SUPPORTING INFORMATION

Contents

Code and datasets

An R script with all analyses and the Swedish and global grassland datasets is available at https://github.com/KevinVanSundert/NutrientMetrics_GCB_2019_KVS/blob/master/NutrientMetrics_GCB_2019.7z.

Collection and pre-processing of soil data

Soil sampling and analyses were performed following the ICP Forests manual on sampling and analysis of soil (Cools & De Vos, 2016). The results of the soil sampling campaigns between 2003 and 2010 of 286 ICP Forests Level II sites have been assembled in the aggregated forest soil condition database (AFSCDB.LII.2.2) (Fleck et al. 2016). This soil database contains the main soil variables of the forest floor, the horizons of the mineral soil, and fixed depth layers (0–10, 10–20, 20–40, and 40–80 cm) for mineral or peat soil. Apart from the aggregated soil data, the database contains variables calculated from raw data (e.g. C:N and C:P ratio, nutrient stocks, etc. - Table S1). Because the metric proposed in Van Sundert et al. (2018) used soil information for the top 0-20 cm of soil (including the organic layer), we also converted the new data to represent this soil depth interval, using the averages weighted by mass for each layer.

Calculation of tree stem volume increment

Tree growth on the ICP Forests Level II sites is assessed approximately every 5 years with standardized methods since the late eighties (Dobbertin & Neumann, 2010). Data used for this study cover the period from 1995 to 2010, thus including three inventory periods. Diameter at breast height (DBH), dead or alive status, and tree height were assessed regularly for every tree (DBH > 5 cm) within a monitoring plot according to the manuals of the ICP Forests Programme (ICP Forests, 2010).

Tree stem volumes were calculated from DBH and height with allometric relationships accounting for species and regional differences (De Vries et al., 2003). The increment of annual stem volume (in m^3 ha^{-1} yr^{-1}) between two inventories was calculated as the sum of increment of standing trees plus ingrowth, and the increment of lost trees until their disappearance. Trees disappeared between two inventories were assumed to have been lost in the middle of the inventory period and stem volumes of these trees at the point of their disappearance was estimated from regressions of stem volumes at the beginning and end of the inventory period of available trees. For our analyses, we averaged volume increments per site over all inventory periods for which increments were available (e.g. 5 to 15 years depending on available data). We considered averaging a proxy for productivity over 5 to 15 years as a more reliable option than averaging over the particular five-year inventory period during which soil sampling occurred, because the five-yearly productivity is still more sensitive to inter-annual variability in weather conditions not taken into account in further analyses (e.g. Anderegg et al., 2015). Furthermore, considerable biases due to a mismatch between soil properties and productivity are unlikely, because even though variables such as soil C:N ratio and pH, and therefore also soil nutrient status have been changing over time in Europe between the first and last inventory periods, temporal trends within sites are relatively small compared to spatial variation (e.g. studies on European forests report multi-decadal pH changes being considerably lower than one unit, while pH in the dataset varied by more than three; soil C:N increased up to a few units, whereas organic and mineral soil C:N varied spatially between 18-65 and 5-54, respectively - Kirk et al., 2010; Jandl et al., 2012; Novotny et al., 2015; Binkley & Högberg, 2016).

Table S1. Overview of variables from the European ICP Forests database used in the current study. Plots for vegetation and soil monitoring (≥ 0.25 ha; *n* = 118) are representative samples of the most important forest types with a risk for acidification in the country they are found. Forests in the inventory are managed for wood production. Note that only variables that are presented in the current study are shown for simplicity, while the original database contained more variables. Abbreviations: MAT = mean annual temperature; MAP = mean annual precipitation; SOC = soil organic carbon concentration; TEB = total exchangeable bases; BS = base saturation.

Data used	location	climate	soil[a]	vegetation
	latitude (° N)	MAT (° C)	organic layer thickness (cm)	age class[c]
	longitude (° E)	MAP (mm)	0-20 cm SOC (%)	tree species[d]
			organic layer C:N ratio	productivity[e] (m^3 ha^{-1} yr^{-1})
			mineral soil C:N ratio	stand density[f] (trees ha^{-1})
			0-20 cm soil C:N ratio	leaf N (‰)
			organic layer pH_{CaCl2}	leaf P (‰)
			mineral soil pH_{CaCl2}	leaf K (‰)
			0-20 cm soil pH_{water}	leaf Ca (‰)
			mineral soil TEB ($cmol_+$ kg^{-1})	leaf Mg (‰)
			organic layer total P (mg kg^{-1})[b]	leaf S (‰)
			mineral soil total P (mg kg^{-1})[b]	leaf N:P ratio
			organic layer C:P ratio[b]	
			mineral soil C:P ratio[b]	

[a] Soil data, collected from 2003 onwards, were taken or derived from the aggregated soil database, provided by the European Forest Soil Coordination Centre (Fleck et al., 2016). The chemical and physical soil analyses results ought to be representative for the whole plot area. The spatial variability is covered by including at least 24 sampling locations in the composite sample. For the laboratory analyses of all solid soil parameters, methodology of the ICP Forests soil manual was followed (Cools & De Vos, 2016).

[b] In this study, we used aqua regia extractable P (ISO 11466, 1995) as the best available proxy for soil total P. Actual total P as derived from the acid digestion method may therefore have been underestimated (Ivanov, 2012).

[c] Tree age was classified into nine 20-year classes (i.e. 0-20, 20-40, etc. up to > 160 in class 9) based on standardized visual assessment of crown conditions. In the current study, we treated the class as a continuous variable.

[d] Homogeneous, managed stands dominated by Common beech (*Fagus* sylvatica L.), Pedunculate oak (*Quercus robur* L.), Scots pine (*Pinus* sylvestris L.) or Norway spruce (*Picea abies* (L.) H. Karst.) were used for the analyses.

[e] Stem volume increment, based on the difference in volume between two five-yearly inventories. Productivity data used in the current study represent averages over all inventory periods between 1995 and 2010 (i.e. three productivity outcomes were averaged when three inventories were performed, but in some cases there were only one or two inventories). We preferred averaging over multiple periods over just using one period (e.g. the last one or the one during soil was sampled) because the five-yearly productivity is still more sensitive to inter-annual variability in weather conditions not taken into account in further analyses (e.g. Anderegg et al., 2015).

