Spatial biases of information influence global estimates of soil respiration:
how can we improve global predictions?

Authors:
Emma Stell¹, Daniel Warner², Jinshi Jian³, Ben Bond-Lamberty³, Rodrigo Vargas¹,⁴,*

Affiliations:
1 Department of Geography, University of Delaware, Newark DE, USA
2 Delaware Geological Survey, University of Delaware, Newark DE, USA
3 Pacific Northwest National Laboratory, Joint Global Change Research Institute, College Park, MD, USA
4 Department of Plant and Soil Sciences, University of Delaware, Newark DE, US

*Corresponding author
531 South College Ave., 152 Townsend Hall
Department of Plant and Soil Sciences
University of Delaware
Newark, DE 19716
E-mail address: rvargas@udel.edu
Phone: 001-302-831-1386

Citation:
https://doi.org/10.1111/gcb.15666
Abstract

Soil respiration (Rs), the efflux of CO₂ from soils to the atmosphere, is a major component of the terrestrial carbon cycle, but is poorly constrained from regional to global scales. The global soil respiration database (SRDB) is a compilation of in-situ Rs observations from around the globe that has been consistently updated with new measurements over the past decade. It is unclear whether the addition of data to new versions has produced better-constrained global Rs estimates. We compared two versions of the SRDB (v3.0 n=5173 and v5.0 n=10366) to determine how additional data influenced global Rs annual sum, spatial patterns and associated uncertainty (1 km spatial resolution) using a machine learning approach. A quantile regression forest model parameterized using SRDBv3 yielded a global Rs sum of 88.6 Pg C yr⁻¹, and associated uncertainty of 29.9 (mean absolute error) and 57.9 (standard deviation) Pg C yr⁻¹, whereas parameterization using SRDBv5 yielded 96.5 Pg C yr⁻¹ and associated uncertainty of 30.2 (mean average error) and 73.4 (standard deviation) Pg C yr⁻¹. Empirically estimated global heterotrophic respiration (Rh) from v3 and v5 were 49.9-50.2 (mean 50.1) and 53.3-53.5 (mean 53.4) Pg C yr⁻¹, respectively. SRDBv5’s inclusion of new data from underrepresented regions (e.g., Asia, Africa, South America) resulted in overall higher model uncertainty. The largest differences between models parameterized with different SRDBv versions were in arid/semi-arid regions. The SRDBv5 is still biased towards northern latitudes and temperate zones, so we tested an optimized global distribution of Rs measurements, which resulted in a global sum of 96.4 ± 21.4 Pg C yr⁻¹ with an overall lower model uncertainty. These results support current global estimates of Rs but highlight spatial biases that influence model parameterization and interpretation and provide insights for design of environmental networks to improve global-scale Rs estimates.
Keywords: soil CO₂ efflux; machine learning; heterotrophic respiration; carbon cycle; network design; network representativeness
1. Introduction

Soil respiration (Rs) is the total efflux of CO₂ from soils to the atmosphere as a result of below-ground processes. Rs comprises the second largest flux in the carbon cycle, after gross primary productivity (GPP) (Raich & Schlesinger 1992), and is a major process involved in land-atmosphere CO₂ exchange (Beer et al. 2010). Roughly 10% of the global atmospheric CO₂ cycles through the soils each year (Raich & Potter 1995), an order of magnitude greater than anthropogenic emissions (Raich & Tufekcioglu 2000). Soil respiration is the sum of autotrophic and heterotrophic components. Autotrophic respiration (Ra) is the efflux of CO₂ due to plant roots, while heterotrophic respiration (Rh) is the efflux of CO₂ due to microbial decomposition of soil organic matter. Rh is an important component of Rs because it is strongly influenced by changes in soil temperature (Wang et al. 2014) and moisture (Moyano et al. 2012), and long-term changes could lead to increases of CO₂ in the atmosphere. A rising Rh:Rs ratio could mean more soil organic carbon is being lost, but the sensitivity of Rh to a changing climate is largely unknown (Bond-Lamberty et al. 2018). It is therefore important to also investigate how Rh varies in comparison to Rs for a better overall understanding of global Rs processes.

Scientists have estimated Rs at the global scale for decades (Schlesinger 1977) and recent estimates converge to around 91-94 Pg C yr⁻¹ (Bond-Lamberty 2018). These estimates are higher than global sums from earlier research that range from 68-76.5 Pg C yr⁻¹ (i.e., Schlesinger 1977, Raich & Schlesinger 1992, Raich & Potter 1995, Raich et al. 2002). There is large variability in new estimates of Rs fluxes using different approaches, ranging from 66.6 - 108.6 Pg C yr⁻¹ (Jian et al. 2018a, Hursh et al. 2017), which introduces substantial uncertainty into climate forecasts. Over 40 Pg C yr⁻¹ separate the highest and lowest global Rs estimates. The lowest estimate of 66.6 Pg C yr⁻¹ was generated through an empirical model (Jian et al. 2018a), while the highest
estimate, of 108.6 Pg C yr⁻¹, was generated through a mechanistic model (Hursh et al. 2017). Historically, most global Rs estimates have been generated through empirical models (e.g., Schlesinger 1977, Raich & Schlesinger 1992, Raich & Potter 1995, Raich et al. 2002), with only some recent examples of mechanistic (e.g., Hursh et al. 2017) and machine learning (ML) approaches (e.g., Zhao et al. 2017, Jian et al. 2018a, Warner et al. 2019).

It is technically impossible to directly measure large-scale Rs, but statistical models and ML offer a potential solution (Vargas et al. 2011, Thessen 2016). Recent studies have used ML (e.g., Zhao et al. 2017, Jian et al. 2018a, Vargas et al. 2018, Warner et al. 2019) which can help improve local to large-scale estimates by relating Rs observations to ancillary vegetation, terrain, climate, and survey data. Some ML algorithms are particularly effective in modeling biogeophysical factors due to their ability to account for nonlinear relationships and low sensitivity to autocorrelation amongst predictor variables relative to more traditional linear models (Reichstein et al. 2019; Guevara et al. 2018; Vargas et al. 2018). ML approaches have been shown to outperform mechanistic and semi-empirical models in modeling carbon processes in which point-based observations are used for regional to global estimates (Reichstein et al. 2019), which shows promise for the use of ML algorithms in future environmental modeling and benchmarking (Bond-Lamberty 2018). Although ML offers a powerful tool for environmental prediction, it can be challenging to evaluate results, especially within ecological contexts (Basu et al. 2018).

Communicating uncertainty is an important aspect of properly reporting global Rs estimates. Uncertainties in land carbon cycle processes account for most of the uncertainty in overall CO₂ projections (Friedlingstein et al. 2006, 2014), but these models could be better constrained with more accurate estimates of global Rs. Many studies estimating global Rs report
uncertainty, but not all. Out of early estimates, only one reports uncertainty, a standard deviation of ±4 Pg C yr\(^{-1}\) (Raich & Schlesinger 1992). The practice of reporting uncertainties became more commonplace starting in the early 2000s (Raich et al. 2002), and most studies of global Rs estimates since then have reported uncertainty (Bond-Lamberty & Thomson 2010, Chen et al. 2013, Hashimoto et al. 2012, 2015, Xu & Shang 2016, Zhao et al. 2017, Hursh et al. 2017, Warner et al. 2019). Current estimates of uncertainty vary widely, and one reason is that uncertainty is reported in different ways, such as root mean square error (RMSE) (Hursh et al. 2017, Zhao et al. 2017), standard deviation (SD) (Raich & Schlesinger 1992), standard error (Bond-Lamberty & Thomson 2010, Chen et al. 2014), 95% confidence level (Xu & Shang 2016, Hashimoto et al. 2012, 2015), or a mixture of different measures (Warner et al. 2019, Chen et al. 2013, Raich et al. 2002). These may represent different uncertainty sources (e.g., year-to-year variability, measurement error, parametric uncertainty), so interpretation and comparison among these estimates should be done carefully (Enting et al. 2012).

