9*, *10*, *15*), and the role of land-cover change was rarely investigated. In this study, global annual R_s estimates were obtained from satellite data–driven statistical models that were well constrained by 1292 site years of in situ measurements of R_s at a per-biome scale. Consistent with previous studies, we found an increasing trend in global R_s due to climate change (*9*, *11*, *15*), and our results supported earlier findings that increased temperatures were negatively associated with the terrestrial C sink (*33*). However, unlike earlier analyses, our study, using statistical models and 1 km by 1 km remote sensing data, showed that land-cover change played a major role in driving changes in annual R_s in boreal and temperate regions. This result further confirmed that considering only climate change could not adequately explain the spatial and temporal variations in R_s, especially in temperate and boreal regions. We caution that there are uncertainties in the estimations of global R_s, such as sampling error, uncertainties in remotely sensed input data, ignoring potentially important factors, and the lack of subannual data. Further efforts should be made toward reducing these uncertainties.

MATERIALS AND METHODS

The data, statistical prediction models, and analytical approaches used in this study are described here in sufficient detail to understand the analysis. Additional methodological details needed to reproduce the results are presented in the Supplementary Materials.

Data

Integration of field-based R_s databases

Three available R_s databases covering global and regional scales (*51*–*53*) were collected to generate a centralized field database. These existing datasets were derived on the basis of comprehensive literature surveys of field (not laboratory) R_s measurements. To guarantee data quality for our analysis and to eliminate the mismatch between field-measured R_s and satellite remote sensing data, we applied the following selective criteria to the available records to create the integrated field database: (i) The database strictly focused on annual (i.e., year-round) R_s measurements, rather than seasonal (e.g., during the growing season) measurements. (ii) Only studies using an infrared gas analyzer or gas chromatograph were selected, and annual R_s values collected from nonagricultural biomes were limited to those without experimental manipulation. (iii) Some low-quality data records were removed, such as those from studies where the annual R_s value was extrapolated from low-frequency measurements or obtained from other sources (not original measurements). (iv) When more than one set of year-round R_s measurements were made at one site and for 1 year, those data were averaged to estimate the mean annual R_s for that site and year. As a general rule, averaging was avoided among different vegetation types. For instance, when the records included data from more than one location at the same site (i.e., site with different vegetation cover types), each location corresponding to a vegetation type was incorporated separately into our database. In addition to the annual R_s estimates, other site variables, such as latitude, longitude, study year, TEM, PRE, and vegetation type, were also included in the integrated database. The final database used in this study contained 1292 annual R_s observations collected from 701 sites from 2000 to 2014. These sites were distributed from 43°12′S to 78°10.20′N and from 155°56.40′W to 175°46.20′E, covering most of the global biomes and climate zones (fig. S1).

Remote sensing data products

We collected the global LST, gross primary productivity (GPP), VI, ET, and LCT products from the available MODIS data. Among these products, the 1-km MOD11A2 LST product (https://lpdaac.usgs.gov/products/mod11a2v006/), 500-m MOD17A2H GPP product (https://lpdaac.usgs.gov/products/mod17a2hv006/), 500-m MOD13A1 VI product (https://lpdaac.usgs.gov/products/mod13a1v006/), and 500-m MOD16A2 ET product (https://lpdaac.usgs.gov/products/mod16a2v006/) provided 8-day average values for every year between 2000 and 2014. The MCD12Q1 LCT product with a 500-m spatial resolution (https://lpdaac.usgs.gov/products/mcd12q1v006/) provided annual values for 2001 to 2014. We collected root-zone soil moisture (RZSM) products from the assimilation of the NASA Advanced Microwave Scanning Radiometer (AMSR-E) data and a two-layer Palmer water balance model (*23*), with global coverage from June 2002 to December 2010 at a spatial resolution of 25 km and a temporal resolution of 1 day. We included the Tropical Rainfall Measuring Mission (TRMM) 3B43 gridded global precipitation product with a monthly temporal resolution and 0.25° by 0.25° spatial resolution for 2000 to 2014 (*54*).

The LCT data in 2000 were missing in the LCT product, and for this, we substituted LCT data for 2001. The spatial resolution of the remote sensing data products is diverse, ranging from 500 m to 25 km. For site-scale analyses, we selected the pixels centered on each site from the remote sensing data products by a nearest-neighbor algorithm to match the geographic coordinates (latitude and longitude) of the R_s observation sites. Low-quality data were removed using the available quality flags in each data product. For the spatial analysis at the global scale, we created a 1 km by 1 km spatial resolution dataset from the original remote sensing data products based on nearest-neighbor resampling. All satellite remote sensing data products with 1-day, 8-day, and monthly temporal resolutions were averaged to obtain the corresponding values at annual, spring, and summer time scales.