[f] Stand density (eventually only used in exploratory analyses, see 'removing confounding effects') was expressed as trees ha^{-1} in this study because the alternative Reineke stand density index (which offers the advantage of being orthogonal to age – e.g. Solberg et al., 2009) was only available for 55 out of the 118 forests initially used in our study. In the database we used, the number of trees ha^{-1} did not significantly correlate with age (*r* = -0.09; *P* = 0.34), and species-specific regression models that explained productivity by climate, age and stand density did not result in problematic variance inflation (variance inflation factor < 4).

Table S2 Overview of variables from the Swedish forest and soil database used in the present study, and in Van Sundert et al. (2018) for development of a first soil nutrient status metric. Each plot for soil and vegetation analyses had a 10 m radius and was sampled once during the period 2003-2012. The (mostly managed) forests in the inventory represent a random sample of Swedish forests. Note that only variables presented in the current study are shown for simplicity, while the original database contained more variables. Abbreviations: TSUM = growing season temperature sum; MAP = mean annual precipitation; SOC = soil organic carbon concentration; TEB = total exchangeable bases. Data for southern Sweden (*n* = 1061; this is the region in Sweden where most variation in soil variables and productivity occurs, hence data from southern Sweden were used to evaluate performance of the metrics in the present study) are available at https://github.com/KevinVanSundert/NutrientMetrics_GCB_2019_KVS/blob/master/NutrientMetrics_GCB_2019.7z

Data used	location	climate	soil[b]	vegetation
	latitude (° N)	TSUM[a] (° C days)	organic layer thickness (cm)	age[c] (yrs)
	longitude (° E)	MAP (mm)	0-20 cm SOC (%)	dominant tree species[d]
	elevation (m)		organic layer C:N ratio	productivity[e]
			0-20 cm soil C:N ratio	(m^3 ha^{-1} yr^{-1})
			organic layer pH_{CaCl2}	
			mineral soil pH_{CaCl2}	
			0-20 cm soil pH_{water}	
			mineral soil TEB ($cmol_+$ kg^{-1})	

[a]TSUM was calculated for each data point based on its latitude, longitude and elevation.
[b]*n* = 3; soil variables were determined using standard sampling and laboratory procedures (e.g. Olsson et al., 2009; Stendahl et al., 2010).
[c]Stand age ranged between 1 and 350 years, with an average of 65 years.
[d]Sites were classified as either spruce or pine forests if ≥ 50% of the basal area consisted of spruce, resp. pine.
[e]Productivities (site quality) or mean annual volume increments (MAI) over a full rotation were estimated based on height development curves. *In situ* productivities may be lower, depending on the management.

Table S3. Overview of variables from the global grassland database used in the current study. Abbreviations: MAT = mean annual temperature; GSP = growing season precipitation (mm). Data were collected by D. Radujkovic by combining published data from fertilizer experiments on grasslands, and unpublished results provided by the principal investigators. For the collected published data, search terms “grasslands”, “soil”, “ANPP” and “biomass” were used in Web of Science. Only studies containing information on ANPP, soil C:N ratio and pH were retained. Data and references per site (*n* = 68) are available at https://github.com/KevinVanSundert/NutrientMetrics_GCB_2019_KVS/blob/master/NutrientMetrics_GCB_2019.7z .

Data used	location	climate	soil	vegetation
	latitude (° N)	MAT (° C)	soil C (%)[a]	Productivity[b]
	longitude (° E)	GSP (mm)	soil C:N ratio	(g C m^{-2} yr^{-1})
			soil pH	
			soil total P (ppm)	
			soil C:P ratio	
			soil total K + Ca + Mg (ppm)	

[a] Since SOC data were not available for most grassland sites, total C was used instead of SOC to calculate the nutrient metric. Grasslands on calcareous soils were omitted from the analysis to ensure total C approximated SOC.
[b] Aboveground net primary productivity.

METHODOLOGY: removing confounding effects

Because factors other than the nutrient status also strongly influence plant growth, especially species, age and climate, evaluations of soil characteristics, plant stoichiometry and nutrient metrics with plant growth data require removing these confounding effects. In Van Sundert et al (2018), this normalization procedure was based on a regression approach, which removed the effect of climate and species (stand age was accounted for by expressing forest productivity over a rotation period). This regression approach was appropriate for the Swedish dataset, where the normalization for climate did not confound the relationship between aboveground tree growth and soil properties (e.g. Van Sundert et al., 2018). For the ICP Forests dataset, however, this approach was more problematic because especially the soil C:N ratio correlated with MAT (Table S5). Also, variation in SOM showed a considerable degree of collinearity with climate, especially with MAP (Table S5). For these reasons, we applied a more advanced normalization procedure; instead of a simple regression approach, we applied structural equation modeling (R package lavaan – Rosseel, 2012) to distinguish direct effects of soil C:N ratio and SOM that correlate with climate (Fig. S1). As a test, we applied this procedure also to the Swedish dataset, but this did not alter the results for that dataset (see Table S9).

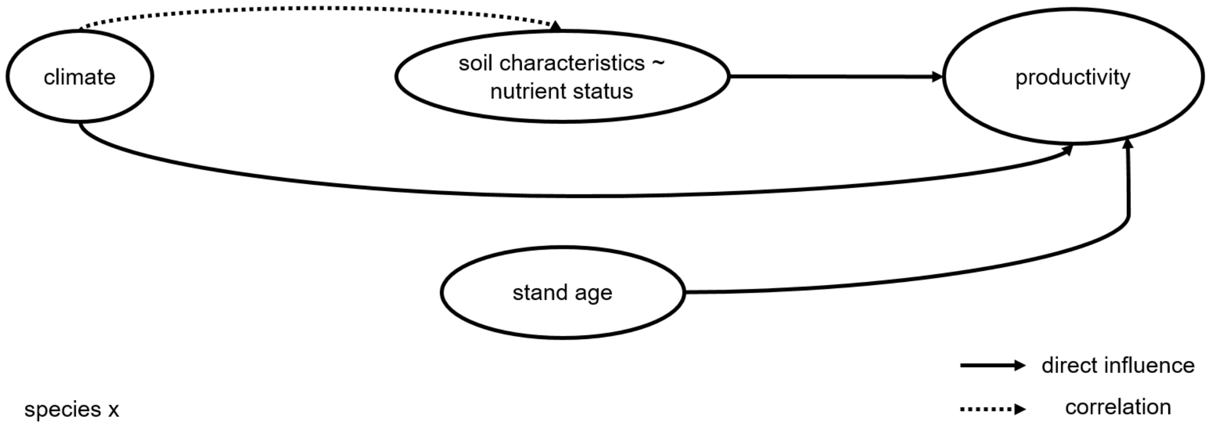

Figure S1 Generalized path diagram showing the main factors influencing productivity. Structural Equation Modeling (SEM) was applied on the path diagram to estimate effects of stand age, climate and soil characteristics on productivity per species, while taking into account correlations between climate and soil. Equation parameters for all four species were estimated in one single SEM.

