A data-driven global soil heterotrophic respiration dataset and the drivers of its inter-annual variability

Yitong Yao, Philippe Ciais, Nicolas Viovy, Wei Li, Fabio Cresto-aleina, Hui Yang, Emilie Joetzjer, Ben Bond-lamberty

▶ **To cite this version:**

Yitong Yao, Philippe Ciais, Nicolas Viovy, Wei Li, Fabio Cresto-aleina, et al.. A data-driven global soil heterotrophic respiration dataset and the drivers of its inter-annual variability. Global Biogeochemical Cycles, 2021, 35 (8), pp.e2020GB006918. 10.1029/2020GB006918 . hal-03317912

HAL Id: hal-03317912
https://hal.science/hal-03317912

Submitted on 16 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Check for updates

Global Biogeochemical Cycles

RESEARCH ARTICLE
10.1029/2020GB006918

A Data-Driven Global Soil Heterotrophic Respiration Dataset and the Drivers of Its Inter-Annual Variability

Yitong Yao[1], Philippe Ciais[1], Nicolas Viovy[1], Wei Li[2], Fabio Cresto-Aleina[1], Hui Yang[1], Emilie Joetzjer[3], and Ben Bond-Lamberty[4]

[1]LSCE/IPSL, CEA-CNRS-UVSQ, Laboratoire des Sciences du Climat et de l'Environnement, Université Paris-Saclay, Gif-sur-Yvette, France, [2]Department of Earth System Science, Tsinghua University, Beijing, China, [3]CNRS, CNRM, Université de Toulouse, Météo-France, Toulouse, France, [4]Pacific Northwest National Laboratory, Joint Global Change Research Institute at the University of Maryland-College Park, College Park, MD, USA

Key Points:

- A new data-driven global soil heterotrophic respiration (SHR) product benefits our understanding of its spatiotemporal dynamics
- Water availability plays an important role in driving inter-annual fluctuation of SHR
- Ecosystem-level drivers for SHR anomalies vary between water-limiting and non-water-limiting regions

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Yao,
yitong.yao@lsce.ipsl.fr

Citation:

Yao, Y., Ciais, P., Viovy, N., Li, W., Cresto-Aleina, F., Yang, H., et al. (2021). A data-driven global soil heterotrophic respiration dataset and the drivers of its inter-annual variability. *Global Biogeochemical Cycles*, *35*, e2020GB006918. https://doi.org/10.1029/2020GB006918

Received 14 DEC 2020
Accepted 23 JUL 2021

© 2021. American Geophysical Union. All Rights Reserved.

Abstract Soil heterotrophic respiration (SHR) is important for carbon-climate feedbacks because of its sensitivity to soil carbon, climatic conditions and nutrient availability. However, available global SHR estimates have either a coarse spatial resolution or rely on simple upscaling formulations. To better quantify the global distribution of SHR and its response to climate variability, we produced a new global SHR data set using Random Forest, up-scaling 455 point data from the Global Soil Respiration Database (SRDB 4.0) with gridded fields of climatic, edaphic and productivity. We estimated a global total SHR of $46.8^{56.3}_{38.6}$ Pg C yr^{-1} over 1985–2013 with a significant increasing trend of 0.03 Pg C yr^{-2}. Among the inputs to generate SHR products, the choice of soil moisture datasets contributes more to the difference among SHR ensemble. Water availability dominates SHR inter-annual variability (IAV) at the global scale; more precisely, temperature strongly controls the SHR IAV in tropical forests, while water availability dominates in extra-tropical forest and semi-arid regions. Our machine-learning SHR ensemble of data-driven gridded estimates and outputs from process-based models (TRENDYv6) shows agreement for a strong association between water variability and SHR IAV at the global scale, but ensemble members exhibit different ecosystem-level SHR IAV controllers. The important role of water availability in driving SHR suggests both a direct effect limiting decomposition and an indirect effect on litter available from productivity. Considering potential uncertainties remaining in our data-driven SHR datasets, we call for more scientifically designed SHR observation network and deep-learning methods making maximum use of observation data.

1. Introduction

Soil heterotrophic respiration (SHR), the CO_2 flux produced by free-living microbial heterotrophs and soil fauna feeding on soil organic matter (Carbone et al., 2016; Hanson et al., 2000), constitutes a key ecosystem-to-atmosphere carbon flux that affects soil carbon storage and carbon-climate feedbacks. Since the magnitude of SHR is roughly four times of global annual anthropogenic fossil fuel emission (Le Quéré et al., 2018) and SHR can regulate the net ecosystem carbon exchange variability in some regions (Liu, Ballantyne, et al., 2018), even small changes in this flux can cause carbon redistribution between soil and atmosphere, and modify the carbon sink. Enhanced microbial dynamics in soil organic matter decomposition have been detected as the dominant factor in an increasing imbalance between higher CO_2 loss rate and CO_2 uptake by plants (Bond-Lamberty et al., 2018). Therefore, a detailed understanding of the SHR spatial and temporal dynamics under changing climate conditions is pivotal to improve projections of the carbon–climate feedback (Ballantyne et al., 2017; Bradford et al., 2019).

However, unlike other components of the terrestrial carbon cycle like gross primary productivity (GPP) that can be measured through eddy covariance flux tower at plot scale, SHR observations mainly come from small-scale chambers, combined with intrusive methods (trenching, root exclusion, root extraction), or non-intrusive methods of isotope labeling with uncertainty in ^{14}C measurements (Bond-Lamberty et al., 2004; Hanson et al., 2000) to partition the heterotrophic and autotrophic soil fluxes. Due to the considerable uncertainty underlying these measurements, SHR is the most poorly constrained ecosystem and global carbon flux (Ciais et al., 2020; Konings et al., 2019).

Table 1
Comparison With Previously Published Respiration Datasets

Data type	Reference	Target variable	Number of site data	Method	Period	Spatial resolution	Uncertainty
Site-level data	Bond-Lamberty and Thomson (2010)	Rs, SHR	3,379 records of Rs; 333 records of SHR (if we use same filtering criterion)	Collecting published studies	1961–2007	-	-
Global map	Hashimoto et al. (2015)	Rs, SHR	1,638 records of Rs	Monthly Rs = f(monthly temperature, monthly precipitation, precipitation of the previous month) A globally constant coarse ratio is used to transform Rs to SHR: ln(annual SHR) = 1.22 + 0.73*ln(annual Rs)	1965–2012	0.5° × 0.5°	-
	Tang et al. (2020)	SHR	504 records of SHR (with update from Chinese publications)	Random Forest	1980–2016	0.5° × 0.5°	-
	Warner et al. (2019)	Rs; SHR	2,657 records of Rs	Quantile regression forest on Rs, and empirical equation on SHR.	-	1 km	Sampling bias uncertainty
	Adachi et al. (2017)	Rs	Only five sites were used in validation	Empirical equation	2001–2009	4 km	-
	Konings et al. (2019) – top-down	SHR	-	Inverting land surface carbon balance	2010–2012	4° × 5°	Inversion uncertainty
	This study	SHR	455 records of SHR (after filtering)	Random Forest	1985-2013 (overlapped period affected by the explanatory variables)	0.5° × 0.5°	Extrapolation uncertainty and uncertainty from alternative gridded input datasets

Note. Rs: soil respiration.
Abbreviation: SHR, soil heterotrophic respiration.

Most existing SHR-related research has focused on testing the sensitivities of SHR to environmental variations through multi-factorial manipulation experiments at small scales like soil warming (Noh et al., 2016; Schindlbacher et al., 2009), rainfall exclusion (Hinko-Najera et al., 2015; S. Huang et al., 2018), water addition (Liu, Lü, et al., 2018; Zou et al., 2018), and nitrogen fertilization experiments (Z. Chen et al., 2018; Peng et al., 2018). Although these studies enable us to understand local SHR responses under different environmental conditions, large-scale spatial information of SHR is still limited, and contains large uncertainty. For example, global SHR estimates from Earth System Models (ESMs) range from 40 to 70 PgC yr^{-1} during 1965–2004 (Hashimoto et al., 2015; Shao et al., 2013) with large inter-model differences. Apart from these ESM estimates, we know of only four global SHR maps available at present (Table 1). The first data-driven SHR map was generated by Hashimoto et al. (2015) who upscaled *in situ* measurements of soil respiration with a modified version of functional relations from Raich and Potter (1995) and Raich et al. (2002) to calculate total soil respiration using monthly temperature and precipitation, and then used a constant partitioning ratio to scale down total soil respiration to SHR. One limitation acknowledged by this study is that only 53 sites from Bond-Lamberty et al. (2004) were used to derive that coarse ratio. Warner et al. (2019) used a

similar approach, first computing total soil respiration and then SHR from a partitioning ratio, to generate a static high resolution global SHR map. Tang et al. (2020) used Random Forest model to produce a data-driven global SHR data set but this lacked uncertainty evaluation. Finally, Konings et al. (2019) employed a constrained carbon balance framework using atmospheric inversion based on net ecosystem productivity (NEP), solar-induced fluorescence based GPP, and modeled carbon use efficiency to produce a top-down global SHR estimation; this is available at relatively coarse spatial resolution (4° × 5°) and for a short time period (2010–2012), which limits its validation against *in situ* measurements. Considering the limitations of the above-mentioned datasets, there is a clear need for spatially explicit data-driven global-scale SHR maps at a finer resolution and for a longer time period with full uncertainty accounting, which can be used for evaluation and optimization of process-based models (N. Huang et al., 2020; J. Li et al., 2016).

A decade ago, Bond-Lamberty and Thomson (2010) compiled a large soil respiration database — the Global Soil Respiration Database (SRDB), with soil respiration observations from peer-reviewed literature studies. The number of soil respiration records has reached 6,634 (4,111 valid values) in the latest SRDB 4.0 (Bond-Lamberty & Thomson, 2018) although SHR observations are far fewer (674 valid values before filtering); very recently, the SRDB 5.0 (Jian et al., 2021, not used here) was released with 1,147 annual SHR values. A data-driven SHR estimate from site to regional even global scale is thereby becoming feasible now, and Bond-Lamberty et al. (2016) suggested that machine learning could be an ideal tool toward the large-scale data-driven SHR estimation. Machine learning algorithms are powerful tools for data-driven up-scaling estimation of a target variable in ecological studies (Jung et al., 2011; Steidinger et al., 2019; Yao et al., 2018). Research on predictions of carbon or water flux (Jung et al., 2010; Z. Zeng et al., 2014) and crop yield (Cai et al., 2019; Feng et al., 2019) have affirmed the utility of those algorithms. Estimation with such data-oriented techniques gives us a new opportunity to evaluate SHR and its dynamics in response to environmental variations globally.

In this study, we apply Random Forest (RF) algorithms to estimate global annual SHR at 0.5° × 0.5° spatial resolution over the period 1985–2013 with meteorological, edaphic factors and GPP as explanatory variables. With this method, we can produce an ensemble of different data-driven SHR gridded data set at global scale over last three decades, and we are also able to examine the contribution of dynamic climate drivers to SHR inter-annual variability (IAV), including annual temperature, precipitation or soil moisture, and radiation.

2. Materials and Methods

2.1. Soil Respiration Database

The Global SRDB, is composed of soil respiration measurements from peer-reviewed studies. It was first released in Bond-Lamberty and Thomson (2010) and updated in Bond-Lamberty and Thomson (2018) as SRDB version 4.0. We used 455 site-year observations after data filtering. There are 290 sites in total and most sites have records less than 3. Our data filtering criteria included: (a) removing records without detailed temporal, coordinates and annual SHR information, and (b) excluding observations from manipulation experiments and soda lime measurements, which tend to underestimate soil CO_2 fluxes (Haynes & Gower, 1995). Records using isotope, gas chromatography or other measurements were retained. Figure 1 shows that current SHR observations after filtering mainly distributed in temperate zones, with higher sample density in East Asia, Europe and North America. These available observations belong to seven ecosystem types (Friedl et al., 2010).

Here we introduce different climatic, soil moisture and GPP datasets that are used to produce large ensemble global SHR datasets by upscaling SRDB point data, which improves on previous studies with only few available SHR members resulting in incomplete uncertainty accounting.

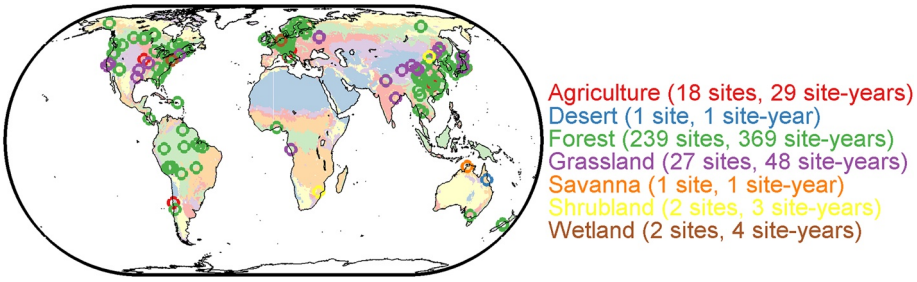

Figure 1. The spatial distribution of selected soil heterotrophic respiration observation sites (bold circles) by ecosystem type. The land cover used to define ecosystems from MCD12Q1 is shown in a color lighter than the circles in same category. Numbers in brackets denote the available sites and records for a given ecosystem.

2.2. Climatic Datasets

2.2.1. Temperature, Precipitation and Radiation

2.2.1.1. CRUNCEP

The climatic variables of annual temperature, annual precipitation and annual short-wave radiation used in this study are obtained from the CRUNCEP v6.1 (Viovy, 2015) covering the period 1901 to 2015. CRUNCEP is a combination of two existing datasets: the Climate Research Unit (CRU) TS3.2 observation-based monthly climatological data in spatial resolution of 0.5° × 0.5°, and the National Center for Environmental Prediction (NCEP) reanalysis product in 2.5° × 2.5° and 6-hour temporal step. The latter is used to define the diurnal and daily variation of the climate forcing.

2.2.1.2. CRUJRA

CRUJRA v1.1 is based on the same methodology as CRUNCEP but uses the Japanese Reanalysis Data (JRA) produced by the Japanese Meteorological Agency (JMA) with a spatial resolution of 0.5° × 0.5°adjusted to match the CRU TS 3.26 data (Harris et al., 2014; Kobayashi et al., 2015). These data are available at a 6-hour time-step from 01/1901 to 12/2017. The annual temperature, annual precipitation and annual short-wave radiation from CRUJRA are also used in this study.

