Journal of Geophysical Research - Biogeosciences

Supporting Information for

Increasing functional diversity in a global land surface model illustrates uncertainties related to parameter simplification

Ethan E. Butler[1], Kirk R. Wythers[1], Habacuc Flores-Moreno[1,2], Daniel M. Ricciuto[3], Abhirup Datta[4], Arindam Banerjee[5], Owen K. Atkin[6,7], Jens Kattge[8,9], Peter E. Thornton[3], Madhur Anand[10], Sabina Burrascano[11], Chaeho Byun[12], J.H.C. Cornelissen[13], Estelle Forey[14], Steven Jansen[15], Koen Kramer[16,17], Vanessa Minden[18,19], and Peter B. Reich[1,20,21]

[1]Department of Forest Resources, University of Minnesota, St. Paul, MN 55108
[2]Biological Sciences, George Washington University, Washington, DC, 20052, USA
[3]Environmental Sciences Division, Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN.
[4]Department of Biostatistics, Johns Hopkins University, Baltimore, MD 21205
[5]Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455
[6]Australian Research Council Centre of Excellence in Plant Energy, Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia;
[7]Division of Plant Sciences, Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia
[8]Max Planck Institute for Biogeochemistry, Jena, Germany.
[9]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Germany
[10]School of Environmental Sciences, University of Guelph 50 Stone Road East, N1G 2W1
[11]Department of Environmental Biology, Sapienza University of Rome, Rome, Italy
[12]Department of Biological Sciences and Biotechnology, Andong National University, Andong, 36729, Korea
[13]Systems Ecology, Department of Ecological Science, Vrije Universiteit, 1081 HV Amsterdam, The Netherlands
[14]Normandie University, UNIROUEN, INRAE, ECODIV, 76000 Rouen, France.
[15]Institute of Systematic Botany and Ecology, Ulm University, Ulm, Germany
[16]Wageningen University, Forest Ecology and Forest Management group, Wageningen, the Netherlands
[17]Land Life Company, Amsterdam, the Netherlands
[18]Department of Biology, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
[19] Institute of Biology and Environmental Sciences, Carl von Ossietzky-University Oldenburg, Carl von Ossietzky-Straße 9-11, 26111 Oldenburg, Germany
[20]Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales 2753, Australia.
[21]Institute for Global Change Biology, and School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109, USA

Contents of this file

Tables S1 to S3
Text S1
Figures S1 to S11

Introduction

This supplemental contains additional details regarding the traits used in this analysis, the differences with flux tower data, the statistical methods employed, and the relationship between traits and the carbon cycle.

SI Tables

Table S1. PFT trait means

PFT	SLA (m^2 kgC^{-1})	LCN (gC gN^{-1})	LLS (years)
Needleleaf evergreen tree - temperate	10.2	44.2	3.6
Needleleaf evergreen tree - boreal	10.6	40.6	3.9
Needleleaf deciduous tree - boreal	17.1	30.6	0.5
Broadleaf evergreen tree – tropical	23.7	26.4	1.7
Broadleaf evergreen tree - temperate	17.7	29.2	1.8
Broadleaf deciduous tree - tropical	28.8	22.6	1.0
Broadleaf deciduous tree - temperate	28.4	24.2	0.6
Broadleaf deciduous tree - boreal	32.9	21.2	0.4
Broadleaf evergreen shrub - temperate	15.3	31.6	1.4
Broadleaf deciduous shrub - temperate	27.5	23.4	0.5
Broadleaf deciduous shrub - boreal	29.5	23.0	0.4
C_3 arctic grass	33.8	21.1	0.2
C_3 grass	37.5	22.2	0.5
C_4 grass	27.8	34.4	0.4

Table S2. Trait covariance matrix

	SLA	LCN	LLS
SLA	0.30	-0.10	-0.21
LCN	-0.10	0.16	0.21
LLS	-0.21	0.21	0.89

Table S3. Mean GPP offset at each flux site (Flux Tower – Model). The model configuration with the best match to the flux tower data are in **bold** in each row.

Site	Default (gC m^2 yr^{-1})	Distribution (gC m^2 yr^{-1})	Updated Mean (gC m^2 yr^{-1})
Shrub-Bo1	76.8	**-20.1**	-25.5
Shrub-Bo2	**-257.5**	-448.5	-263.9
Tree-Bo1	**-143.2**	246.5	251.2
Tree-Bo2	**-21.6**	411.5	408.0
Tree*-Te1	-411.6	**-238.7**	-250.5
Tree*-Te2	171.2	**49.6**	171.8
Tree-Te1	181.5	189.2	**153.3**
Tree-Te2	**1869.1**	1982.4	1884.6
Tree-Te3	-240.5	**-155.2**	-206.9
Tree*-Te3	**-12.7**	80.8	39.9
Grass-Ar	-28.8	**-24.8**	-29.6
Shrub-Ar	**55.6**	56.9	60.7
Grass*-Ar	**392.8**	428.6	434.7
Tree-Tr1	**27.2**	890.1	1138.3
Tree-Tr2	-549.2	**319.4**	586.5
Best Matches	**8**	**6**	**1**

Text S1: Supplemental Methods

Determination of the size of the trait distribution

Subsamples of trait values drawn from TRY (Adler et al. 2004, Adriaenssens 2012, Atkin, et al. 1997, 1999, Auger and Shipley 2012, Bahn et al. 1999, Bakker et al 2005, 2006, Baraloto et al 2010, Beckmann et al 2012, Blonder et al 2012, 2013, 2015, Bond-Lamberty et al 2002, Brown et al 2011, Burrascano et al 2015, Butterfield and Briggs 2011, Byun et al. 2012, Campbell et al. 2007, Campetella et al 2011, Carswell et al 2000, Cavender-Bares et al 2006, 2007, Cerabolini et al. 2010, Cernusca et al, Chen et al 2011, Choat et al 2012, Coomes et al 2008, Cornelissen 1996, Cornelissen et al 1996, 2004, 2006, Cornwell et al 2007, 2008, Craine et al. 2005, 2009, 2011, 2012, Craen et al. 2007, Demey et al 2013, Díaz et al 2004, Domingues et al. 2007, 2010, Dunbar-Co et al. 2009, Fitter and Peat 1994, Fonseca et al 2000, Frenette-Dussault et al. 2012, Freschet et al. 2010, Fyllas et al. 2009, Garnier et al. 2007, Givnish et al. 2004, Gutiérrez and Huth 2012, Guy et al 2013, Han et al 2005, Hao et al 2010, Hickler 1999, Hoof et al. 2008, Kattge et al. 2009, Kazakou et al. 2006, Kerkhoff et al. 2006, Kichenin et al. 2013, Kleyer et al. 2008, Kraft et al. 2008, Kurokawa and Nakashizuka 2008, Laughlin et al. 2010, 2011, Louault et al. 2005, Loveys et al. 2003, Martin et al. 2007, Medlyn et al. 1999, Meir et al 2002, Meir and Levy 2007, Messier et al. 2010, Meziane and Shipley 1999, Milla and Reich 2011, Minden et al. 2012, Minden and Kleyer 2011, Nakahashi et al. 2005, Niinemets 2001, Ogaya and Penuelas 2003, Onoda et al. 2011, Ordonez et al. 2010, Pahl et al. 2013, Peco et al. 2005, Penuelas et al. 2010, Pierce et al. 2007a, 2007b, 2012, 2013, Pillar and Sosinski 2003, Poorter et al. 2009, Powers and Tiffin 2012, Prentice et al. 2011, Preston et al. 2006, Price and Enquist 2007, Pyankov et al. 1999, Quero et al. 2008, Quested et al. 2003, Reich et al. 2008, 2009, Sack 2004, Sack et al. 2003, 2005, 2006, Sandel et al. 2011, Scherer-Lorenzen et al. 2007, Scoffoni et al. 2008, Shiodera et al. 2008, Shipley 1995, 2002, Shipley and Lechowicz 2000, Shipley and Vu 2002, Spasojevic and Suding 2012, Swaine 2007, Tucker et al. 2011, van Bodengom et al. 2008, Vergutz et al. 2012, Vile 2005, Waite and Sack 2010, Han et al. 2011, Williams et al. 2012, Willis et al. 2010, Wilson et al. 2000, Wirth and Lichstein 2009, Wright et al. 2004, 2007, 2011, Wright and Sutton-Grier 2012, Xu and Baldocchi 2003, Yguel et al. 2011) and used in ELM site runs were drawn from the joint trait distribution using a Latin Hypercube design (LHD) (Carnell 2018) with 100 draws used per PFT. Each draw comprises a triplet of trait values, referred to here as a subPFT to indicate that trait values are nested within the extant PFT categories. The number of draws (100) and the LHD design were determined following a simulation analysis of two sites: University of Michigan Biological Station (with PFT type Temperate Broadleaf Deciduous Trees) and Santarem, Brazil [km 67] (Tropical Broadleaf Evergreen Trees). At each of these sites 1000 subPFT samples were taken from the joint trait distribution and the model was run using each of these 1000 samples as input values. These trait combinations (x) and the temporal mean GPP output (f(x)) from the model were then used to emulate the model f() via a Gaussian Process emulator (Sacks et al. 1989) using the DiceKriging R package (Roustant et al. 2012). This allows us to estimate $f(x_0)$ for any new trait combination (subPFT) value x_0 without a full run of ELM, see Eqn. 1. This practice of using GPs to emulate deterministic computer models such as climate models or similar physics/chemistry transport models of environmental processes is well documented (Kennedy and O'Hagan 2001, Rougier et al. 2009, Castruccio et al. 2014). The advantage of emulation is that the time-intensive climate model only runs with the 1000 subPFTs from which we can estimate $f(x_i)$ for a large number (N=100,000) of trait combinations and obtain a simulated GPP based on the much larger distribution

$$GPP_{dist} = E\big(f(X)\big) \approx \frac{1}{N}\sum_{i=1}^{N} f(x_i)\,. (1)$$

Using the larger GPP_{dist} as an estimate of the "true" GPP we can evaluate the different trait modeling strategies explored in the paper. In Figures S2 and S3, we compare the GPP_{dist} (yellow line) to the GPP of the updated mean (black line), and the distribution mean of GPP values using the sample of trait values. The latter was done using subsets of trait draws of sizes 20, 50, 100, 200 and 500. We also evaluated the use of Latin Hypercube design (LHD) draws against strictly random draws from a multivariate normal. The mean GPP of both the LHD and random draws are centered around the simulated truth (Figs. S2-S3), while the LHD design has better precision (less variation) than the random draws, as is expected (Stein 1987). Across years, at both sites (see Fig. S2-S3), 100 subPFTs using the LH design were consistently within a 5% bound of the GPP as predicted by the runs with 1000 draws, this was taken as adequate accuracy and a computationally manageable sample size to be used for the remainder of the analysis.