Strong correlations among climate and soil variables were included in the SEMs of the Swedish, ICP and grassland datasets. Based on Tables S4-6, correlations between climate and soil variables > 0.50 (or < -0.50) were included in the SEMs (criterion 1), and also soil variables exhibiting a correlation > 0.50 with soil variables that met criterion 1 were included.

Based on the SEMs, species-specific equations for productivity (P_i) were derived (Table S7):

$$P_i = \textit{direct climate effect} + \textit{soil effect (influenced by climate)} + \textit{age effect} \; [+ \textit{stand density}] + \varepsilon \qquad (1)$$

For each species, productivity was then normalized for climate and age, such that only soil and residual effects remained. We did eventually not correct productivity for stand density (the number of trees ha^{-1}, which had a positive influence on productivity of pine only), because this worsened SEM fit measures, therefore making the normalization procedure less reliable (not shown). Whether stand density was included in the analyses or not did not bias our results, since productivity normalized for climate and age was highly correlated with productivity normalized for stand density in addition (Pearson's r = 0.97). Species-specific normalized productivity for a species (P_{nsi}) was thus calculated as:

$$P_{nsi} = P_i - \textit{direct climate effect} - \textit{age effect} \qquad (2)$$

Finally, to make normalized productivity comparable among species within a dataset, the final normalized productivity (P_N) was computed by setting the averages for each species to zero (number of data points per species = *n*):

$$P_N = P_{nsi} - \mathrm{sum}(P_{nsi})/n \qquad (3)$$

To test whether the results from Van Sundert et al (2018) depended on the normalization procedure, we tested the SEM-based normalization also for the Swedish dataset used in Van Sundert et al (2018). This SEM-based normalized productivity was exactly the same as the normalized productivity derived from regression, presented in Van Sundert et al. (2018) (Table S9), hence indicating that possible correlation between soil properties and climate did not cause any artifacts. For the datasets considered in the current manuscript (ICP Forests, and to a lesser extent the grassland dataset), on the other hand, the parameter estimates obtained from the SEM approach did differ from those obtained through the regression approach.

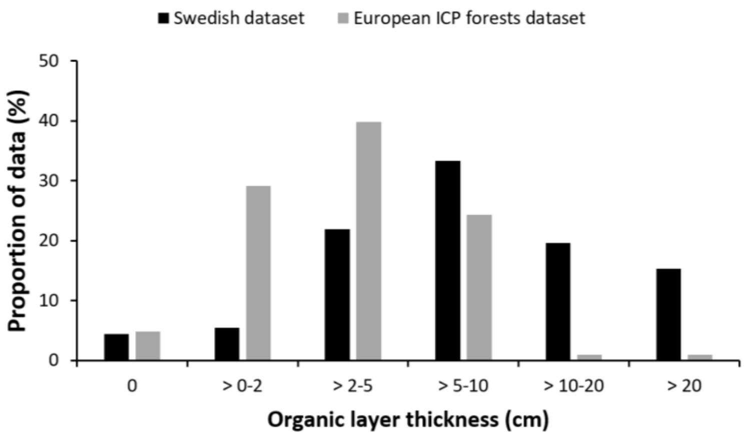

Figure S2 Histograms for the thickness of the organic layer for forests in Sweden (*n* = 1061) and for the European ICP Forests sites (*n* = 103).

Table S4 Matrix showing correlations (Pearson's *r*) among key climate and soil variables in the Swedish conifer forest dataset. These soil variables in particular were chosen because of their link with the soil nutrient status (e.g. Van Sundert et al., 2018), and our observation during exploratory analyses that organic layer characteristics in particular explain variation in normalized productivity across both the Swedish and ICP Forests datasets (e.g. Table S22). Abbreviations: TSUM = growing season temperature sum (°C days); MAP = mean annual precipitation (mm); C:N = soil carbon to nitrogen ratio; pH_{CaCl2} = soil pH, measured in $CaCl_2$ solution; TEB = total exchangeable bases ($cmol_+ kg^{-1}$); SOC = soil organic carbon concentration (%); org. = organic layer; min. = upper mineral soil; 0-20 cm = upper 20 cm of the soil, starting on top of the organic layer. Variables were log-transformed in case of positive skewness. Underlined correlations were significant (*P* < 0.05).

Climate or soil variable	MAP	C:N org.	pH_{CaCl2} org.	ln(TEB) min.	ln(SOC) 0-20 cm
TSUM	+0.11	-0.42	+0.04	+0.08	+0.28
MAP		-0.11	-0.05	-0.09	+0.13
C:N org.			-0.53	-0.32	-0.12
pH_{CaCl2} org.				+0.60	-0.01
ln(TEB) min.					+0.23

Table S5 Matrix showing correlations (Pearson's *r*) among key climate and soil variables in the European ICP Forests dataset. These soil variables in particular were chosen because of their link with the soil nutrient status (e.g. Van Sundert et al., 2018), and our observation during exploratory analyses that organic layer characteristics in particular explain variation in normalized productivity across both the Swedish and ICP Forests datasets (e.g. Table S22). Abbreviations: MAT = mean annual temperature (°C); MAP = mean annual precipitation (mm); C:N = soil carbon to nitrogen ratio; pH_{CaCl2} = soil pH, measured in $CaCl_2$ solution; C:P = soil carbon to phosphorus ratio (aqua regia extractable phosphorus was used as the best proxy for soil total P); TEB = total exchangeable bases (cmol+ kg^{-1}); SOC = soil organic carbon concentration (%); org. = organic layer; min. = upper mineral soil; 0-20 cm = upper 20 cm of the soil, starting on top of the organic layer. Variables were log-transformed in case of positive skewness. Underlined correlations were significant (P < 0.05).

