Strong biotic influences on regional patterns of climate regulation services

Serna-Chavez, H.M.; Swenson, N.G.; Weiser, M.D.; van Loon, E.E.; Bouten, W.; Davidson, M.D.; van Bodegom, P.M.

⊘AGUPUBLICATIONS

Supporting Information for

Strong biotic influences on regional patterns of climate regulation services

H. M. Serna-Chavez[1], N. G. Swenson[2], M. D. Weiser[3], E. E. van Loon[1], W. Bouten[1], M. D. Davidson[1,4], P. M. van Bodegom[5]

[1]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, 1090 GE, The Netherlands

[2]Department of Biology, University of Maryland, College Park, MD, 20742, USA.

[3]Department of Biology, University of Oklahoma, Norman, OK, 73069 USA.

[4]International Centre for Integrated assessment and Sustainable development, Maastricht University, Maastricht, 6200 MD, The Netherlands.

[5]Department of Conservation Biology, Leiden University, Leiden, 2333 CC, The Netherlands.

Contents of this file

Additional Supporting Information (Files uploaded separately)

Captions for Table S1. Descriptive statistics of unstandardized input variables.

Introduction

This document presents information on the processing of input variables, statistical analyses, and supplementary findings supporting the main conclusions of our research.

Text S1. Aggregating gridded data

The original grid-cell resolution of MODIS tree cover, and forest understory, dead wood, litter, live tree above and belowground carbon data was finer than 1×1 km. To homogenize the grid-cell resolution of all input data in our study, we aggregated tree cover and all forest carbon products from 250 × 250 m to 1×1 km by calculating the average of values included in the new resolution. The aggregation of gridded data to a coarser resolution did not apply any distance-weighting criteria as in resampling techniques.

   To evaluate if the aggregation of gridded data introduced biases into our analyses, we regressed the values extracted for each forest plot before and after aggregating gridded data to 1 × 1 km. A 1:1 relationship, or close to 1:1, would indicate no bias in the aggregated data. The slopes (β) of all bivariate regressions between the values extracted from the high-resolution (250 × 250 m) and the aggregated data (1×1 km) were ≥0.90 (Fig. S1). In all cases, more than 90% of the variance in the original data was accounted by the aggregated data (Fig. S1). The procedure to homogenize grain size did not introduce strong biases in our input data and, hence, in the results from model sets and hierarchical partitioning of variance.

Text S2. Stand age data

We chose to represent successional stage with forest carbon pool data (soil litter and dead carbon, for example) and functional traits (seed mass, maximum height, for example), and not with the stand age data available from FIA databases. We supported this decision by conducting a supplementary analysis on the gains of incorporating FIA stand age data into our analyses.

   We obtained and processed FIA data tables to match plot locations with their stand age estimates. However, not all the 16,955 forest plots used in our analyses had stand age data reported for the 2000-2009-time span. For over half our study plots (about 8,804 plots) stand age estimates were missing. We opted to use stand age estimates registered in inventory years 2000 until 2016. This way only 3,974 of our plots missed a stand age estimate. When stand age was reported for several inventory years we took the average over the 2000-2016 period. Stand age data ranged from 1 to 220 years. The majority of stand age estimates came from the 2005 inventory year.

   We set to answer two questions on the influence of stand age in our models:
>   1) how strongly is stand age correlated to climate regulation processes
>   2) how does the total explained variance change when including stand age in biotic
>   models

For the first question, we constructed a correlogram matrix of Pearson's correlation coefficients with climate regulation processes (albedo, evapotranspiration, gross and net primary productivity, and soil organic carbon), stand age, and the variables related to forest structure and successional stage (soil litter carbon, tree cover, dead carbon, maximum height, seed mass, wood density) (Fig. S2).

   The correlation of stand age with albedo, evapotranspiration, and soil organic carbon was considerably smaller than that of soil litter and maximum height, two factors linked to forest structure and successional stage [*De Deyn et al.,* 2008]. The correlation between stand age and gross and net productivity was similar to that of maximum height traits, and greater than those of soil litter. It is, however, important to note that, in our study, forest structure (as represented by understory carbon, for example) was a stronger influence on gross and net primary productivity, than soil litter (as described in the main manuscript).

For the second question, we included stand age in the biotic models reported in the main manuscript. We compared the changes in total explained variance, and then conducted hierarchical partitioning analyses to quantify the amount of variance independently explained by stand age in comparison to that solely attributed to other variables. The changes in explanatory power of biotic models were small, ranging from +2% to -3% in R^2_{adj} (Fig. S3). Moreover, the independent influence of stand age was in all cases small, especially when compared to that of forest litter and understory carbon (Fig. S2).

Surprisingly, the correlations between stand age and the variables from our analyses related to forest structure and successional stage was small. We explain this by two factors: 1) Our selected biotic variables reflect effects of both forest structure and successional stage; and, 2) stand age, as defined by FIA, has been criticized in a number of publication as imperfect measure for successional stage.