Fig. S1. Combined relationship of inverse SLA and LCN with GPP. Note the r^2 reported for each site is the squared Pearson's correlation, which varies from 0.03 to 0.91 with a mean value of 0.49. Panels are ordered from coldest to hottest by mean annual temperature in Table S1.

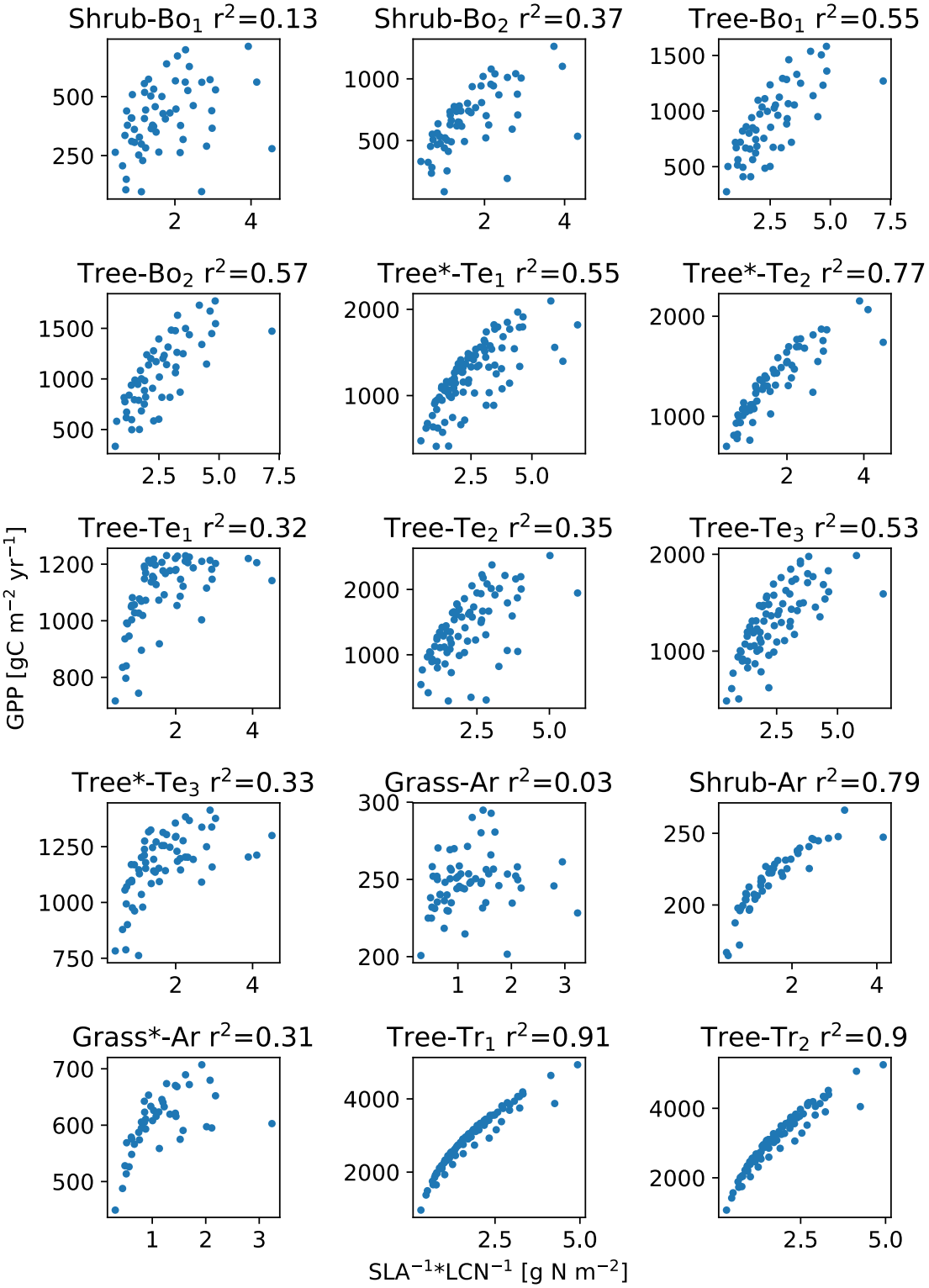
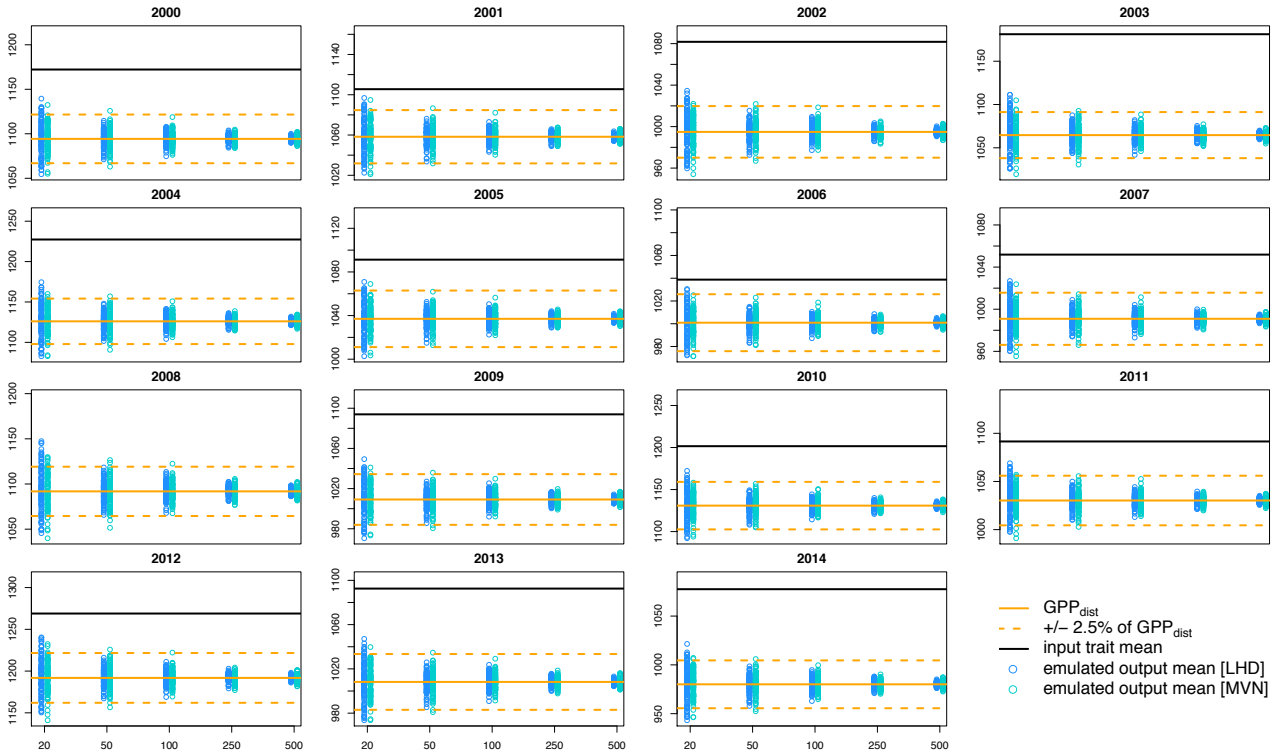


Fig. S2. Gaussian Process emulated GPP at the University of Michigan Biological Station Site. Each panel is a single year of the simulation. The yellow solid line is GPP_{dist} from the full 100,000 sample simulation with the dashed line indicating +/- 2.5%. The solid black line is the single simulation using the updated trait mean as input. The blue and green circles represent resampling of the trait distribution using the latin hypercube (LHD) and random multivariate normal (Rand) schemes and the output GPP from the emulated model. Note that after 20 samples the latin hypercube samples are consistently closer to the full sample mean and that the 100 samples are within the +/- 2.5% bounds.



Note that the GPP of the input mean was well away from the GPP_{dist} and even outside the 5% band around GPP_{dist}. Interestingly, the nature of this deviation was different across sites. In Michigan, the mean GPP overestimated the true GPP (Fig. S1), whereas in Brazil it underestimated (Fig. S2).

Fig. S3. Gaussian Process emulated GPP at the Brazil, Santarem [km 67] site. Each panel is a single year of the simulation. The yellow solid line is GPP_{dist} from the full 100,000 sample simulation with the dashed line indicating +/- 2.5%. The solid black line is the single simulation using the trait mean as input. The blue and green circles represent resampling of the distribution using the latin hypercube (LHD) and rand om multivariate normal (Rand) schemes and the output GPP from the emulated model. Note that after 20 samples the latin hypercube samples are consistently closer to the full sample mean and that the 100 samples are within the +/- 2.5% bounds.