Climate or soil variable	ln(MAP)	ln(C:N) org.	pH_{CaCl2} org.	ln(C:P) min.	ln(TEB) min.	ln(SOC) 0-20 cm
MAT	+0.07	-0.70	+0.26	+0.37	+0.31	+0.18
ln(MAP)		-0.23	+0.27	+0.24	+0.46	+0.56
ln(C:N) org.			-0.31	-0.33	-0.42	-0.35
pH_{CaCl2} org.				-0.11	+0.81	+0.11
ln(C:P) min.					+0.19	+0.54
ln(TEB) min.						+0.46

Table S6 Matrix showing correlations (Pearson's r) among key climate and soil variables in the global grasslands dataset. These soil variables in particular were chosen because of their link with the soil nutrient status (e.g. Van Sundert et al., 2018), and our observation during exploratory analyses that organic layer characteristics in particular explain variation in normalized productivity across both the Swedish and ICP Forests datasets (e.g. Table S8). In this dataset, only mineral soil data were available. Abbreviations: MAT = mean annual temperature (°C); GSP = growing season precipitation (mm); C:N = soil carbon to nitrogen ratio; pH = soil pH; P = soil total phosphorus (ppm); K + Ca + Mg = sum of total base cations relevant to plant nutrition (ppm); SC = soil organic carbon concentration (%); min. = upper mineral soil. Variables were log-transformed in case of positive skewness. Underlined correlations were significant (P < 0.05).

Climate or soil variable	ln(GSP)	ln(C:N) min.	pH min.	ln(P) min.	ln(K + Ca + Mg) min.	ln(SC) min.
MAT	+0.38	+0.11	-0.06	-0.02	+0.17	-0.22
ln(GSP)		+0.20	-0.63	+0.12	-0.02	+0.35
ln(C:N) min.			-0.49	-0.03	+0.08	+0.53
pH min.				+0.05	+0.31	-0.65
ln(P) min.					+0.58	-0.10
ln(K + Ca + Mg) min.						-0.21

Table S7 Structural equation model (SEM) parameters applied per species and dataset on the path diagram presented in Fig. S1, taking into account strongly correlated climate and soil variables (Pearson's r > 0.50; Tables S4-6). Abbreviations: MAI = mean annual volume increment over a rotation period (m³ ha^{-1} yr^{-1}); CAI = current (five-year averaged) annual volume increment; ANPP = aboveground net primary productivity (g C m^{-2} yr^{-1}); on TSUM = growing season temperature sum (°C days); MAP = mean annual precipitation (mm); GSP = growing season precipitation (mm); age class = forest age, discretized into seven categories; S(O)C = soil (organic) carbon concentration (%); C:N = soil carbon to nitrogen ratio; pH = soil pH; org. = organic layer; min. = upper mineral soil; 0-20 cm = upper 20 cm of the soil, starting on top of the organic layer; n/a = not applicable. Variables were log-transformed in case of positive skewness. Note that the SEMs were merely used for extracting direct climate, age and species influence as a better alternative to regression, and not for explaining the path diagram in Fig. S1. For the latter, too much variance remained unexplained by the SEMs applied on the Swedish forest and global grassland data (Table S8). Errors represent the s.e.m.

Dataset	Response variable	Direct climate effect	Age effect	Soil effect	Correlations taken into account
Swedish spruce forests (*n* = 1099)	MAI	$(+8.1 \pm 0.6) * 10^{-6}$ * TSUM² $(-6 \pm 1) * 10^{-3}$ * TSUM $(+1.0 \pm 0.4) * 10^{-3}$ * MAP	n/a[1]	(-0.37 ± 0.04) * $\ln(SOC_{0\text{-}20cm})$ (-0.039 ± 0.005) * $C:N_{org.}$	$\ln(SOC_{0\text{-}20cm})$ ~ MAP $C:N_{org.}$ ~ TSUM
Swedish pine forests (*n* = 1422)	MAI	$(-3.0 \pm 0.4) * 10^{-6}$ * TSUM² $(+11.2 \pm 0.9) * 10^{-3}$ * TSUM $(+0.2 \pm 0.4) * 10^{-3}$ * MAP	n/a[1]	(-0.35 ± 0.03) * $\ln(SOC_{0\text{-}20cm})$ (-0.030 ± 0.004) * $C:N_{org.}$	$\ln(SOC_{0\text{-}20cm})$ ~ MAP $C:N_{org.}$ ~ TSUM

[1]Age was already accounted for in the response variable by averaging annual volume increment over a rotation period.
[2]In contrast to forests, age in grassland was not taken into account.

Table S7 (continued).

Dataset	Response variable	Direct climate effect	Age effect	Soil effect	Correlations taken into account