In FIA, forest stand age is estimated from a few dominant trees from the overstory [*USDA, 2016*]. A plot can thus be assigned a stand age based on the average age of a subset of trees within the plot or near it, which were selected on basis of diameter at breast height size class [*USDA, 2016*]. The age of these trees is measured using ring counts on increment cores taken at breast height [*USDA, 2016*]. However, it is also possible that stand age is assigned to a plot using documentary evidence or expert judgment based on productivity or tree diameter [*Stevens et al., 2016*]. This does not necessarily include the oldest trees in the plot. Moreover, FIA stand age was not designed the time since a stand-replacing disturbance [*Purves et al., 2008; Stevens et al., 2016*].

Considering the problematic ecological interpretation of stand age (as outlined above), the small improvement in our results (Fig. S3), and the considerable reduction in our sample size due to stand age data unavailable, we decide not to include forest stand age data in our main analyses.

Text S3. Cross-validation and multi-model selection

We performed 10-fold cross-validation to assess model accuracy and to select the best model set. We first randomly split the dataset into ten equal folds. We fitted all possible model combinations with up to five explanatory variables. We fitted all models using nine folds of the data (training set); we predicted the response (y) using the tenth fold of data (validation set) and calculated the root-mean squared error for all models (RMSE; [*James et al., 2013*]). After ten runs, we averaged the RMSE values for each model in the set and used this average to select the set of linear models that best explained climate regulation processes: albedo, evapotranspiration, soil organic carbon, gross and net primary productivity. We defined the best model set as the multivariate models in the lowest 10% of RMSE values.

We did not select or average coefficients in best model sets with Akaike's Information Criteria (AIC) because our large sample size and the number of model combinations returned an artifact of no model-selection uncertainty [*Murray and Conner, 2009*]. The large sample size and the large number of model combinations lead to great differences in AIC values and AIC-weights between the top 1% model and all other candidate models [*Burnham and Anderson, 2002; Hoeting et al., 2006*]. This artifact also prevented us from using AIC-weights to rank explanatory variables according to their importance in the best model sets.

Text S4. Linear regression assumptions and spatial autocorrelation

We fitted the full abiotic, biotic, and combined models to evaluate ordinary least-squares (OLS) assumptions and spatial autocorrelation (SAC) in model residuals. Full models are those fitted using all the variables resulting from the two-step variable selection process (adapted after *Murray and Conner* [2009]). To evaluate OLS assumptions, we fitted a smoother with 95% confidence bands to residual plots [*Zuur et al.,* 2007]. We considered there were no considerable deviations from the null pattern if zero was within the 95% confidence bands of the smoother in residuals-versus-fitted-values and residuals-versus-variables plots [*Zuur et al.,* 2007]. We added a quadratic term to the model if the 95% confidence bands did not contain zero in residuals-versus-variable plots. We added a cubic term if after the addition of the quadratic term the residuals-versus-variable plot still deviated from the null pattern. No higher-order terms were necessary for the smoother to approach the null pattern in residuals plots. In the construction of all possible combinations of multiple linear regressions (1-5 regressors per model), polynomial terms could only appear in combination with the linear term [*Zuur et al.,* 2007].

We assessed SAC in Pearson residuals from full abiotic, biotic and combined models. (Polynomial terms were included in full models when the relationships were non-linear – as described above). We did not consider SAC to be a problem if for the first 500 distance intervals $0.15 \leq$ Moran's I $\geq$ -0.15. Residuals from most full models showed no strong SAC (Fig. S4). We did not assess the significance of Moran's I at any distance class using permutations because our large sample size always returned an artifact of highly significant p-values. If the full models were a good fit for the data and showed no considerable SAC, we assumed the models in the upper 10% RMSE would also be a good fit [*Rhodes et al.,* 2006].

Text S5. Hierarchical partitioning of variance

Hierarchical partitioning analyses (HP) did not include polynomial terms because, in trial runs, effects of the linear terms were always far greater than those of polynomials. Additionally, there is a limitation of maximum 12 regressors in the HP algorithm [*Walsh and Mac Nally,* 2013]. That is why we only report the independent and joint effects of linear terms.