Regardless of methodology, most global Rs modeling efforts rely on \textit{in situ} measurements of Rs for calibration and/or validation (see however Konings et al. 2019). There have been several efforts to integrate point-based Rs measurements from studies around the world into publicly accessible databases (Bond-Lamberty & Thomson 2010; Jian et al. 2018a; Jian et al. 2018b; Bond-Lamberty et al. 2020). The Soil Respiration Database (SRDB) is a database of annual Rs measurements from around the globe (Bond-Lamberty & Thomson 2010). Version 1.0 of the SRDB was released in May 2010 with over 3370 observations from over 800 studies, and Version 5.0 was released in April 2020 (Jian et al. 2020) with over 10,000 observations from over 2000 studies ranging from 1961-2017, making it a valuable resource for regional to global analysis.
In addition, the Hourly and Daily Timescale Global Soil Respiration Database (HGRsD) is a database of hourly and daily Rs measurements also compiled from studies around the world (Jian et al. 2018c). Finally, COSORE is a community database of continuous Rs measurements first released in June 2019, in which automated Rs measurements from individual sites are compiled into one database (Bond-Lamberty et al. 2020). With each subsequent version released, the spatiotemporal coverage of global Rs sites increases, facilitating additional insight to global Rs processes.

However, even with the expanded SRDB v5.0, there still exists an uneven distribution of Rs observations around the globe. There are substantially more measurements of Rs in the Northern Hemisphere, especially in Europe, China, and North America (Xu & Shang 2016). Measurements are especially lacking in Africa (Epule et al. 2015), central Asia, and South America (especially in the Amazon). Most available measurements are located in temperate regions, with few data points in tropical and dryland ecosystems. This may lead to high uncertainty in spatial upscaling, given the role of tropical and dryland regions in regulating much of the global carbon cycle (Malhi & Grace 2000; Poulter et al. 2014). The uneven distribution of measurements could bias regional-to-global representation and consequently the benchmarking (i.e., comparing known observations to models) of earth system models (ESMs). Improving ESMs requires extensive comparison between modeled and observed data and updated observational data of Rs has been identified as a critical need (Hoffman et al. 2017). If the addition of more measurements decreases uncertainty in global estimates, then benchmarks in ESMs could be better constrained, enhancing the ability of researchers to accurately quantify global carbon fluxes. A critical question, then, is how much ‘value’ additional Rs observations (i.e., recently added to SRDBv5) are bringing to global Rs estimation in terms of decreasing
uncertainty and spatial bias. While previous research has argued for more observations to constrain models, there has been little investigation into how new Rs observations, and new observation locations, influence global Rs estimates.

The overarching goal of the study was to predict spatial patterns and global estimates of Rs using two different versions of the SRDB (version 3.0 (v3) vs version 5.0 (v5)) in order to determine if adding new Rs measurements could better constrain global estimates of Rs. Our objectives are: 1) model and map mean annual global Rs with its associated uncertainty at a 1 km resolution; 2) estimate the spatial variability of annual Rh and Rh:Rs; 3) determine if the addition of recent Rs sites to the SRDB decreases the uncertainty of model predictions of Rs at the global scale; 3) investigate how model performance varies based on the distribution of sampling locations. We hypothesize that SRDBv5 will have lower uncertainty in the overall global prediction of Rs as it has more sites and better distribution of information available for model parameterization. As the SRDB grows and expands, this is the first study comparing different versions of this global database to investigate the effects of inclusion of additional measurements, especially with new additions across underrepresented regions of the world (e.g., Africa, South America, Asia).

2. Methods

2.1 Rs training data

We used data from two versions of the global soil respiration database (SRDB): v3 (downloaded September 5, 2018; http://dx.doi.org/10.3334/ORNLDAAC/1235) and v5 (downloaded April 9, 2020; https://github.com/bpbond/srdb). The SRDB is an initiative to compile field studies of Rs from around the globe into one harmonized database (Bond-
Lamberty and Thomson (2010). Information selected from the database includes record number, study number, latitude, longitude, manipulation, measurement method, annual Rs, biome, and study midyear. Data were only included if certain criteria were met: Rs_annual < 10,000 g C m$^{-2}$ yr$^{-1}$ (accounting for outliers); there were no field manipulations; and the measurement methods included either an infrared gas analyzer (IRGA) or gas chromatography.

A total of 2321 data points from 618 studies for v3, and 4115 data points from 1036 studies for v5 of the SRDB, met these requirements (Figure 1). Consequently, our analyses using v5 included 44% more training data points. In comparison to v3, the expanded v5 database has increased spatiotemporal coverage, newly included studies within Chinese and Russian scientific literature, and enhanced interoperability with other carbon datasets (Jian et al., 2020).

2.2 Prediction factors for global Rs

Our initial predictors are a set of 15 globally-distributed predictors that were represented by climatological, vegetative, and geophysical predictive factors related to Rs. Prediction factors are represented by environmental information from worldgrids.org (last accessed Jan. 2018), an initiative of ISRIC Soil Information (Reuter & Hengl 2016) that has since been taken offline, and from worldclim.org (Fick & Hijmans 2017). The worldgrids and worldclim predictors come from three sources: remote sensing, climate surfaces, and digital terrain analysis (see complete list with abbreviations and description of layers in Table S1). These were mean annual temperature, mean annual precipitation, mean of the monthly MODIS Enhanced Vegetation Index time series data, mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%), artificial surfaces and associated areas (Urban areas >50%), geological ages based on the surface geology, land cover classes based on the Medium Resolution Imaging Spectrometer
(MERIS) full resolution images, mean potential incoming solar radiation derived in the System for Automated Geoscientific Analyses (SAGA) GIS, mean value of the 8-day MODIS LAI time series data, Long-term precipitation for periods Nov/Dec/Jan, Feb/Mar/Apr, May/Jun/Jul, and Aug/Sep/Oct based on WorldClim.org data (version 2.1; average from 1970-2000). Slope map in percent derived using the DEMSRE3, and SAGA Topographic Wetness Index derived using the DEMSRE3. All the predictors have the same spatial resolution and projection and were subsequently organized into a regression matrix for further analysis. The regression matrix was used for variable selection and subsequent modeling as described in section 2.3.

2.3 Prediction of global Rs and model evaluation

We used VSURF (Variable Selection Using Random Forests) in order to refine and select the most relevant environmental variables for prediction. The two main steps composing VSURF are interpretation and prediction. First, predictor variables are evaluated for their individual impact on model performance, then predictors are analyzed for redundancy to find a smaller set of variables sufficient for prediction (Genuer et al. 2010, 2015). In the interest of comparing the same model across the two versions of the SRDB, we used the variables selected for v5 of the SRDB by VSURF for predicting Rs using SRDB v3 and v5.