2.2.1.3. Princeton Climate Data Set

This global meteorological forcing data set is a blend of NCEP/NCAR reanalysis data and a series of global observations to form global 0.25°×0.25° daily temperature and precipitation datasets from 1948 to 2016 (Sheffield et al., 2006). We aggregate daily data to annual values first.

2.2.1.4. WFDEI Meteorological Forcing Data Set

This WFDEI meteorological forcing data set is a combination of ERA-interim re-analysis data with daily variability and monthly in-situ observation. There are two precipitation products available from WFDEI: one is corrected using CRU observations, and the other using the Global Precipitation Climatology Centre (GPCC) data set (Weedon et al., 2014). WFDEI-CRU and WFDEI-GPCC are both used in this study. These two datasets span from 1979 to 2016 and are provided at 0.5°×0.5° spatial resolution.

2.2.1.5. Climate Prediction Center (CPC)

This CPC climatological data set includes daily maximum and minimum temperature (t_{max} and t_{min}) and daily precipitation. Daily average temperature is generated by averaging t_{max} and t_{min}. Then the daily temperature and daily precipitation are aggregated to annual values. This data set starts from 1979, has been updated through 2021, and is available at 0.5° × 0.5°.

2.2.2. Soil Moisture Datasets

2.2.2.1. CPC Soil Moisture

The CPC monthly soil moisture data set is estimated by a one-layer leaky-bucket model (J. Huang et al., 1996; Van den Dool et al., 2003). The driving fields for the model include temperature and precipitation from CPC

precipitation reconstruction over Land (M. Chen et al., 2002) and CPC Global Land Surface Air Temperature Analysis (Fan & Van den Dool, 2008). This data set starts from 1948 and updates in real time (currently to 2021), with a spatial resolution of 0.5° × 0.5° (Fan & van den Dool, 2004). The CPC moisture is calculated with a soil depth of 1.6 m.

2.2.2.2. Gravity Recovery and Climate Experiment Terrestrial Water Storage Reconstruction (GRACE-Rec TWS)

Humphrey et al. (2017) constructed a statistical model by linking the anomalies of the main meteorological drivers to the terrestrial water storage (TWS) anomalies observed by the GRACE satellite after 2002. The pre-2002 TWS anomalies are reconstructed based on this calibrated statistical model, driven by precipitation and temperature. This reconstructed TWS data set covers 1985–2015 and is provided at a spatial resolution of 0.5° × 0.5°. It should be noted that this variable is not equal to soil moisture but also includes change in land ice, and free water (Humphrey et al., 2017). In the results section below, "GRACE-rec" is used as an abbreviation for GRACE reconstruction.

2.2.2.3. Global Land Data Assimilation System (GLDAS) Version 2

GLDAS version 2 data is composed by GLDAS 2.0, which uses the Princeton meteorological data as forcing data for 1948–2010, and GLDAS 2.1, which is forced by a combination of model and observation based datasets from 2000 to present (Beaudoing & Rodell, 2016; Rodell et al., 2004). The output soil moisture is at 0.25° × 0.25° and has four soil layers, 0–0.1 m, 0.1–0.4 m, 0.4–1.0 m, and 1.0–2.0 m. We take the sum of soil moisture in these four layers.

2.3. Soil Properties Datasets

2.3.1. Soil Carbon Content

Soil carbon content is extracted from Harmonized World Soil Database (HWSD), which is produced by FAO and IIASA by combining existing global regional and national inventories for soil information in over 15,000 different soil mapping units (Nachtergaele et al., 2010; Wieder et al., 2014). Top-soil (0–0.3 m) and sub-soil (0.3–1.0 m) organic carbon content are provided and employed.

2.3.2. Soil Nitrogen Density

We also used a soil nitrogen density data set (unit: g N m^{-2}) from Global Gridded Surfaces of Selected Soil Characteristics developed by the International Geosphere-Biosphere Program Data and Information System (IGBP-DIS). A statistical bootstrapping approach is applied by the SoilData System to link the global pedon records to the FAO/UNCESCO digital soil map. The total soil nitrogen content is for a soil depth of 0–1.0 m. Soil nitrogen density was treated as explanatory variable for SHR estimate.

2.4. Land Cover Data Set

The MODIS land cover type product (MCD12Q1) is derived using a supervised decision tree classification algorithm (Friedl et al., 2010). We group land cover types in the MODIS product to correspond with the classification in SRDB following Table S1. In this study, we do not consider the effect of land use/cover change (cf. N. Huang et al., 2020), and we use a static land cover map in 2001 as input for the estimation model.

2.5. Gross Primary Productivity Datasets

2.5.1. FLUXCOM

GPP drives net primary production and litterfall, thus it is used as a predictor of SHR. We used an ensemble of gridded GPP products generated by training three machine learning algorithms (Random Forest, Multivariate Adaptive Regression Spline, and Artificial Neural Network) on daily GPP estimates from 224 flux towers (Jung et al., 2017; Tramontana et al., 2016). The combinations between three available algorithms and two GPP estimates (due to two partitioning methods) produce 6 GPP ensembles spanning 1980–2013 in monthly intervals. Each GPP member is used as one GPP data source for SHR estimation.

2.5.2. P-Model GPP

P-model is a Light Use Efficiency (LUE) model (Wang et al., 2017), in which monthly LUE is predicted on the basis of changing environmental conditions and an optimality criterion with respect to stomatal behaviors and other related traits (Prentice et al., 2014). Daily GPP, which is then calculated by monthly LUE and daily varying absorbed photosynthetically active radiation, can be further aggregated to annual time step. This product is released at a spatial resolution of 0.25×0.25°and over a time span of 1982–2016 (Stocker et al., 2019). This GPP data set includes the soil moisture effects on LUE.

2.6. Random Forest Algorithm and Its Performance

To generate a data-driven SHR estimation, we implemented a Random Forest ensemble machine learning algorithm, as in former carbon research studies (Baccini et al., 2012; Buermann et al., 2018; Jung et al., 2017; Schwalm et al., 2017; Zhu et al., 2017). Random Forest, as one of the widely used and fast running algorithms, is suited for handling non-linear relationship between the target and the corresponded independent variables, without requiring predefined functional forms or a normal sample distribution (Breiman, 2001). A RF model consists of multiple uncorrelated regression trees, each of which uses a subset of all the training samples with replacement (~63%) to reach the same total sample size and random subset of explanatory variables (Breiman, 2001). This bootstrapping procedure can decrease the influence of noise and outliers, and raise the stability of model predictions, by averaging over all constructed trees. In our study, the number of trees is set to be 1,000. In addition to the mean value output from 1,000 trees, we also obtain 95% confidence interval of outputs from these trees. The feature importance is provided by assessing the difference in prediction error (Mean Squared Error for this study) on out-of-bag data before and after variable permutation.

To get a robust evaluation of model performance, we use Leave-One-Out Cross Validation (hereinafter LOOCV). The goal of cross-validation is to test the model's ability to work on independent samples, in order to identify problems like over-fitting. Each time a sample is excluded and the remaining samples are used to train a RF model, the predicted value of the excluded one is estimated by that fitted model. Instead of taking average of out-of-bag R^2 from all trained RF models, we use the LOOCV-based predicted value of each observation to get R^2 as metric. LOOCV is used both at the sample and site levels.

Since the observed SHR data are provided at annual timescale, we need to aggregate the high temporal resolution to coarser time scale (annual) as inputs to the RF model. The RF model was first fitted on our filtered 455 site-year observation data set, and then applied to predict annual SHR for each $0.5° \times 0.5°$ grid cell driven by meteorological factors and other environmental indicators (Table 2). We obtain 126 (6*7*3 = 126) global SHR members in total derived from crossing combinations of six climate datasets, seven GPP datasets (six members from FLUXCOM and one member from P-model) and three soil moisture/ TWS datasets (CPC, GRACE-rec and GLDAS) used as gridded inputs to RF model. Considering that different climate datasets (temperature and precipitation), soil moisture datasets, and GPP datasets served as our efficient explanatory variables and their different time period coverage, we choose 1985–2013 as the common time length for estimation.

2.7. Soil Heterotrophic Respiration Datasets Used for Comparison

2.7.1. Soil Respiration Data Set From Hashimoto et al. (2015)

Hashimoto et al. (2015) applied climate-driven functions modified from Raich et al. (2002) to fit soil respiration from 1,638 data points. The soil respiration was solely driven by temperature and precipitation, and the partitioning between its autotrophic and heterotrophic part followed two fixed parameters summarized from the data in Bond-Lamberty et al. (2004) (an approach also used by Warner et al., 2019). Two parameters in the model were assumed to be globally constant, so that the spatial heterogeneity in the Hashimoto et al. (2015) data set is provided only by variations in climatic drivers. This data-driven model was extrapolated to 1901–2012. We downloaded this data set from http://cse.ffpri.affrc.go.jp/shojih/data/index.html.

Table 2
Explanatory Variables Used for Soil Heterotrophic Respiration Upscaling and Their Data Access

Variable name	Variable variability state	Data access
Annual temperature	Inter-annual	CPC: https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globaltemp.html
Annual precipitation	Inter-annual	https://www.esrl.noaa.gov/psd/data/gridded/data.cpc.globalprecip.html
Annual short-wave radiation	Inter-annual	CRUNCEP: http://dods.extra.cea.fr/store/p529viov/cruncep/ CRUJRA: http://dx.doi.org/10.5285/13f3635174794bb98cf8ac4b0ee8f4ed Princeton: http://hydrology.princeton.edu/data.pgf.php WFDEI: http://www.eu-watch.org/data_availability
Annual soil moisture	Inter-annual	CPC: https://www.esrl.noaa.gov/psd/data/gridded/data.cpcsoil.html GRACE reconstruction: http://doi.org/10.5905/ethz-1007-85 GLDAS: https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH025_M_V2.1/summary
Annual gross primary productivity	Inter-annual	FLUXCOM: https://www.bgc-jena.mpg.de/geodb/projects/Home.php P-model: https://doi.org/10.5281/zenodo.1423484
Annual nitrogen deposition	Inter-annual	https://daac.ornl.gov/NACP/guides/NACP_MsTMIP_Model_Driver.html
Soil carbon content	Static	https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=1247
Total nitrogen density	Static	https://daac.ornl.gov/SOILS/guides/igbp-surfaces.html
Land cover classification	Static	https://modis.gsfc.nasa.gov/data/dataprod/mod12.php

Abbreviations: CPC, climate prediction center; GLDAS, global land data assimilation system; GRACE, gravity recovery and climate experiment.

2.7.2. Soil Heterotrophic Respiration Data From Tang et al. (2020)

Tang et al. (2020) upscaled site-level observations in SRDB4.0 using Random Forest model to produce an annual global SHR data set spanning 1980–2016 in spatial resolution of 0.5° × 0.5°. The explanatory variables include mean annual temperature, mean annual precipitation, diurnal temperature range, nitrogen deposition, Palmer Drought Severity Index, shortwave radiation, soil carbon content, soil nitrogen content and soil water content. We downloaded this data set from https://doi.org/10.6084/m9.figshare.8882567.

2.7.3. Top-Down SHR Data Set From Konings et al., (2019)

Konings et al., (2019) produced a top-down SHR estimate from 2010 to 2012 in spatial resolution of 4° × 5°. Total ecosystem respiration was first derived from carbon balance method by the difference between GPP and NEP and carbon use efficiency from model-data fusion framework was used to partition the autotrophic and heterotrophic respiration. NEP is from atmospheric inversion of NASA Carbon Monitoring System-Flux. GPP is based on satellite-observed solar-induced fluorescence. Monthly top-down SHR data was aggregated to an annual time step.

2.7.4. TRENDY Global Models

We use simulations of 12 dynamic global vegetation models from the project "Trends and drivers of the regional scale sources and sinks of carbon dioxide" (TRENDY) v6 for the period of 1985–2013. These models used a common set of observed climate, atmospheric CO_2 concentration, land-use change, and experimental protocols. Our analysis uses model SHR outputs (named as Rh in models) under the TRENDY "S3" experiment (which included the effects of time-varying CO_2 concentrations, climate change, and land use change). The 12 models are CABLE (Haverd et al., 2018), CLASS-CTEM (Melton & Arora, 2016), CLM4.5

(Oleson et al., 2013), ISAM (Jain et al., 2013), JSBACH (Reick et al., 2013), JULES (Clark et al., 2011), LPX (Keller et al., 2017), OCN (Zaehle & Friend, 2010), ORCHIDEE-MICT (Guimberteau et al., 2018), ORCHIDEE (Krinner et al., 2005), VEGAS (N. Zeng et al., 2005) and VISIT (Kato et al., 2013).

2.8. Attribution Analysis for SHR Inter-Annual Variability (IAV)

We apply the carbon flux anomaly decomposition approach proposed in Jung et al. (2017) to diagnose the contribution of different climate variables on SHR IAV, including annual temperature (TEMP), annual water availability proxies (precipitation or soil moisture, PREC or SMC) and annual short-wave radiation (RAD). We first obtain detrended annual SHR anomalies and climate indicators anomalies in each pixel by removing their linear trend on annual basis (least-squares fitting). Then we implement multiple linear regressions with zero-intercept between anomalies in SHR and all variables to separate their contribution. The estimated SHR sensitivity to climate anomaly (shown as the linear regression slopes a_s^{var} in Equation 1) multiplied by the corresponding forcing anomaly (shown as $\Delta \mathrm{var}_{s,y}$ in Equation 1) defines the SHR anomaly component $\Delta \mathrm{SHR}_{s,y}^{\mathrm{var}}$ driven by each forcing variable. These components can be added together to form a reconstructed SHR anomaly ($\Delta \mathrm{SHR}_{s,y}^{*}$ in Equation 2). A correlation coefficient is calculated to demonstrate the consistency between the climate-reconstructed SHR anomalies ($\Delta \mathrm{SHR}_{s,y}^{*}$) and the detrended estimates ($\Delta \mathrm{SHR}_{s,y}$). We also implement similar procedures on SHR from the Hashimoto and Tang datasets and TRENDY models. In all attribution analyses (except Hashimoto), both precipitation and soil moisture are considered as water-related proxies in the regression processes (Equations 1–4).

$$\Delta \mathrm{SHR}_{s,y,e} = a_{s,e}^{\mathrm{TEMP}} \times \Delta \mathrm{TEMP}_{s,y,e} + a_{s,e}^{\mathrm{PREC}} \times \Delta \mathrm{PREC}_{s,y,e} + a_{s,e}^{\mathrm{RAD}} \times \Delta \mathrm{RAD}_{s,y,e} + \in_{s,y,e} \quad (1)$$

$$\Delta \mathrm{SHR}^{*}{}_{s,y,e} = \Delta \mathrm{SHR}_{s,y,e}^{\mathrm{TEMP}} + \Delta \mathrm{SHR}_{s,y,e}^{\mathrm{PREC}} + \Delta \mathrm{SHR}_{s,y,e}^{\mathrm{RAD}} \quad (2)$$

$$\Delta \mathrm{SHR}_{s,y,e} = a_{s,e}^{\mathrm{TEMP}} \times \Delta \mathrm{TEMP}_{s,y,e} + a_{s,e}^{\mathrm{SMC}} \times \Delta \mathrm{SMC}_{s,y,e} + a_{s,e}^{\mathrm{RAD}} \times \Delta \mathrm{RAD}_{s,y,e} + \in_{s,y,e} \quad (3)$$

$$\Delta \mathrm{SHR}^{*}{}_{s,y,e} = \Delta \mathrm{SHR}_{s,y,e}^{\mathrm{TEMP}} + \Delta \mathrm{SHR}_{s,y,e}^{\mathrm{SMC}} + \Delta \mathrm{SHR}_{s,y,e}^{\mathrm{RAD}} \quad (4)$$

Subscripts *s*, *y,* and *e* refer to index of grid cells, every year from 1985 to 2013, and one member from the SHR ensemble, respectively.