The HP algorithm reports a small rounding error when more than nine variables are introduced [*Walsh and Mac Nally,* 2013]. Thus, when full abiotic, biotic or combined models for each climate regulation process included more than nine non-spurious, non-highly collinear variables (but less than 12), we ran the HP algorithm ten times, randomizing the order of variables at each turn. We then reported the average of independent and joint effects in the main text. Repeating the HP analyses ten times and averaging effects did not show considerable deviations: in all cases, standard deviations did not surpass 4×10^{-3}.
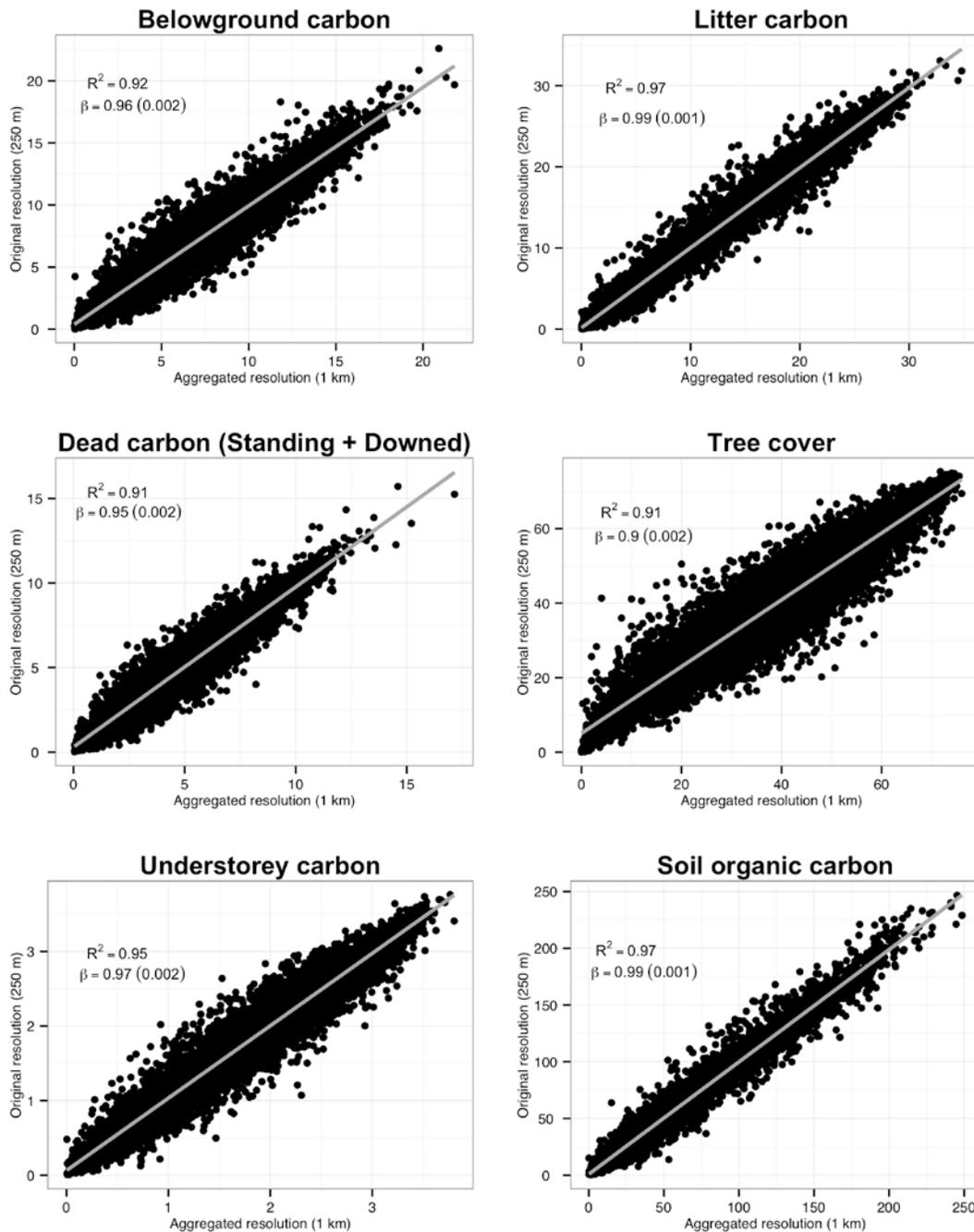
Figure S1. Comparison between original and aggregated data. Dead wood represents standing and downed deadwood. R^2 represents the amount of variance explained in a bivariate linear regression (*y~x*, all R^2≥0.91%). Solid grey line represents the regression coefficient from the bivariate regression (all β ≥0.90). The β coefficients are presented with one standard error between parentheses. All forest carbon data are on Mg Carbon ha^{-1}, tree cover is in percentage (0-100%).
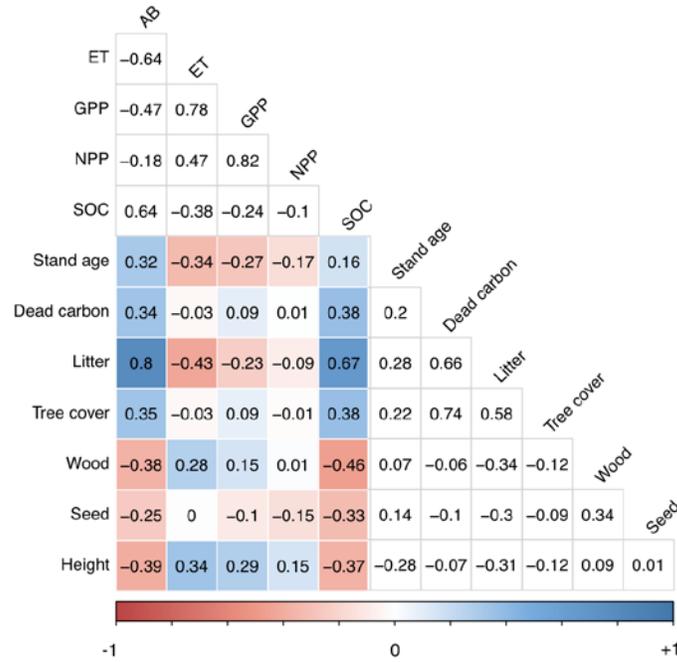
Figure S2. Pearson's correlation coefficients. The correlation coefficient of stand age with climate regulation processes was smaller or similar than that of soil litter carbon (Litter). Only for NPP, stand age had a higher correlation than soil litter. Panels with secondary information are set to white background. AB=albedo, ET=evapotranspiration, GPP and NPP=gross and net primary productivity, SOC=soil organic carbon. All variables other than AB, ET, GPP, NPP, and SOC, were standardized (mean=0, SD=1).
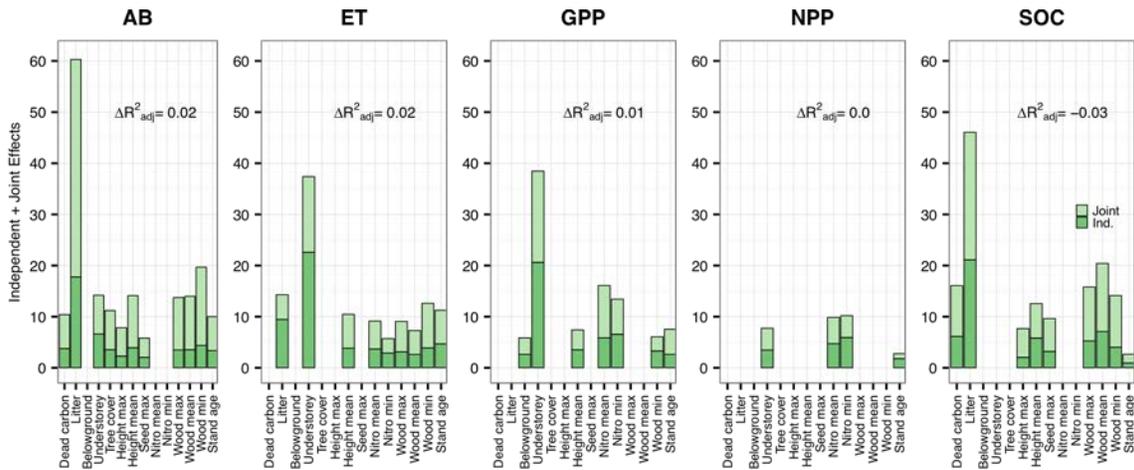