Climate and land-cover change data

We extracted gridded monthly air temperature and precipitation time series data at a resolution of 0.5° by 0.5° from the website of the Center for Climate Research, University of Delaware (http://climate.geog.udel.edu/~climate/html_pages/download.html#ghcn_T_P_clim2). We also used global standardized precipitation-ET index data (http://SPEI.csic.es/database.html) as a measure of drought intensity (spatial resolution, 0.5° by 0.5°). Three climatic variables—TEM, PRE, and SPEI—were calculated for each year and each pixel from 2000 to 2014 over the globe to analyze the response of R_s to climate change. In addition, we used an annual global land-cover product by Song *et al.* (*35*) to analyze the response of R_s to land-cover change. This land-cover product provided proportional estimates of TC cover, SV cover, and BG cover at a 0.05° by 0.05° spatial resolution over the globe for each year during 2000–2014. Trees are defined as all vegetation taller than 5 m in height. SV refers to all vegetation other than trees, including shrubs, herbaceous vegetation, and mosses, while BG represents the proportion of the land surface not covered by vegetation.

Available predictive variables for annual R_s estimation

Our model development assumes that temperature, moisture, and plant productivity are the predominant drivers of the spatiotemporal variation in global R_s. Remotely sensed LST has shown great potential in estimating R_s at the site and regional scales (*19*). In this study, temperature variables include MOD11A2 LST measurements during the daytime (LST_day) and nighttime (LST_night). Moisture availability, as it relates to temperature, may be better addressed using data representing actual ET or the ratio of actual to potential ET (ET_PET) (*16*) and can be accounted for via the difference between LST_day and LST_night (*18*). Thus, four parameters are used as a surrogate for soil moisture status, including the remotely sensed precipitation from TRMM 3B43 (RS_pre), difference between LST_day and LST_night (LST_diff), ET and ET_PET from MOD16A2, and RZSM, which were extracted from our collected remote sensing data products. GPP and VI can be used as quantitative indicators of plant productivity to estimate R_s (*19*). Thus, this study used the MOD17A2H GPP, enhanced VI (EVI), and normalized difference VI (NDVI) from MOD13A1 as measures of plant productivity.

Because annual R_s is affected by various biotic and abiotic factors at annual and seasonal time scales (*55*), this study used predictive variables from remotely sensed temperature, moisture, and plant productivity variables at annual, spring, and summer time scales for the estimation of global R_s. In this study, spring and summer correspond to March to May and June to August, respectively. In addition, the remotely sensed temperature variables used in this study include LST_day_annual, LST_day_spring, LST_day_summer, LST_night_annual, LST_night_spring, and LST_night_summer. The moisture variables include RS_pre_annual, RS_pre_spring, RS_pre_summer, LST_diff_annual, LST_diff_spring, LST_diff_summer, ET_annual, ET_spring, ET_summer, ET_PET_annual, ET_PET_spring, ET_PET_summer, RZSM_annual, RZSM_spring, and RZSM_summer. The plant productivity variables include GPP_annual, GPP_spring, GPP_summer, EVI_annual, EVI_spring, EVI_summer, NDVI_annual, NDVI_spring, and NDVI_summer.

R_s rates vary among major biomes because of the interactions between climate and vegetation at the global scale (*14*, *15*). Therefore, it is necessary to establish the models at the biome scale for accurately estimating global R_s. Furthermore, the large dataset available from field measurements and satellite remote sensing technology at specific biomes enable the development of biome-specific models, which could be applied to provide global coverage. Thus, comprehensively considering the similarities in climate and vegetation types and the minimum number of data points needed for constructing a statistical model, we divided the globe into 10 biomes: temperate evergreen needleleaf forest, temperate evergreen broadleaf forest, temperate deciduous needleleaf forest, temperate deciduous broadleaf forest, temperate mixed forest, boreal vegetation region, tropical forest, grasslands, croplands, and shrublands. To achieve this analysis, we used global climatic data from the European Soil Data Centre (https://esdac.jrc.ec.europa.eu/) defined on the basis of the climatic classification of the Intergovernmental Panel on Climate Change (2006). The vegetation type data used are MCD12Q1 LCT data from 2000 to 2014.