We used quantile regression forests (QRF) to predict Rs at 1 km resolution at the global scale (see Objective 1) following a previously published approach (Warner et al. 2019). QRF is a generalization of Random Forests (RF), which is a commonly used ML algorithm (Biau & Scornet 2016; Breiman 2001) that has been used for studying Rs (Vargas et al. 2010). Unlike RF, in QRF the conditional prediction distributions are preserved, allowing users to infer quantiles and other distribution statistics beyond the conditional mean (Meinshausen 2006). Furthermore,
QRF has been shown to outperform Regression-Kriging in predictions of uncertainty and can be used when data are sparse (Vaysse & Lagacherie 2017). The QRF model was trained using the environmental covariates identified by VSURF, and using 2321 training Rs data for v3 and 4115 observations for v5. Model parameters \textit{ntree} (the number of trees grown in the QRF model) and \textit{mtry} (the number of variables considered at each split in the prediction trees) were calibrated using five repetitions of a 10-fold cross validation. The resulting model was then extrapolated to the set of global environmental covariates at 1 km resolution in order to generate conditional prediction distributions of Rs for each pixel. The mean of the conditional distribution of the QRF was used to make the final global mean Rs prediction map using SRDB v3 or v5. We used the conditional SD (standard deviation) and C.V. (coefficient of variation) for looking at model prediction consistency within each pixel.

Global Rs estimates were made as the sum of the mean of the conditional prediction distribution at each pixel multiplied by its surface area and compared among v3 and v5. Uncertainty was reported as the sum of the SD of the conditional prediction distribution at each pixel multiplied by its surface area (\(\varepsilon_{SD}\)), the sum of the model RMSE multiplied by the area of each pixel (\(\varepsilon_{RMSE}\)), and the sum of the model MAE (mean absolute error) multiplied by the area of each pixel (\(\varepsilon_{MAE}\)) (see Objective 2). Model prediction consistency at each pixel was expressed as C.V., which is the standard deviation of the conditional distribution of the QRF model divided by the mean. To illustrate the discrepancies between the model outputs, we created difference maps by subtracting Rs mean or C.V. of v3 from v5. Zonal statistics were extracted from our annual Rs prediction raster for 16 IGBP vegetation classes based on MODIS data (Channan et al. 2014) in order to compare the mean and total Rs sums among them.
We examined the effects of each predictor variable on model output through partial dependence plots to better understand the inner workings of our QRF model. Partial dependence plots provide insights on the general response of model output (i.e., predicted Rs) across the range of an individual independent variable over the marginal distribution of all other independent variables (Goldstein et al. 2015). Partial dependence plots are useful for interpreting relationships between each individual predictor and the model output, while accounting for the effects of the other predictors within the model (Greenwell 2017). This partially accounts for the drawback that ML methods can be more difficult to interpret relative to more traditional statistical methods. Within the plots, positive or negative trends indicate a general positive or negative influence on predicted output among training instances, while a flat trend indicates that the effects of the predictor on the model output are weak, or inconsistent across individual training instances.

2.4 Optimizing global measurement locations

We applied a framework to identify potential sites that better represent the multivariate covariate space across the world. This analysis aims to inform potential biases in the SRDB and identify measurement locations around the world to better represent global Rs spatial variability (see Objective 3). Individual measurements within the SRDB were never planned to be representative of global Rs patterns, but because of the wealth of information compiled, the SRDB is applicable to address research at the global scale (Bond-Lamberty et al. 2018; Bond-Lamberty and Thomson 2010). That said, the database is biased towards studies in northern latitudes and temperate zones (Jian et al. 2020).
We used a conditioned Latin Hypercube Sampling (cLHS) analysis to identify alternative potential sites across the globe, modeled after similar analyses (e.g., improving the design and representation of environmental networks (Villarreal et al. 2019)). cHLS is a sampling strategy in which \( x \) new sites are chosen such that the multivariate distribution is maximally stratified (Minasny & McBratney 2006). While cLHS has mostly been used within digital soil mapping (e.g., Chu et al. 2010, Mulder et al. 2013, Silva et al. 2015, Stumpf et al. 2016), some studies have utilized cLHS for other environmental applications (e.g., Lin et al. 2009, Yin et al. 2016, Yin et al. 2017, Contreras et al. 2019, Villarreal et al. 2019). We applied the cLHS algorithm to identify 3000 sites around the globe (because SRDBv5 has >3000 data entries) in locations that maximize the representation of the multivariate space of the previously selected environmental covariates used to model Rs.

The ability of statistical models to improve prediction of a given variable relies not only on the quantity of available data, but also on how well that data represents the variability of the space in which predictions are being made. We repeated our QRF approach, but the training data came from the hypothetical optimized 3000 locations (identified by the cLHS analysis) extracted from the global output using a QRF parameterized with SRDBv5. The resulting maps (annual mean and C.V.) are useful to test how the spatial distribution of the training data influence Rs estimates and uncertainty around the world. We clarify that these maps are not intended to be accurate predictions of global Rs, as the input are the previous model outputs of our global Rs estimates using v5. Information about the influence of spatial distribution on global Rs estimates is important for interpretation of the model outputs, and provides insights about where new measurements are needed, data-model mismatches, and how uncertainty could be constrained.
2.5 Calculation of Rh

We calculated heterotrophic respiration (Rh) on a pixel-by-pixel basis from our rasters of global estimated Rs. Rh represents the portion of Rs contributed by microbes and other heterotrophic organisms within the soil (Chapin et al. 2006). The role of Rh is directly relevant to issues of carbon balance, soil carbon stability, and climate feedbacks (Bond-Lamberty et al. 2018). Therefore, Rh was calculated from global Rs (at 1x1 km pixels) from SRDB v3 and v5 for comparison. The following equations were used to estimate Rh:

\[
\ln(Rh) = 1.22 + 0.73\ln(Rs) \quad \text{equation 1}
\]

(Bond-Lamberty et al. 2004)

\[
\frac{Rh}{Rs} = -0.138\ln(Rs) + 1.482 \quad \text{equation 2}
\]

(Subke et al. 2006)

These simple functions provide a useful and generally robust (for undisturbed systems) approximation of Rh, although we note that the increasing availability of direct Rh measurements means that global-scale upscaling for this flux is now possible (Tang et al. 2020).

We additionally calculated the global Rh:Rs ratio for v3 and v5, as well as the Rh:Rs ratio for the different land cover types as a proxy for testing potential differences in functional relationships between SRDB versions. For this objective, we used the mean global Rh estimates derived from equation 1.

All analyses were completed using R (R Core Team 2019). Packages raster, VSURF, quantregForest, pdp, and cLHS were used.
3. Results

3.1 Prediction factors for global Rs

Six variables had the highest impact on the prediction of Rs. These variables were mean annual temperature (MAT), mean annual MODIS enhanced vegetation index (EVI), mean precipitation from November to January (WMP), mean precipitation from February to April (SpMP), mean precipitation from May-July (SMP), and mean precipitation from August-October (FMP). These six variables were used for the prediction of Rs with both the v3 and v5 SRDB.

3.2 Prediction of global Rs and model evaluation

We calculated the mean sum as well as the $\varepsilon_{\text{SD}}$, $\varepsilon_{\text{RMSE}}$, and $\varepsilon_{\text{MAE}}$ of both v3 and v5. The best $n_{\text{tree}}$ and $m_{\text{try}}$ QRF parameters were 210 and 1, respectively. For v3 we found a final model $R^2$, RMSE, and MAE of 0.61, 368.8 g C m$^{-2}$ yr$^{-1}$, and 200.8 g C m$^{-2}$ yr$^{-1}$, respectively, while for v5, we found a final model $R^2$, RMSE, and MAE of 0.61, 357.0 g C m$^{-2}$ yr$^{-1}$, and 202.4 g C m$^{-2}$ yr$^{-1}$, respectively. The sum of the mean of the conditional distribution of the QRF model for global Rs with v3 data was 88.6 Pg C yr$^{-1}$, while the sum of the mean for v5 was 96.5 Pg C yr$^{-1}$ (Figures 2a and 2c). For v3, the $\varepsilon_{\text{SD}}$, $\varepsilon_{\text{RMSE}}$, and $\varepsilon_{\text{MAE}}$ was 57.9, 54.9, and 29.9 Pg C yr$^{-1}$, respectively. For v5, the $\varepsilon_{\text{SD}}$, $\varepsilon_{\text{RMSE}}$, and $\varepsilon_{\text{MAE}}$ was 73.4, 53.2, and 30.2 Pg C yr$^{-1}$, respectively. The mean global C.V. was 72.2 +- 21.8% for v3, and 89.8 +- 40.4% for v5.