Figure S3. Forest stand age in biotic models and hierarchical partitioning. Independent and joint effects of biotic variables on climate regulation processes. ΔR^2_{adj} indicates the change in explanatory power between biotic models with stand age and the ones reported in our main manuscript. In all cases, the joint effect of stand age was larger than its independent effect. Ind.=independent effect, Joint=Joint effect, AB=albedo, ET=evapotranspiration, GPP and NPP=gross and net primary productivity, SOC=soil organic carbon.
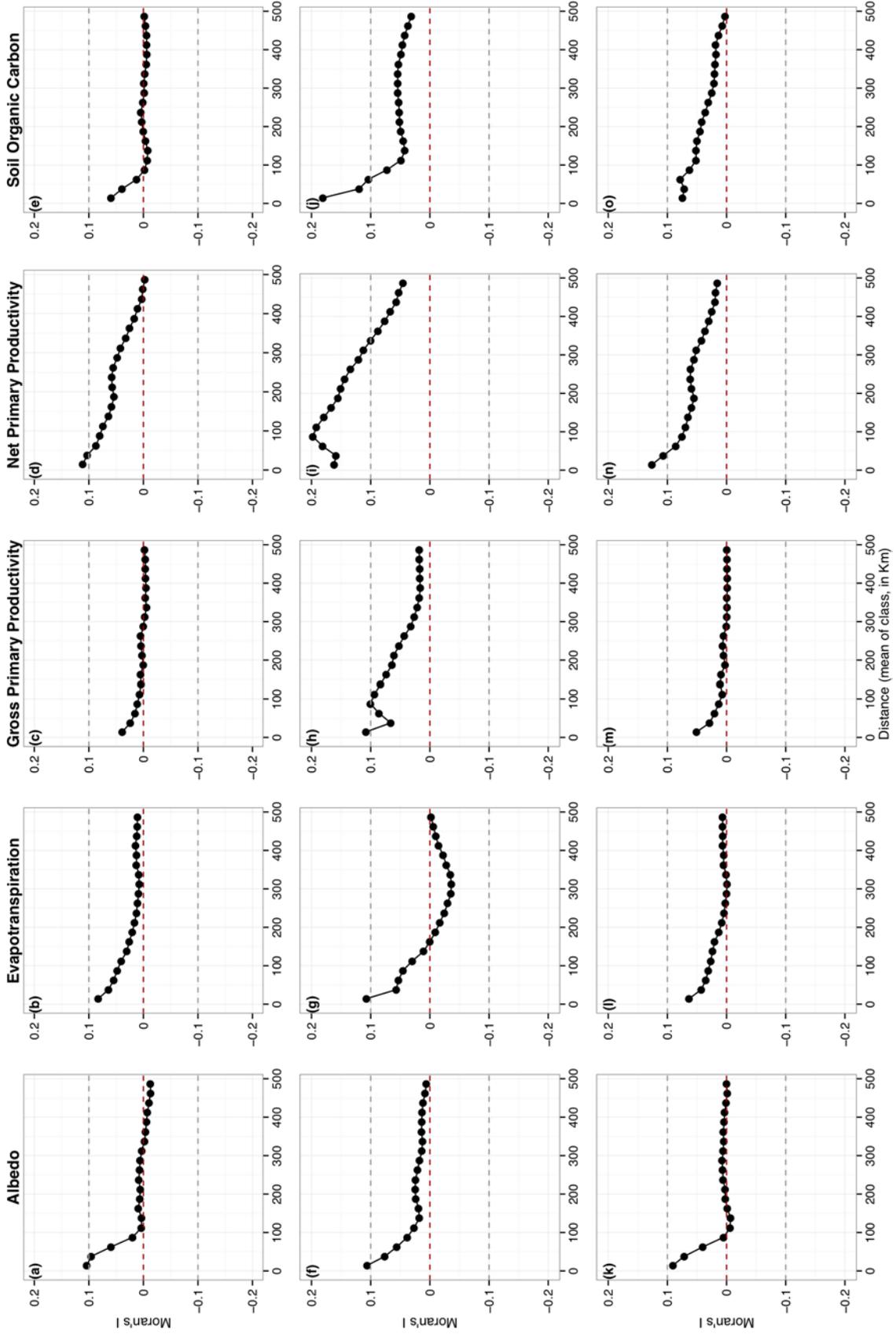
Figure S4. Spatial correlograms of Moran's I. Panels in the first row (a-e) present the results for full abiotic models. Panels in the second row (f-j) present the results