2.9. Factorial Analysis for SHR Trend

To understand the controlling variables for the regional variation in SHR trend map, we performed factorial estimation by removing the inter-annual variation of each dynamic explanatory variable, that is, keeping each variable static during 1985–2013. Dynamic explanatory variables of annual temperature, annual precipitation, annual radiation, annual GPP, annual soil moisture and annual nitrogen deposition were tested.

2.10. Statistical Software

All data processing and statistical analysis were performed in R statistical software (R Development Core Team, 2019) version 3.5.0 using packages "randomForest" version 4.6–14 (Liaw & Wiener, 2018), "raster" version 3.0–7 (Hijmans et al., 2019), "pracma" version 2.2.5 (Borchers, 2019).

3. Results

3.1. Performance of Random Forest Via Cross Validation

After testing on all combinations of these temperature and precipitation datasets, soil moisture and GPP choices, we derived the LOOCV R^2 of these RF models, which is 0.57 ± 0.01 (0.01 is the standard deviation of LOOCV R^2 from all combinations). We can see that at least at site-level, different climate datasets sources, soil moisture and GPP variable choices do not result in any R^2 or RMSE difference (Figure 2a). We

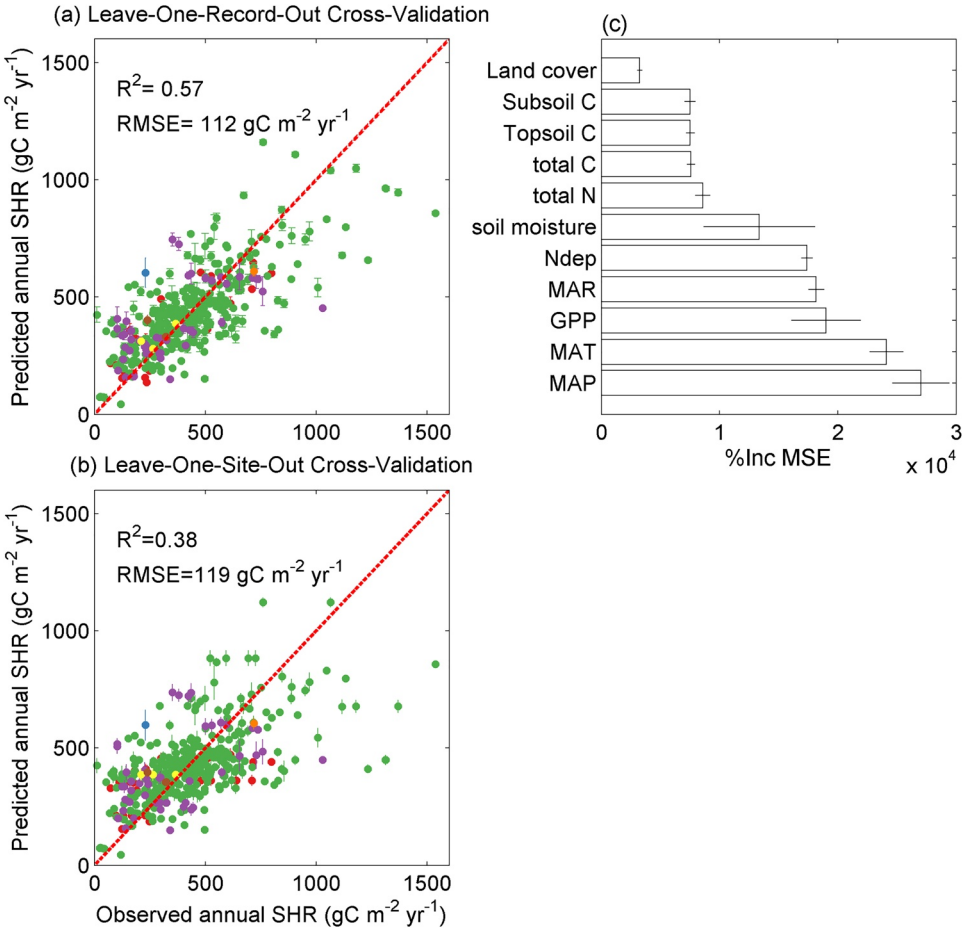

Figure 2. The performance of Random Forest evaluated by (a) Leave-One-Record-Out Cross Validation and (b) Leave-One-Site-Out Cross Validation. The error bar on each point denotes the standard deviation of soil heterotrophic respiration predictions estimated on the basis of different temperature and precipitation, soil moisture and gross primary productivity (GPP) datasets combinations. Different colors show the land cover types of the observation data (same color setting as Figure 1). (c) Feature importance scores ranking deduced from increase of mean squared error (MSE) after permuting variable. MAT (annual temperature), MAP (annual precipitation), MAR (annual radiation), Ndep (nitrogen deposition), GPP, soil moisture are time specific values depending on the observation year.

also try leave-one-site-out evaluation, which resulted in a lower R^2 (Figure 2b). Tests of model performance with more input explanatory variables are shown in Figures S1–S3. We show the importance of different variables in Figure 2c. Annual precipitation and annual temperature are the two most important variables, whereas static variables such as land cover type, soil carbon and nitrogen content contribute relatively less to model performance.

3.2. Spatiotemporal Pattern of Global SHR

The spatial pattern of mean annual SHR at 0.5°×0.5° spatial resolution during 1985–2013 is shown in Figure 3a, and follows the geographic GPP variations to a large degree (see Figure S4). Annual SHR decreases from tropics to high-latitude area, with the highest values in wet tropics of exceeding 800 gC m^{-2} yr^{-1} and the lowest in northern boreal area including Alaska, northern Canada and Siberia area (less than 200 gC m^{-2} yr^{-1}). Such a latitudinal SHR gradient particularly appears in Australia (from coastal area to inland), Africa and South America. Similar spatial gradients appear among all members of our data-driven SHR estimates.

The global mean of our data-driven SHR is 46.8 Pg C yr^{-1} over 1985–2013, (95% confidence interval: 38.6–56.3 Pg C yr^{-1} based on the estimates of 1,000 Random Forest model trees), with an increasing trend of

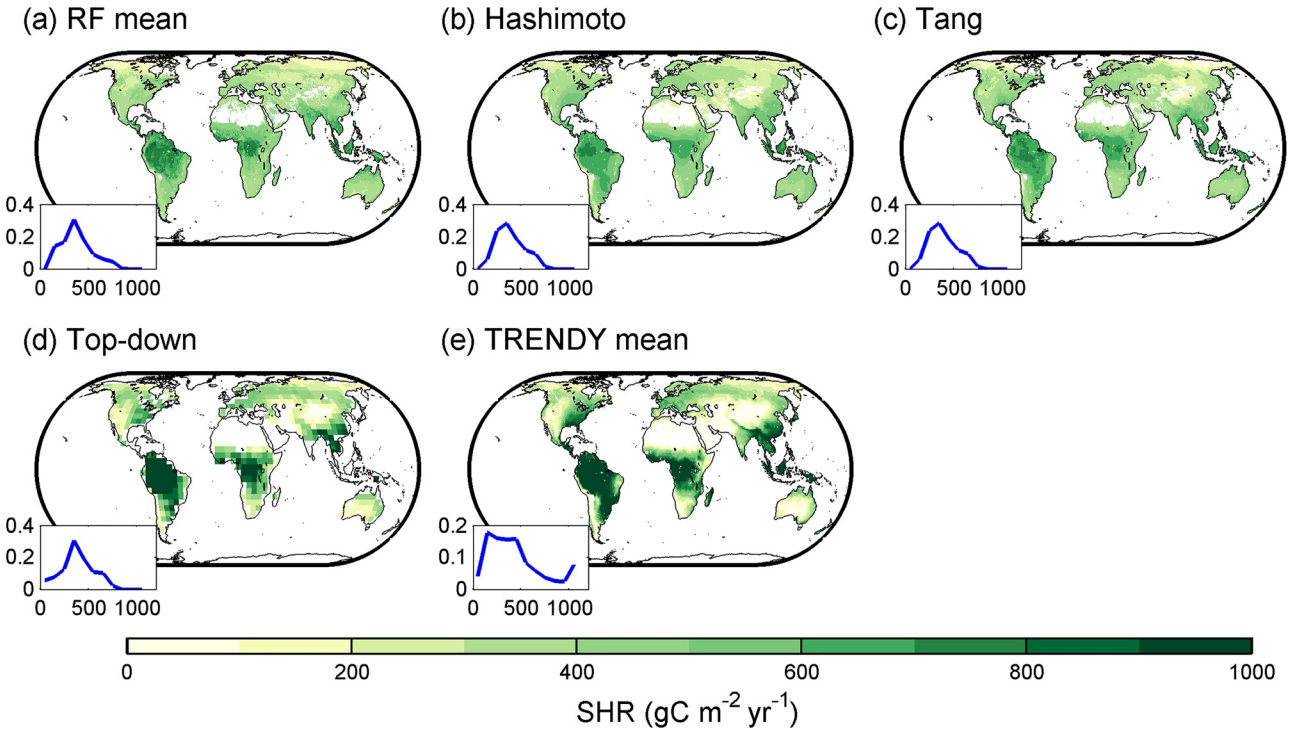

Figure 3. Spatial pattern of (a) ensemble mean of our data-driven soil heterotrophic respiration (SHR) members, (b) Hashimoto et al. (2015), (c) Tang et al. (2020), (d) top-down and (e) TRENDY model ensemble mean. Frequency distribution of SHR in different levels for each SHR data set is shown in the bottom left of each panel. It should be noted that the studied period of Hashimoto product is up to 2012 and the studied period of top-down SHR data set is between 2010 and 2012.

0.03 Pg C yr^{-2} (0.015–0.044 Pg C yr^{-2}, $P < 0.01$), which is similar to the 1.2% increase of SHR detected in Bond-Lamberty et al. (2018) over a similar timespan. The global total SHR is close to global total NPP (48.8 Pg C yr^{-1} during 2001–2015, Table S2) when we applied the same spatial mask. We do expect SHR to be globally smaller than NPP, because of lateral export by ecosystems that reduce considerably the fraction of NPP given to soil as litter for SHR (Ciais et al., 2020).

The choice of soil moisture variables contributes more to the difference among these data-driven SHR members rather than that of GPP, temperature and precipitation data sources (Figure S5). For example, using CPC soil moisture data produces an almost 1 Pg C yr^{-1} lower global SHR value than other data-driven estimates using GLDAS soil moisture (Figure S5c). In spite of these differences, the total amount of all our data-driven SHR members consistently displays a drop in 1992 and a peak in 2010 across all members (Figure S5).

3.3. Comparison With Previous Data Products and Models

Our data-driven SHR estimation is close to that of Hashimoto et al. (2015), who predicted a global flux of 46.5 Pg C yr^{-1} and temporal trend of 0.05 Pg C yr^{-2} (Figure 4). These two sets of gridded SHR products also show similar frequency distributions, although the spatial variation of Hashimoto et al. (2015) as smoother, as SHR in that data set depends only on climate (Figure 3b). Compared to our data-driven estimate, Hashimoto et al. (2015) presented higher SHR values in boreal area and lower values in tropics (Figures S6a, S6b and S7). Conversely, Tang et al. (2020) found lower values in Amazon and central Africa and higher values in temperate areas; their global total SHR amount of 46.9 Pg C yr^{-1} after applying same land mask was similar (Figures S6c and S6d). Global total amount of top-down SHR was about 40.8 Pg C yr^{-1} during 2010–2012 (Konings et al., 2019). The top-down SHR shows more distinct spatial gradients (Figure 3d) and larger values than this study in both tropics and boreal area (Figures S6e and S6f).

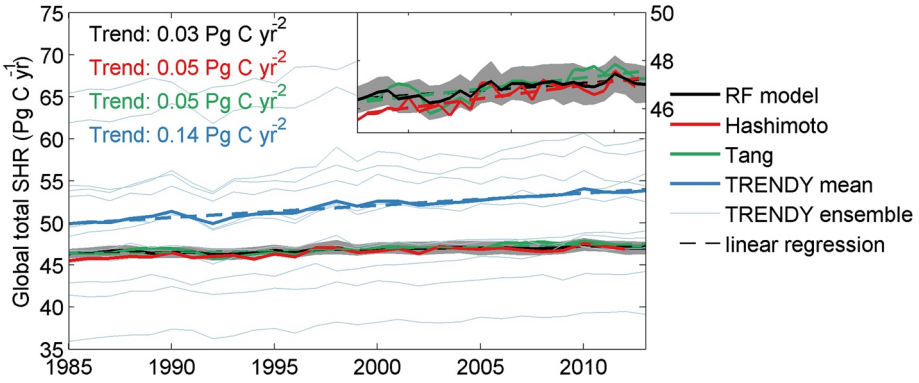

Figure 4. Global total amount of soil heterotrophic respiration (SHR) over 1985–2013. The shaded area in gray indicates the spread over all members from Random Forest models (RF). The zoomed plot in the top right shows the global mean of our data-driven ("RF"), Tang and Hashimoto SHR estimates more clearly. The dashed lines refer to fitted global total SHR time series against time. The global total amount of SHR in each TRENDY model is shown as thin blue line.