Nonlinear regression models

The well-known Q_{10} model (Eq. 1) has been frequently used to describe the relationship between annual R_s (g C m^{-2} $year^{-1}$) and temperature. The simple exponential temperature model has been widely applied to the modeling of the temperature sensitivity of R_s at the global scale (*10*, *12*). However, a common criticism of the Q_{10} model is the fixed temperature sensitivity (Q_{10} value, the factor by which the respiration rate increases for a temperature increase of 10°C), which violates the fact that the Q_{10} value decreases with an increase in temperature. In this study, we used a more flexible second-order

polynomial temperature model (Eq. 2), which has been reported to behave better than the simple exponential temperature model in modeling global R_s (*9*)

$$R_s = \theta_1 \times e^{\theta_2 \times T} \quad (1)$$

$$R_s = \theta_1 \times e^{(\theta_2 \times T + \theta_3 \times T^2)} \quad (2)$$

A quadratic model (Eq. 3) was used to quantify the dependence of annual R_s on moisture availability, which considered a reduction in microbial decomposition at very low or very high soil moisture content (*16*)

$$R_s = \theta_1 \times e^{(\theta_2 \times T + \theta_3 \times T^2)} \times e^{(\theta_4 \times M + \theta_5 \times M^2)} \quad (3)$$

A linear model (Eq. 4) was used to describe the dependency of the annual R_s on plant productivity (*7*)

$$R_s = \theta_1 \times e^{(\theta_2 \times T + \theta_3 \times T^2)} \times e^{(\theta_4 \times M + \theta_5 \times M^2)} + \theta_6 \times P + \theta_7 \quad (4)$$

where R_s is the annual soil respiration (g C m^{-2} $year^{-1}$), T is a temperature-related variable, and M is a variable referring to moisture availability. P is a quantitative indicator of plant productivity. θ_1,…, θ_n are model parameters and differ among models. The parameters in nonlinear regression models are estimated by nonlinearly minimizing the sum of the squared residuals.

Predictive variables selected for estimating annual R_s

First, the abovementioned three types of variables (i.e., those related to temperature, moisture, and plant productivity) from remote sensing data products were used to implement nonlinear regressions (i.e., Eqs. 2 to 4). There could be a high correlation among the variables within the same type. As the spatiotemporal coverage and data quality of these global remote sensing data products were different, the available observation data for the fitting models (i.e., Eqs. 2 to 4) were different when one model from Eqs. 2 to 4 was driven by different input variables. To select the optimal variables for R_s estimations in Eqs. 2 to 4 for each biome, a performance metric (ρ) for model evaluation (Eq. 5) was used

$$\rho = 0.25 \times \frac{\mathrm{RMSE}_{\min}}{\mathrm{RMSE}} + 0.25 \times \frac{R^2}{R^2_{\max}} + 0.5 \times \frac{n}{n_{\max}} \quad (5)$$

where n is the number of observation data. $\mathrm{RMSE}_{\min}$ is the minimum RMSE from one model driven by different input variables (i.e., temperature, moisture, or plant productivity variables) for one biome, and $R^2_{\max}$ and $n_{\max}$ are the corresponding maximum R^2 and n, respectively. ρ is a measure of the model's explanatory power $\left(\text{i.e., } \frac{\mathrm{RMSE}_{\min}}{\mathrm{RMSE}} \text{ and } \frac{R^2}{R^2_{\max}}\right)$ and data representation $\left(\frac{n}{n_{\max}}\right)$ for one parameterized model for one biome. The weighting factor for data representation (i.e., 0.5) is greater than that for the model's explanatory power (i.e., 0.25) because we have more confidence in the performance of a model with large n and such a model will tend to have a better spatiotemporal representation.

For each model, with Eqs. 2 to 4, each variable within the same type (temperature, moisture, or plant productivity) was used as one predictor, and the variable with the maximum ρ was selected for each biome. When the optimal temperature variable in Eq. 2 was determined, it was used in Eq. 3. The same data processing was used for Eq. 3 to select the optimal moisture variable. Then, the selected temperature and moisture variables were used in an MNLR model (Eq. 4), and the optimal plant productivity variable was determined for each biome. By gradually adding the optimal inputs from Eqs. 2 to 4, to what extent additional predictor variables would improve the model performance at the biome scale, and the importance of remotely sensed variables, could be determined. To ensure comparisons with the machine learning algorithm models, the same predictive variables (the optimal temperature, moisture, and plant productivity) were maintained for the analysis of the machine learning methods.

Applying machine learning algorithm models

Random forest regression

RFR is an ensemble machine learning algorithm that predicts a response from a set of predictors by creating multiple decision trees and aggregating their results (*56*). Each tree is constructed by a randomly selected subset of training data. The remaining training data, which are called out-of-bag data, are used to estimate the prediction error and variable importance.

Support vector regression

SVR, a regression version of support vector machine algorithms, with an improved generalization ability, uses unique and globally optimal architecture and can be rapidly trained (*57*). SVR projects the input space data into a feature space with a much larger dimension, enabling linearly nonseparable data to become separable in the feature space. It identifies optimum hyperplanes by using kernel functions and arrives at an optimum solution by iteratively adjusting the hyperplanes on the basis of their associated errors.