for full biotic models. Panels in the third row (k-p) present the results for full combined models. The full biotic model for NPP (i) included only three linear terms (forest understory, mean and min. leaf nitrogen content) and a quadratic term. For all panels, the *x*-axis represents average of the distances within each class.
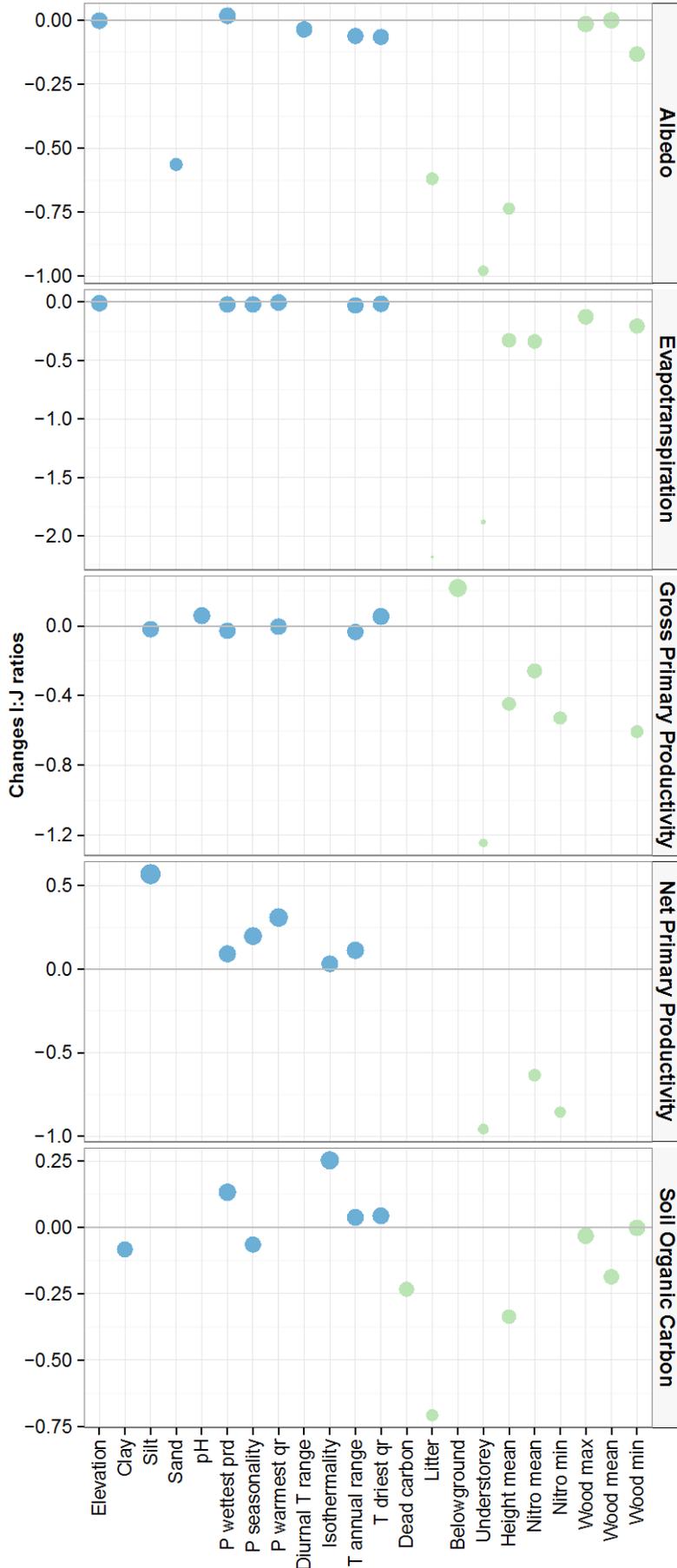
Figure S5. Changes in independent-to-join effects ratios. Changes in independent-to-join effects ratio (ΔI:J) calculated as the I:J ratio in the combined set minus that of the variable in the abiotic or biotic set. Size of filled circles is relative to the ΔI:J. Blue circles represent abiotic variables, and green represents biotic variables. Abbreviations: CEC=soil cation exchange capacity; T= temperature; P=precipitation; qr.=quarter of a year.