The ensemble mean SHR over all TRENDY models is larger than empirically estimated SHR products, increasing from 49.9 Pg C yr^{-1} to 53.8 Pg C yr^{-1} during our studied period, due to a four times larger temporal trend (0.14 Pg C yr^{-2} vs. 0.03 Pg C yr^{-2} in our data-driven products) (Figure 4). The TRENDY multi-model ensemble mean shows more drastic geographic contrasts across the globe (Figure 3e), with a different frequency distribution of SHR from our data-oriented estimate (Figures 3a and 3e). Large inter-model discrepancies are present among TRENDY models both for global totals and spatial details; the standard deviation can be as high as 8–9 Pg C yr^{-1} in each year and both underestimation and overestimation of SHR appear in different models in comparison to our data-oriented products (Figures S6g, S6h and s8–s9).

With respect to the spatial distribution of temporal trend of SHR, we can find that there is no uniform trend in our data-driven products (Figure S10a). To understand the controlling variables for the regional variation in SHR trend map, here we ran factorial estimation for one SHR member only for efficiency (see Section 2.9). Through comparison between Figures S11 and S12 we can see that change in soil moisture dominates the negative SHR trend in South America and central Africa, as well as the positive trend in Arctic tundra, and the increase in atmospheric nitrogen deposition contributes to patterns of positive SHR trend in Asia near urbanized areas. Tang et al. (2020) found clearly positive SHR trends in Arctic tundra areas (Figure S10c). Hashimoto et al. (2015) and TRENDY ensemble mean also produced positive trends in most areas, matching their increasing global trends (Figures 4, S10b, and S10d).

3.4. SHR Anomalies in Relation to Meteorological Factors

We attribute the factors contributing to SHR IAV using a linear decomposition approach described in Section 2.8. We first verified that the SHR anomaly reconstructed with climate factors can correctly reproduce the detrended SHR time series spatially (evaluated by the correlation coefficients as shown in Figure S13). It should be noted that both soil moisture and precipitation are used in upscaling of SHR but only one water proxy is used in the reconstruction of SHR anomalies. The quality of the reconstruction is impacted by the choice of water-related variables in linear decomposition process. For example, when CPC soil moisture is employed in the up-scaling of SHR, including it allows for better reconstruction of the SHR IAV than when using precipitation (Figure S13). Conversely, when using GLDAS soil moisture as input, precipitation has better performance in reconstructing SHR anomaly than the soil moisture variable. When the reconstructed anomalies are integrated to the global scale, we find that the correlation coefficient between reconstructed and detrended SHR anomalies can reach 0.9 when annual precipitation is used as a predictor, and 0.94 when soil moisture is used.

In general, we can see from Figure 5 that the globally integrated SHR IAV is mainly controlled by water availability, with higher correlation coefficients between reconstructed SHR and the water related SHR anomalies (SHR^{PREC} or SHR^{SMC}). However, the strength of this control varies depending on the water

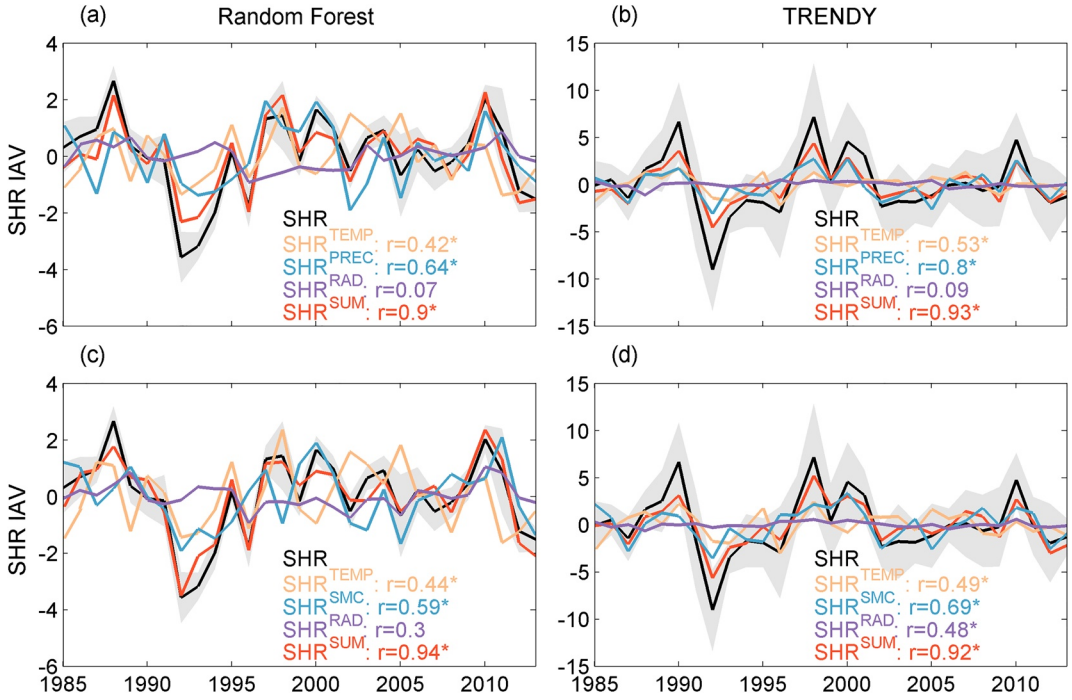

Figure 5. Climatic variables controls on soil heterotrophic respiration (SHR) inter-annual variability (IAV) at global scale. The black lines indicate detrended SHR anomalies. The colored lines indicate SHR anomalies driven by different climate factors. 'SUM' means the sum of SHR IAV driven by three climatic factors. The top panels include regressions using precipitation as a water proxy; in the bottom panels, the water proxy is soil moisture. The correlation coefficient between SHR anomaly and the component driven by a forcing is labeled. An asterisk denotes the significance of the corresponding correlation coefficient ($P < 0.05$). Shaded areas represent the spread among ensembles of Random Forest or TRENDY products. Due to the difference between soil moisture and water storage variables, only SHR estimated using soil moisture (CPC and GLDAS) is shown here. Regression for SHR estimated from GRACE-rec TWS is shown in Figure S14.

availability variables that are used in the SHR estimation and subsequent regression. For SHR up-scaled using GRACE-rec TWS, the correlation coefficient between SHR^{SMC} and SHR IAV is two times higher than when using precipitation as a regressor, although the overall reconstruction efficiency does not differ greatly (Figure S14). This should be expected since TWS anomaly is not always comparable to soil moisture and precipitation. Specifically, there are two most prominent anomalies in all data-driven SHR estimation members in 1992 and 2010 consistently (Figure S15). According to the similarity between SHR IAV driven by each climatic factor and the detrended series locally, we found that the negative SHR anomaly in 1992 is particularly driven by water-related proxies under cooler and drier climate after the Mount Pinatubo eruption (Figures S16–S18). The contribution of precipitation variability to the positive SHR anomaly in 2010 is larger than the one from temperature anomaly (Figure 5). We also identified the climatic drivers for the Tang et al. (2020) SHR product, and found that, similarly, precipitation or soil moisture anomaly dominate its SHR IAV (Figure S19).

We analyzed the drivers for SHR IAV from TRENDY model ensembles using the same approach. For TRENDY models, their simulated soil moisture is also used in the regression. In general, climate variables had a good capacity for reconstructing TRENDY SHR anomalies (Figures S20 and S21). SHR IAV from TRENDY ensemble mean displays dominance by water availability at global scale (Figures 5b and 5d). The driving factors for SHR IAV differ widely among TRENDY models (Figure S22), with six models showing dominance of water availability fluctuation on SHR IAV at the global scale, and five exhibiting dominance by temperature. The TRENDY models capture the negative SHR anomaly in 1992 and precipitation variability can better account for that. With regard to the distinctly positive anomaly in 2010, TRENDY ensemble mean also shows an apparent dominance of precipitation variability (Figure 5b). We repeated the analysis

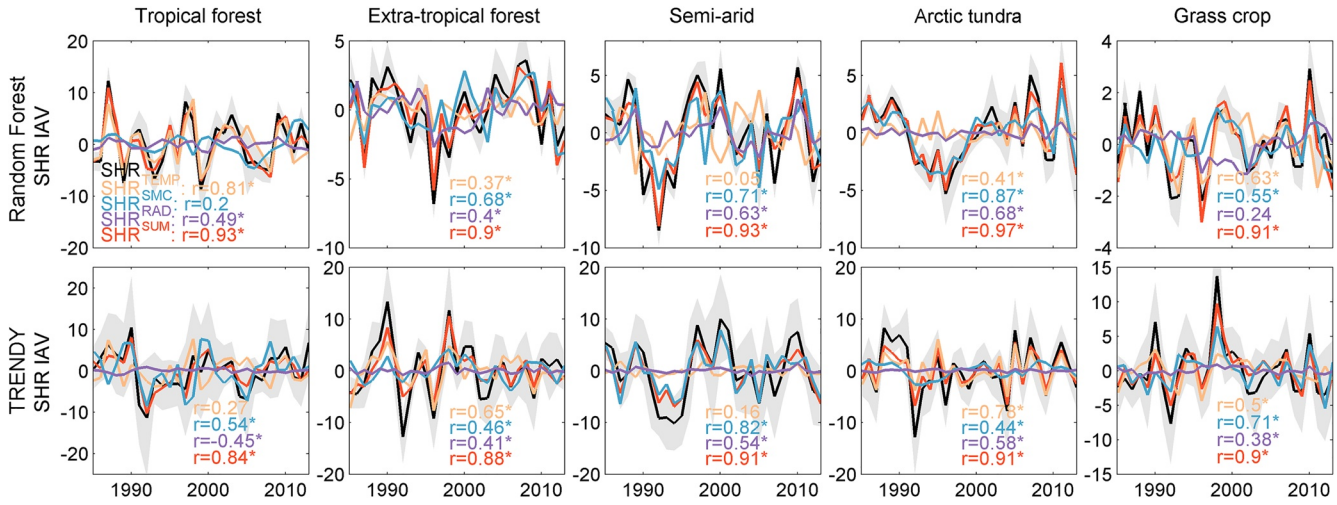

Figure 6. Environmental controls on soil heterotrophic respiration (SHR) inter-annual variability (IAV) in different ecosystems. Panels in top row are derived using our data-driven SHR products. Panels in the bottom row are derived using TRENDY multi-ensemble mean. The correlation coefficients between SHR IAV and SHR anomalies driven by a given factor are labeled in each panel. An asterisk denotes the significance of the corresponding correlation coefficient ($P < 0.05$). The decomposition using precipitation anomaly is shown in Figure S25.

on the Hashimoto data set and we show the result in the supporting information section of this paper (Text S1).

3.5. Biome-Scale SHR Anomalies Attribution

To check whether the dominance of water availability on SHR IAV differs between biome types and regions, we aggregate SHR anomalies to tropical forest, extra-tropical forest, semi-arid region, Arctic tundra, grass and crop area (Figure S23, adapted from Ahlström et al. (2015)). To confirm whether the uncertainty affects the relationship between SHR and climatic factors, here we show the pattern of correlation between SHR anomaly driven by each climatic factor and SHR anomaly in Figure S24. We can find that precipitation and soil moisture significantly control the SHR IAV in Arctic tundra, semi-arid regions and extra-tropical forests. In tropical forests, as well as grass and cropland, fluctuations in both temperature and precipitation show significant effects on SHR IAV in more than half of our SHR ensembles (63 of 126 members). Such a statistically significant response to climatic factors across multiple members allows us to analyze the regional drivers of SHR IAV.

In tropical forests, SHR IAV is mainly driven by temperature variability, whereas in extra-tropical forests and semi-arid areas, precipitation variability plays a dominant role (Figure S25). Precipitation variability is comparable to temperature in driving SHR anomalies in grass and crop areas. In Arctic tundra, the dominant factor for SHR IAV is the fluctuation in soil water content, regardless of the moisture variable being used, while fluctuation in temperature seems to be less important (Figures 6, S25, and S26). When considering regression against soil moisture, we notice that the ranking of factors is shuffled to some extent (Figures 6 and S25), which may be due to the difference between variability in precipitation and soil moisture regionally (Figures S17 and S18). For example, in extra-tropical areas, the importance of water availability increases greatly when regression is carried out against soil moisture (CPC or GLDAS, Figures 6 and S25). For the Tang data set, we found that a dominance of precipitation or soil moisture also appears in tropical and extra-tropical forests, semi-arid areas, and Arctic tundra (Figure S27). In grass and cropland, temperature is also important for SHR IAV regionally (Figure S27). Our attribution of climatic drivers for SHR IAV differs from Tang et al. (2020) mainly in tropical forests (temperature is dominant in our study vs. precipitation in Tang et al., 2020).