European spruce forests (*n* = 23)	CAI	(+0.1 ± 0.2) * MAT (+5 ± 2) * ln(MAP)	(–1.1 ± 0.2) * age class	(-0.1 ± 0.9) * ln(SOC$_{0-20cm}$) (-8 ± 2) * ln(C:N$_{org.}$)	ln(SOC$_{0-20cm}$) ~ ln(MAP) ln(C:N$_{org.}$) ~ MAT
European pine forests (*n* = 22)	CAI	(+0.1 ± 0.2) * MAT (-5 ± 2) * ln(MAP)	(-0.8 ± 0.3) * age class	(+1 ± 1) * ln(SOC$_{0-20cm}$) (-6 ± 2) * ln(C:N$_{org.}$)	ln(SOC$_{0-20cm}$) ~ ln(MAP) ln(C:N$_{org.}$) ~ MAT
European beech forests (*n* = 24)	CAI	(+0.2 ± 0.3) * MAT (–1 ± 3) * ln(MAP)	(-0.7 ± 0.3) * age class	(-3.7 ± 0.7) * ln(SOC$_{0-20cm}$) (+6 ± 3) * ln(C:N$_{org.}$)	ln(SOC$_{0-20cm}$) ~ ln(MAP) ln(C:N$_{org.}$) ~ MAT
European oak forests (*n* = 8)	CAI	(+1 ± 1) * MAT (–19 ± 22) * ln(MAP)	(+0 ± 1) * age class	(+2 ± 6) * ln(SOC$_{0-20cm}$) (+7 ± 7) * ln(C:N$_{org.}$)	ln(SOC$_{0-20cm}$) ~ ln(MAP) ln(C:N$_{org.}$) ~ MAT
Grasslands worldwide (*n* = 68)	ln(ANPP)	(-0.01 ± 0.02) * MAT (+0.4 ± 0.2) * ln(GSP)	n/a[2]	(-0.21 ± 0.06) * ln²(SC$_{min.}$) (+0.6 ± 0.1) * ln(SC$_{min.}$) (0.0 ± 0.4) * ln(C:N$_{min.}$) (-0.1 ± 0.1) * pH$_{min.}$	ln²(SC$_{min.}$) ~ ln(GSP) ln(SC$_{min.}$) ~ ln(GSP) ln(C:N$_{min.}$) ~ ln(GSP) pH$_{min}$ ~ ln(GSP) ln²(SOC$_{min.}$) ~ ln(C:N$_{min.}$) ln(SC$_{min.}$) ~ ln(C:N$_{min.}$) pH$_{min}$ ~ ln(C:N$_{min.}$) ln²(SC$_{min.}$) ~ pH$_{min}$ ln(SC$_{min.}$) ~ pH$_{min.}$ ln²(SC$_{min.}$) ~ ln(SC$_{min.}$)

[1]Age was already accounted for in the response variable by averaging annual volume increment over a rotation period.
[2]In contrast to forests, age in grassland was not taken into account.

Table S8 Fit measures for structural equation models (SEMs) applied on the path diagram of Fig. S1, with parameters presented in Table S7 (blue = OK; orange = borderline case; red = not OK). Explanation of fit measures: implied vs observed = comparison of SEM implied vs observed covariance matrix (OK when $P > 0.05$); robust CFI = robust Comparative Fit Index (OK when ≥ 0.9); robust TLI = robust Tucker Lewis Index (OK when ≥ 0.9); robust RMSEA = robust Root Mean Square Error of Approximation (OK when ≤ 0.06); SRMR = Standardized

Root Mean Square Residual (OK when ≤ 0.05). Note that the SEMs were merely used for extracting direct climate, age and species influence as a better alternative to regression, and not for explaining the path diagram in Fig. S1. For the latter, too much variance remained unexplained by the SEMs applied on the Swedish forest and global grassland data. One single model for Sweden, and one single model for Europe were used for estimating parameters for multiple species at once.

Dataset	Implied vs observed	Robust CFI	Robust TLI	Robust RMSEA	SRMR	Model fit on diagram in Fig. S1 sufficient to explain observed values ?
Swedish conifer forests (*n* = 2521)	$\chi^2_{16} = 10421.58$ $P < 0.001$ ***	0.25	-0.41	0.748	0.643	No
European forests (*n* = 77)	$\chi^2_{32} = 39.76$ $P = 0.16$	0.94	0.89	0.118	0.138	+/-
Grasslands worldwide (*n* = 68)	$\chi^2_{5} = 30.92$ $P < 0.001$ ***	0.91	0.61	0.276	0.092	No

Table S9 Comparison of regression and structural equation model (SEM) parameters for equations describing spatial variation in productivity by climate, age, and soil (for b and c) across the Swedish and European ICP Forests databases. Correlations are given between productivity "normalized" for climate, age and species, following cases (a)-(c). Normalization consisted of subtracting the climate and age part of the equation from productivity, and setting the average result to zero per species, such that the average normalized productivity was species-independent. Abbreviations: TSUM = growing season temperature sum (°C days), MAP = mean annual precipitation (mm), GSP = growing season precipitation (mm), age class = forest age, discretized into seven categories, S(O)C = soil (organic) carbon concentration (%); C:N = soil carbon to nitrogen ratio; pH = soil pH; org. = organic layer; min. = upper mineral soil; 0-20 cm = upper 20 cm of the soil, starting on top of the organic layer. Variables were log-transformed in case of positive skewness. Errors represent the s.e.m.

Dataset	Productivity ~ climate + age regression derived direct climate effect (a)	Productivity ~ climate + age + soil regression derived direct climate effect (b)	Productivity ~ climate + age + soil SEM derived direct climate effect (c)	Correlations between climate- and age-normalized productivity from alternative models
Swedish spruce forests (*n* = 1099)	(+9.0 ± 0.6) * 10^{-6} * $TSUM^2$ (-8 ± 1) * 10^{-3} * TSUM (+0.3 ± 0.3) * 10^{-3} * MAP	(+8.1 ± 0.6) * 10^{-6} * $TSUM^2$ (-7 ± 1) * 10^{-3} * TSUM (+1.2 ± 0.4) * 10^{-3} * MAP	(+8.1 ± 0.6) * 10^{-6} * $TSUM^2$ (-6 ± 1) * 10^{-3} * TSUM (+1.0 ± 0.4) * 10^{-3} * MAP	a vs b: r = 0.99 a vs c: r = 1.00 b vs c: r = 0.99
Swedish pine forests (*n* = 1422)	(-3.0 ± 0.6) * 10^{-6} * $TSUM^2$ (+11.2 ± 0.9) * 10^{-3} * TSUM (+0.3 ± 0.3) * 10^{-3} * MAP	(-3.0 ± 0.6) * 10^{-6} * $TSUM^2$ (+11 ± 1) * 10^{-3} * TSUM (+0.6 ± 0.4) * 10^{-3} * MAP	(-3.0 ± 0.4) * 10^{-6} * $TSUM^2$ (+11.2 ± 0.9) * 10^{-3} * TSUM (+0.2 ± 0.4) * 10^{-3} * MAP	

Table S9 (continued).