Artificial neural network

An ANN is a parallel-distributed information processing system that simulates the actions of neurons in the human brain and is able to learn from examples (*58*). In an ANN, information flows in a unidirectional forward mode from an input layer to an output layer via hidden layer(s). Network connection weights are adjusted if the separation of inputs incurs an error during training, and convergence proceeds until the reduction in error between iterations reaches a decay threshold.

Parameter optimizations of the three machine learning algorithm models

The RFR, SVR, and ANN parameters within a certain range (i.e., different numbers or types) were adjusted and tested to ensure that their performance was optimal. For the RFR model, the parameter set included the number of trees (NumTrees), a predictor-splitting algorithm, and the number of predictors to select at random for each split (NumPredictorsToSample). For SVR, the parameter set included half of the width of the epsilon-insensitive band (Epsilon), the flag to standardize the predictor data (Standardize), an optimization routine (Solver), a kernel scale parameter (KernelScale), and a kernel function (KernelFunction). For ANN, the parameter set included training algorithms (trainFcn), the number of hidden layer (N), and hidden layer size (hiddenSizes). An internationally recognized uniform method for parameter optimization for these machine learning algorithm models has not been established. In this study, a trial-and-error process was used to select the optimal parameters by grid search method, which results in the best performance model with the highest R^2 and lowest RMSE.

Global R_s estimation

A 10-fold cross-validation (*59*) was used to evaluate the performance of the four statistical models for annual R_s estimation of each biome. One measured R_s dataset for each biome type was randomly partitioned into 10 equal-sized subsets. During each model fitting, one of the partitions was reserved for validation, while the other nine were used for training. This modeling process was repeated 10 times, and the performance metrics including R^2 and RMSE were averaged to describe the final performance. For each biome, the model with the highest R^2 and lowest RMSE was selected as the optimal model for annual R_s estimation. To confirm the uncertainties associated with the structure and parameters of the selected models, a Monte Carlo approach was used to propagate the model errors to the global estimates. A normal distribution with a 10% coefficient of variance was assumed for each quantitative parameter (*13*). Random sampling from all candidate values was used to analyze the uncertainty of qualitative parameters. For each trial (N = 500), new parameters were chosen from the uncertainty for each parameter, and a group of new models for the 10 biomes were generated to compute the annual global R_s for each year from 2000 to 2014. The means and 95% CIs of all the estimates of the annual global R_s were calculated to represent the model uncertainties. Notably, our "global" study area included only the regions with vegetation according to the definitions of the remotely sensed LCT products. Bare lands, water bodies, urban lands, and built-up lands were not included in our study area.

Trend and attribution analysis

The temporal variation in R_s from 2000 to 2014 was characterized by combining the results of a Mann-Kendall test and a Theil-Sen median trend analysis. The trend uncertainty was estimated by a 1000-bootstrap analysis. If the Mann Kendall test result was statistically significant ($P < 0.05$), then we applied the Theil-Sen estimator to derive the slope (annual change) of trend and provide the estimate of R_s net change between 2000 and 2014 (i.e., slope times 14 years). The direct response of R_s to climate and land-cover changes was examined using the partial correlation analysis between R_s and potential driving factors [i.e., climate (TEM, PRE, and SPEI) and land-cover (TC, SV, and BG cover) factors]. The partial correlation coefficient represents the correlation of each pair of variables after statistically controlling for all the other variables. Before conducting the partial correlation analysis, we first detrended all time series data, including R_s, the climate variables and the land-cover fractions, to avoid autocorrelation and to make the time series stationary. To achieve this, the linear trend derived from the least-squares method was removed for both R_s and its driving factors. All these analyses were performed at the global, regional, and pixel scales. To compare with previous studies on global and regional R_s estimates, we defined three regions based on annual air temperature following the method of Bond-Lamberty and Thomson (*11*): tropical [annual air temperature (T) > 17°C], temperate (2°C ≤ T ≤ 17°C), and boreal (T < 2°C).