Attribution of the TRENDY ensemble shows consistent water-driven estimation in semi-arid areas but differs from our data-driven results mainly in the tropical forest (Figure 6): in these regions, TRENDY SHR IAV is mainly driven by precipitation or soil moisture variability rather than temperature variability (Figures 6

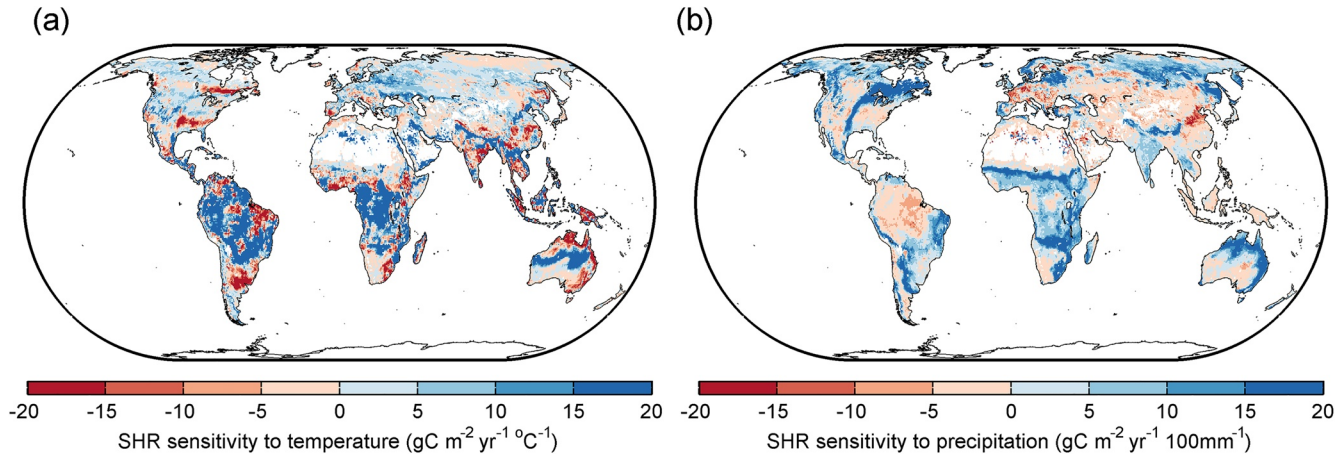

Figure 7. Spatial distribution of the sensitivities of our data-oriented soil heterotrophic respiration (SHR) to anomalies in (a) temperature and (b) precipitation. A negative temperature sensitivity means that when temperature increases, SHR decreases.

and S28). In extra-tropical forest and Arctic tundra, the controlling effect of temperature on TRENDY SHR anomaly is also non-negligible compared to that of water availability (Figure 6). In grass and crop area, TRENDY ensemble mean shows that both precipitation and soil water effects outbalance temperature effects, which does not perfectly match our estimation. Furthermore, it also should be noted that the different TRENDY models disagree on the attribution of SHR anomalies between water versus temperature across different ecosystems (Figures S28 and S29).

During the reconstruction process, we also computed the distinct sensitivities of SHR to climate factors, especially temperature and water availability (Figures 7 and S30). Regional heterogeneity of the sensitivity of SHR to temperature and precipitation/ soil moisture may alter the overall trajectory of SHR because of coexistence of regional deceleration and acceleration of SHR and/or compensatory effects of temperature and precipitation. For our data-driven SHR members, for example, in the tropics, positive temperature sensitivity in Amazon, central Africa, as well as negative sensitivity to precipitation in these areas lead to opposite temperature and precipitation driven SHR anomalies in the cooler and drier year of 1992. In Siberia, negative SHR sensitivity to temperature and positive sensitivity to precipitation contribute to additive climate driven sub-components in IAV under warm but dry condition. The sensitivity of SHR to soil moisture depends on the data set used, especially in boreal regions and central Africa (Figure S30). More negative sensitivity to soil moisture availability was found in wet tropics areas when using GLDAS soil moisture (0–2.0 m) as a predictor than when using CPC (0–1.6 m), indicating that soil moisture increases in deeper layers can be more negatively related to SHR than changes in shallow layers if we assume little difference of soil moisture in their common depth intervals.

4. Discussion

4.1. Implications and Future Directions of Data-Driven SHR estimation

In this study, we generated a newly up-scaled SHR data set using a Random Forest algorithm and explanatory variables of climatological indicators, GPP as well as soil properties. We carefully evaluated the model performance with cross-validation and assessed the estimated uncertainty, which has generally not been done in previous studies. Our products can be utilized to evaluate the sensitivities of process-based models and constrain their performances in both spatial and temporal scales (Ichii et al., 2017).

Our results exhibited both similarities and differences compared to results from previous analyses. The similarities in spatial distribution with Hashimoto et al. (2015) could be due to two reasons. First, both studies used observation records from SRDB, which ensured their similar SHR range (although a more recent version of SRDB, with much more data, was used here). Second, both two studies are vulnerable to uneven sampling and potential extrapolation problem to some extent. In terms of the Tang SHR data, both differences in explanatory variables selection and sample data set affect regional SHR variation between Tang

et al. (2020) and our SHR datasets, and the inclusion of records with annual SHR above 1,100 gC m^{-2} yr^{-1} influences the magnitude of SHR in tropical regions (Figures S31a and S31d). When we exclude GPP from explanatory variables sets, SHR differs especially in central Africa and India, where inclusion of GPP helps produce higher SHR in central Africa and lower SHR in India (Figures S31a and S31c). It also should be noted that Tang et al. (2020) included more sample data in China, which also contributes to the difference of their SHR estimation in temperate climate regions like China and Europe. RMSE from LOOCV was 20% smaller in our study (112 gC m^{-2} yr^{-1}) compared to Tang et al. (2020), with RMSE of 143 gC m^{-2} yr^{-1}.

We found that TRENDY models underestimate boreal region SHR, and overestimate SHR especially in tropics, with a conspicuous spatial contrast along latitudes (Figure S8). A misrepresentation of nitrogen constraints, and oversimplified treatment of processes like microbial dynamics and the climatological dependence of decomposition in different soil texture types, may be responsible for this tendency to produce unrealistic SHR values compared to observations, in turn suggesting areas for focused model research and improvement (Shao et al., 2013; Yan et al., 2018). Our newly derived estimation can thus help improve model ability to accurately predict response of soil carbon to future climate change scenarios.

Nevertheless, limitations undoubtedly remain in our data-driven products, which we suggest can be split into uncertainties in the (a) underlying soil respiration observation data, (b) the driving climate and GPP gridded data, and (c) the up-scaling process. First of all, the uncertainty introduced by partitioning of autotrophic and heterotrophic (i.e., SHR) components of the measured soil respiration flux is poorly understood but potentially large. Approaches such as isotope labeling bring less disturbance to the root-soil system, but introduce their own uncertainties from for example, mixing model assumptions. Root extraction methods neglect the contribution of priming mechanism (Kuzyakov & Larionova, 2005), the amplified effect of which is ~12% in permafrost ecosystem (Keuper et al., 2020). Trenching approaches feature significant disturbance but subsequent simpler inference. Better partitioning will reduce uncertainties entered into the estimation model (Carbone et al., 2016), and we suggest that it is the right time for a meta-analysis examining the potential biases of these different approaches.

Second, observation data of SHR is provided at coarse temporal resolution of a year. Our upscaled SHR data set thus incorporates responses to environmental variations only annually, which mask different sensitivities (possibly different signs of sensitivities as well) at the seasonal scales (Shao et al., 2013). More SHR observations at daily, monthly, and seasonal measurements will help us understand more of its responses to varying environmental conditions like seasonal lagged responses (Vargas et al., 2010) and also benefit model evaluations.

Third, the biased sampling of observations at the global scale (Schimel et al., 2015; Xu & Shang, 2016) forces up-scaling approaches to extrapolate to under-sampled area, affecting the accuracy of model prediction. For example, the tropics are predicted to have the largest SHR fluxes, but are greatly under-sampled, leading to the largest uncertainty in our data-driven SHR data set. The climate space sampled by all filtered observations ranges from 11.5 to 5,302 mm in mean annual precipitation and from −10 to 31.5°C in mean annual temperature. Relatively few reported data exist from cold areas, warm dry areas, and warm humid areas (Figure S32). The space of GPP, soil moisture, soil carbon content and annual nitrogen deposition under-sampled by measurement records are shown in Figure S33. The spread of 1,000 trees in Random Forest can reflect the uncertainty from such extrapolation to under-sampled domain, from which the 95% CI is 38.6–56.3 Pg C yr^{-1}. This uncertainty measure has also been used in J. Zeng et al. (2020) for net ecosystem productivity and Warner et al. (2019) as well as Stell et al. (2021) for soil respiration. Through the comparison of two sources of uncertainties (alternative gridded explanatory data and uncertainty from extrapolation to under-sampled domain by spread of trees), we find that uncertainty from individual trees in Random Forest model is far larger than that from different explanatory variables datasets. Our similarity with Hashimoto data set in spatial gradient could also be partly explained by the under-sampling issue in both studies. In the Hashimoto work, the two globally fixed parameters to derive SHR from soil respiration can possibly narrow the real range of SHR compared to broader coverage of soil respiration point data in climate space. Therefore, the priority should be more data collection in under-sampled areas to constrain the current wider prediction distributions, as an optimized network design has been demonstrated to decrease uncertainty in global estimates (Stell et al., 2021).

In addition, unlike eddy-covariance records extending more than 20 years, most available SHR observations are of short duration (the longest continuous observation lasted 6 years in SRDB 4.0). The "space substitution for time" concept is common under such circumstances but should be treated with caution since the prediction accuracy of such substitution requires consistency between temporal and spatial variation in climate space (Blois et al., 2013). Therefore, we call for more valid sub-sampling, which can be indispensable to fill the gap between spatial and temporal climatic gradients. Further data compilation and integration are extremely valuable for upscaling and improving performances of SHR estimation models (Bond-Lamberty, 2018; Harden et al., 2018).

With regard to uncertainty emerged during the up-scaling process, it is worth noting that some variables are poorly represented within our procedure. For example, land cover is found to contribute the least to change in model error as shown in Figure 2c. Although our way of handling this categorical variable, one-hot encoding (Lantz, 2013), is good at tackling discrete variables, land cover's effect may covary with other environmental gradients and climatic proxies. In addition, climatological and other physiological inputs are less certain, especially in the tropics. Besides, some key explanatory variables remain unavailable currently, for example, other nutrients like phosphorus availability, microbial activity, mycorrhizal types (Crowther et al., 2019), as well as disturbance information including land use/ cover change (LUCC). The uncertainties in inputs and the lack of consideration of other potential factors can be responsible for the underestimation in prediction for observed SHR exceeding 1000 gC m^{-2} yr^{-1} in tropics (Figure 2a). Disturbances affect how SHR varies over vegetation succession and then lead to changes of ecosystem soil carbon pools (Harmon et al., 2011), but limited mapping of this information, and the highly uncertain consequences of SHR responses to LUCC, impede our careful consideration of its effect. The treatment of disturbance history has been attempted in the case of taking stand age to account for disturbance effects on forest carbon dynamics (Xiao et al., 2014), and a similar approach could be used to extend our estimation framework toward better SHR prediction.

Recently, scientists have started to explore the application of deep-learning in data-driven earth system science (Reichstein et al., 2019). Recurrent Neural Network (RNN) may have great potential in modeling dynamic time series like net ecosystem productivity, with advantages over other regression methods in considering legacy effects or lagged indicators (Reichstein et al., 2018). Considering that time-lag responses of SHR to climate anomalies are important in evaluating terrestrial carbon cycle feedbacks to climate warming (X. Zhou et al., 2010), and photosynthesis can also influence soil respiration with hysteresis regionally (Kuzyakov & Gavrichkova, 2010), deep learning frameworks like RNN are ideal tools for improving our SHR estimation. Given the large demand for observation samples in deep learning neural networks, we highlight again that more SHR observations should be implemented and compiled. Since eddy covariance data provides ecosystem respiration records (auto-plus heterotrophic respiration) as well, a reasonable model as partition tool from autotrophic parts can also be an ideal solution for collecting more available SHR data (Koerber et al., 2016).

4.2. Climatic Drivers of SHR Anomalies

Since many studies emphasize the controls of water availability on ecosystem carbon fluxes (Humphrey et al., 2018; Jung et al., 2017), it is not surprising that we found that global SHR IAV is mainly controlled by the fluctuation in water condition. The dominance of water variation exists, yet the sensitivity of SHR IAV to water depends on the choice of proxy for water content (using soil moisture or precipitation in linear decomposition of SHR anomalies). Similar findings have been identified in Ballantyne et al. (2017) and Yan et al. (2018). This can be expected since SHR depends on soil water content (Skopp et al., 1990), although this relationship can be influenced by soil property (Moyano et al., 2012, 2013), microbial diversity (Zhang & Zhang, 2016), historical rainfall condition (Hawkes et al., 2017), and background water condition (Hinko-Najera et al., 2015; Matteucci et al., 2015). In addition, water content also affects the microbial community composition (Zhao et al., 2016), substrates availability, activities of extracellular enzymes (Schindlbacher et al., 2012), and even temperature sensitivity to SHR (Suseela et al., 2012).

Regionally speaking, our result of dominant water controls of SHR in extra-tropical forest and semi-arid regions are consistent with plot-scale experimental tests (Hursh et al., 2017; W. Liu et al., 2009). In semi-arid areas, a small response to temperature fluctuation is possibly restrained by soil moisture and substrate

availability (N. Zhang et al., 2013). As for Arctic tundra, such colder high-latitude areas can be more responsive to warming (Carey et al., 2016) and warmer temperature can cause shifts from heterotrophic to autotrophic respiration (Hicks Pries et al., 2015). Nevertheless, our diagnosed soil moisture control on tundra SHR IAV is not unrealistic, as some researchers have also reported that soil moisture affects ground thaw and the magnitudes of carbon loss is driven by soil moisture (Natali et al., 2015). There are also few records from Arctic tundra, meaning that incorrect spatial extrapolation might explain this particular positive response of tundra SHR to water availability.

Given that large-scale droughts will likely to happen by the end of the 21st century (Lu et al., 2019), our examined response of SHR to variability in water condition is a first step toward making reliable projections of soil carbon loss. However, we only consider the nonlag (yearly) response of SHR to environmental variability in current year; in the real world, lagged or adapted responses are non-negligible (Arnone Iii et al., 2008; de Nijs et al., 2019; Göransson et al., 2013). As noted above, applying deep learning techniques such as RNN (Kraft et al., 2019) hold great promise to robustly deal with such dynamics.

Despite the prevalent positive water controlling effects of SHR in most ecosystems, we also note that our ecosystem-level IAV attribution found that temperature anomalies dominate tropical forest SHR variability. Warming stimulating soil respiration has been demonstrated in many previous studies (Hursh et al., 2017; O'Connell et al., 2018), although thermal acclimation of heterotrophic microbes occurs (Carey et al., 2016; Crowther & Bradford, 2013). Since warming could also lead to soil water loss, primary positive effects of temperature mainly distribute in non-water-limited areas. Tropical forests account for the largest fraction of global SHR, and current tropical temperatures seem still lower than the physiological optima for respiration (M. Huang et al., 2019; Liu, He, et al., 2018). In other words, a stimulation effect from warming to SHR still exists, which may greatly affect the local soil-atmospheric carbon fluxes fluctuations and place tropical soil carbon storage at risk given the positive sensitivity to temperature. In addition, any interplay among different factors is not accounted for in our multiple linear regression formula. Therefore, further research is required to clarify the interaction effects between two associated factors of temperature and water availability.