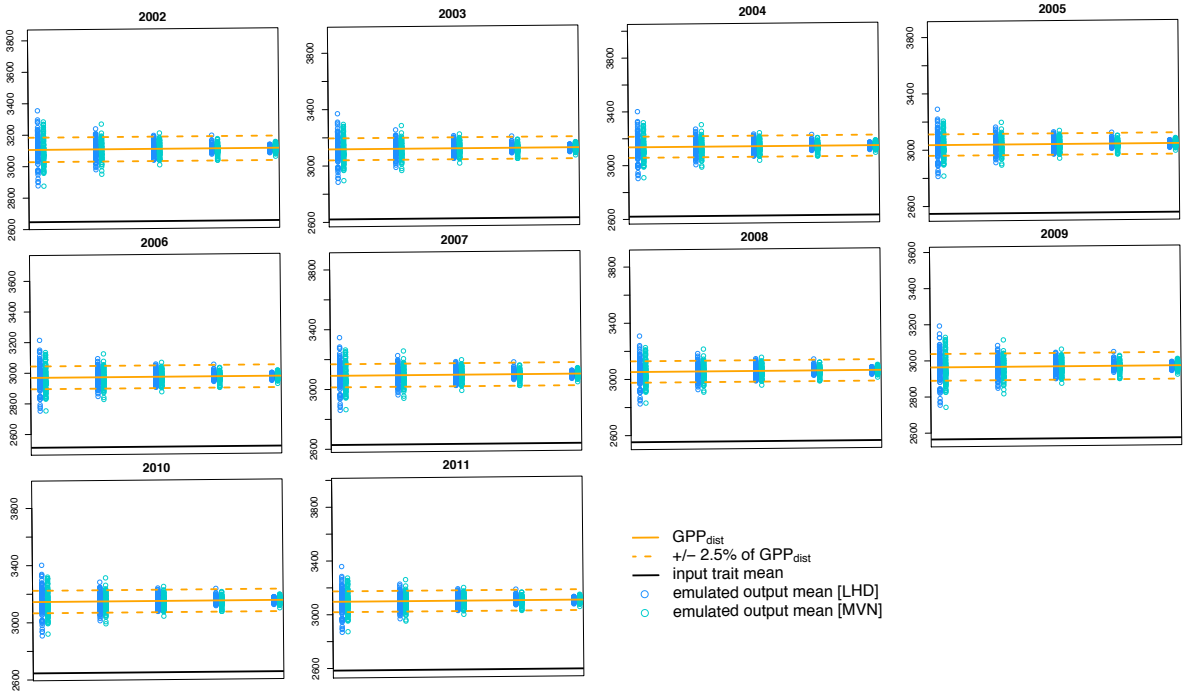


Fig. S4. Site Dominant PFT trait distributions. The full trait distribution estimated for each sites' dominant PFT is indicated in blue. The implemented distribution, in red, has had too long leaf lifespans and runs which produced 0 GPP, often due to low leaf lifespan, removed. We focus on the implemented distribution as we recognize that the longest leaf lifespans generated are implausibly long (>10 years), however even when the response is restricted to shorter lifespans (i.e. ≈4 or fewer years) the LAI values can still reach implausible proportions (LAI ≈ 100). There are several plausible sources for this breakdown: a disconnect in the distribution of carbon and nitrogen within the canopy, allocation routines, or the SLA-LAI gradient within the canopy.