Difference maps illustrate the discrepancy between the model outputs of v3 vs v5 (Figure 3). For the mean difference (Figure 3a), values are especially high in northern central Asia, eastern Europe, western semi-arid United States, portions of central Africa, South America and Australia.
The mean percent changes between v3 and v5 were highest in arid/semi-arid regions (Figure 4a). Within different regions of the world, the same numerical change can be a large percentage change within a region with typically low values, but a small percentage change within a region with higher values. The values of the mean percent change range from -638 to 66.4% following a normal distribution. A subsample of percent differences was plotted within climate space (i.e., along the global distributions of MAP and MAT; Figure 4b) demonstrating higher values of Rs in SRDBv5 in wet areas (>200 mm of annual precipitation) around the world.

The models derived from v3 and v5 produced similar partial dependence plots for each variable, with the exception of FMP (mean fall precipitation), which showed a stronger response to v5, and SpMP (mean spring precipitation), which showed a stronger positive effect to v3 within lower values (Figure 5). MAT shows a continual strong positive response to the model, indicating large dependency for the model over the range of its values. Only higher EVI values (>3000) exhibited a strong dependency on the model. The four precipitation variables show a stronger response in low-mid amounts of precipitation, with the response dropping or leveling out for values >300-400mm.

Model estimated Rs varied widely across the 16 different IGBP classes throughout the world, but Evergreen Broadleaf Forests (EBF) show the highest mean Rs. Table 1 shows the Rs mean, total sum, and C.V. for each IGBP class for v3 and v5 of the SRDB. Mean Rs in EBF was 1317.66 g C m⁻² for v3, and 1305.22 g C m⁻² for v5. Savannas had the next largest mean Rs, at 947.73, and 1076.73 g C m⁻² for v3 and v5, respectively.

3.3 Prediction of Rs using an optimized network design
The cLHS analysis evenly distributed 3000 points around the globe to represent an optimized hypothetical network (Figure 6). The locations from this optimized network were selected to maximize the representation of the covariate multivariate space, and we predicted global patterns of Rs using values extracted at these optimized locations from Figure 2c (Figure 7). Using this approach, we found a global Rs sum of 96.4 Pg C yr\(^{-1}\), but with lower uncertainty represented by a global SD of 21.4 Pg C yr\(^{-1}\), and a mean C.V. of 24.8 ± 10.1%.

3.4 Calculation of global Rh and Rh:Rs

For v3, Rh was calculated to be 49.9-50.2 Pg C yr\(^{-1}\) (mean 50.1 Pg C yr\(^{-1}\)), while for v5 Rh was calculated to be 53.3-53.5 Pg C yr\(^{-1}\) (mean 53.4 Pg C yr\(^{-1}\); see Figure S1). There was an increase in Rh from v3 to v5 in northern temperate regions, central and eastern Asia, Australia, and dryland areas in North and South America. Many areas of the world that experienced increases in Rh fall in arid/semi-arid areas. We calculated the zonal statistics for v3 and v5 Rh based on the Bond-Lamberty (2004) equation (Table 3). The largest Rh sums for both v3 and v5 were found in Evergreen Broadleaf Forest, Grasslands, and Open Shrublands. The lowest Rh sums were in Closed Shrublands, Urban and built-up areas, and Deciduous Needleleaf Forest. We additionally investigated the Rh:Rs ratio between v3 and v5.

The global Rh:Rs ratio was similar between versions, with v3 having a ratio of 0.64, v5 a ratio of 0.62 (Figure S2). However, differences arise within the zonal statistics investigating different land cover types (Table 3). Significant changes were > 4% for percent mean change, and >6% for the percent SD change. The four land cover types with significant changes in percent mean change all had negative changes, i.e., the Rh:Rs ratio in v3 was substantially larger than in v5. Changes in percent SD change were even more significant in many cases, with the
largest change being a decrease of 37.95% within grasslands. However, in this case, there were both positive and negative significant changes. Investigating the Rh:Rs ratio of the cLHS model reveals that compared to v5, the global mean is also 0.60, but differences arise within the land cover types (Table S2; Figure S3). The largest percentage mean change among land cover types between v5 and the cLHS model is 7.13% within open shrublands. However, the percentage SD change is much larger, with the lowest percentage change value being 18.58%, and the largest 60.61%, a substantial increase from comparing between v3 and v5.

4. Discussion

4.1 Global Rs and Rh predictions

The global Rs and Rh sum increased from v3 to v5. The global Rs and Rh sums for v3 were 88.6 and 50.1 Pg C yr\(^{-1}\), and for v5 were 96.5 and 53.4 Pg C yr\(^{-1}\), respectively. There was an 8.9% increase in Rs and a 6.9% increase in Rh between v3 and v5 of the SRDB. These predictions hover just outside the range of other recent global estimates, which have been found to be converging on a global Rs of 91-94 Pg C yr\(^{-1}\) (Bond-Lamberty 2018). Our Rs predictions agree well with other RF implementations (Jian et al. 2018a; Warner et al. 2019) based on the SRDB (Bond-Lamberty & Thomson 2010). Our Rh predictions compare well to other empirical global Rh estimates of 50 Pg C yr\(^{-1}\) (Raich & Schlesinger 1992), 57.2 Pg C yr\(^{-1}\) (Tang et al. 2020), and 49.7 Pg C yr\(^{-1}\) (Warner et al. 2019), while falling within the range estimated by ESMs (i.e., 41.3-71.6 Pg C yr\(^{-1}\)) (Shao et al. 2013). However, it is worth noting a more recent study facilitated via a top-down approach found a global Rh of only 39 Pg C yr\(^{-1}\) (Ciais et al. 2021), although this is a lower estimate in comparison to aforementioned studies.
The six overall predictors selected (using an unsupervised approach) to predict Rs agree well with other studies. We highlight that it is critical to properly select environmental predictors within a ML framework to improve model interpretability, reduce model overparameterization, and reduce computing costs (i.e., learning time). Our unsupervised approach (VSURF) maximized model parameterization while reducing the covariate space to improve model interpretability. A previous machine learning-based study found that using less (and interpretable) covariates result in a model with comparable accuracy but higher interpretability and lower computational costs than a model using >100 covariates (Guevara et al. 2020). We highlight that selection of covariates for proper implementation of machine learning approaches is a matter of debate and affects model interpretability, reproducibility and applicability.

Overall, temperature and precipitation are well-known predictors of local-to-global Rs, and other studies have shown the importance/influence of productivity variables on Rs (e.g., Hursh et al. 2017). We investigated the effect of each predictor on modeled Rs values via partial dependence plots. MAT showed a weak response for lower temperatures, but an overall positive response throughout the rest of its range of values, only dropping off in the highest temperatures. MAT is known to be a suitable predictor of annual Rs across different ecosystems (Bahn et al. 2010). EVI showed little effect on Rs within lower values, but an increasingly positive trend emerged in the middle to upper ranges of its values. This could be because higher values of vegetation indices (such as EVI) likely represent potential increased in substrate supply (Hill et al. 2020), which directly influences Rs (Davidson et al. 2006).