The responses of SHR from TRENDY models to climatic variability differ among models at biome and global scales (Figures S22 and S28). This could be due to different specific formula forms of SHR parameterizations with climatic factors among models, like Arrhenius, hill or monotonic type (Todd-Brown et al., 2013) and to different degrees of couplings between productivity and respiration. For example, CABLE, ISAM and VEGAS exhibit dominant temperature controls on SHR IAV, which are different from the water effects identified in nearly all data-driven products in our study. The underestimation of water effects on year-to-year fluctuations of net carbon fluxes in process-based models has already been highlighted (Green et al., 2017; Humphrey et al., 2018; Liu, Ballantyne, et al., 2018). Therefore, the internal parameterization schemes should be rigorously constrained and improved with the aid of causal perspectives in sensitivities of SHR to climate variability from our data-driven estimation. Using such causal statistics as a benchmark can better constrain the modeled carbon cycle feedback (Claessen et al., 2019), reducing the uncertainties in climate projections (Friedlingstein et al., 2014).

4.3. Effects of Other Environmental Factors on SHR

Apart from the climate variables that have been used in our attribution analyses, other explanatory variables like GPP (affected by climate) and nitrogen availability are non-negligible factors affecting SHR through substrate availability and/ or microbial activity (Figure 2c). The spatial gradient reflected in our data-driven SHR resembles that of GPP. This is expected, as decomposition of fresh or recent organic matter like leaves and fine roots contributes to the main component of SHR (Janssens et al., 2001). However, the association between productivity and substrate availability is not constant or always predictable (Peterson & Lajtha, 2013), and the dependence of SHR on productivity can be confounded by temperature effects (P. Li et al., 2013).

We also found that nitrogen deposition influences the change in mean squared error of predicted SHR (Figure 2c). Nitrogen content may exert a major control on soil microbial activity (Janssens et al., 2010), indirectly expressed as soil acidification (Wang et al., 2019), limitation on substrate sources supply through harm on specific enzymes (Y. Li et al., 2015), and functional changes in the microbial community (Allison

et al., 2008). Higher carbon to nitrogen ratio in soil organic matter decreases decomposer carbon use efficiency and often impedes decomposition (Manzoni et al., 2017; H. Zhang et al., 2018), but not always (Bowden et al., 2004; Guo et al., 2017; Olsson et al., 2005). Controversy still exists on the stimulatory or suppressive effects of nitrogen addition on SHR (Z. Chen et al., 2018), which also vary depending on original nitrogen availability (X. Liu et al., 2017; Sun et al., 2014), the amount of addition (Gao et al., 2014) and even ecosystem types (Cusack et al., 2010; Maaroufi et al., 2019; Mo et al., 2008; L. Zhou et al., 2014). Since any SHR reduction caused by nitrogen addition could reach the same order of magnitude of forest carbon sink (Janssens et al., 2010), further fundamental research is needed to understand how SHR responds to nitrogen deposition in different ecosystems, and whether possible negative effect on SHR could offset the warming induced increase in carbon loss, to better predict SHR changes given increasing nitrogen deposition in the future.

5. Conclusion

Our study integrates a large number of in-situ SHR measurements, satellite and meteorological observations using Random Forest models to produce an ensemble of data-driven global SHR products. This ensemble data set is independent of process-based model outputs and is expected to be beneficial for the model sensitivity parameter calibration. Our attribution analysis provides evidence for an important role of water availability in impacting year-to-year fluctuations of carbon fluxes, although of which can be mediated to some extent by choice of water content proxies in attribution process. To reduce SHR uncertainty and advance our ability to diagnose the state of SHR, we argue that more evenly distributed SHR observations—especially from cold, warm dry and warm humid areas—and more powerful deep learning methods should be considered in further global SHR mapping tasks.

Data Availability Statement

The detailed R codes for producing the soil heterotrophic respiration data can be obtained from http://doi.org/10.5281/zenodo.4321758. Soil heterotrophic respiration data from 1985 to 2013 (all 126 combinations using Random Forest) are available at https://doi.org/10.6084/m9.figshare.11340770.

Acknowledgments
This work was financially supported by the CLAND Convergence Institute funded by ANR (16-CONV-0003). Y. Yao also acknowledges support from Make Our Planet Great Again (MOPGA) Scholarship.

References

Adachi, M., Ito, A., Yonemura, S., & Takeuchi, W. (2017). Estimation of global soil respiration by accounting for land-use changes derived from remote sensing data. *Journal of Environmental Management*, *200*, 97–104. https://doi.org/10.1016/j.jenvman.2017.05.076

Ahlström, A., Raupach, M. R., Schurgers, G., Smith, B., Arneth, A., Jung, M., et al. (2015). The dominant role of semi-arid ecosystems in the trend and variability of the land CO_2 sink. *Science*, *348*, 895–899. https://doi.org/10.1126/science.aaa1668

Allison, S. D., Czimczik, C. I., & Treseder, K. K. (2008). Microbial activity and soil respiration under nitrogen addition in Alaskan boreal forest. *Global Change Biology*, *14*, 1156–1168. https://doi.org/10.1111/j.1365-2486.2008.01549.x

Arnone Iii, J. A., Verburg, P. S., Johnson, D. W., Larsen, J. D., Jasoni, R. L., Lucchesi, A. J., et al. (2008). Prolonged suppression of ecosystem carbon dioxide uptake after an anomalously warm year. *Nature*, *455*, 383–386. https://doi.org/10.1038/nature07296

Baccini, A., Goetz, S., Walker, W., Laporte, N. T., Sun, M., Sulla-Menashe, D., et al. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, *2*, 182–185. https://doi.org/10.1038/nclimate1354

Ballantyne, A., Smith, W., Anderegg, W., Kauppi, P., Sarmiento, J., Tans, P., et al. (2017). Accelerating net terrestrial carbon uptake during the warming hiatus due to reduced respiration. *Nature Climate Change*, *7*, 148–152. https://doi.org/10.1038/nclimate3204

Beaudoing, H., & Rodell, M. (2016). GLDAS Noah Land Surface Model L4 monthly 0.25 x 0.25 degree V2. 1. In *NASA/GSFC/HSL: Greenbelt, Maryland, USA*. Goddard Earth Sciences Data and Information Services Center (GES DISC).

Blois, J. L., Williams, J. W., Fitzpatrick, M. C., Jackson, S. T., & Ferrier, S. (2013). Space can substitute for time in predicting climate-change effects on biodiversity. *Proceedings of the National Academy of Science of the United States of America*, *110*, 9374–9379. https://doi.org/10.1073/pnas.1220228110

Bond-Lamberty, B. (2018). New techniques and data for understanding the global soil respiration flux. *Earth's Future*, *6*, 1176–1180. https://doi.org/10.1029/2018EF000866

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

Bond-Lamberty, B., Epron, D., Harden, J., Harmon, M. E., Hoffman, F., Kumar, J., et al. (2016). Estimating heterotrophic respiration at large scales: Challenges, approaches, and next steps. *Ecosphere*, *7*, e01380. https://doi.org/10.1002/ecs2.1380

Bond-Lamberty, B., & Thomson, A. (2010). A global database of soil respiration data. *Biogeosciences*, *7*, 1915–1926. https://doi.org/10.5194/bg-7-1915-2010

Bond-Lamberty, B., & Thomson, A. (2018). *A Global Database of Soil Respiration Data*. Version 4.0. ORNL DAAC.

Bond-Lamberty, B., Wang, C., & Gower, S. T. (2004). A global relationship between the heterotrophic and autotrophic components of soil respiration? *Global Change Biology*, *10*, 1756–1766.

Borchers, H. W. (2019). *Pracma: Practical numerical math functions*. R package version 2.2.5. 8 April 2019. Retrieved from https://cran.r-project.org/web/packages/pracma/

Bowden, R. D., Davidson, E., Savage, K., Arabia, C., & Steudler, P. (2004). Chronic nitrogen additions reduce total soil respiration and microbial respiration in temperate forest soils at the Harvard Forest. *Forest Ecology and Management*, *196*, 43–56. https://doi.org/10.1016/j.foreco.2004.03.011

Bradford, M. A., McCulley, R. L., Crowther, T. W., Oldfield, E. E., Wood, S. A., & Fierer, N. (2019). Cross-biome patterns in soil microbial respiration predictable from evolutionary theory on thermal adaptation. *Nature Ecology & Evolution*, *3*, 223–231. https://doi.org/10.1038/s41559-018-0771-4

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5–32. https://doi.org/10.1023/a:1010933404324

Buermann, W., Forkel, M., O'Sullivan, M., Sitch, S., Friedlingstein, P., Haverd, V., et al. (2018). Widespread seasonal compensation effects of spring warming on northern plant productivity. *Nature*, *562*, 110–114. https://doi.org/10.1038/s41586-018-0555-7

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., et al. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricutural and Forest Meteorology*, *274*, 144–159. https://doi.org/10.1016/j.agrformet.2019.03.010

Carbone, M. S., Richardson, A. D., Chen, M., Davidson, E. A., Hughes, H., Savage, K. E., & Hollinger, D. Y. (2016). Constrained partitioning of autotrophic and heterotrophic respiration reduces model uncertainties of forest ecosystem carbon fluxes but not stocks. *Journal of Geophysical Research: Biogeosciences*, *121*, 2476–2492. https://doi.org/10.1002/2016jg003386

Carey, J. C., Tang, J., Templer, P. H., Kroeger, K. D., Crowther, T. W., Burton, A. J., et al. (2016). Temperature response of soil respiration largely unaltered with experimental warming. *Proceedings of the National Academy of Science of the United States of America*, *113*, 13797–13802. https://doi.org/10.1073/pnas.1605365113

Chen, M., Xie, P., Janowiak, J. E., & Arkin, P. A. (2002). Global land precipitation: A 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology*, *3*, 249–266. https://doi.org/10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2

Chen, Z., Xu, Y., He, Y., Zhou, X., Fan, J., Yu, H., & Ding, W. (2018). Nitrogen fertilization stimulated soil heterotrophic but not autotrophic respiration in cropland soils: A greater role of organic over inorganic fertilizer. *Soil Biology and Biochemistry*, *116*, 253–264. https://doi.org/10.1016/j.soilbio.2017.10.029

Ciais, P., Yao, Y., Gasser, T., Alessandro, B., Yilong, W., Ronny, L., et al. (2020). Empirical estimates of regional carbon budgets imply reduced global soil heterotrophic respiration. *National Science Review*, *8*. https://doi.org/10.1093/nsr/nwaa145

Claessen, J., Molini, A., Martens, B., Detto, M., Demuzere, M., & Miralles, D. (2019). Global biosphere–climate interaction: A multi-scale appraisal of observations and models. *Biogeosciences Discussions*. https://doi.org/10.5194/bg-2019-212

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., et al. (2011). The Joint UK Land Environment Simulator (JULES), model description–Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, *4*, 701–722. https://doi.org/10.5194/gmd-4-701-2011

Crowther, T., Van Den Hoogen, J., Wan, J., Mayes, M. A., Keiser, A. D., Mo, L., et al. (2019). The global soil community and its influence on biogeochemistry. *Science*, *365*, eaav0550. https://doi.org/10.1126/science.aav0550

Crowther, T. W., & Bradford, M. A. (2013). Thermal acclimation in widespread heterotrophic soil microbes. *Ecology Letters*, *16*, 469–477. https://doi.org/10.1111/ele.12069

Cusack, D. F., Torn, M. S., McDOWELL, W. H., & Silver, W. L. (2010). The response of heterotrophic activity and carbon cycling to nitrogen additions and warming in two tropical soils. *Global Change Biology*, *16*, 2555–2572. https://doi.org/10.1111/j.1365-2486.2009.02131.x

de Nijs, E. A., Hicks, L. C., Leizeaga, A., Tietema, A., & Rousk, J. (2019). Soil microbial moisture dependences and responses to drying–rewetting: The legacy of 18 years drought. *Global Change Biology*, *25*, 1005–1015. https://doi.org/10.1111/gcb.14508

Fan, Y., & van den Dool, H. (2004). Climate Prediction Center global monthly soil moisture data set at 0.5 resolution for 1948 to present. *Journal of Geophysical Research*, *109*, D10102. https://doi.org/10.1029/2003JD004345

Fan, Y., & Van den Dool, H. (2008). A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research*, *113*, D01103. https://doi.org/10.1029/2007JD008470

Feng, P., Wang, B., Li Liu, D., Waters, C., & Yu, Q. (2019). Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. *Agricultural and Forest Meteorology*, *275*, 100–113. https://doi.org/10.1016/j.agrformet.2019.05.018

Friedl, M. A., Sulla-Menashe, P., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., & Huang, X. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, *114*, 168–182. https://doi.org/10.1016/j.rse.2009.08.016

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, *27*, 511–526. https://doi.org/10.1175/jcli-d-12-00579.1

Gao, Q., Hasselquist, N. J., Palmroth, S., Zheng, Z., & You, W. (2014). Short-term response of soil respiration to nitrogen fertilization in a subtropical evergreen forest. *Soil Biology and Biochemistry*, *76*, 297–300. https://doi.org/10.1016/j.soilbio.2014.04.020

Göransson, H., Godbold, D. L., Jones, D. L., & Rousk, J. (2013). Bacterial growth and respiration responses upon rewetting dry forest soils: Impact of drought-legacy. *Soil Biology and Biochemistry*, *57*, 477–486. https://doi.org/10.1016/j.soilbio.2012.08.031

Green, J. K., Konings, A. G., Alemohammad, S. H., Berry, J., Entekhabi, D., Kolassa, J., et al. (2017). Regionally strong feedbacks between the atmosphere and terrestrial biosphere. *Nature Geoscience*, *10*, 410–414. https://doi.org/10.1038/ngeo2957

Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., et al. (2018). ORCHIDEE-MICT (v8. 4.1), a land surface model for the high latitudes: Model description and validation. *Geoscientific Model Development*, *11*, 121–163. https://doi.org/10.5194/gmd-11-121-2018

Guo, H., Ye, C., Zhang, H., Pan, S., Ji, Y., Li, Z., et al. (2017). Long-term nitrogen & phosphorus additions reduce soil microbial respiration but increase its temperature sensitivity in a Tibetan alpine meadow. *Soil Biology and Biochemistry*, *113*, 26–34. https://doi.org/10.1016/j.soilbio.2017.05.024

Hanson, P., Edwards, N., Garten, C. T., & Andrews, J. A. (2000). Separating root and soil microbial contributions to soil respiration: A review of methods and observations. *Biogeochemistry*, *48*, 115–146. https://doi.org/10.1023/a:1006244819642