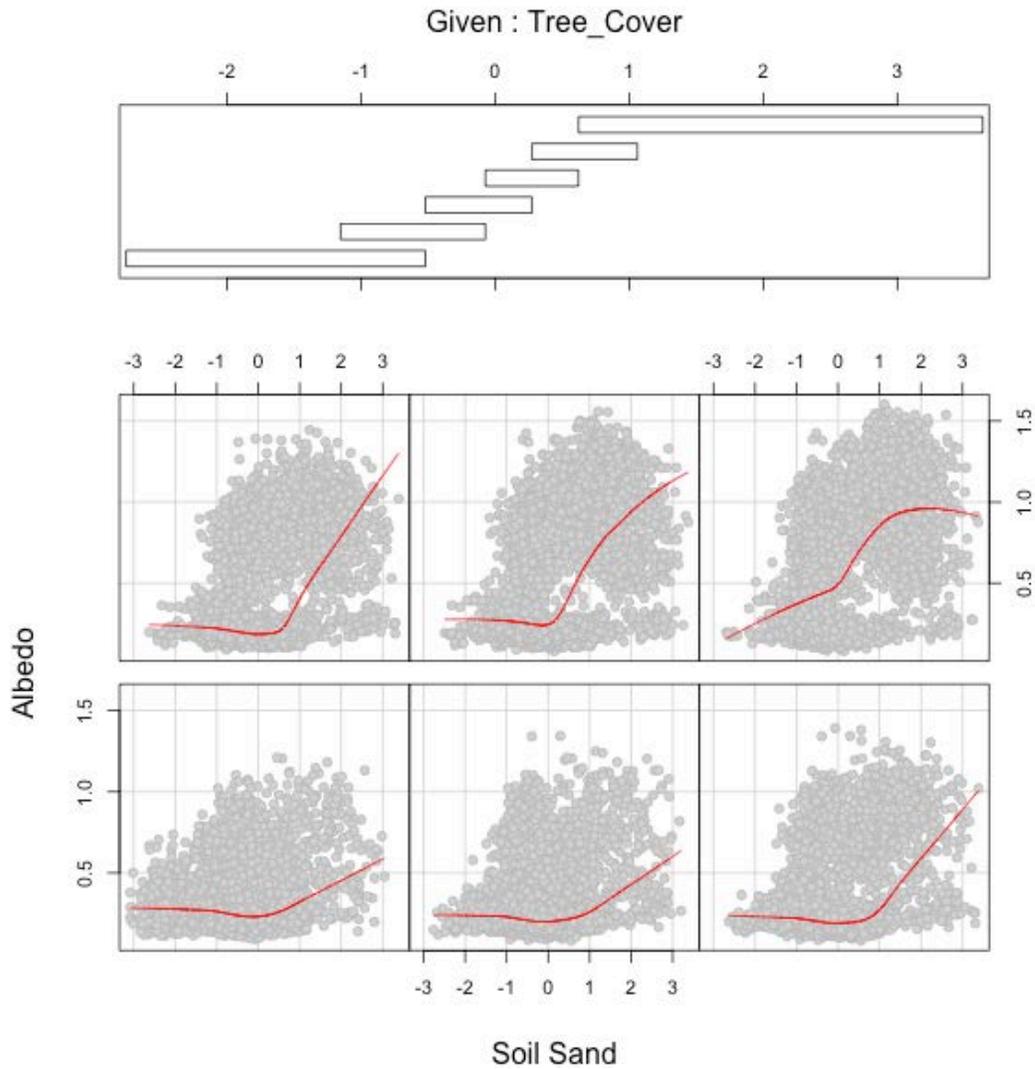
Figure S6. Conditional plot: Effect of soil sand on Albedo is likely mediated through changes in tree cover. Soil Sand and Tree cover are standardized (mean=0, SD=1). Red line is the loess smoother.
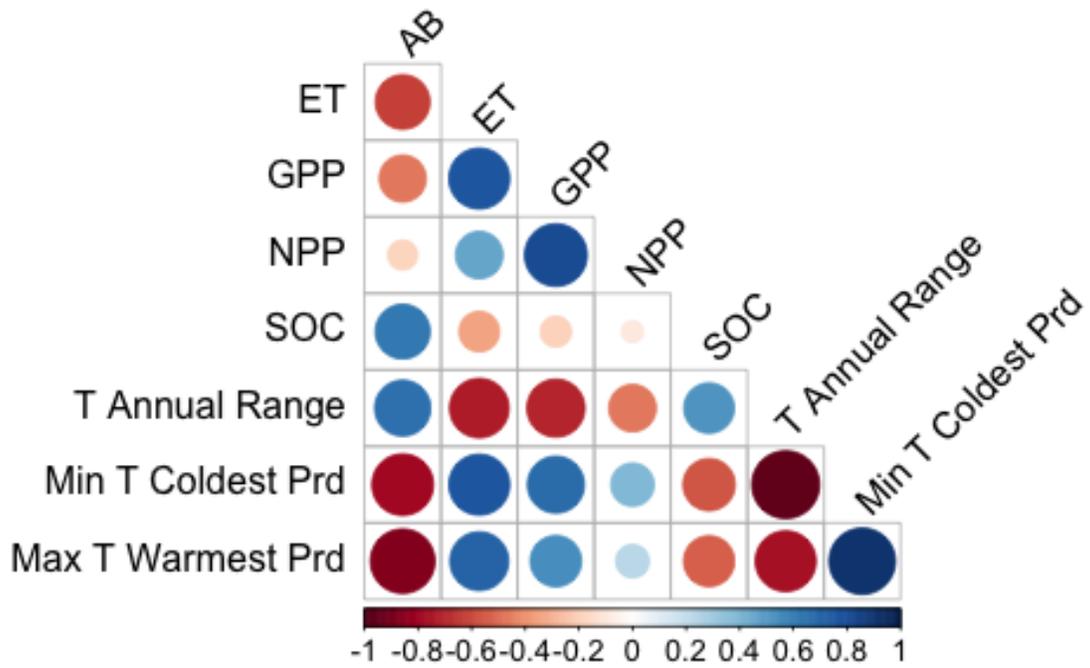
Figure S7. Correlogram matrix: Effects of temperature annual range, minimum temperature of the coldest period, and maximum temperature of the warmest period. Temperature annual range was a strong influence on all climate regulation services. It is calculated as the difference between maximum temperature of the warmest period and minimum temperature of the coldest period.