Dataset	Productivity ~ climate + age regression derived direct climate effect (a)	Productivity ~ climate + age + soil regression derived direct climate effect (b)	Productivity ~ climate + age + soil SEM derived direct climate effect (c)	Correlations between climate- and age-normalized productivity from alternative models
European spruce forests (*n* = 23)	(+0.7 ± 0.2) * MAT (+5 ± 2) * ln(MAP) (–1.0 ± 0.3) * age class	(+0.6 ± 0.4) * MAT (+1 ± 3) * ln(MAP) (-0.7 ± 0.4) * age class	(+0.1 ± 0.2) * MAT (+5 ± 2) * ln(MAP) (–1.1 ± 0.2) * age class	a vs b: *r* = 0.90 a vs c: *r* = 0.91 b vs c: *r* = 0.96
European pine forests (*n* = 22)	(+0.5 ± 0.1) * MAT (-4 ± 2) * ln(MAP) (-0.5 ± 0.3) * age class	(+0.1 ± 0.2) * MAT (-4 ± 3) * ln(MAP) (-0.7 ± 0.4) * age class	(+0.1 ± 0.2) * MAT (-5 ± 2) * ln(MAP) (-0.8 ± 0.3) * age class	
European beech forests (*n* = 24)	(+0.5 ± 0.1) * MAT (-4 ± 2) * ln(MAP) (-0.5 ± 0.3) * age class	(-0.1 ± 0.4) * MAT (–1 ± 4) * ln(MAP) (-0.5 ± 0.4) * age class	(+0.2 ± 0.3) * MAT (–1 ± 3) * ln(MAP) (-0.7 ± 0.3) * age class	
European oak forests (*n* = 8)	(+1 ± 1) * MAT (-6 ± 10) * ln(MAP) (-0.3 ± 0.4) * age class	(+1 ± 3) * MAT (-20 ± 37) * ln(MAP) (+0 ± 2) * age class	(+1 ± 1) * MAT (–19 ± 22) * ln(MAP) (+0 ± 1) * age class	
Grasslands worldwide (*n* = 68)	(-0.02 ± 0.01) * MAT (+0.7 ± 0.1) * ln(GSP)	(-0.01 ± 0.02) * MAT (+0.4 ± 0.2) * ln(GSP)	(-0.01 ± 0.02) * MAT (+0.4 ± 0.2) * ln(GSP)	a vs b: *r* = 0.95 a vs c: *r* = 0.95 b vs c: *r* = 1.00

SOIL OR PLANT DATA TO ASSESS NUTRIENT STATUS?

Table S10 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by key soil characteristics across European beech forests. The selected model is marked in gray. Abbreviations and symbols: pH = soil pH, measured in $CaCl_2$ solution; C:P = soil carbon to phosphorus ratio (aqua regia extractable phosphorus was used as the best proxy for soil total P); SOC = soil organic carbon concentration (%) in the upper 20 cm of the soil, starting on top of the organic layer; org. = organic layer; min. = upper mineral soil. Variables were log-transformed in case of positive skewness. par X = parameter X, corresponding to element no. X in the variables column; ":" = ratio. For simplicity, only models including C:P, SOC and pH are shown, as preliminary analyses indicated importance of these variables. Only models with up to two explanatory variables were tested because of limited sample size. Errors represent the s.e.m.

Variables in model	Overall ms	Regression statistics
ln(C:P) min., ln(SOC)	12.7	par 1 = -3 ± 1 par 2 = -2.0 ± 0.9 intercept = 16 ± 4 $P < 0.001$ *** $R^2 = 0.57$ $n = 22$
ln(C:P) min., pH org.	11.8	
pH org., ln(SOC)	11.2	
pH^2 org., pH org.	12.3	
ln^2(SOC), ln(SOC)	12.0	
ln(C:P) min.	14.7	

Table S11 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by key soil characteristics across European spruce forests. The selected model is marked in gray. Abbreviations and symbols: C:N = soil carbon to nitrogen ratio, pH = soil pH, measured in $CaCl_2$ solution; C:P = soil carbon to phosphorus ratio (aqua regia extractable phosphorus was used as the best proxy for soil total P); TEB = total exchangeable bases ($cmol_+ \ kg^{-1}$); SOC = soil organic carbon concentration (%) in the upper 20 cm of the soil, starting on top of the organic layer; org. = organic layer; min. = upper mineral soil. Variables were log-transformed in case of positive skewness. par X = parameter X, corresponding to element no. X in the variables column; ":" = ratio. For simplicity, only models including C:P, SOC and pH are shown, as preliminary analyses indicated importance of these variables. Only models with up to two explanatory variables were tested because of limited sample size. Errors represent the s.e.m.

Variables in model	Overall ms	Regression statistics
ln(C:N) org., ln(SOC)	6.81	
ln(C:N) org., pH org.	7.43	
ln(C:N) org.	6.67	slope = -9 ± 2 intercept = 31 ± 8 *P* < 0.001 *** R^2 = 0.43 *n* = 22
ln(C:N) org., ln(C:P) min.	6.87	
ln(C:N) org., ln(TEB) min.	7.00	

Table S12 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by key soil characteristics across European pine forests. The selected model is marked in gray. Abbreviations and symbols: C:N = soil carbon to nitrogen ratio, pH = soil pH, measured in $CaCl_2$ solution; C:P = soil carbon to phosphorus ratio (aqua regia extractable phosphorus was used as the best proxy for soil total P); TEB = total exchangeable bases ($cmol_+ \ kg^{-1}$); SOC = soil organic carbon concentration (%) in the upper 20 cm of the soil, starting on top of the organic layer; org. = organic layer; min. = upper mineral soil. Variables were log-transformed in case of positive skewness. par X = parameter X, corresponding to element no. X in the variables column; ":" = ratio. For simplicity, only models including C:N, are shown, as preliminary analyses indicated importance of this variable. Only models with up to two explanatory variables were tested because of limited sample size. Errors represent the s.e.m.