Harden, J. W., Hugelius, G., Ahlström, A., Blankinship, J. C., Bond-Lamberty, B., Lawrence, C. R., et al. (2018). Networking our science to characterize the state, vulnerabilities, and management opportunities of soil organic matter. *Global Change Biology*, *24*, e705–e718. https://doi.org/10.1111/gcb.13896

Harmon, M. E., Bond-Lamberty, B., Tang, J., & Vargas, R. (2011). Heterotrophic respiration in disturbed forests: A review with examples from North America. *Journal of Geophysical Research*, *116*, G00K04. https://doi.org/10.1029/2010JG001495

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations–the CRU TS3. 10 Dataset. *International Journal of Climatology*, *34*, 623–642. https://doi.org/10.1002/joc.3711

Hashimoto, S., Carvalhais, N., Ito, A., Nishina, K., & Reichstein, M. (2015). Global spatiotemporal distribution of soil respiration modeled using a global database. *Biogeosciences*, *12*, 4121–4132. https://doi.org/10.5194/bg-12-4121-2015

Haverd, V., Smith, B., Nieradzik, L., Briggs, P. R., Woodgate, W., Trudinger, C. M., et al. (2018). A new version of the CABLE land surface model (Subversion revision r4601) incorporating land use and land cover change, woody vegetation demography, and a novel optimisation-based approach to plant coordination of photosynthesis. *Geoscientific Model Development*, *11*, 2995–3026. https://doi.org/10.5194/gmd-11-2995-2018

Hawkes, C. V., Waring, B. G., Rocca, J. D., & Kivlin, S. N. (2017). Historical climate controls soil respiration responses to current soil moisture. *Proceedings of the National Academy of Science of the United States of America*, *114*, 6322–6327. https://doi.org/10.1073/pnas.1620811114

Haynes, B. E., & Gower, S. T. (1995). Belowground carbon allocation in unfertilized and fertilized red pine plantations in northern Wisconsin. *Tree Physiology*, *15*, 317–325. https://doi.org/10.1093/treephys/15.5.317

Hicks Pries, C. E., van Logtestijn, R. S., Schuur, E. A., Natali, S. M., Cornelissen, J. H., Aerts, R., & Dorrepaal, E. (2015). Decadal warming causes a consistent and persistent shift from heterotrophic to autotrophic respiration in contrasting permafrost ecosystems. *Global Change Biology*, *21*, 4508–4519. https://doi.org/10.1111/gcb.13032

Hijmans, R. J., van Etten, J., Sumner, M., Cheng, J., Baston, D., Bevan, A., et al. (2019). *raster: Geographic Data Analysis and Modeling*. R package version 3.0-7. 22 September 2019. Retrieved from https://CRAN.R-project.org/package=raster

Hinko-Najera, N., Fest, B., Livesley, S. J., & Arndt, S. K. (2015). Reduced throughfall decreases autotrophic respiration, but not heterotrophic respiration in a dry temperate broadleaved evergreen forest. *Agricultural and Forest Meteorology*, *200*, 66–77. https://doi.org/10.1016/j.agrformet.2014.09.013

Huang, J., van den Dool, H. M., & Georgarakos, K. P. (1996). Analysis of model-calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. *Journal of Climate*, *9*, 1350–1362. https://doi.org/10.1175/1520-0442(1996)009<1350:AOMCSM>2.0.CO;2

Huang, M., Piao, S., Ciais, P., Peñuelas, J., Wang, X., Keenan, T. F., et al. (2019). Air temperature optima of vegetation productivity across global biomes. *Nature Ecology & Evolution*, *3*, 772–779. https://doi.org/10.1038/s41559-019-0838-x

Huang, N., Wang, L., Song, X. P., Black, T. A., Jassal, R. S., Myneni, R. B., et al. (2020). Spatial and temporal variations in global soil respiration and their relationships with climate and land cover. *Science advances*, *6*, eabb8508. https://doi.org/10.1126/sciadv.abb8508

Huang, S., Ye, G., Lin, J., Chen, K., Xu, X., Ruan, J. H., et al. (2018). Autotrophic and heterotrophic soil respiration responds asymmetrically to drought in a subtropical forest in the Southeast China. *Soil Biology and Biochemistry*, *123*, 242–249. https://doi.org/10.1016/j.soilbio.2018.04.029

Humphrey, V., Gudmundsson, L., & Seneviratne, S. I. (2017). A global reconstruction of climate-driven subdecadal water storage variability. *Geophysical Research Letters*, *44*, 2300–2309. https://doi.org/10.1002/2017gl072564

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., & Seneviratne, S. I. (2018). Sensitivity of atmospheric CO_2 growth rate to observed changes in terrestrial water storage. *Nature*, *560*, 628–631. https://doi.org/10.1038/s41586-018-0424-4

Hursh, A., Ballantyne, A., Cooper, L., Maneta, M., Kimball, J., & Watts, J. (2017). The sensitivity of soil respiration to soil temperature, moisture, and carbon supply at the global scale. *Global Change Biology*, *23*, 2090–2103. https://doi.org/10.1111/gcb.13489

Ichii, K., Ueyama, M., Kondo, M., Saigusa, N., Kim, J., Alberto, M. C., et al. (2017). New data-driven estimation of terrestrial CO_2 fluxes in Asia using a standardized database of eddy covariance measurements, remote sensing data, and support vector regression. *Journal of Geophysical Research: Biogeosciences*, *122*, 767–795. https://doi.org/10.1002/2016jg003640

Jain, A. K., Meiyappan, P., Song, Y., & House, J. I. (2013). CO_2 emissions from land-use change affected more by nitrogen cycle, than by the choice of land-cover data. *Global Change Biology*, *19*, 2893–2906. https://doi.org/10.1111/gcb.12207

Janssens, I., Dieleman, W., Luyssaert, S., Subke, J. A., Reichstein, M., Ceulemans, R., et al. (2010). Reduction of forest soil respiration in response to nitrogen deposition. *Nature Geoscience*, *3*, 315–322. https://doi.org/10.1038/ngeo844

Janssens, I., Lankreijer, H., Matteucci, G., Kowalski, A. S., Buchmann, N., Epron, D., et al. (2001). Productivity overshadows temperature in determining soil and ecosystem respiration across European forests. *Global Change Biology*, *7*, 269–278. https://doi.org/10.1046/j.1365-2486.2001.00412.x

Jian, J., Vargas, R., Anderson-Teixeira, K., Stell, E., Herrmann, V., Horn, M., et al. (2021). A restructured and updated global soil respiration database (SRDB-V5). *Earth System Science Data*, *13*, 255–267. https://doi.org/10.5194/essd-13-255-2021

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., et al. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, *467*, 951–954. https://doi.org/10.1038/nature09396

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., et al. (2011). Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *Journal of Geophysical Research*, *116*, G00J07. https://doi.org/10.1029/2010JG001566

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., et al. (2017). Compensatory water effects link yearly global land CO_2 sink changes to temperature. *Nature*, *541*, 516–520. https://doi.org/10.1038/nature20780

Kato, E., Kinoshita, T., Ito, A., Huntingford, C., Sitch, S., Ahlström, A., et al. (2013). Evaluation of spatially explicit emission scenario of land-use change and biomass burning using a process-based biogeochemical model. *Journal of Land Use Science*, *8*, 104–122. https://doi.org/10.1080/1747423x.2011.628705

Keller, K., Lienert, S., Bozbiyik, A., Stocker, T. F., Frank, D. C., Klesse, S., et al. (2017). 20th century changes in carbon isotopes and water-use efficiency: Tree-ring-based evaluation of the CLM4. 5 and LPX-Bern models. *Biogeosciences*, *14*, 2641–2673. https://doi.org/10.5194/bg-14-2641-2017

Keuper, F., Wild, B., Kummu, M., Beer, C., Blume-Werry, G., Fontaine, S., et al. (2020). Carbon loss from northern circumpolar permafrost soils amplified by rhizosphere priming. *Nature Geoscience*, *13*, 560–565. https://doi.org/10.1038/s41561-020-0607-0

Kobayashi, S., Ota, Y., Harada, Y., Beer, C., Blume-Werry, G., Fontaine, S., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan Ser. II*, *93*, 5–48. https://doi.org/10.2151/jmsj.2015-001

Koerber, G. R., Meyer, W. S., SUN, Q, Cale, P., & Ewenz, C. M. (2016). Under a new light: Validation of eddy covariance flux with light response functions of assimilation and estimates of heterotrophic soil respiration. *Biogeosciences Discussions*, 1–28. https://doi.org/10.5194/bg-2016-182

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

Kraft, B., Jung, M., Körner, M., Requena, M. C., Cortés, J., & Reichstein, M. (2019). Identifying Dynamic Memory Effects on Vegetation State Using Recurrent Neural Networks. *Frontiers in Big Data*, *2*, 31. https://doi.org/10.3389/fdata.2019.0003

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., et al. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*, GB1015. https://doi.org/10.1029/2003GB002199

Kuzyakov, Y., & Gavrichkova, O. (2010). Time lag between photosynthesis and carbon dioxide efflux from soil: A review of mechanisms and controls. *Global Change Biology*, *16*, 3386–3406. https://doi.org/10.1111/j.1365-2486.2010.02179.x

Kuzyakov, Y., & Larionova, A. A. (2005). Root and rhizomicrobial respiration: A review of approaches to estimate respiration by autotrophic and heterotrophic organisms in soil. *Journal of Plant Nutrition and Soil Science*, *168*, 503–520. https://doi.org/10.1002/jpln.200421703

Lantz, B. (2013). *Machine learning with R*. Packt Publishing Ltd.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., et al. (2018). Global carbon budget 2018. *Earth System Science Data*, *10*, 2141–2194.

Li, J., Wang, Y. P., Duan, Q., Lu, X., Pak, B., Wiltshire, A., et al. (2016). Quantification and attribution of errors in the simulated annual gross primary production and latent heat fluxes by two global land surface models. *Journal of Advances in Modeling Earth Systems*, *8*, 1270–1288. https://doi.org/10.1002/2015ms000583

Li, P., Yang, Y., & Fang, J. (2013). Variations of root and heterotrophic respiration along environmental gradients in China's forests. *Journal of Plant Ecology*, *6*, 358–367. https://doi.org/10.1093/jpe/rtt009

Li, Y., Liu, Y., Wu, S., Niu, L., & Tian, Y. (2015). Microbial properties explain temporal variation in soil respiration in a grassland subjected to nitrogen addition. *Scientific Reports*, *5*, 18496. https://doi.org/10.1038/srep18496

Liaw, A., & Wiener, M. (2018). *Classification and regression based on a forest of trees using random inputs, based on Breiman (2001) R package version 4* (pp. 6–14). 22 March 2018. Retrieved from https://www.stat.berkeley.edu/~breiman/RandomForests/

Liu, W., Lü, X., Xu, W., Shi, H., Hou, L., Li, L., & Yuan, W. (2018). Effects of water and nitrogen addition on ecosystem respiration across three types of steppe: The role of plant and microbial biomass. *Science of the Total Environment*, *619*, 103–111. https://doi.org/10.1016/j.scitotenv.2017.11.119

Liu, W., Zhang, Z., & Wan, S. (2009). Predominant role of water in regulating soil and microbial respiration and their responses to climate change in a semiarid grassland. *Global Change Biology*, *15*, 184–195. https://doi.org/10.1111/j.1365-2486.2008.01728.x

Liu, X., Yang, Z., Lin, C., Giardina, C. P., Xiong, D., Lin, W, et al. (2017). Will nitrogen deposition mitigate warming-increased soil respiration in a young subtropical plantation? *Agricultural and Forest Meteorology*, *246*, 78–85. https://doi.org/10.1016/j.agrformet.2017.06.010

Liu, Y., He, N., Wen, X., Xu, L., Sun, X., Yu, G., et al. (2018). The optimum temperature of soil microbial respiration: Patterns and controls. *Soil Biology and Biochemistry*, *121*, 35–42. https://doi.org/10.1016/j.soilbio.2018.02.019

Liu, Z., Ballantyne, A. P., Poulter, B., Anderegg, W. R., Li, W., Bastos, A., & Ciais, P. (2018). Precipitation thresholds regulate net carbon exchange at the continental scale. *Nature Communications*, *9*, 3596. https://doi.org/10.1038/s41467-018-05948-1

Lu, J., Carbone, G. J., & Grego, J. M. (2019). Uncertainty and hotspots in 21st century projections of agricultural drought from CMIP5 models. *Scientific Reports*, *9*, 4922. https://doi.org/10.1038/s41598-019-41196-z

Maaroufi, N. I., Nordin, A., Palmqvist, K., Hasselquist, N. J., Forsmark, B., Rosenstock, N. P., et al. (2019). Anthropogenic nitrogen enrichment enhances soil carbon accumulation by impacting saprotrophs rather than ectomycorrhizal fungal activity. *Global Change Biology*. https://doi.org/10.1111/gcb.14722

Manzoni, S., Čapek, P., Mooshammer, M., Lindahl, B. D., Richter, A., & Šantrůčková, H. (2017). Optimal metabolic regulation along resource stoichiometry gradients. *Ecology Letters*, *20*, 1182–1191. https://doi.org/10.1111/ele.12815

Matteucci, M., Gruening, C., Ballarin, I. G., Seufert, G., & Cescatti, A. (2015). Components, drivers and temporal dynamics of ecosystem respiration in a Mediterranean pine forest. *Soil Biology and Biochemistry*, *88*, 224–235. https://doi.org/10.1016/j.soilbio.2015.05.017

Melton, J. R., & Arora, V. K. (2016). Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (CTEM) v. 2.0. *Geoscientific Model Development*, *9*, 323–361. https://doi.org/10.5194/gmd-9-323-2016

Mo, J., Zhang, W., Zhu, W., Gundersen, P. E. R., Fang, Y., Li, D., & Wang, H. U. I. (2008). Nitrogen addition reduces soil respiration in a mature tropical forest in southern China. *Global Change Biology*, *14*, 403–412. https://doi.org/10.1111/j.1365-2486.2007.01503.x

Moyano, F. E., Manzoni, S., & Chenu, C. (2013). Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models. *Soil Biology and Biochemistry*, *59*, 72–85. https://doi.org/10.1016/j.soilbio.2013.01.002

Moyano, F. E., Vasilyeva, N. A., Bouckaert, L., Cook, F., Craine, J., Curiel Yuste, J., et al. (2012). The moisture response of soil heterotrophic respiration: Interaction with soil properties. *Biogeosciences*, *9*, 1173–1182. https://doi.org/10.5194/bg-9-1173-2012

Nachtergaele, F., van Velthuizen, H., Verelst, L., Batjes, N. H., Dijkshoorn, K., van Engelen, V. W. P., et al. (2010). The harmonized world soil database. In *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World*, Brisbane, Australia, 1–6 August 2010 (pp. 34–37).