The effects of precipitation variables were more varied. SpMP showed the most consistent positive trend across the range of its values before its influence declined after ~400mm. FMP only showed influence between ~50-200mm, while WMP showed the most
influence from 150-250mm. SMP only showed consistent influence between 300-350mm. The contrasting influences of the precipitation variables across different ranges of their values could explain why VSURF selected seasonal precipitation variables instead of annual mean precipitation. The biggest differences in the partial dependence plots between v3 and v5 are in SpMP and FMP. This is possibly due to the large increase in observations in regions with different precipitation seasonality, such as in regions of Asia that are influenced by monsoons. It is worth noting that just because these plots do not show a steep slope does not mean that there is no effect; there could be high variability among model instances within these predictor ranges across the different model trees that cancel each other out.

Contrary to our hypothesis, model uncertainty increased when using SRDB v5 data over v3 data. Between the two versions of the SRDB, there was an increase of 1794 usable training data points (~77%). We expected that with an increased number of measurements, uncertainty should decrease in v5. The model uncertainty metrics (RMSE and MAE) were similar between v3 and v5. However, both SD and C.V. increased between v3 and v5. The SD sum increased by 15.5 Pg C yr\(^{-1}\) (a 26.8% increase), while the mean C.V. increased by 17.6%. We clarify that this is not a reflection of the accuracy of the model, but is instead a result of the effect on consistency in model runs by adding more data points, especially across underrepresented regions around the world (i.e., areas with few data points). In other words, the model is biased with information from northern latitudes but has inconsistencies (resulting in larger uncertainty) when new but limited information is added in underrepresented areas. Thus, spatial bias influences model consistency and therefore uncertainty in the estimates. For v3, the C.V. is lower because the model has consistency across some areas even without data assuming results from a similar multivariate space. While the SD and C.V. saw substantial increases, the MAE and RMSE were
comparable between v3 and v5. MAE and RMSE express the overall error between model predictions and observed values as a means to estimate uncertainty across extrapolated predictions. In the case of the QRF model, the SD and CV reflect the consistency or consensus of the conditional prediction distribution for each pixel. For a given pixel, a higher CV indicates that the set of local covariate data produces highly inconsistent predictions across the individual trees within the regression forest. This may reflect a poor representation of that pixel’s unique geographic and climate characteristics by the training dataset. We concluded that the training performance using v3 and v5 are comparable, but the effect of increasing data in v5 resulted in decreased model consistency with greater variability among the regression forest predictions, likely as a result of a spatial unbalance between training points (i.e., dataset biased toward northern latitudes).

Many large differences in mean Rs between v3 and v5 were located in arid/semi-arid regions. The difference maps (Figure 3) and percent change map (Figure 4a) visually illustrate the differences between v3 and v5. Many arid/semi-arid areas, especially in the eastern hemisphere, experienced decreases in Rs from v3 to v5, including the Middle East, northern Australia, Saharan Africa, and southern Africa. Of these regions, only southern Africa had an increase in data points from v3 to v5. Even with the decreases observed in Figure 4a, only two IGBP land cover classifications saw decreases in mean and sum Rs, evergreen broadleaf forests and barren or sparsely vegetated land cover. In contrast to the eastern hemisphere, in the western hemisphere there was an increase in Rs in many dryland ecosystems, such as in the western United States, and South American arid/semi-arid regions. There was an increase in data points in both of these regions. Other regions with an increase in Rs include Central America, eastern Europe, Greenland, southern Australia, and in sub-Saharan African regions dominated by
dryland ecosystems (i.e., grasslands, savannas, and woody savannas; Figure 3a). It is possible
that the large areas of contiguous woody savannas in sub-Saharan Africa could explain the large
increase in mean and sum Rs (134.48 g C m⁻² yr⁻¹ and 1.72 Pg C yr⁻¹, respectively) between v3
and v5. While no data points were added in this region with v5, data points were added in other
regions (i.e., eastern China) that are within woody savannas. Both the mean and sum Rs in
savannas increased from v3 to v5 (by 128.99 g C m⁻² yr⁻¹ and 1.24 Pg C yr⁻¹, respectively),
possibly because of the increase in the number of points within underrepresented savannas in
Africa and South America.

The difference map for C.V. confirmed that there were large differences within dryland
areas, including northern Africa, southern Africa, the Middle East, Australia, Pacific coastal
areas in South America, and high latitudes. Many of these areas are located within barren or
sparsely vegetated land cover, which corresponds to the 66.24% increase in C.V. Given the large
increase in C.V. within this land cover type, more measurements are needed in these areas. There
was a decrease in C.V. values from v3 to v5 in most of Greenland, central Europe, the Amazon,
and central Africa within tropical regions (Figure 3b). The only land cover classification with a
decrease in C.V. was evergreen broadleaf forest, corresponding to the aforementioned tropical
regions. Within IGBP classifications, there were higher differences in C.V. within permanent
wetlands, barren areas, and open shrublands. Dryland environments such as the latter two are
large drivers of interannual variability in Rs, with the majority of variability being attributed to
the Southern Hemisphere (Poulter et al. 2014, Ahlstrom et al. 2015). Accurately accounting for
large scale Rs fluxes in these ecosystems remains a challenge for global carbon cycle research.

The mean percent change (Figure 4a) had the largest increases in grasslands, open
shrublands, savannas, and in Asian barren regions, reinforcing that drylands are large sources of
variability. As expected, the percent change highlighted regions that were not elucidated within the mean difference map, mainly within arid/semi-arid regions. Within the global climate space (Figure 4b), v5 had consistently higher values than v3 in warmer, wetter (precipitation above ~200-250mm yr\(^{-1}\)) climates. On the other hand, v5 had lower values than v3 in colder, arid (precipitation below ~200-250mm yr\(^{-1}\)) climates. These differences on opposite ends of the climate spectrum indicate that new measurements should be focused across these climate spaces.

Regarding global estimates of Rh, we found that there was an 6.9% increase in Rh between v3 and v5 of the SRDB. These predictions hover just outside the range of other recent global Rh estimates, which have been found to be converging on a global Rh of about 50 Pg C yr\(^{-1}\) (Hashimoto et al. 2015, Bond-Lamberty et al. 2016), although estimates as low as 39 Pg C yr\(^{-1}\) (Ciais et al. 2021) and as high as 57.2 Pg C yr\(^{-1}\) (Tang et al. 2020) have been made. Among land cover types, the highest mean and sum Rh was predictably in evergreen broadleaf forests. The largest mean and sum change was in grasslands, which is consistent with Rs as our approach derives Rh from Rs, but this result needs to be corroborated with other analyses.