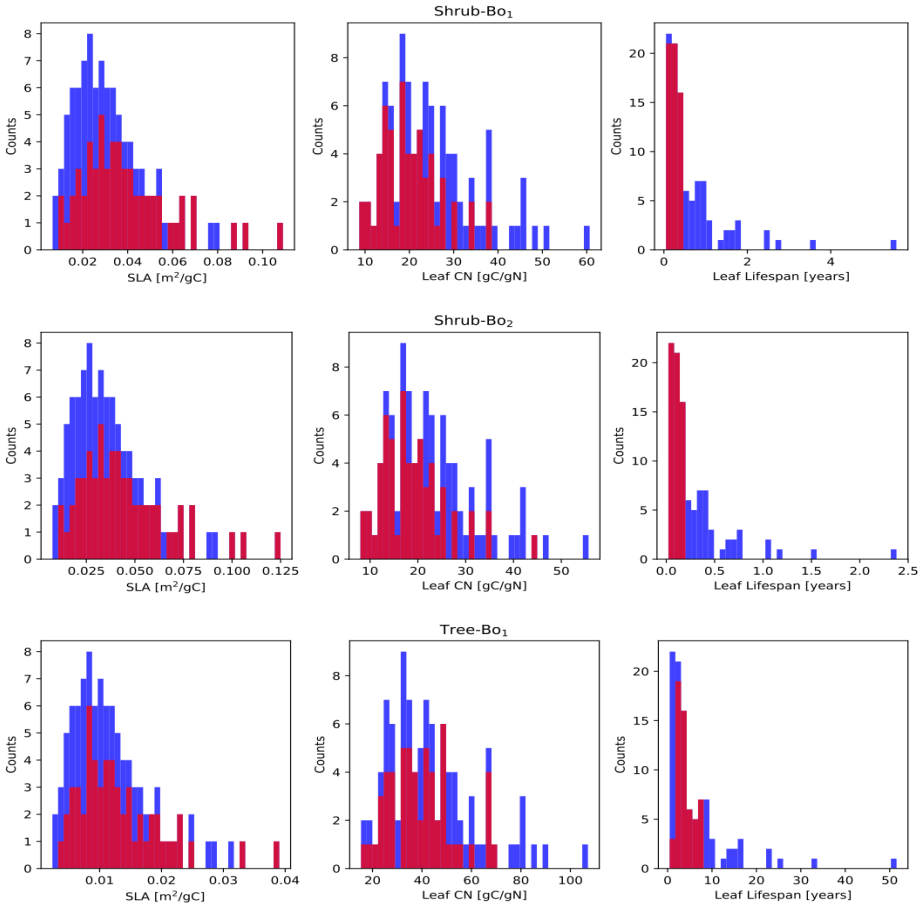


Fig. S4. continued

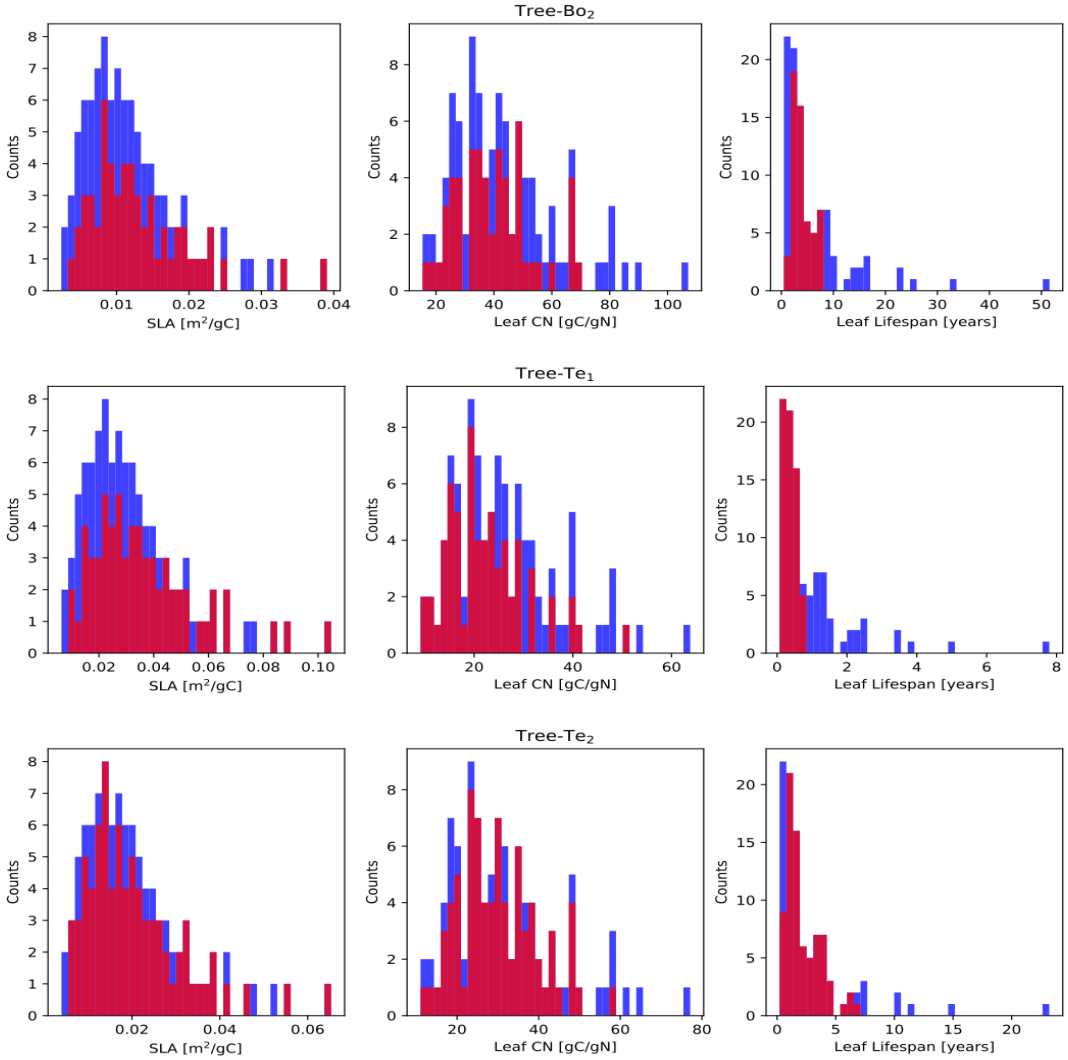


Fig. S4. continued

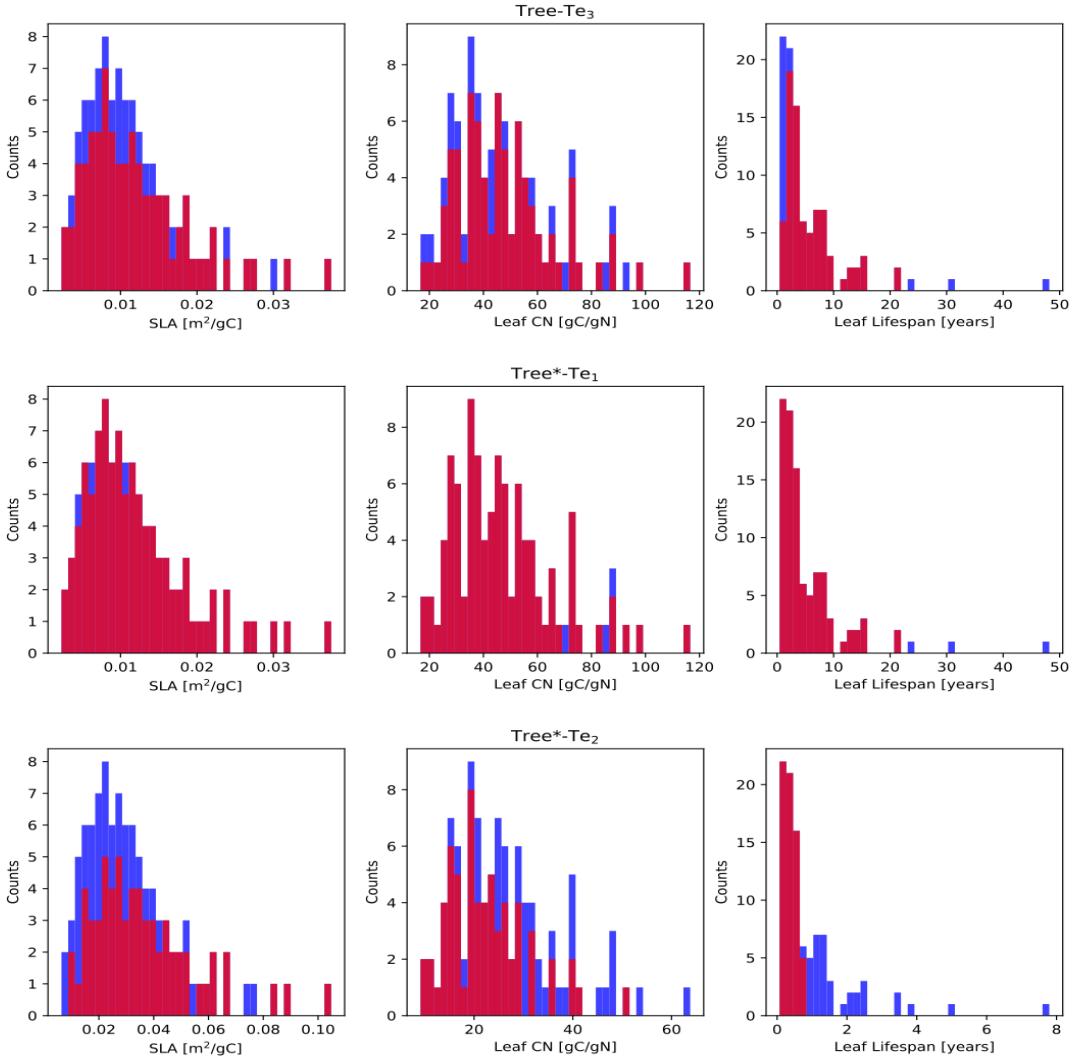


Fig. S4. continued

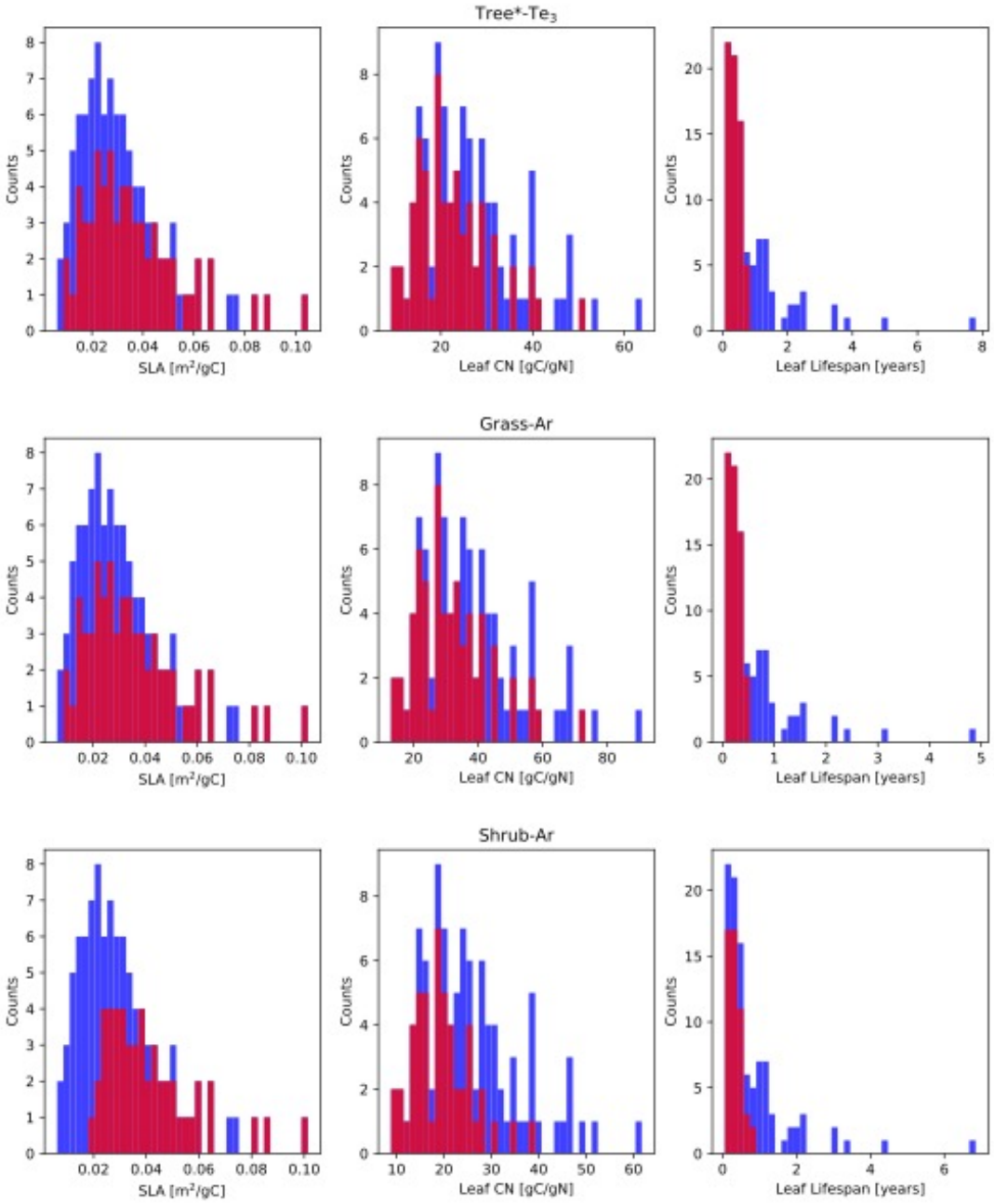


Fig. S4. continued

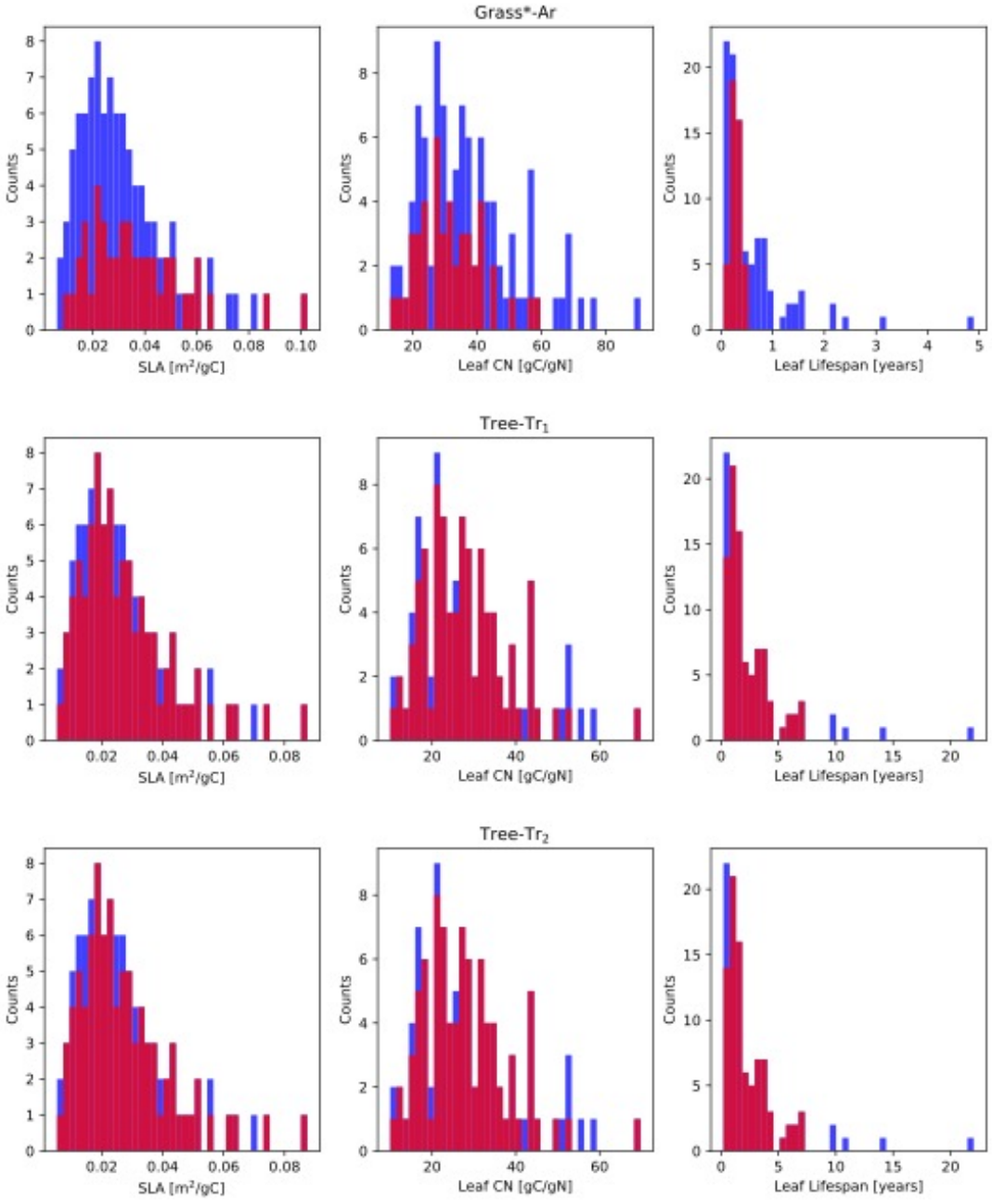


Fig. S5. LAI at tropical sites is unrealistic for long leaf life spans. The top row shows the relationship between LAI and the three traits: SLA, LCN, and Leaf Lifespan at the Brazil, Santarem [km 67] site (Tree-Tr_2) while the bottom row shows the same at the French Guiana site (Tree-Tr_1). The X's indicate values that were removed by the censoring exercise. Red and green points are the model default and mean of traits, respectively. Overall, this analysis supports ELM's ability to simulate C cycle processes using a wide range of trait input values (Fig. 3), but there are issues with leaf lifespan that will need to be corrected before implementing the full range of trait space. Note the unrealistic LAI for long leaf lifespans marked by a nearly linear relationship between LLS and LAI (note the y axis limit below). Further, models with short-lived leaves would often fail to grow. These values did not cause the model to crash, but to equilibrate to zero values for GPP, NPP, etc. The trait combinations that produced zero output may be a result of not varying the trait distributions in