Variables in model	Overall ms	Regression statistics
ln(C:N) org., ln(SOC)	3.29	
ln(C:N) org., ln(C:P) min.	3.70	
ln(C:N) org., ln(TEB) min.	3.84	
ln(C:N) org., pH org.	8.06	
ln(C:N) org.	3.27	slope = -6 ± 2 intercept = 22 ± 6 *P* < 0.001 *** R^2 = 0.42 *n* = 21

Table S13 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by key soil characteristics across European ICP Forests dataset. The selected model is marked in gray. Abbreviations and symbols: C:N = soil carbon to nitrogen ratio, pH = soil pH, measured in $CaCl_2$ solution; C:P = soil carbon to phosphorus ratio (aqua regia extractable phosphorus was used as the best proxy for soil total P); TEB = total exchangeable bases ($cmol_+$ kg^{-1}); SOC = soil organic carbon concentration (%) in the upper 20 cm of the soil, starting on top of the organic layer; org. = organic layer; min. = upper mineral soil. Variables were log-transformed in case of positive skewness. par X = parameter X, corresponding to element no. X in the variables column; ":" = ratio; * = interaction in the model, suggested by a regression tree. Errors represent the s.e.m.

Variables in model	Overall ms	Regression statistics
ln(C:N) org., ln(C:P) min., ln(TEB) min., ln²(SOC), ln(SOC)	8.16	
ln(C:N) org., ln(C:P) min., ln(TEB) min., ln(SOC)	7.54	
ln(C:N) org., ln(TEB) min., ln(SOC)	7.34	
ln(C:N) org., ln(SOC)	7.30	
ln(C:N) org., ln(SOC), ln(C:N) org. * ln(SOC)	7.23	par 1 = -7 ± 3 par 2 = –10 ± 7 par 3 = 2 ± 2 intercept = 27 ± 9 $P = 0.001$ ** $R^2 = 0.17$ $n = 72$
ln(C:N) org., ln(C:P) min.., ln²(SOC), ln(SOC), pH² org., pH org.	8.49	
ln(C:N) org., ln(C:P) min.., ln(SOC), pH² org., pH org.	7.86	
ln(C:N) org., ln(C:P) min.., ln(SOC), pH org.	7.53	
ln(C:N) org., ln(SOC), pH org.	7.35	

Table S14 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by foliar stoichiometry across European beech forests. The selected model is marked in gray. Symbol: ":" = ratio. P and N:P were never included in the same model because of multicollinearity (variance inflation factor > 4). Only models with up to two explanatory variables were tested because of limited sample size.

Variables in model	Overall ms	Regression statistics
N	12.7	
P	11.8	
K	11.2	*P* = 0.93 *n* = 22
Ca	12.3	
Mg	12.0	
S	14.7	
N:P	12.1	
N, P	14.1	
N, K	12.9	
N, Ca	14.0	
N, Mg	13.7	
N, S	15.2	
N, N:P	14.1	
P, K	12.0	
P, Ca	12.7	
P, Mg	12.8	
P, S	15.0	
K, Ca	12.5	
K, Mg	12.6	
K, S	16.3	
K, N:P	12.3	
Ca, Mg	14.8	
Ca, S	15.9	
Ca, N:P	13.3	
Mg, S	14.8	
Mg, N:P	12.7	
S, N:P	14.8	

Table S15 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by foliar stoichiometry across European spruce forests. The selected model is marked in gray. Abbreviations and symbols: par X = parameter X, corresponding to element no. X in the variables column; ":" = ratio. P and N:P were not included in the same model because of multicollinearity (variance inflation factor > 4). For simplicity, only models including N:P are shown, as preliminary analyses indicated importance of this variable. Only models with up to two explanatory variables were tested because of limited sample size. Errors represent the s.e.m.

Variables in model	Overall ms	Regression statistics
N:P, N	6.66	par 1 = 0.7 ± 0.3 par 2 = 0.6 ± 0.4 intercept = -15 ± 5 *P* = 0.009 ** R^2 = 0.32 *n* = 22
N:P, K	8.44	
N:P, Ca	8.07	
N:P, Mg	7.36	
N:P, S	8.01	
N:P	7.29	

Table S16 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by foliar stoichiometry across European pine forests. The selected model is marked in gray. Abbreviations and symbols ":" = ratio. N and S were not included in the same model because of multicollinearity (variance inflation factor > 4). Only models with up to two explanatory variables were tested because of limited sample size. Errors represent the s.e.m.

Variables in model	Overall ms	Regression statistics
N	7.88	
P	5.76	
K	6.48	
Ca	5.08	
Mg	6.20	
S	7.17	
N:P	7.21	
N, P	8.14	
N, K	8.51	
N, Ca	7.38	
N, Mg	9.52	
N, N:P	7.85	
P, K	6.93	
P, Ca	5.60	$P = 0.20$ $R^2 = 0.07$ $n = 21$
P, Mg	6.52	
P, S	7.52	
P, N:P	7.74	
K, Ca	5.79	
K, Mg	6.40	
K, S	8.23	
K, N:P	8.00	
Ca, Mg	7.90	
Ca, S	6.52	
Ca, N:P	7.32	
Mg, S	8.49	
Mg, N:P	9.73	
S, N:P	7.57	

Table S17 Overall mean squared error (ms) after 10-fold cross-validation for candidate model structures that explain variation in normalized productivity by foliar stoichiometry across the European ICP Forests dataset. The selected model is marked in gray. Symbol: ":" = ratio. N/P and N:P, and S and N were never included in the same model because of multicollinearity (variance inflation factor > 4). Errors represent the s.e.m.

Variables in model	Overall ms	Regression statistics
N, P, K, Ca, Mg	10.4	
N, K, Ca, Mg	9.66	
K, Ca, Mg	9.25	
Ca, Mg	8.71	
Ca	8.69	P = 0.26 n = 72
N:P, K, Ca, Mg	9.97	
S, P, K, Ca, Mg	9.90	
S, K, Ca, Mg	9.35	
S, K, Ca	9.09	
S, K	8.88	
S	9.12	

Table S18 Selected regression models explaining variation in normalized productivity by foliar stoichiometry across a more elaborate version of the European ICP Forests dataset (with data on key soil variables lacking for the principal analyses in this study, but including data on stoichiometry). Abbreviations and symbols: par X = parameter X, corresponding to element no. X in the variable's column; ":" = ratio. Errors represent the s.e.m.