We investigated the global Rh:Rs ratio between v3 and v5 as a proxy to test a functional relationship and to provide further information for model benchmarking to interpreting results from different SRDB versions. A previous study found that the global Rh:Rs ratio had increased from 0.54 to 0.63 from 1990 to 2014, an increase of 14.29% (Bond-Lamberty et al. 2018). Comparing results from different versions of SRSB we found a mean global Rh:Rs ratio of 0.60 for v3 and 0.62 for v5, only a -3.31% difference. These estimates suggest consistency that the mean global Rh:Rs falls within a range of ~0.61 when incorporating all the information from the SRDB. That said, detailed studies comparing measured Rs with measured Rh are needed to clearly identify annual changes from local-to-global scales (Bond-Lamberty et al. 2018). We
analyzed the changes of Rh:Rs (between v3 and v5) for each and found statistical differences between versions of SRDB for all land cover types. Changes of Rh:Rs ranged from <1 to >6% and changes in the standard deviation of Rh:Rs ranged from <1 to >37% (Table 3). Many land cover types had decreases in the percentage change in both mean and SD, indicating that the v5 Rh:Rs ratio is becoming more conservative when using all available data in v5. We highlight that larger changes in Rh:Rs may suggest the need for future research in grasslands, deciduous needleleaf forests and open shublands; while large changes in the SD points towards deciduous needleleaf and broadleaf forests, open shrublands, grasslands, soils from urban areas, and areas with sparse vegetation.

**4.2 Improving global representativeness of Rs**

Even with improvements in spatial coverage with each subsequent version of the SRDB, it is still biased toward the northern hemisphere and temperate ecosystems (Jian et al. 2020). The majority of the world’s landmass is located in the northern hemisphere, along with more economically developed nations. The bias toward the northern hemisphere and more temperate regions affects global Rs estimates by influencing estimations of Rs across multivariate spaces with little data to parameterize models. In other words, the model will assume a value of Rs relatively similar to the closest multivariate space (likely from the northern hemisphere) when no data are available. Our results highlight that we as a community should be strategically targeting specific areas for measurement and better constrain global estimates.

Our results suggest that it is possible to decrease uncertainty in global Rs by using an optimized network design. Global Rs predictions using locations derived from the cLHS model is comparable with the global estimate from using SRDBv5. However, predictions using the
cLHS locations substantially decreased the global SD sum and mean C.V. These results open further discussion about the interpretation of regional-to-global estimates of terrestrial carbon dynamics derived from spatially biased environmental network designs (Villarreal et al. 2018, 2019, 2021). Not every location identified by the cLHS algorithm is able to realistically be measured. Many regions are poorly accessible due to lack of infrastructure, extreme climate, funds, remoteness, or a combination of these. However, the hypothetical optimization of locations indicates that better spatially distributed measurement sites would better represent the global covariate space.

Investigating the cLHS model Rh:Rs relationship reinforces that we may need a better distribution of measurements among certain land cover types to reduce uncertainty in our description of functional relationships. Comparing v3 and v5, it is evident that annual estimates, relationships, and the spatial distribution of uncertainty, should be considered for data interpretation and benchmarking purposes. As benchmarking is imperative to the validation of ESMs, we must embrace rigorous analyzes of representatives of environmental observatory networks as well as how datasets evolve with time and the implication in new estimates and uncertainties.

A previous study suggests that if Rs measurements were randomly distributed around the globe, only 591 measurements are needed to have a 5% relative error at a 95% confidence level—even less for a 10% relative error (Xu & Shang 2016), but this estimate may need to be reassessed with our current understanding of the spatial variability of Rs. Many of the points identified from the cLHS analysis were located in water-limited ecosystems such as the Sahara Desert and within the Australian outback. Although arid lands have lower Rs fluxes, they cover 30% of the world, representing a large multivariate space within the earth system and therefore
represent relevant locations for the cLHS analysis. Furthermore, there is strong evidence that hot-spots and hot-moments regulate the spatial and temporal variation of Rs in water-limited ecosystems (Leon et al 2014). Ultimately, our cLHS analysis should be considered as a guideline to compare how an unorganized bottom-up approach (i.e., the SRDB) compares to an optimized global design aimed to optimize the representation the global multivariate environmental space.

We highlight that is still unknown whether more points are needed in regions with less spatial variability, even if not adding points could lead to increased bias within the multivariate space. It has been proposed that Rs measured when soil temperature or air temperature reaches its MAT is able to well represent annual Rs (Bahn et al 2010; Jian et al. 2020). This suggests that for those inaccessible regions a combination of meteorological data with Rs rates at MAT could decrease uncertainty in global Rs estimates, potentially overcoming some of the logistical challenges associated with data collection in data sparse areas.

Moving forward, it is important to reduce uncertainty in regional-to-global Rs models. Our model output using a hypothetical optimized network shows that improving global results in lower model uncertainty. Consequently, we propose that it is important to have a better distributed network design, but it is particularly important to promote measurements within underrepresented biomes. Having better representation of global biomes can illuminate patterns and responses not currently well described in the scientific literature. Increased field measurements of Rh are also paramount to increasing our understanding of global Rs dynamics. Not only does this benefit bottom-up processes using ML, but also process-based models and other estimation methods. There is consensus that there is a pressing need to decrease uncertainty in Rs estimates, but the question remains as to which sources are driving the majority of uncertainty. That is, whether the spatial bias of the Rs observational data is driving the majority
of uncertainty, or if it is due to other factors such as sub-pixel heterogeneity or uncertainty in covariate data layers. However, these potential sources of error cannot be separated without additional measurements in targeted locations. Constraining the global Rs flux will require further intercomparisons and collaborations between scientists monitoring in situ Rs, the remote sensing community, and process-based modelers. This underscores the need for stronger interoperability, transparency, collaboration, and data sharing in environmental research.

**Acknowledgements**

This research was supported in part through the use of Information Technologies (IT) resources at the University of Delaware, specifically the high-performance computing resources. RV acknowledges support from a NASA Carbon Monitoring System grant (80NSSC18K0173).
References


Chapin, F. S., Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D.


from the C4MIP Model Intercomparison. J Climate 19:3337–3353. doi: 
10.1175/JCLI3800.1

Projections due to Carbon Cycle Feedbacks. J Climate 27:511–526. doi: 10.1175/JCLI-
D-12-00579.1

Pattern Recognition Letters, 31(14), 2225–2236. 
https://doi.org/10.1016/j.patrec.2010.03.014


Visualizing Statistical Learning with Plots of Individual Conditional Expectation. Journal 
of Computational and Graphical Statistics 24:1, 44-65, DOI: 
10.1080/10618600.2014.907095


Guevara, M., Olmedo, G. F., Stell, E., Yigini, Y., Aguilar Duarte, Y., Arellano Hernández, C., 
Arévalo, G. E., Arroyo-Cruz, C. E., Bolivar, A., Bunning, S., Bustamante Cañas, N., 
Cruz-Gaistardo, C. Ó., Rey Brina, J. C., Ramos, I. A., Pietro, P., Pereira, G., Osorio 
digital soil mapping: Country-specific soil organic carbon estimates across Latin 
America. https://doi.org/10.5194/soil-4-173-2018

Guevara, M., Arroyo, C., Brunsell, N., Cruz, C. O., Domke, G., Equihua, J., Etchevers, J., Hayes, 
D., Hengl, T., Ibelles, A., Johnson, K., Jong, B. de, Libohova, Z., Llamas, R., Nave, L., 

Data-Oriented Model. 
Nov 2018

doi: 10.5194/bg-12-4331-2015

Hoffman, F. M., Koven, C. D., Keppel-Aleks, G., Lawrence, D. M., Riley, W. J., Randerson, J. 


Tables

Table 1. Rs zonal statistics for soil respiration using SRDBv3 and SRDBv5.