accord with the local environment, a promising area for future work. The leaf lifespan issues were restricted to evergreen PFTs because deciduousness is governed by a distinct phenology module that largely overrides the lifespan trait value.

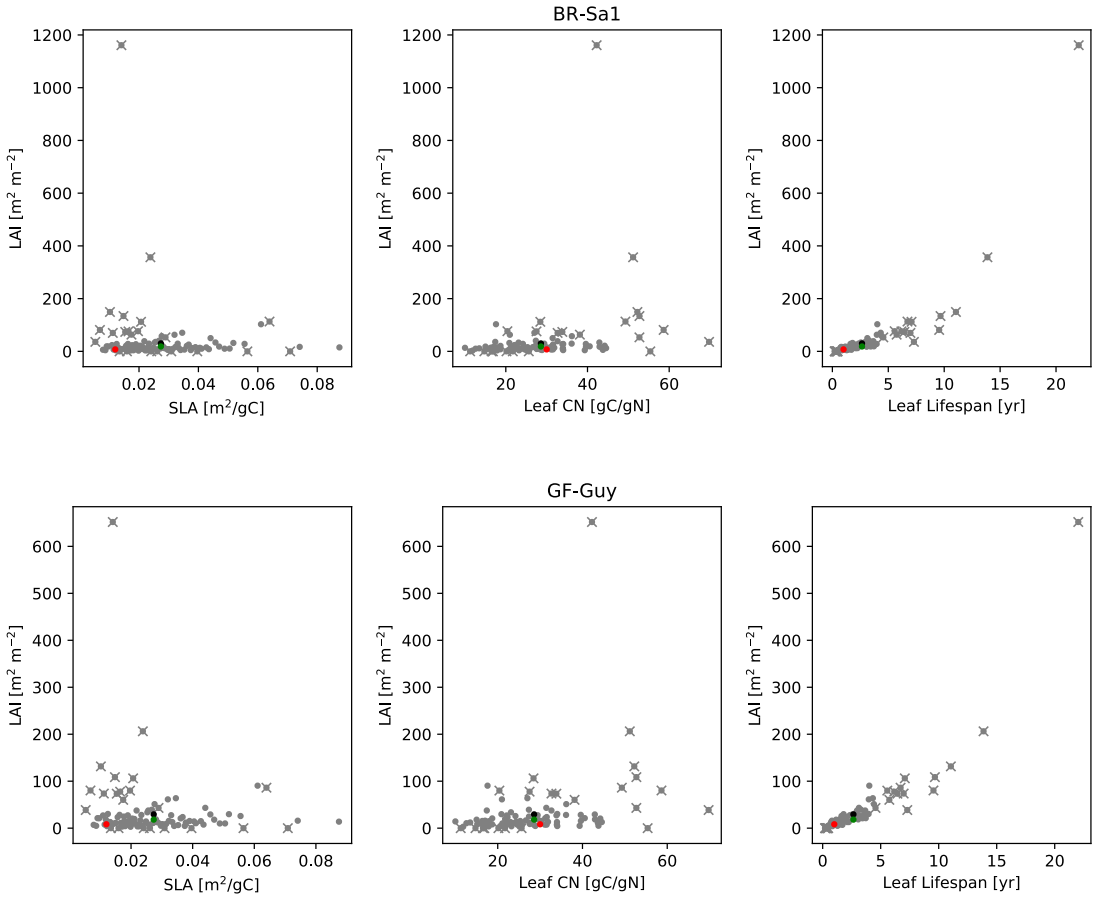


Fig. S6. Productivity density plots. For comparison, the three primary elements of the terrestrial carbon cycle are plotted together with gross primary production in black, net primary production in green and maintenance respiration in red. Each histogram is standardized into a z-score by removing the mean and dividing by the standard deviation. Most sites show very similar structure across the distributions, but temperate sites and arid sites show notable variation across the three metrics. Distributions that are non-normal according to the D'Agostino K^2 test are shown as dashed lines, though note that the test, which evaluates skewness and kurtosis, may miss some non-normal distributions among the arid sites due to bi-modality.

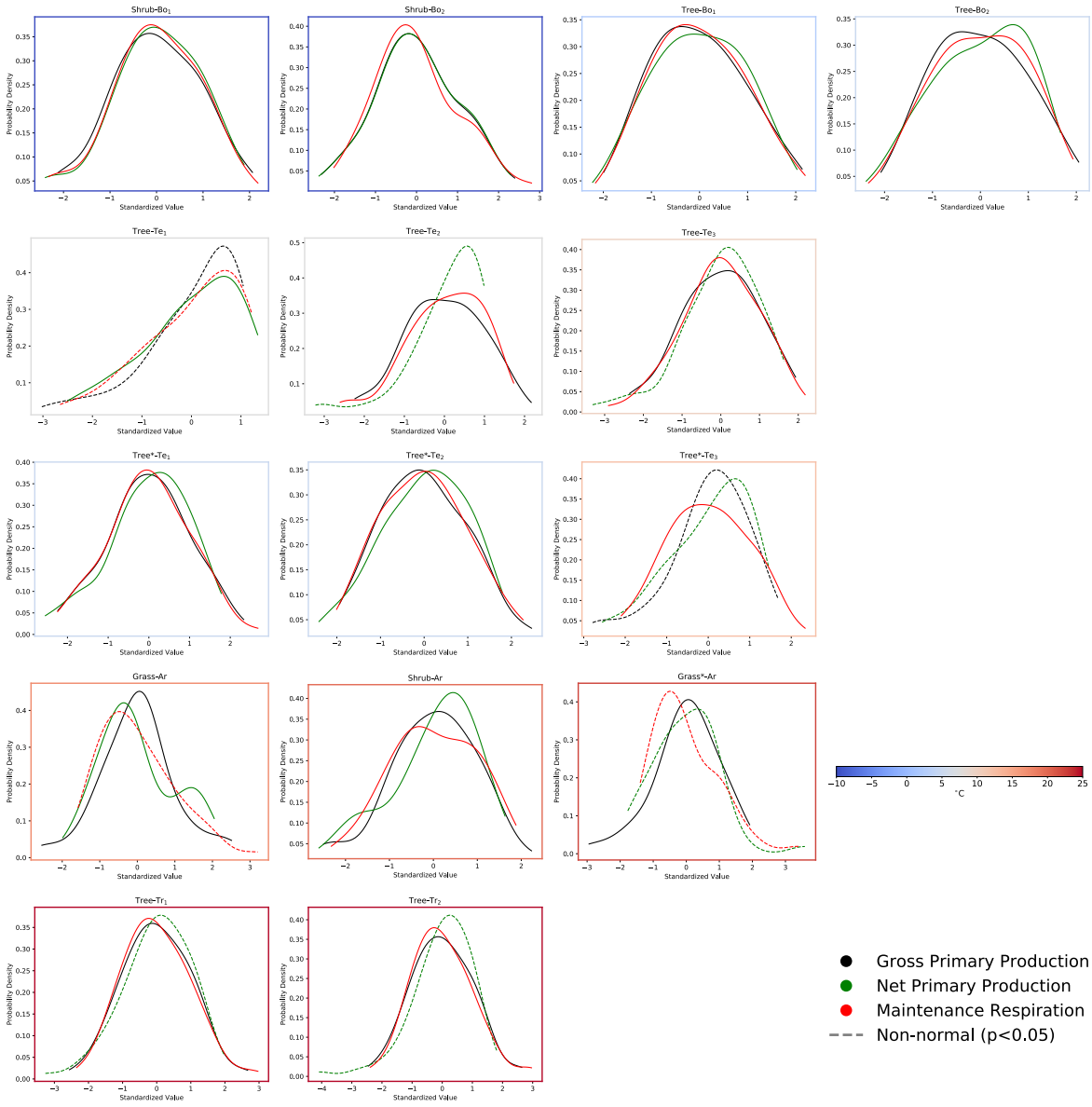


Fig. S7. Net Primary Production (NPP) distribution and central estimates at the 15 sites ordered from coldest to hottest within climatic space. The range of NPP values displayed on the x-axis is determined by the climatic region: boreal/temperate, arid, or tropical. Compare with the GPP and MR distributions (Fig. 2, S2, S4).