Table S1. Descriptive statistics of unstandardized variables. (See Excel spreadsheet)

Table S2. How do the best 10% RMSE model sets compare to the corresponding full models?

	Full model			Best 10% model set		
	p	R^2_{adj}	RMSE* (±1SD)	*n*	R^2_{adj}	RMSE* (±1SD)
AB						
Abiotic	13	0.747	0.171 (±0.003)	96	0.681	0.192 (±0.003)
Biotic	14	0.798	0.153 (±0.004)	111	0.784	0.158 (±0.002)
Combined	14	0.830	0.140 (±0.003)	206	0.795	0.154 (±0.003)
ET						
Abiotic	15	0.658	96.14 (±1.68)	164	0.614	102.0 (±0.723)
Biotic	9	0.557	109.2 (±3.12)	29	0.545	110.7 (±0.340)
Combined	12	0.658	96.03 (±1.83)	159	0.633	99.51 (±1.293)
GPP						
Abiotic	15	0.676	237.7 (±2.80)	164	0.596	265.0 (±2.522)
Biotic	7	0.452	308.5 (±5.54)	9	0.442	311.1 (±1.857)
Combined	16	0.650	247.0 (±3.01)	239	0.609	260.8 (±2.628)
NPP						
Abiotic	13	0.421	178.9 (±3.93)	96	0.340	190.9 (±2.003)
Biotic	4	0.159	215.4 (±4.67)	1	0.159	215.4 (±0.000)
Combined	14	0.387	184.0 (±3.69)	77	0.326	192.8 (±1.718)
SOC						
Abiotic	12	0.667	20.64 (±0.59)	82	0.605	22.44 (±0.255)
Biotic	11	0.627	21.84 (±0.49)	44	0.588	22.94 (±0.308)
Combined	17	0.847	14.01 (±0.32)	265	0.746	18.01 (±0.757)