<table>
<thead>
<tr>
<th>Land cover</th>
<th>Abbreviation</th>
<th>% land cover</th>
<th>v5 Mean (g C m(^{-2}))</th>
<th>v5 Sum (Pg C yr(^{-1}))</th>
<th>v3 Mean (g C m(^{-2}))</th>
<th>v3 Sum (Pg C yr(^{-1}))</th>
<th>v5 CV Mean (%)</th>
<th>v3 CV Mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evergreen Needleleaf Forest</td>
<td>ENF</td>
<td>3.59</td>
<td>565.18</td>
<td>2.17</td>
<td>545.08</td>
<td>2.09</td>
<td>71.92</td>
<td>61.70</td>
</tr>
<tr>
<td>Evergreen Broadleaf Forest</td>
<td>EBF</td>
<td>7.72</td>
<td>1305.22</td>
<td>18.01</td>
<td>1317.66</td>
<td>18.19</td>
<td>55.78</td>
<td>58.22</td>
</tr>
<tr>
<td>Deciduous Needleleaf Forest</td>
<td>DNF</td>
<td>1.64</td>
<td>388.65</td>
<td>0.59</td>
<td>323.41</td>
<td>0.49</td>
<td>85.89</td>
<td>69.40</td>
</tr>
<tr>
<td>Deciduous Broadleaf Forest</td>
<td>DBF</td>
<td>0.88</td>
<td>884.54</td>
<td>1.17</td>
<td>862.70</td>
<td>1.14</td>
<td>55.55</td>
<td>52.97</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>MF</td>
<td>8.39</td>
<td>687.52</td>
<td>6.9</td>
<td>631.90</td>
<td>6.42</td>
<td>59.23</td>
<td>53.55</td>
</tr>
<tr>
<td>Closed Shrublands</td>
<td>CSh</td>
<td>0.05</td>
<td>707.98</td>
<td>0.06</td>
<td>706.21</td>
<td>0.06</td>
<td>86.14</td>
<td>71.19</td>
</tr>
<tr>
<td>Open Shrublands</td>
<td>OSh</td>
<td>19.01</td>
<td>401.98</td>
<td>9.38</td>
<td>352.00</td>
<td>8.41</td>
<td>103.30</td>
<td>82.06</td>
</tr>
<tr>
<td>Woody Savannas</td>
<td>WSv</td>
<td>7.47</td>
<td>858.75</td>
<td>10.25</td>
<td>724.27</td>
<td>8.53</td>
<td>79.83</td>
<td>68.23</td>
</tr>
<tr>
<td>Savannas</td>
<td>Sv</td>
<td>5.53</td>
<td>1076.73</td>
<td>10.44</td>
<td>947.73</td>
<td>9.2</td>
<td>84.77</td>
<td>69.43</td>
</tr>
<tr>
<td>Grasslands</td>
<td>GL</td>
<td>14.84</td>
<td>544.07</td>
<td>10.78</td>
<td>438.51</td>
<td>8.94</td>
<td>91.47</td>
<td>78.19</td>
</tr>
<tr>
<td>Permanent Wetlands</td>
<td>PW</td>
<td>0.89</td>
<td>587.54</td>
<td>0.71</td>
<td>554.03</td>
<td>0.66</td>
<td>90.20</td>
<td>67.09</td>
</tr>
<tr>
<td>Croplands</td>
<td>Crop</td>
<td>8.37</td>
<td>757.52</td>
<td>9.07</td>
<td>670.42</td>
<td>8.01</td>
<td>67.06</td>
<td>60.29</td>
</tr>
<tr>
<td>Urban and Built-Up</td>
<td>Urb</td>
<td>0.26</td>
<td>820.63</td>
<td>0.3</td>
<td>720.40</td>
<td>0.27</td>
<td>71.06</td>
<td>65.81</td>
</tr>
<tr>
<td>Cropland/Natural Mosaic</td>
<td>Cr/Nat</td>
<td>4.90</td>
<td>955.38</td>
<td>7.34</td>
<td>833.56</td>
<td>6.4</td>
<td>63.12</td>
<td>60.75</td>
</tr>
<tr>
<td>Permanent Snow and Ice</td>
<td>Snow</td>
<td>4.83</td>
<td>482.92</td>
<td>1.38</td>
<td>377.56</td>
<td>1.08</td>
<td>103.29</td>
<td>96.34</td>
</tr>
<tr>
<td>Barren or Sparsely Vegetated</td>
<td>Bare</td>
<td>11.64</td>
<td>368.60</td>
<td>6.88</td>
<td>401.47</td>
<td>7.66</td>
<td>149.94</td>
<td>83.71</td>
</tr>
</tbody>
</table>
Table 2. Rh zonal statistics. We used the Rh derived from the Bond-Lamberty (2004) equation to compare SRDBv3 to SRDBv5.

<table>
<thead>
<tr>
<th>Land Cover</th>
<th>Abbreviation</th>
<th>% land cover</th>
<th>v5 Mean (g C m⁻²)</th>
<th>v5 Sum (Pg C yr⁻¹)</th>
<th>v3 Mean (g C m⁻²)</th>
<th>v3 Sum (Pg C yr⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evergreen Needleleaf Forest</td>
<td>ENF</td>
<td>3.59</td>
<td>342.12</td>
<td>1.3</td>
<td>333.24</td>
<td>1.27</td>
</tr>
<tr>
<td>Evergreen Broadleaf Forest</td>
<td>EBF</td>
<td>7.72</td>
<td>633.26</td>
<td>8.74</td>
<td>637.58</td>
<td>8.80</td>
</tr>
<tr>
<td>Deciduous Needleleaf Forest</td>
<td>DNF</td>
<td>1.64</td>
<td>260.86</td>
<td>0.39</td>
<td>227.86</td>
<td>0.35</td>
</tr>
<tr>
<td>Deciduous Broadleaf Forest</td>
<td>DBF</td>
<td>0.88</td>
<td>477.13</td>
<td>0.63</td>
<td>469.32</td>
<td>0.62</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>MF</td>
<td>8.39</td>
<td>395.93</td>
<td>3.95</td>
<td>371.82</td>
<td>3.74</td>
</tr>
<tr>
<td>Closed Shrublands</td>
<td>CSh</td>
<td>0.05</td>
<td>404.24</td>
<td>0.03</td>
<td>404.06</td>
<td>0.03</td>
</tr>
<tr>
<td>Open Shrublands</td>
<td>OSh</td>
<td>19.03</td>
<td>264.73</td>
<td>5.99</td>
<td>239.07</td>
<td>5.50</td>
</tr>
<tr>
<td>Woody Savannas</td>
<td>WSv</td>
<td>7.47</td>
<td>455.46</td>
<td>5.27</td>
<td>405.51</td>
<td>4.65</td>
</tr>
<tr>
<td>Savannas</td>
<td>Sv</td>
<td>5.54</td>
<td>545.43</td>
<td>5.27</td>
<td>498.29</td>
<td>4.81</td>
</tr>
<tr>
<td>Grasslands</td>
<td>GL</td>
<td>14.80</td>
<td>330.47</td>
<td>6.42</td>
<td>279.10</td>
<td>5.53</td>
</tr>
<tr>
<td>Permanent Wetlands</td>
<td>PW</td>
<td>0.89</td>
<td>342.79</td>
<td>0.39</td>
<td>328.33</td>
<td>0.37</td>
</tr>
<tr>
<td>Croplands</td>
<td>Crop</td>
<td>8.38</td>
<td>424.24</td>
<td>5.04</td>
<td>388.88</td>
<td>4.62</td>
</tr>
<tr>
<td>Urban and Built-Up</td>
<td>Urb</td>
<td>0.25</td>
<td>448.78</td>
<td>0.17</td>
<td>411.35</td>
<td>0.15</td>
</tr>
<tr>
<td>Cropland/Natural Mosaic</td>
<td>Cr/Nat</td>
<td>4.90</td>
<td>498.77</td>
<td>3.79</td>
<td>453.09</td>
<td>3.44</td>
</tr>
<tr>
<td>Permanent Snow and Ice</td>
<td>Snow</td>
<td>4.82</td>
<td>302.30</td>
<td>0.85</td>
<td>253.32</td>
<td>0.71</td>
</tr>
<tr>
<td>Barren or Sparsely Vegetated</td>
<td>Bare</td>
<td>11.66</td>
<td>250.23</td>
<td>4.68</td>
<td>266.62</td>
<td>5.06</td>
</tr>
</tbody>
</table>
Table 3. Rh:Rs Zonal Statistics. The Rh used was derived from the Bond-Lamberty 2004 equation. Abbreviation was marked with an asterisk (*) if the 95% confidence interval did not overlap between v3 and v5.