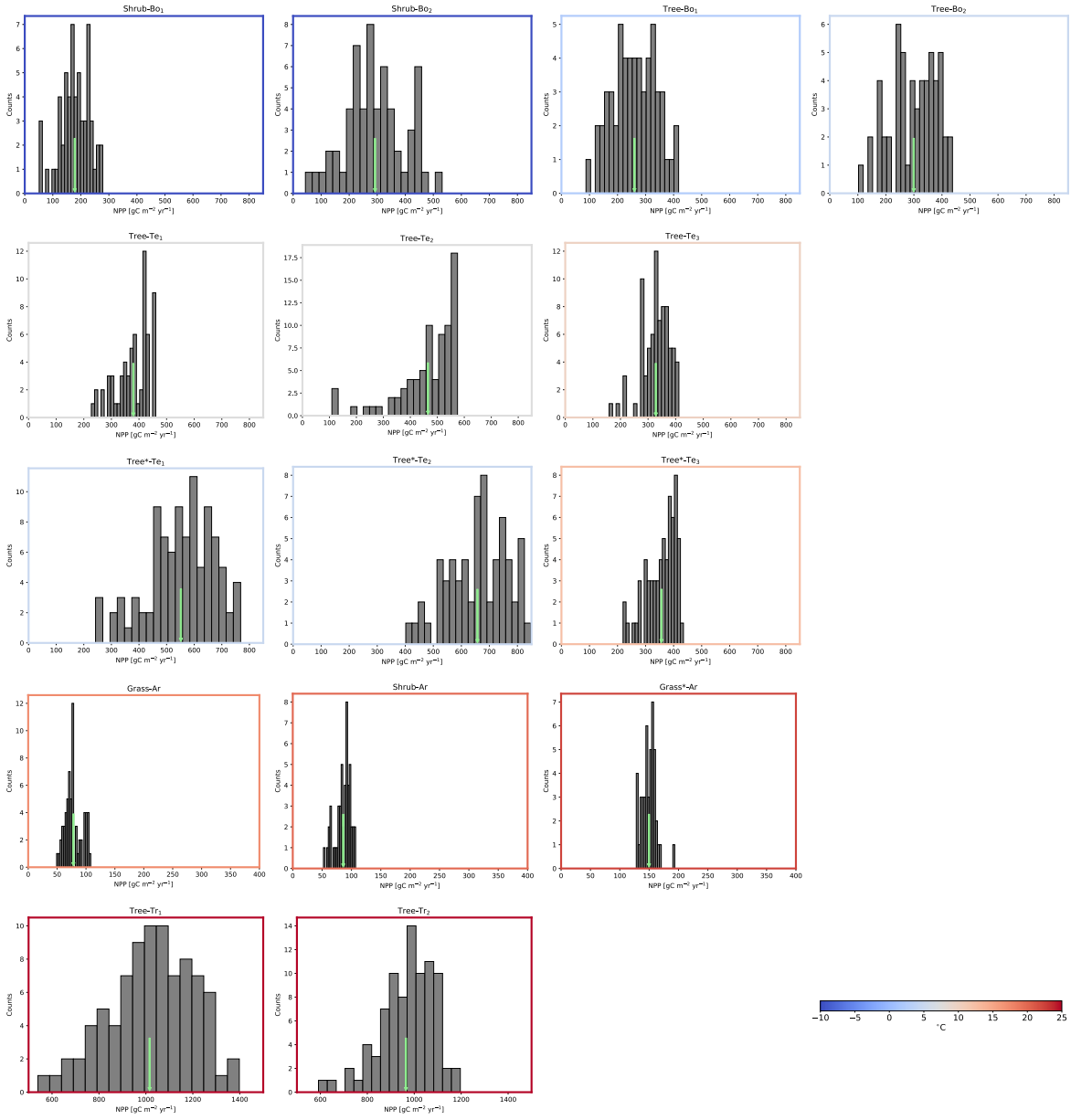


Fig. S8. Maintenance Respiration (MR) distribution and central estimates at the 15 sites ordered from coldest to hottest within climatic space. The range of NPP values displayed on the x-axis is determined by the climatic region: boreal/temperate, arid, or tropical. Compare with the GPP and MR distributions (Fig. 2, S2, S3).

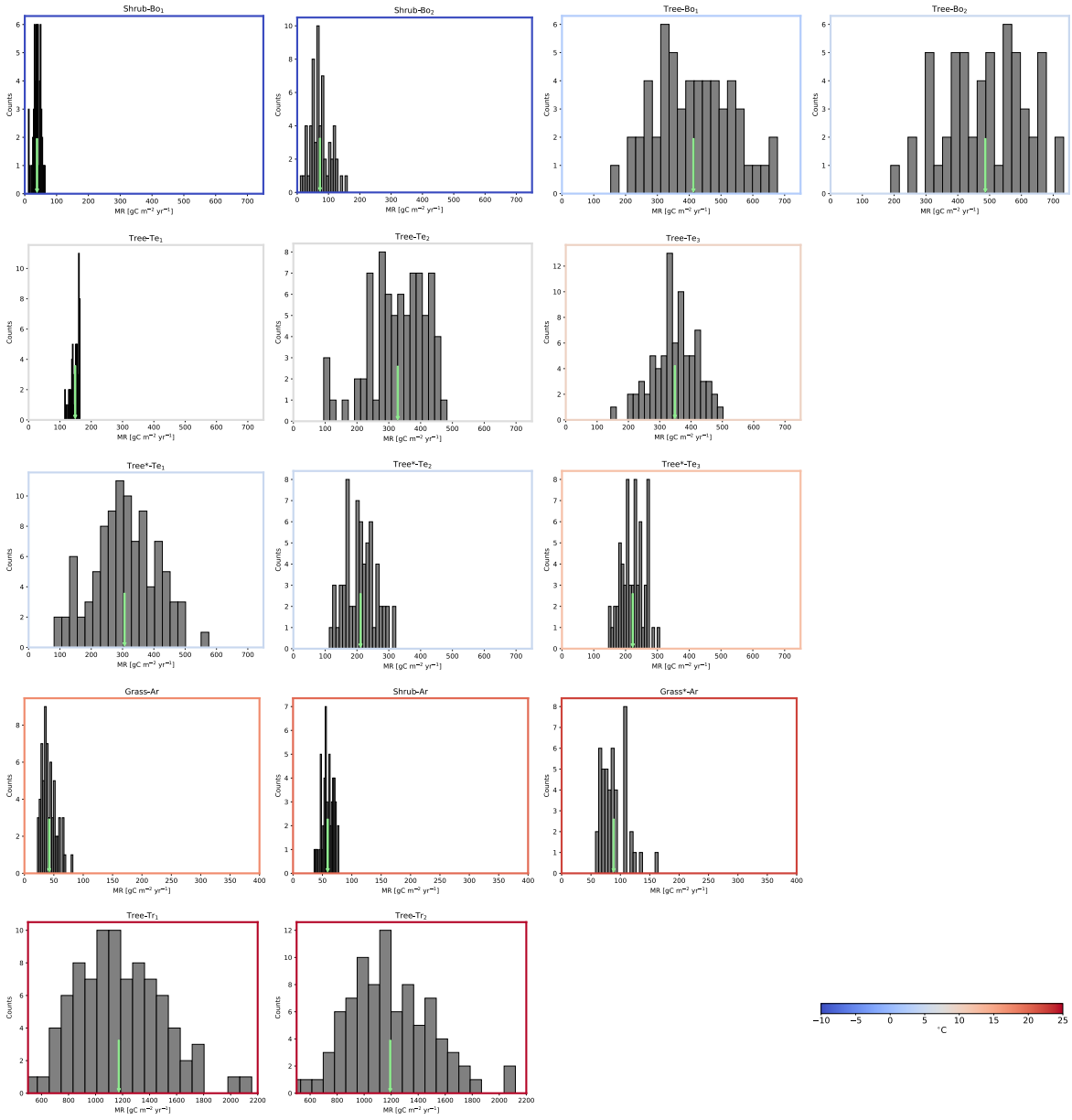


Figure S9. Nonlinear trait-GPP relationships at all sites. Black points are individual trait simulations, red filled circles are the distribution mean, blue stars are the updated mean, and green stars are the default trait values. The red lines are the nonlinear fit described by Eqn. 2. The r^2 values above each panel are a pseudo-R^2 calculated as the squared Pearson's correlation.