*Average from 10-fold cross-validation, as explained in the "Cross-validation and model selection" section.
p: number of explanatory variables, including polynomial terms
n: number of models in the best set.
Albedo (AB), evapotranspiration (ET), gross primary productivity (GPP), net primary productivity (NPP), soil organic carbon (SOC).
Best 10% model set is defined as the collection of multivariate models in the upper 10th percentile of RMSE after 10-fold cross-validation.

Table S3. Average model coefficients in the best 10% abiotic and biotic model sets. The best model sets are the model combinations in the upper 10% RMSE after 10-fold cross-validation. Standard errors of averaged coefficients are reported between parentheses.

	Averaged model coefficients for polynomial terms	n	RMSE	R^2_{adj}
Abiotic model sets	**AB** T annual range^2×-0.007(±0.009) Elevation^2×-0.028(±0.015), Elevation^3×-0.0016(±0.003)	96	0.192	68.07
	ET P wettest period^2×-4.54(±2.003) P warmest qr^2×-3.52(±2.245) T annual range^2×10.4(±3.263), T annual range^3×6.50(±1.014)	164	102.0	61.45
	GPP P wettest period^2×-10.6(±9.58) P warmest qr^2×-5.63(±7.88), P warmest qr^3×-8.34 T annual range^2×35.5(±27.60), T annual range^3×19.6(±2.91)	115	265	59.64
	NPP P wettest period^2×5.06(±13.81) T annual range^2×24.2(±24.60), T annual range^3×13.2(±0.717)	96	190.0	34.00
	SOC T annual range^2×6.10(±2.712), T annual range^3×1.94(±5.740)	82	22.4	60.55
Biotic model sets	**AB** Litter^2×-0.033(±0.003) Height mean^2× 0.001(±0.0003) Wood mean^2×0.0015(±0.001), Wood mean^3×0.002	111	0.158	78.36
	ET Height mean^2×-1.97(±0.194)	29	110.7	54.53
	GPP Height mean^2×-5.94(±0.48)	9	311	44.24
	NPP Nitro mean^2×15.6	1	215.4	15.93
	SOC Litter^2×-7.62(±0.686), Height mean^2×1.07(±0.017), Wood mean^2×3.22(±0.340)	44	22.94	58.77

Variables without SE were only included in one model of the best 10% model sets, hence, no standard errors. *n*, the number of models included in the upper 10% RMSE considered in calculating the mean of each coefficient. *RMSE*, root-mean-squared-error, average over the *n* models. R^2_{adj}, in %, provides the average explained variance over the *n* models. Abbreviations: T= temperature; P=precipitation; qr= quarter of a year

Table S4. Average model coefficients found in the best 10% combined model sets. Best model sets are the model combinations in the upper 10% RMSE after 10-fold cross-validation. Combined models were constructed using the biotic and the abiotic variables with the highest independent effects in hierarchical partitioning analyses. Standard errors of averaged coefficients are reported between parentheses.

	Averaged model coefficients for polynomial terms	n	RMSE	R^2_{adj}
AB	Elevation^2×-0.01(±0.003) Wood mean^2×0.002(±0.002)	206	0.154	79.53
ET		159	99.51	63.28
GPP	P wettest period^2×8.33(±6.72) P warmest qr^2×10.1(±8.91), P warmest qr^3×-6.46 T annual range^2×24.8(±11.35)	239	261	60.89
NPP	P wettest period^2×12.2(±8.441) P warmest qr^2×13.3(±4.931), P warmest qr^3×-5.63(±0.440) T annual range^2×6.78(±19.26), T annual range^3×2.00	77	192.8	32.64
SOC	Litter^2×-6.47(±0.259) T annual range^2×3.63(±4.082), T annual range^3×2.30(±1.210) Height mean^2×0.09(±0.053), Wood mean^2×0.74(±0.259)	265	18.01	74.58

Variables without SE were only included in one model of the best 10% model sets, hence, no standard errors. *n*, the number of models included in the upper 10% RMSE considered in calculating the mean of each coefficient. *RMSE*, root-mean-squared-error, average over the *n* models. R^2_{adj} in %, provides the average explained variance over the *n* models. Abbreviations: T= temperature; P=precipitation; qr= quarter of a year