<table>
<thead>
<tr>
<th>Land Cover</th>
<th>Abbreviation</th>
<th>v5 Mean</th>
<th>v5 SD</th>
<th>v3 Mean</th>
<th>v3 SD</th>
<th>% change</th>
<th>% change SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evergreen Needleleaf Forest</td>
<td>ENF*</td>
<td>0.62</td>
<td>0.057</td>
<td>0.63</td>
<td>0.053</td>
<td>-0.81</td>
<td>7.24</td>
</tr>
<tr>
<td>Evergreen Broadleaf Forest</td>
<td>EBF*</td>
<td>0.49</td>
<td>0.032</td>
<td>0.49</td>
<td>0.032</td>
<td>0.23</td>
<td>-1.08</td>
</tr>
<tr>
<td>Deciduous Needleleaf Forest</td>
<td>DNF*</td>
<td>0.69</td>
<td>0.055</td>
<td>0.72</td>
<td>0.063</td>
<td>-5.35</td>
<td>-14.62</td>
</tr>
<tr>
<td>Deciduous Broadleaf Forest</td>
<td>DBF*</td>
<td>0.55</td>
<td>0.031</td>
<td>0.55</td>
<td>0.025</td>
<td>-0.39</td>
<td>17.79</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>MF*</td>
<td>0.59</td>
<td>0.044</td>
<td>0.60</td>
<td>0.047</td>
<td>-2.49</td>
<td>-6.87</td>
</tr>
<tr>
<td>Closed Shrublands</td>
<td>CSh</td>
<td>0.58</td>
<td>0.047</td>
<td>0.58</td>
<td>0.046</td>
<td>0.01</td>
<td>2.23</td>
</tr>
<tr>
<td>Open Shrublands</td>
<td>OSh*</td>
<td>0.69</td>
<td>0.078</td>
<td>0.72</td>
<td>0.094</td>
<td>-4.64</td>
<td>-19.76</td>
</tr>
<tr>
<td>Woody Savannas</td>
<td>WSv*</td>
<td>0.58</td>
<td>0.100</td>
<td>0.60</td>
<td>0.097</td>
<td>-3.49</td>
<td>3.07</td>
</tr>
<tr>
<td>Savannas</td>
<td>Sv*</td>
<td>0.53</td>
<td>0.075</td>
<td>0.55</td>
<td>0.075</td>
<td>-3.12</td>
<td>-0.29</td>
</tr>
<tr>
<td>Grasslands</td>
<td>GL*</td>
<td>0.64</td>
<td>0.081</td>
<td>0.69</td>
<td>0.111</td>
<td>-8.47</td>
<td>-37.95</td>
</tr>
<tr>
<td>Permanent Wetlands</td>
<td>PW*</td>
<td>0.64</td>
<td>0.094</td>
<td>0.65</td>
<td>0.086</td>
<td>-0.91</td>
<td>7.80</td>
</tr>
<tr>
<td>Croplands</td>
<td>Crop*</td>
<td>0.58</td>
<td>0.049</td>
<td>0.59</td>
<td>0.045</td>
<td>-2.96</td>
<td>8.30</td>
</tr>
<tr>
<td>Urban and Built-Up</td>
<td>Urb*</td>
<td>0.57</td>
<td>0.058</td>
<td>0.58</td>
<td>0.049</td>
<td>-2.70</td>
<td>15.82</td>
</tr>
<tr>
<td>Cropland/Natural Mosaic</td>
<td>Cr/Nat*</td>
<td>0.55</td>
<td>0.056</td>
<td>0.56</td>
<td>0.054</td>
<td>-3.27</td>
<td>3.63</td>
</tr>
<tr>
<td>Permanent Snow and Ice</td>
<td>Snow*</td>
<td>0.66</td>
<td>0.087</td>
<td>0.71</td>
<td>0.093</td>
<td>-6.58</td>
<td>-6.91</td>
</tr>
<tr>
<td>Barren or Sparsely Vegetated</td>
<td>Bare*</td>
<td>0.70</td>
<td>0.066</td>
<td>0.69</td>
<td>0.074</td>
<td>2.04</td>
<td>-13.40</td>
</tr>
</tbody>
</table>
**Figure Legends**

**Figure 1.** Measurement sites of both the SRDB v3 (dark circles) annual Rs observations, and the new measurement locations included within v5 (white circles) that were utilized in this study. The background is a 2012 land cover classification map at a spatial resolution of 5 minutes from the International Geosphere-Biosphere Program (IGBP).

**Figure 2.** Comparison of the model-derived mean and uncertainty between v3 and v5 of the SRDB. 2a and 2b are the mean and C.V. of v3, while 2c and 2d are the mean and C.V. of v5.

**Figure 3.** Difference maps comparing both the a) mean and b) C.V. of v3 vs v5. v3 values were subtracted from v5 values to get the resulting maps. Values less than zero indicate that v5 is underestimating in an area compared to v3, while values greater than zero indicate that v5 is overestimating in comparison to v3.

**Figure 4.** Global mean percent change (a) derived from the mean Rs difference map, along with the mean percent change plotted within climate space, i.e., the mean annual temperature and mean annual precipitation (b).

**Figure 5.** Partial dependence plots of the six covariates used to model Rs. The x-axis represents the values of the respective covariates, while the y-axis represents the response of each variable within the QRF model.

**Figure 6.** Visualization of 3000 points chosen by the cLHS analysis to represent an optimized hypothetical network of global Rs measurements. These points represent the optimized locations for observations within a multivariate environmental space used to model global Rs.
Figure 7. Modelled global Rs (a) and global C.V. (b) using a hypothetical optimized global distribution of Rs measurements.
Figure 1

- SRDB V5 New Points
- SRDB V3 Rs Observations
Figure 2

(a) Predicted Annual Rs

(b) C.V. (%)
Figure 3

(a) Rs Mean Difference

-1120
-25
0
20
71
1653

(g C m$^{-2}$ yr$^{-1}$)

(b) Rs C.V. Difference

-177
-4
0
9
31
358

(%)
Figure 4
Figure 5

(a) Relationship between MAT (°C) and f(MAT) for V5 SRDB, V3 SRDB, and the line of regression.

(b) Relationship between EVI and f(EVI) for V5 SRDB, V3 SRDB, and the line of regression.

(c) Relationship between WMP (mm) and f(WMP) for V5 SRDB, V3 SRDB, and the line of regression.

(d) Relationship between SpMP (mm) and f(SpMP) for V5 SRDB, V3 SRDB, and the line of regression.

(e) Relationship between SMP (mm) and f(SMP) for V5 SRDB, V3 SRDB, and the line of regression.

(f) Relationship between FMP (mm) and f(FMP) for V5 SRDB, V3 SRDB, and the line of regression.
Figure 6

- Idealized cLHS points