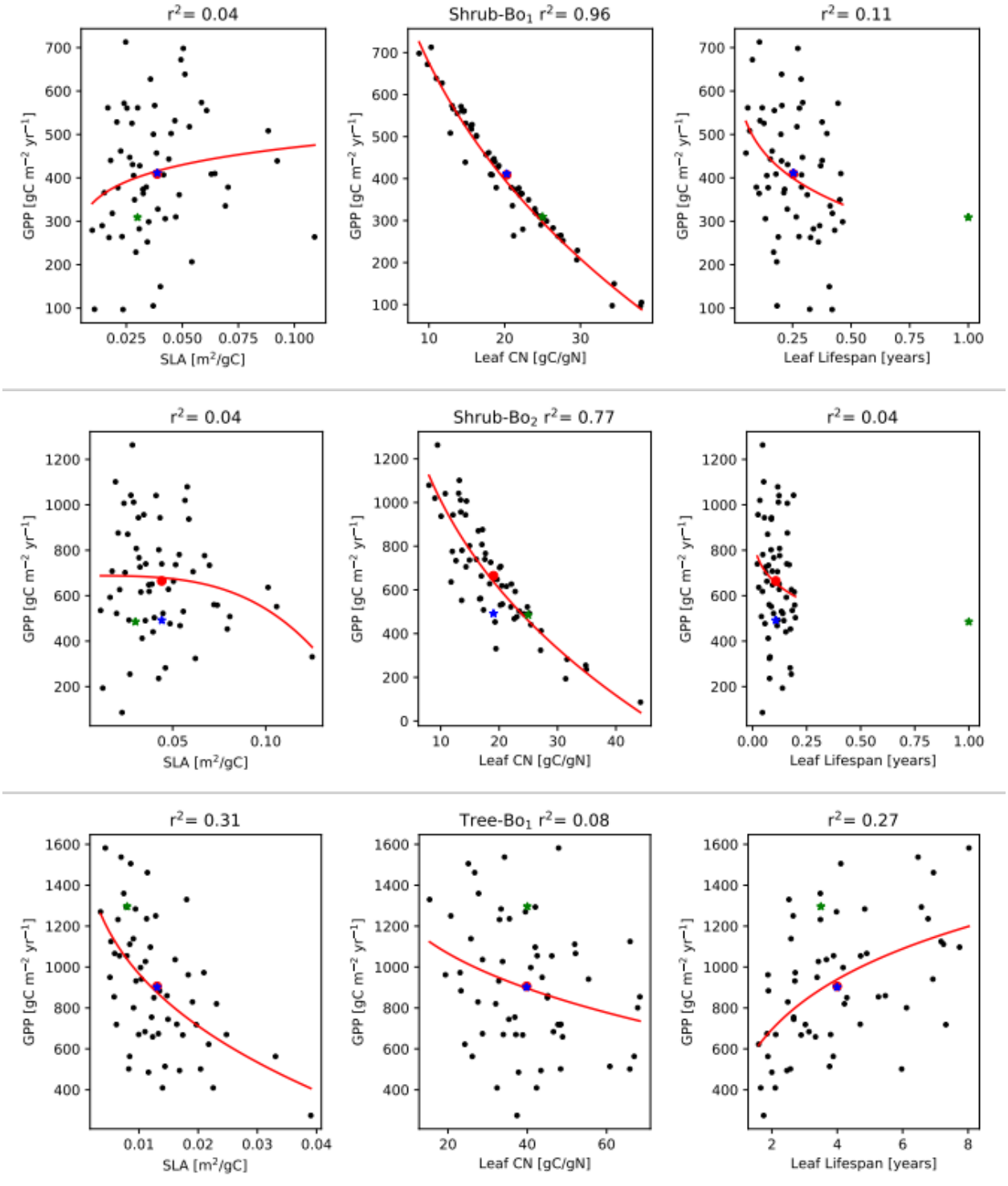


Fig. S9. continued

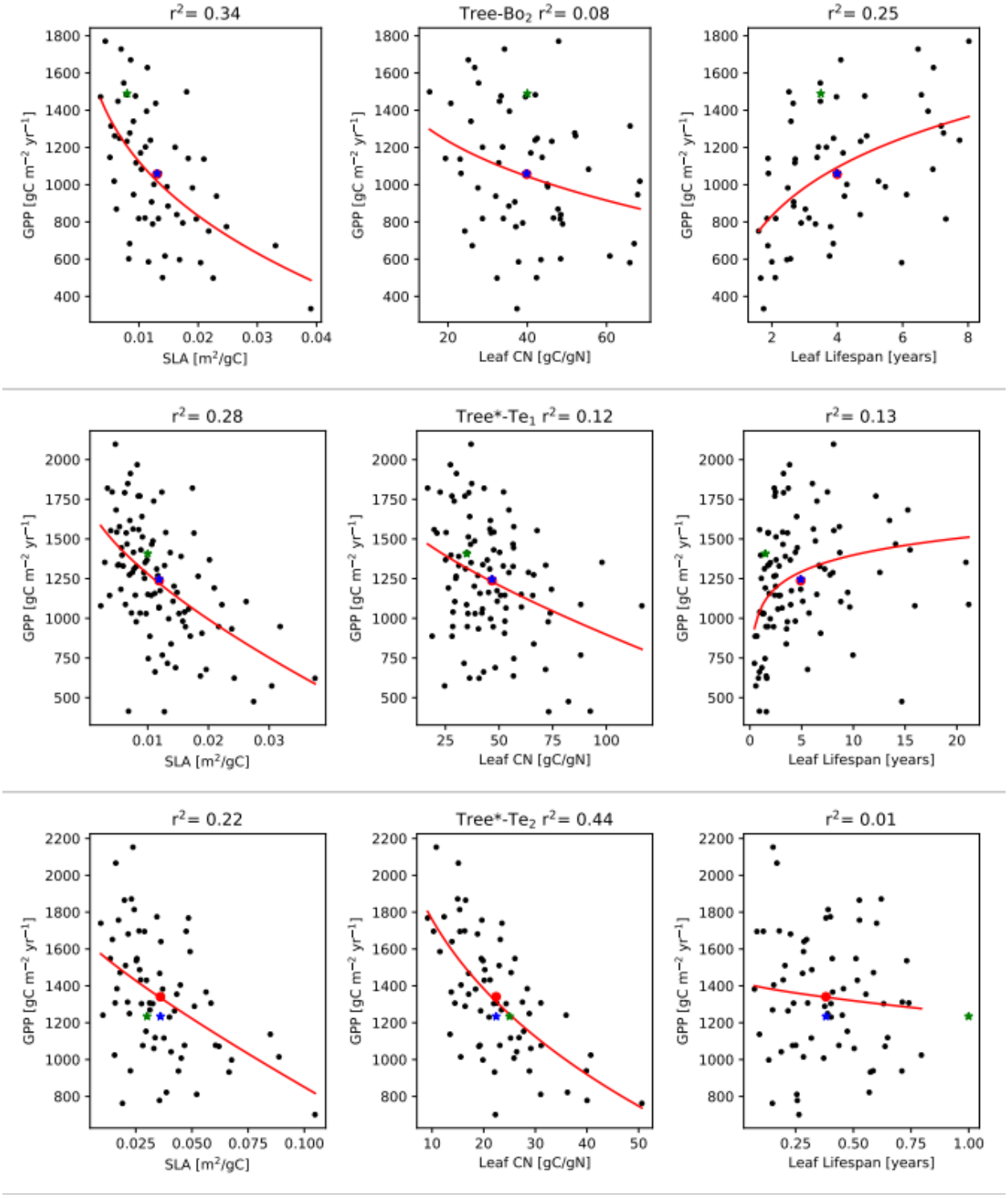


Fig. S9 continued

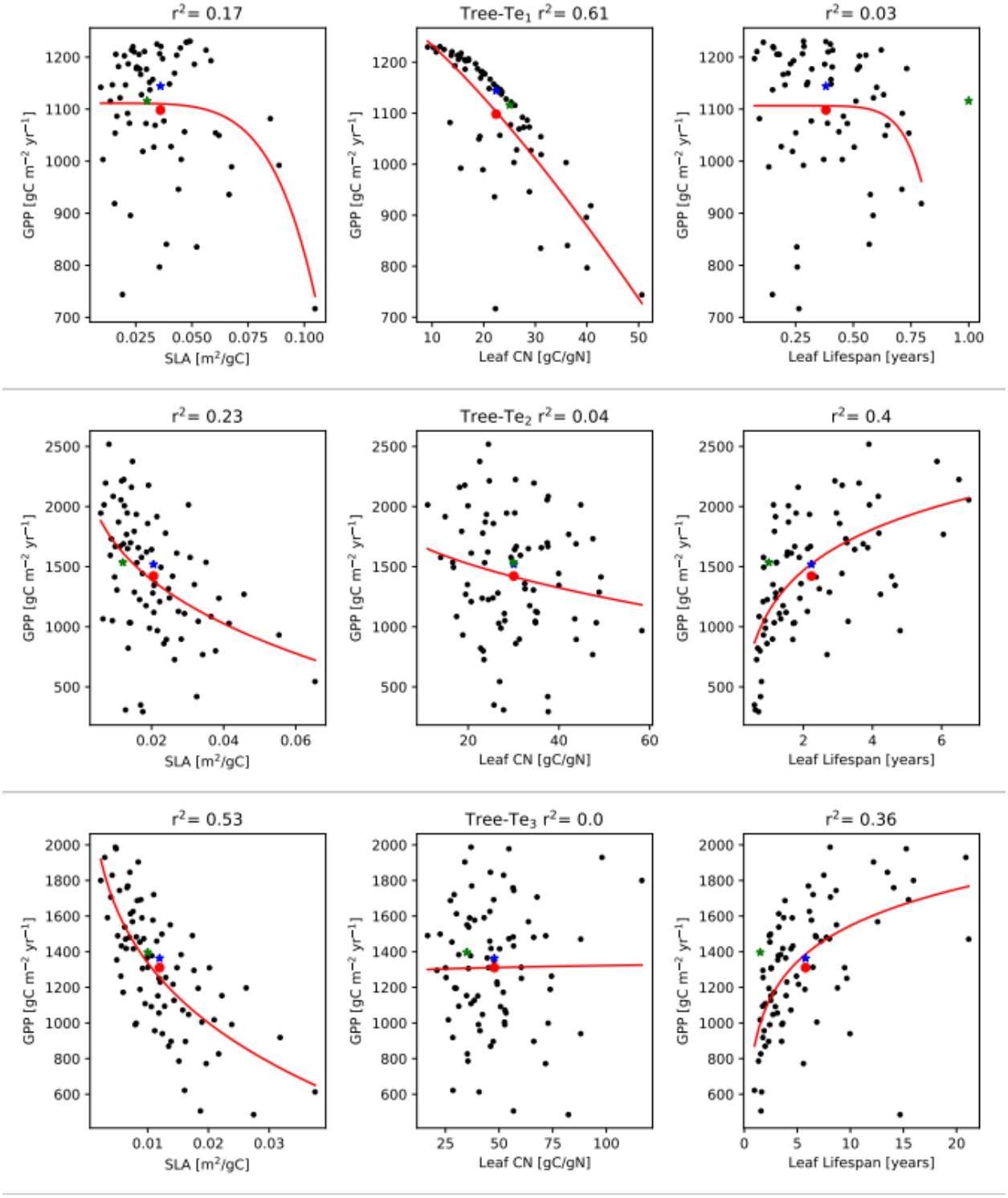


Fig. S9 continued

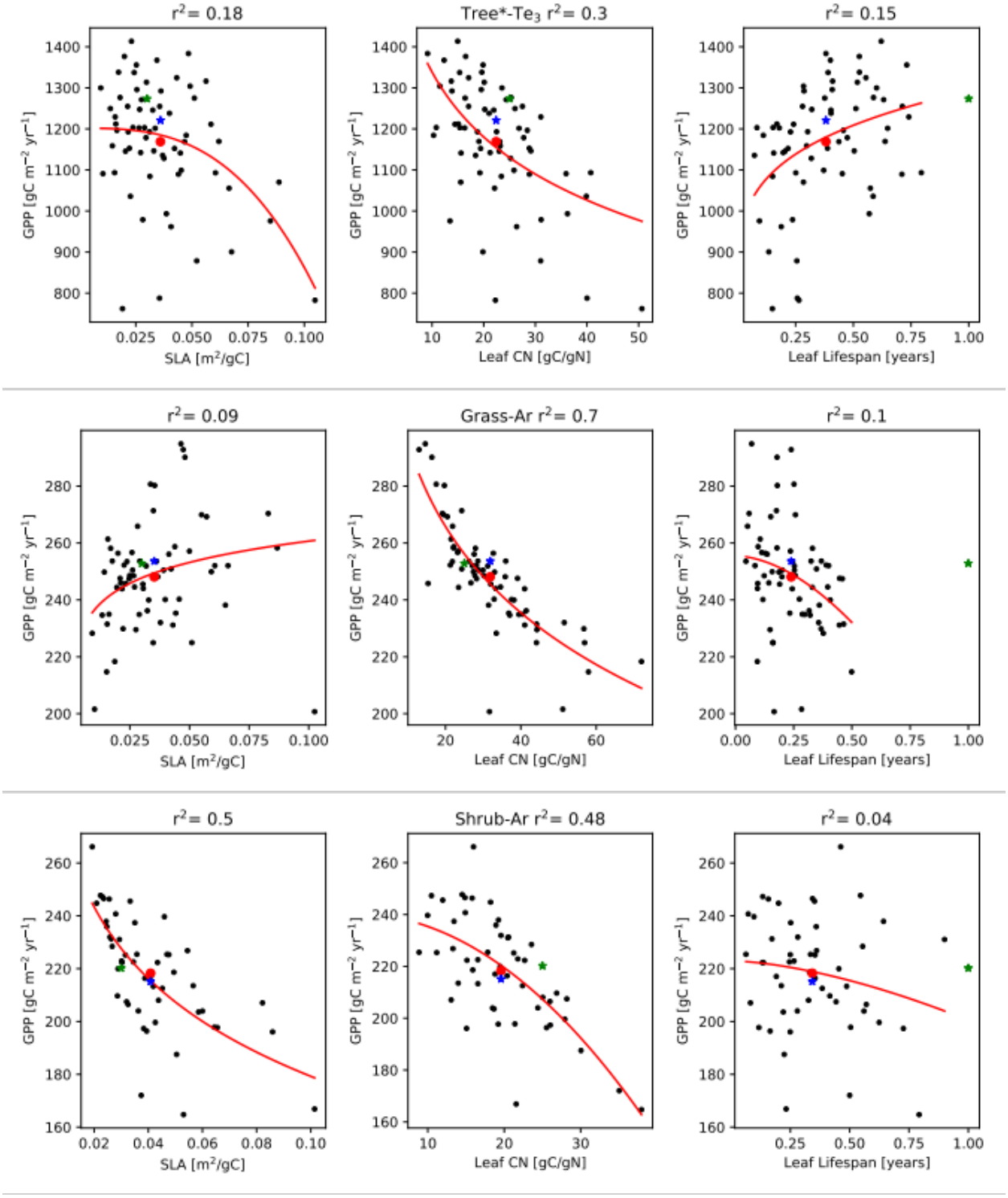


Fig. S9 continued

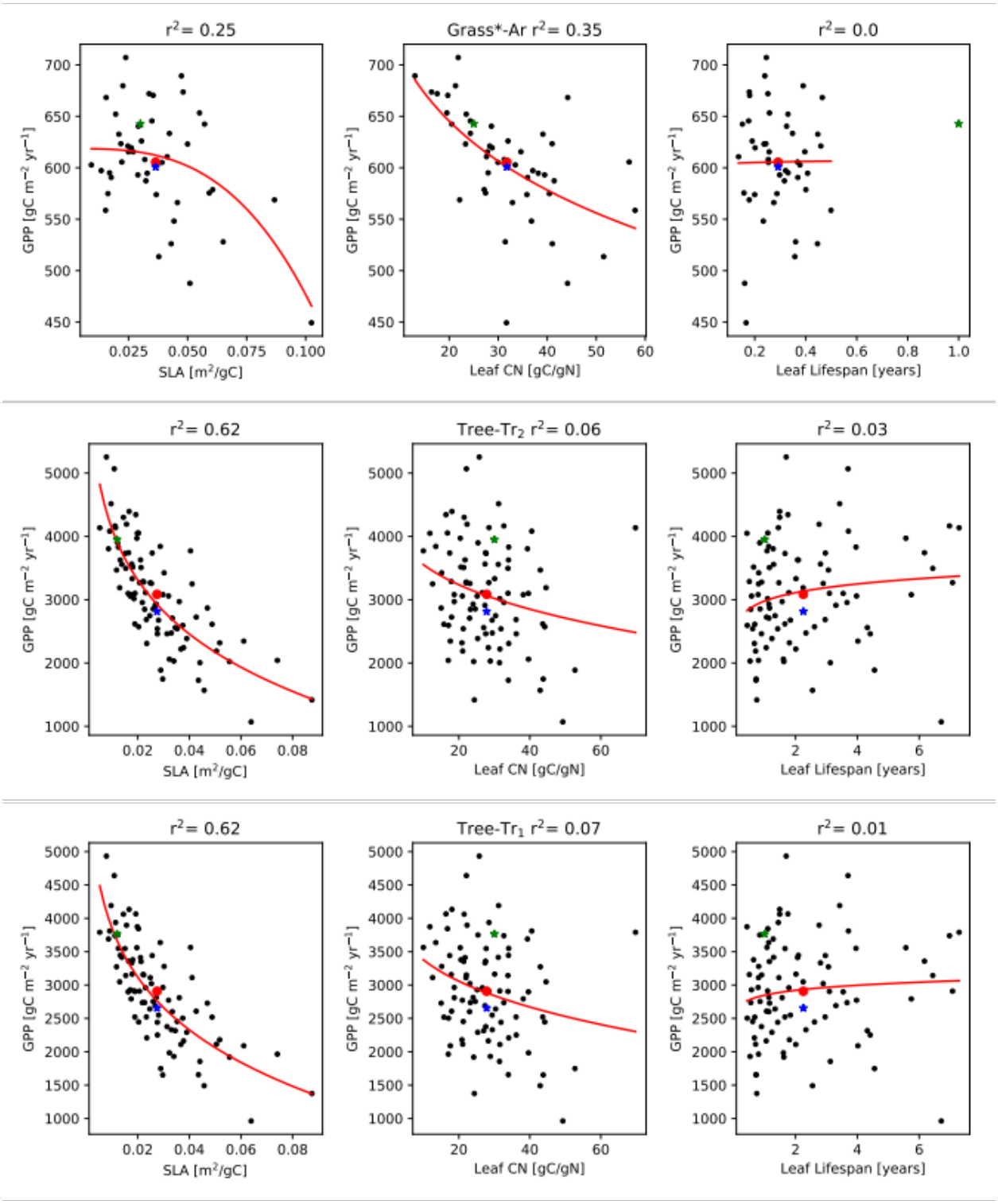


Figure S10. Example of Jensen's Inequality. On the left is a convex function, $f(x)= x^3$, which demonstrates the separation of the input (black point) from the output (green point) mean. The magnitude of this difference will grow with larger input values. On the right