The independent effects of climate and land-cover changes on the spatial and temporal variations in R_s were difficult to separate because these variables covaried at different scales (*60*). For instance, climate change is known to control R_s because of its close link with environmental factors (*3*, *10*, *12*), such as soil temperature, soil moisture, and substrate quality, while land-cover change (e.g., grassland to woodland) can be expected to change some or all of these environmental determinants of R_s (*24*, *25*). To analyze the contributions of climate and land-cover changes to the observed global R_s change, land-cover changes between 2000 and 2014 were first derived using Mann-Kendall test and Theil-Sen median trend analysis, following the method of Song *et al.* (*35*). To approximately quantify the impact of climate change and land-cover change, we assumed that changes in R_s colocated with significant land-cover change (two-sided Mann-Kendall test, $P < 0.05$) were attributable to land-cover change, whereas R_s changes that occurred without land-cover change were attributable to climate change. R_s net change at different spatial scale was calculated by summing the per-pixel absolute R_s net change at the corresponding spatial scale. For example, R_s net change for one land-cover change type was calculated by summing the per-pixel absolute R_s net change over this land-cover change type. Last, the overall impact of land-cover change was defined as the proportion of the R_s net change in all significant land-cover change types. The overall climate change impact was defined as the residual of the land-cover change impact. All these analyses were conducted at global and regional scales.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/41/eabb8508/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. P. Friedlingstein, M. W. Jones, M. O'Sullivan, R. M. Andrew, J. Hauck, G. P. Peters, W. Peters, J. Pongratz, S. Sitch, C. Le Quéré, D. C. E. Bakker, J. G. Canadell, P. Ciais, R. B. Jackson, P. Anthoni, L. Barbero, A. Bastos, V. Bastrikov, M. Becker, L. Bopp, E. Buitenhuis, N. Chandra, F. Chevallier, L. P. Chini, K. I. Currie, R. A. Feely, M. Gehlen, D. Gilfillan, T. Gkritzalis, D. S. Goll, N. Gruber, S. Gutekunst, I. Harris, V. Haverd, R. A. Houghton, G. Hurtt, T. Ilyina, A. K. Jain, E. Joetzjer, J. O. Kaplan, E. Kato, K. Klein Goldewijk, J. I. Korsbakken, P. Landschützer, S. K. Lauvset, N. Lefèvre, A. Lenton, S. Lienert, D. Lombardozzi, G. Marland, P. C. McGuire, J. R. Melton, N. Metzl, D. R. Munro, J. E. M. S. Nabel, S. Nakaoka, C. Neill, A. M. Omar, T. Ono, A. Peregon, D. Pierrot, B. Poulter, G. Rehder, L. Resplandy, E. Robertson, C. Rödenbeck, R. Séférian, J. Schwinger, N. Smith, P. P. Tans, H. Tian, B. Tilbrook, F. N. Tubiello, G. R. van der Werf, A. J. Wiltshire, S. Zaehle, Global carbon budget 2019. *Earth Syst. Sci. Data* **11**, 1783–1838 (2019).
2. D. L. Warner, B. Bond-Lamberty, J. Jian, E. Stell, R. Vargas, Spatial predictions and associated uncertainty of annual soil respiration at the global scale. *Glob. Biogeochem. Cycles* **33**, 1733–1745 (2019).
3. X. Tang, S. Fan, M. Du, W. Zhang, S. Gao, S. Liu, G. Chen, Z. Yu, W. Yang, Spatial-and temporal-patterns of global soil heterotrophic respiration in terrestrial ecosystems. *Earth Syst. Sci. Data* **12**, 1037–1051 (2020).
4. B. Bond-Lamberty, New techniques and data for understanding the global soil respiration flux. *Earth Future* **6**, 1176–1180 (2018).
5. A. Anav, P. Friedlingstein, C. Beer, P. Ciais, A. Harper, C. Jones, G. Murray-Tortarolo, D. Papale, N. C. Parazoo, P. Peylin, S. Piao, S. Sitch, N. Viovy, A. Wiltshire, M. Zhao, Spatiotemporal patterns of terrestrial gross primary production: A review. *Rev. Geophys.* **53**, 785–818 (2015).
6. A. G. Konings, A. A. Bloom, J. Liu, N. C. Parazoo, D. S. Schimel, K. W. Bowman, Global satellite-driven estimates of heterotrophic respiration. *Biogeosciences* **16**, 2269–2284 (2019).
7. M. Xu, H. Shang, Contribution of soil respiration to the global carbon equation. *J. Plant Physiol.* **203**, 16–28 (2016).
8. C. L. Phillips, B. Bond-Lamberty, A. R. Desai, M. Lavoie, D. Risk, J. W. Tang, K. Todd-Brown, R. Vargas, The value of soil respiration measurements for interpreting and modeling terrestrial carbon cycling. *Plant Soil* **413**, 1–25 (2017).
9. S. Hashimoto, N. Carvalhais, A. Ito, M. Migliavacca, K. Nishina, M. Reichstein, Global spatiotemporal distribution of soil respiration modeled using a global database. *Biogeosciences* **12**, 4121–4132 (2015).
10. J. W. Raich, C. S. Potter, D. Bhagawati, Interannual variability in global soil respiration, 1980-94. *Glob. Chang. Biol.* **8**, 800–812 (2002).
11. B. Bond-Lamberty, A. Thomson, Temperature-associated increases in the global soil respiration record. *Nature* **464**, 579–582 (2010).
12. S. Chen, Y. Huang, W. Xie, J. Zou, Y. Lu, Z. Hu, A new estimate of global soil respiration from 1970 to 2008. *Chin. Sci. Bull.* **58**, 4153–4160 (2013).
13. S. Hashimoto, A new estimation of global soil greenhouse gas fluxes using a simple data-oriented model. *PLOS ONE* **7**, e41962 (2012).