Natali, S. M., Schuur, E. A., Mauritz, M., Schade, J. D., Celis, G., Crummer, K. G., et al. (2015). Permafrost thaw and soil moisture driving CO_2 and CH4 release from upland tundra. *Journal of Geophysical Research: Biogeosciences*, *120*, 525–537. https://doi.org/10.1002/2014jg002872

Noh, N.-J., Kuribayashi, M., Saitoh, T. M., Nakaji, T., Nakamura, M., Hiura, T., & Muraoka, H. (2016). Responses of soil, heterotrophic, and autotrophic respiration to experimental open-field soil warming in a cool-temperate deciduous forest. *Ecosystems*, *19*, 504–520. https://doi.org/10.1007/s10021-015-9948-8

O'Connell, C. S., Ruan, L., & Silver, W. L. (2018). Drought drives rapid shifts in tropical rainforest soil biogeochemistry and greenhouse gas emissions. *Nature Communications*, *9*, 1348. https://doi.org/10.1038/s41467-018-03352-3

Oleson, K., Lawrence, D. M., Bonan, G. B., & Drewniak, B. (2013). *Technical description of version 4.5 of the Community Land Model (CLM)*. Boulder, CO. NCAR.

Olsson, P., Linder, S., Giesler, R., & Högberg, P. (2005). Fertilization of boreal forest reduces both autotrophic and heterotrophic soil respiration. *Global Change Biology*, *11*, 1745–1753. https://doi.org/10.1111/j.1365-2486.2005.001033.x

Peng, Y., Chen, G., Li, S., Hu, H. L., Hu, T. X., Liu, L., et al. (2018). Nitrogen additions reduce rhizospheric and heterotrophic respiration in a subtropical evergreen broad-leaved forest. *Plant and Soil*, *431*, 449–463. https://doi.org/10.1007/s11104-018-3751-1

Peterson, F. S., & Lajtha, K. J. (2013). Linking aboveground net primary productivity to soil carbon and dissolved organic carbon in complex terrain. *Journal of Geophysical Research: Biogeosciences*, *118*, 1225–1236. https://doi.org/10.1002/jgrg.20097

Prentice, I. C., Dong, N., Gleason, S. M., Maire, V., & Wright, I. (2014). Balancing the costs of carbon gain and water transport: Testing a new theoretical framework for plant functional ecology. *Ecology Letters*, *17*, 82–91. https://doi.org/10.1111/ele.12211

Raich, J. W., & Potter, C. S. (1995). Global patterns of carbon dioxide emissions from soils. *Global Biogeochemical Cycles*, *9*, 23–36. https://doi.org/10.1029/94gb02723

Raich, J. W., Potter, C. S., & Bhagawati, D. (2002). Interannual variability in global soil respiration, 1980–94. *Global Change Biology*, *8*, 800–812. https://doi.org/10.1046/j.1365-2486.2002.00511.x

R Development Core Team (2019). *R: A language and environment for statistical computing*. Vienna, Austria. R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reichstein, M., Besnard, S., Carvalhais, N., Gans, F., Jung, M., Kraft, B., & Mahecha, M. (2018). Modelling Landsurface Time-Series with Recurrent Neural Nets. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 7640–7643).

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N., et al. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*, 195–204. https://doi.org/10.1038/s41586-019-0912-1

Reick, C. H., Raddatz, T., Brovkin, V., & Gayler, V. (2013). Representation of natural and anthropogenic land cover change in MPI-ESM. *Journal of Advances in Modeling Earth Systems*, *5*, 459–482. https://doi.org/10.1002/jame.20022

Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*, 381–394. https://doi.org/10.1175/bams-85-3-381

Schimel, D., Stephens, B. B., & Fisher, J. B. (2015). Effect of increasing CO_2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Science of the United States of America*, *112*, 436–441. https://doi.org/10.1073/pnas.1407302112

Schindlbacher, A., Wunderlich, S., Borken, W., Kitzler, B., Zechmeister-Boltenstern, S., & Jandl, R. (2012). Soil respiration under climate change: Prolonged summer drought offsets soil warming effects. *Global Change Biology*, *18*, 2270–2279. https://doi.org/10.1111/j.1365-2486.2012.02696.x

Schindlbacher, A., ZECHMEISTER-BOLTENSTERN, S., & Jandl, R. (2009). Carbon losses due to soil warming: Do autotrophic and heterotrophic soil respiration respond equally? *Global Change Biology*, *15*, 901–913. https://doi.org/10.1111/j.1365-2486.2008.01757.x

Schwalm, C. R., Anderegg, W. R., Michalak, A. M., Fisher, J. B., Biondi, F., Koch, G., et al. (2017). Global patterns of drought recovery. *Nature*, *548*, 202–205. https://doi.org/10.1038/nature23021

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

Sheffield, J., Goteti, G., & Wood, E. F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, *19*, 3088–3111. https://doi.org/10.1175/jcli3790.1

Skopp, J., Jawson, M., & Doran, J. (1990). Steady-state aerobic microbial activity as a function of soil water content. *Soil Science Society of America Journal*, *54*, 1619–1625. https://doi.org/10.2136/sssaj1990.03615995005400060018x

Steidinger, B. S., Crowther, T. W., Liang, J., Van Nuland, M. E., Werner, G. D., Reich, P. B., et al. (2019). Climatic controls of decomposition drive the global biogeography of forest-tree symbioses. *Nature*, *569*, 404–408. https://doi.org/10.1038/s41586-019-1128-0

Stell, E., Warner, D., Jian, J., Bond-Lamberty, B., & Vargas, R. (2021). Spatial biases of information influence global estimates of soil respiration: How can we improve global predictions? *Global Change Biology*. https://doi.org/10.1111/GCB.15666

Stocker, B. D., Zscheischler, J., Keenan, T. F., Prentice, I. C., Seneviratne, S. I., & Peñuelas, J. (2019). Drought impacts on terrestrial primary production underestimated by satellite monitoring. *Nature Geoscience*, *12*, 264–270. https://doi.org/10.1038/s41561-019-0318-6

Sun, Z., Liu, L., Ma, Y., Yin, G., Zhao, C., Zhang, Y., & Piao, S. (2014). The effect of nitrogen addition on soil respiration from a nitrogen-limited forest soil. *Agricultural and Forest Meteorology*, *197*, 103–110. https://doi.org/10.1016/j.agrformet.2014.06.010

Suseela, V., Conant, R. T., Wallenstein, M. D, & Dukes, J. S. (2012). Effects of soil moisture on the temperature sensitivity of heterotrophic respiration vary seasonally in an old-field climate change experiment. *Global Change Biology*, *18*, 336–348. https://doi.org/10.1111/j.1365-2486.2011.02516.x

Tang, X., Fan, S., Du, M., Zhang, W., Gao, S., Liu, S., et al. (2020). Spatial and temporal patterns of global soil heterotrophic respiration in terrestrial ecosystems. *Earth System Science Data*, *12*, 1037–1051. https://doi.org/10.5194/essd-12-1037-2020

Todd-Brown, K., Randerson, J., Post, W., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., & Allison, S. D. (2013). Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences*, *10*, 1717–1736. https://doi.org/10.5194/bg-10-1717-2013

Tramontana, G., Jung, M., Camps-Valls, G., Schwalm, C. R., Ichii, K., Ráduly, B., et al. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences Discussions*. https://doi.org/10.5194/bg-2015-661

Van den Dool, H., Huang, J., & Fan, Y. (2003). Performance and analysis of the constructed analogue method applied to US soil moisture over 1981–2001. *Journal of Geophysical Research*, *108*(D16), 8617. https://doi.org/10.1029/2002JD003114

Vargas, R., Baldocchi, D. D., Allen, M. F., Bahn, M., Black, T. A., Collins, S. L., et al. (2010). Looking deeper into the soil: Biophysical controls and seasonal lags of soil CO_2 production and efflux. *Ecological Applications*, *20*, 1569–1582. https://doi.org/10.1890/09-0693.1

Viovy, N. (2015). *CRU-NCEP v6.1 Dataset*. Retrieved from http://dods.extra.cea.fr/store/p529viov/cruncep/

Wang, H., Prentice, I. C., Keenan, T. F., Davis, T. W., Wright, I. J., Cornwell, W. K., et al. (2017). Towards a universal model for carbon dioxide uptake by plants. *Nature Plants*, *3*, 734–741. https://doi.org/10.1038/s41477-017-0006-8

Wang, J., Fu, X., Zhang, Z., Li, M., Cao, H., Zhou, X., & Ni, H. (2019). Responses of soil respiration to nitrogen addition in the Sanjiang Plain wetland, northeastern China. *PloS One*, *14*, e0211456. https://doi.org/10.1371/journal.pone.0211456

Warner, D. L., Bond-Lamberty, B., Jian, J., Stell, E., & Vargas, R. (2019). Spatial predictions and associated uncertainty of annual soil respiration at the global scale. *Global Biogeochemical Cycles*, *33*, 1733–1745. https://doi.org/10.1029/2019gb006264

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data. *Water Resources Research*, *50*, 7505–7514. https://doi.org/10.1002/2014wr015638

Wieder, W., Boehnert, J., Bonan, G., & Langseth, M. (2014). *Regridded harmonized world soil database v1. 2*. ORNL DAAC.

Xiao, J., Ollinger, S. V., Frolking, S., Hurtt, G. C., Hollinger, D. Y., Davis, K. J., et al. (2014). Data-driven diagnostics of terrestrial carbon dynamics over North America. *Agricultural and Forest Meteorology*, *197*, 142–157. https://doi.org/10.1016/j.agrformet.2014.06.013

Xu, M., & Shang, H. (2016). Contribution of soil respiration to the global carbon equation. *Journal of Plant Physiology*, *203*, 16–28. https://doi.org/10.1016/j.jplph.2016.08.007

Yan, Z., Bond-Lamberty, B., Todd-Brown, K. E., Bailey, V. L., Li, S., Liu, C., & Liu, C. (2018). A moisture function of soil heterotrophic respiration that incorporates microscale processes. *Nature Communications*, *9*, 2562. https://doi.org/10.1038/s41467-018-04971-6

Yao, Y., Wang, X., Li, Y., Wang, T., Shen, M., Du, M., et al. (2018). Spatiotemporal pattern of gross primary productivity and its covariation with climate in China over the last thirty years. *Global Change Biology*, *24*, 184–196. https://doi.org/10.1111/gcb.13830

Zaehle, S., & Friend, A. D. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, *24*. https://doi.org/10.1029/2009GB003522

Zeng, J., Matsunaga, T., Tan, Z. H., Saigusa, N., Shirai, T., Tang, Y., et al. (2020). Global terrestrial carbon fluxes of 1999–2019 estimated by upscaling eddy covariance data with a random forest. *Scientific data*, *7*, 1–11. https://doi.org/10.1038/s41597-020-00653-5

Zeng, N., Mariotti, A., & Wetzel, P. (2005). Terrestrial mechanisms of interannual CO_2 variability. *Global Biogeochemical Cycles*, *19*, GB1016. https://doi.org/10.1029/2004GB002273

Zeng, Z., Wang, T., Zhou, F., Ciais, P., Mao, J., Shi, X., & Piao, S. (2014). A worldwide analysis of spatiotemporal changes in water balance-based evapotranspiration from 1982 to 2009. *Journal of Geophysical Research: Atmospheres*, *119*, 1186–1202. https://doi.org/10.1002/2013jd020941

Zhang, F. G., & Zhang, Q. G. (2016). Microbial diversity limits soil heterotrophic respiration and mitigates the respiration response to moisture increase. *Soil Biology and Biochemistry*, *98*, 180–185. https://doi.org/10.1016/j.soilbio.2016.04.017

Zhang, H., Goll, D. S., Manzoni, S., Ciais, P., Guenet, B., & Huang, Y. (2018). Modeling the effects of litter stoichiometry and soil mineral N availability on soil organic matter formation using CENTURY-CUE (v1. 0). *Geoscientific Model Development*, *11*, 4779–4796. https://doi.org/10.5194/gmd-11-4779-2018

Zhang, N., Liu, W., Yang, H., Yu, X., Gutknecht, J. L., Zhang, Z., et al. (2013). Soil microbial responses to warming and increased precipitation and their implications for ecosystem C cycling. *Oecologia*, *173*, 1125–1142. https://doi.org/10.1007/s00442-013-2685-9

Zhao, C., Miao, Y., Yu, C., Zhu, L., Wang, F., Jiang, L., et al. (2016). Soil microbial community composition and respiration along an experimental precipitation gradient in a semiarid steppe. *Scientific Reports*, *6*, 24317. https://doi.org/10.1038/srep24317

Zhou, L., Zhou, X., Zhang, B., Lu, M., Luo, Y., Liu, L., & Li, B. (2014). Different responses of soil respiration and its components to nitrogen addition among biomes: A meta-analysis. *Global Change Biology*, *20*, 2332–2343. https://doi.org/10.1111/gcb.12490

Zhou, X., Luo, Y., Gao, C., Verburg, P. S., Arnone, J. A. III, Darrouzet-Nardi, A., & Schimel, D. S. (2010). Concurrent and lagged impacts of an anomalously warm year on autotrophic and heterotrophic components of soil respiration: A deconvolution analysis. *New Phytology*, *187*, 184–198. https://doi.org/10.1111/j.1469-8137.2010.03256.x

Zhu, J., Hu, H., Tao, S., Chi, X., Li, P., Jiang, L., et al. (2017). Carbon stocks and changes of dead organic matter in China's forests. *Nature Communications*, *8*, 151. https://doi.org/10.1038/s41467-017-00207-1

Zou, J., Tobin, B., Luo, Y., & Osborne, B. (2018). Response of soil respiration and its components to experimental warming and water addition in a temperate Sitka spruce forest ecosystem. *Agricultural and Forest Meteorology*, *260*, 204–215. https://doi.org/10.1016/j.agrformet.2018.06.020