is the concave function, $f(x) = - x^3+40$, with a similar difference but in the opposite direction as indicated by Jensen's Inequality. The input (black) and output (green) distributions are displayed to the bottom and right of each function respectively. Note that even when the input and output means do not greatly differ, the tail of the input distribution can make a substantial difference in the output of the function.

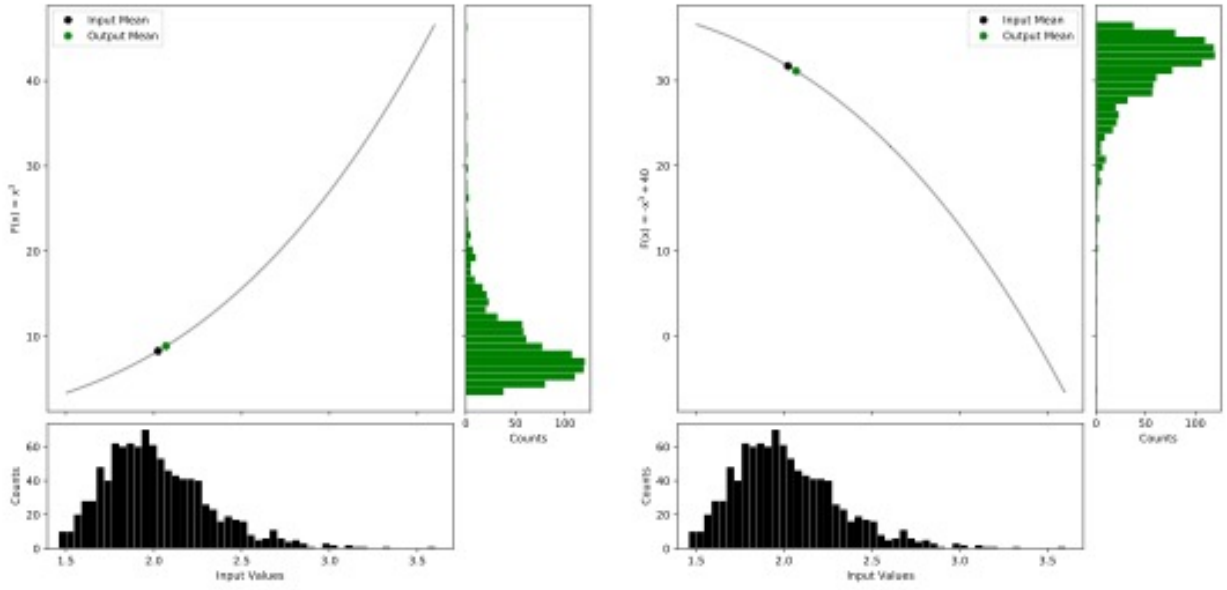


Figure S11. Magnitude of difference between distribution mean and updated mean at each site.