14. M. Adachi, A. Ito, S. Yonemura, W. Takeuchi, Estimation of global soil respiration by accounting for land-use changes derived from remote sensing data. *J. Environ. Manag.* **200**, 97–104 (2017).
15. Z. Zhao, C. Peng, Q. Yang, F.-R. Meng, X. Song, S. Chen, T. Epule, P. Li, Q. Zhu, Model prediction of biome-specific global soil respiration from 1960 to 2012. *Earth Future* **5**, 715–729 (2017).
16. A. Hursh, A. Ballantyne, L. Cooper, M. Maneta, J. Kimball, J. Watts, The sensitivity of soil respiration to soil temperature, moisture, and carbon supply at the global scale. *Glob. Chang. Biol.* **23**, 2090–2103 (2017).
17. J. S. Jian, M. K. Steele, R. Q. Thomas, S. D. Day, S. C. Hodges, Constraining estimates of global soil respiration by quantifying sources of variability. *Glob. Chang. Biol.* **24**, 4143–4159 (2018).
18. J. Jägermeyr, D. Gerten, W. Lucht, P. Hostert, M. Migliavacca, R. Nemani, A high-resolution approach to estimating ecosystem respiration at continental scales using operational satellite data. *Glob. Chang. Biol.* **20**, 1191–1210 (2014).
19. N. Huang, L. H. Gu, T. A. Black, L. Wang, Z. Niu, Remote sensing-based estimation of annual soil respiration at two contrasting forest sites. *J. Geophys. Res.* **120**, 2306–2325 (2015).
20. J. Xiao, F. Chevallier, C. Gomez, L. Guanter, J. A. Hicke, A. R. Huete, K. Ichii, W. J. Ni, Y. Pang, A. F. Rahman, G. Sun, W. Yuan, L. Zhang, X. Zhang, Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sens. Environ.* **233**, 111383 (2019).
21. S. W. Running, P. E. Thornton, R. Nemani, J. M. Glassy, Global terrestrial gross and net primary productivity from the earth observing system. *Methods Ecosyst. Sci.* 44–57 (2000).
22. Z. M. Wan, New refinements and validation of the MODIS Land-Surface Temperature/ Emissivity products. *Remote Sens. Environ.* **112**, 59–74 (2008).
23. J. D. Bolten, W. T. Crow, X. W. Zhan, T. J. Jackson, C. A. Reynolds, Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. *IEEE J. Sel. Top. Appl. Earth Observ.* **3**, 57–66 (2010).
24. A. Alekseev, T. Alekseeva, P. Kalinin, M. Hajnos, Soils response to the land use and soil climatic gradients at ecosystem scale: Mineralogical and geochemical data. *Soil Tillage Res.* **180**, 38–47 (2018).
25. A. D. Thomas, D. R. Elliott, A. J. Dougill, L. C. Stringer, S. R. Hoon, R. Sen, The influence of trees, shrubs, and grasses on microclimate, soil carbon, nitrogen, and CO_2 efflux: Potential implications of shrub encroachment for Kalahari rangelands. *Land Degrad. Dev.* **29**, 1306–1316 (2018).
26. J. M. Chen, W. Ju, P. Ciais, N. Viovy, R. Liu, Y. Liu, X. Lu, Vegetation structural change since 1981 significantly enhanced the terrestrial carbon sink. *Nat. Commun.* **10**, 4259 (2019).
27. S. Piao, X. Wang, T. Park, C. Chen, X. Lian, Y. He, J. W. Bjerke, A. Chen, P. Ciais, H. Tømmervik, R. R. Nemani, R. B. Myneni, Characteristics, drivers and feedbacks of global greening. *Nat. Rev. Earth Environ.* **1**, 14–17 (2020).
28. B. Mitra, G. Miao, K. Minick, S. G. McNulty, G. Sun, M. Gavazzi, J. S. King, A. Noormets, Disentangling the effects of temperature, moisture, and substrate availability on soil CO_2 efflux. *J. Geophys. Res.* **124**, 2060–2075 (2019).
29. S. Piao, M. Huang, Z. Liu, X. Wang, P. Ciais, J. G. Canadell, K. Wang, A. Bastos, P. Friedlingstein, R. A. Houghton, C. L. Quéré, Y. Liu, R. B. Myneni, S. Peng, J. Pongratz, S. Sitch, T. Yan, Y. Wang, Z. Zhu, D. Wu, T. Wang, Lower land-use emissions responsible for increased net land carbon sink during the slow warming period. *Nat. Geosci.* **11**, 739–743 (2018).
30. B. Bond-Lamberty, V. L. Bailey, M. Chen, C. M. Gough, R. Vargas, Globally rising soil heterotrophic respiration over recent decades. *Nature* **560**, 80–83 (2018).
31. J. Lloyd, J. A. Taylor, On the temperature dependence of soil respiration. *Funct. Ecol.* **8**, 315–323 (1994).
32. C. D. Koven, G. Hugelius, D. M. Lawrence, W. R. Wieder, Higher climatological temperature sensitivity of soil carbon in cold than warm climates. *Nat. Clim. Chang.* **7**, 817–822 (2017).
33. M. Fernández-Martínez, J. Sardans, F. Chevallier, P. Ciais, M. Obersteiner, S. Vicca, J. G. Canadell, A. Bastos, P. Friedlingstein, S. Sitch, S. L. Piao, I. A. Janssens, J. Peñuelas, Global trends in carbon sinks and their relationships with CO_2 and temperature. *Nat. Clim. Chang.* **9**, 73–82 (2019).
34. J. Y. Liang, J. Y. Xia, L. L. Liu, S. Q. Wan, Global patterns of the responses of leaf-level photosynthesis and respiration in terrestrial plants to experimental warming. *J. Plant Ecol.* **6**, 437–447 (2013).
35. X.-P. Song, M. C. Hansen, S. V. Stehman, P. V. Potapov, A. Tyukavina, E. F. Vermote, J. R. Townshend, Global land change from 1982 to 2016. *Nature* **560**, 639–643 (2018).
36. L. L. Reynolds, B. R. Johnson, L. Pfeifer-Meister, S. D. Bridgham, Soil respiration response to climate change in Pacific Northwest prairies is mediated by a regional Mediterranean climate gradient. *Glob. Chang. Biol.* **21**, 487–500 (2015).
37. M. L. Thomey, S. L. Collins, R. Vargas, J. E. Johnson, R. F. Brown, D. O. Natvig, M. T. Friggens, Effect of precipitation variability on net primary production and soil respiration in a Chihuahuan Desert grassland. *Glob. Chang. Biol.* **17**, 1505–1515 (2011).
38. S. E. Chadburn, E. J. Burke, P. M. Cox, P. Friedlingstein, G. Hugelius, S. Westermann, An observation-based constraint on permafrost loss as a function of global warming. *Nat. Clim. Chang.* **7**, 340–344 (2017).
39. M. P. Girardin, E. H. Hogg, P. Y. Bernier, W. A. Kurz, X. J. Guo, G. Cyr, Negative impacts of high temperatures on growth of black spruce forests intensify with the anticipated climate warming. *Glob. Chang. Biol.* **22**, 627–643 (2016).
40. A. Tyukavina, A. Baccini, M. C. Hansen, P. V. Potapov, S. V. Stehman, R. A. Houghton, A. M. Krylov, S. Turubanova, S. J. Goetz, Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012. *Environ. Res. Lett.* **10**, 074002 (2015).
41. F. Mapanda, J. Mupini, M. Wuta, J. Nyamangara, R. M. Rees, A cross-ecosystem assessment of the effects of land cover and land use on soil emission of selected greenhouse gases and related soil properties in Zimbabwe. *Eur. J. Soil Sci.* **61**, 721–733 (2010).
42. A. G. Dai, Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900-2008. *J. Geophys. Res.-Atmos.* **116**, D12115 (2011).
43. W. P. Yuan, Y. Zheng, S. L. Piao, P. Ciais, D. Lombardozzi, Y. P. Wang, Y. Ryu, G. X. Chen, W. J. Dong, Z. M. Hu, A. K. Jain, C. Y. Jiang, E. Kato, S. H. Li, S. Lienert, S. G. Liu, J. E. M. S. Nabel, Z. C. Qin, T. Quine, S. Sitch, W. K. Smith, F. Wang, C. Y. Wu, Z. Q. Xiao, S. Yang, Increased atmospheric vapor pressure deficit reduces global vegetation growth. *Sci. Adv.* **5**, eaax1396 (2019).
44. C. Xu, N. Mcdowell, R. A. Fisher, L. Wei, S. Sevanto, B. Christoffersen, E. Weng, R. Middleton, Increasing impacts of extreme droughts on vegetation productivity under climate change. *Nat. Clim. Chang.* **9**, 948–953 (2019).
45. R. S. Jassal, T. A. Black, M. D. Novak, D. Gaumont-Guay, Z. Nesic, Effect of soil water stress on soil respiration and its temperature sensitivity in an 18-year-old temperate Douglas-fir stand. *Glob. Chang. Biol.* **14**, 1305–1318 (2008).
46. D. Schimel, R. Pavlick, J. B. Fisher, G. P. Asner, S. Saatchi, P. Townsend, C. Miller, C. Frankenberg, K. Hibbard, P. Cox, Observing terrestrial ecosystems and the carbon cycle from space. *Glob. Chang. Biol.* **21**, 1762–1776 (2015).
47. I. A. Janssens, W. Dieleman, S. Luyssaert, J.-A. Subke, M. Reichstein, R. Ceulemans, P. Ciais, A. J. Dolman, J. Grace, G. Matteucci, D. Papale, S. L. Piao, E.-D. Schulze, J. Tang, B. E. Law, Reduction of forest soil respiration in response to nitrogen deposition. *Nat. Geosci.* **3**, 315–322 (2010).
48. J. S. King, P. J. Hanson, E. Bernhardt, P. DeAngelis, R. J. Norby, K. S. Pregitzer, A multiyear synthesis of soil respiration responses to elevated atmospheric CO_2 from four forest FACE experiments. *Glob. Chang. Biol.* **10**, 1027–1042 (2004).
49. P. Leinweber, G. Jandl, C. Baum, K. U. Eckhardt, E. Kandeler, Stability and composition of soil organic matter control respiration and soil enzyme activities. *Soil Biol. Biochem.* **40**, 1496–1505 (2008).
50. B. Borkhuu, S. D. Peckham, B. E. Ewers, U. Norton, E. Pendall, Does soil respiration decline following bark beetle induced forest mortality? Evidence from a lodgepole pine forest. *Agric. For. Meteorol.* **214**, 201–207 (2015).
51. B. Bond-Lamberty, A. Thomson, A global database of soil respiration data. *Biogeosciences* **7**, 1915–1926 (2010).
52. S. T. Chen, J. W. Zou, Z. H. Hu, H. S. Chen, Y. Y. Lu, Global annual soil respiration in relation to climate, soil properties and vegetation characteristics: Summary of available data. *Agric. For. Meteorol.* **198**, 335–346 (2014).
53. T. E. Epule, A new compendium of soil respiration data for Africa. *Challenges* **6**, 1–10 (2015).
54. G. J. Huffman, D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, E. F. Stocker, The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **8**, 38–55 (2010).
55. M. T. Domínguez, A. R. Smith, S. Reinsch, B. A. Emmett, Inter-annual variability of soil respiration in wet shrublands: Do plants modulate its sensitivity to climate? *Ecosystems* **20**, 796–812 (2017).
56. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).
57. V. N. Vapnik, An overview of statistical learning theory. *IEEE Trans. Neural Netw.* **10**, 988–999 (1999).
58. R. Rojas, *Neural Networks: A Systematic Introduction* (Springer-Verlag, 1996).
59. R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection. *Int. Joint Conf. Artif. Intell.* **14**, 1137–1145 (1995).
60. N. Linscheid, L. M. Estupinan-Suarez, A. Brenning, N. Carvalhais, F. Cremer, F. Gans, A. Rammig, M. Reichstein, C. A. Sierra, M. D. Mahecha, Towards a global understanding of vegetation–climate dynamics at multiple time scales. *Biogeosciences* **17**, 945–962 (2020).
61. A. Vina, W. J. Mcconnell, H. Yang, Z. Xu, J. Liu, Effects of conservation policy on China's forest recovery. *Sci. Adv.* **2**, e1500965 (2016).

Acknowledgments

Funding: This research was funded by the National Key Research and Development Program of China (no. 2017YFA0603002), the National Natural Science Foundation of China (nos. 41771465 and 41871347), and the Strategic Priority Research Program of the Chinese Academy of Sciences (no. XDA19030404). **Author contributions:** N.H., L.W., and Z.N. designed the study and analysis. N.H. performed the R_s model runs, analyzed the data, and wrote the

manuscript. All authors contributed substantially to the data processing, writing, and discussion of the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data processing and statistical analyses were conducted using MATLAB (MathWorks, Natick, MA). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Code and raster datasets of model output have been deposited into the Dryad Data Repository at https://doi.org/10.5061/dryad.w3r2280nq. Additional data related to this paper maybe requested from the authors.

Submitted 23 March 2020
Accepted 18 August 2020
Published 7 October 2020
10.1126/sciadv.abb8508

Citation: N. Huang, L. Wang, X.-P. Song, T. A. Black, R. S. Jassal, R. B. Myneni, C. Wu, L. Wang, W. Song, D. Ji, S. Yu, Z. Niu, Spatial and temporal variations in global soil respiration and their relationships with climate and land cover. *Sci. Adv.* **6**, eabb8508 (2020).