Species	Variables in selected model	Regression statistics
Beech	P	*P* = 0.23 *n* = 30
Spruce	N:P, K, Mg, ln(S)	par 1 = 0.4 ± 0.2 par 2 = -0.8 ± 0.4 par 3 = 3 ± 2 par 4 = 4 ± 3 intercept = -3 ± 4 *P* = 0.004 ** R^2 = 0.25 *n* = 42
Pine	Ca	slope = 1.0 ± 0.3 intercept = -3 ± 1 *P* = 0.001 ** R^2 = 0.28 *n* = 30
ALL (Beech, Spruce, Pine, Oak)	ln(Ca), ln(Mg)	par 1 = 1.5 ± 0.6 par 2 = –1.7 ± 0.9 intercept = -2 ± 1 *P* = 0.05 * R^2 = 0.04 *n* = 114

Table S19 Matrix showing correlations (Pearson's *r*) among key soil variables in the European ICP Forests dataset. These soil variables in particular were chosen because of their link with the soil nutrient status (e.g. Van

Sundert et al., 2018), and our observation during exploratory analyses that organic layer characteristics in particular explain variation in normalized productivity across both the Swedish and ICP Forests datasets (e.g. Table S22). Note that correlations were not identical to those presented in Table S5 because for Table S5, a more elaborate version of the ICP Forests dataset was used with sufficient data for normalization of productivity, but not for e.g. comparison of soil vs plant data to explain variation in normalized productivity. Abbreviations: C:N = soil carbon to nitrogen ratio; pH_{CaCl2} = soil pH, measured in $CaCl_2$ solution; C:P = soil carbon to phosphorus ratio (aqua regia extractable phosphorus was used as the best proxy for soil total P); TEB = total exchangeable bases ($cmol_+ kg^{-1}$); SOC = soil organic carbon concentration (%); org. = organic layer; min. = upper mineral soil; 0-20 cm = upper 20 cm of the soil, starting on top of the organic layer. Variables were log-transformed in case of positive skewness. Underlined correlations were significant ($P < 0.05$).

Climate or soil variable	pH_{CaCl2} org.	ln(C:P) min.	ln(TEB) min.	ln(SOC) 0-20 cm
ln(C:N) org.	-0.21	-0.31	-0.34	-0.31
pH_{CaCl2} org.		-0.14	+0.82	+0.05
ln(C:P) min.			+0.12	+0.54
ln(TEB) min.				+0.45

Table S20 Matrix showing correlations (Pearson's *r*) among foliar stoichiometry variables in the European ICP Forests dataset. Underlined correlations were significant ($P < 0.05$).

Foliar nutrient concentration or ratio	P	K	Ca	Mg	S	N:P
N	-0.12	+0.80	+0.50	+0.43	+0.90	+0.85
P		+0.12	-0.21	-0.04	-0.17	-0.57
K			+0.48	+0.52	+0.80	+0.61
Ca				+0.65	+0.59	+0.56
Mg					+0.49	+0.44
S						+0.82

EXAMPLE: A SOIL-BASED METRIC OF THE NUTRIENT STATUS

Evaluation of the earlier nutrient metric

Performance of the earlier metric

Table S21 Ability of the original nutrient metric developed by IIASA (IIASA & FAO, 2012; see also Van Sundert et al., 2018 for analysis of the IIASA-metric across all of Sweden) to explain variation in normalized productivity across the southern Swedish and European ICP Forests datasets. Same datasets were used as in Table 2. The IIASA-metric was not evaluated against global grassland data because it requires soil total exchangeable bases, which were not available in that dataset. *n* varied because of data availability.

Dataset	Explanatory power of original IIASA-metric (mineral soil)	Explanatory power of original IIASA-metric (upper 20 cm of soil)
Swedish conifer forests (southern Sweden only)	$P = 0.44$ $n = 331$	$P = 0.55$ $n = 454$
European forests	$P = 0.32$ $n = 79$	$P = 0.59$ $n = 57$
European spruce forests	$P = 0.64$ $n = 28$	$P = 0.25$ $n = 15$
European pine forests	$P = 0.94$ $n = 17$	$P = 0.75$ $n = 15$
European beech forests	$P = 0.26$ $n = 25$	$P = 0.46$ $n = 20$

Adjustment of the earlier metric

Adjusting the earlier metric

Table S22 Ability of mineral vs organic soil C:N ratio and pH to explain spatial variation in normalized productivity of Swedish conifer forests and European forests in the ICP Forests database. Abbreviations: quad = parameter estimate for quadratic term; lin = parameter estimate for the linear term of a quadratic function; ic = intercept. Note that for the Swedish dataset, only data of southern Sweden were used for heteroscedasticity reasons explained in Van Sundert et al. (2018). Moreover, only half of the southern data was used such that this subset served as a calibration set to adjust the nutrient metric, and the other half was used to evaluate the adjusted metric. Parameter estimates shown for pH differ from those in Eq. 7 because the pH optimum was fixed to 4.5 based on the ICP Forests database. *n* for the ICP Forests database was larger here than elsewhere in the manuscript because of missing data on $SOC_{0\text{-}20cm}$, which was needed for the other analyses. Errors represent the s.e.m.

Dataset	Mineral soil C:N ratio[a]	ln(organic soil C:N ratio)	Mineral soil pH_{CaCl2}	Organic soil pH_{CaCl2}
Swedish conifer forests	slope = -0.02 ± 0.01 ic = 0.6 ± 0.2 P = 0.10 (*) R^2 = 0.01 n = 340	slope = –1.8 ± 0.2 ic = 5.7 ± 0.8 P < 0.001 *** R^2 = 0.09 n = 538	P = 0.55 n = 340	quad = -0.8 ± 0.1 lin = 7 ± 1 ic = –13 ± 2 P < 0.001 *** R^2 = 0.12 n = 264
European forests	slope = –1.6 ± 0.8 ic = 5 ± 2 P = 0.05 * R^2 = 0.03 n = 118	slope = -3.2 ± 0.9 ic = 11 ± 3 P < 0.001 *** R^2 = 0.09 n = 112	P = 0.86 n = 116	quad = -0.7 ± 0.4 lin = 6 ± 3 ic = –13 ± 6 P = 0.05 * R^2 = 0.04 n = 111

[a] The frequency distribution of mineral soil C:N was right-skewed for the ICP Forests database, but not for the Swedish database, such that C:N was log-transformed for ICP but not for Sweden.

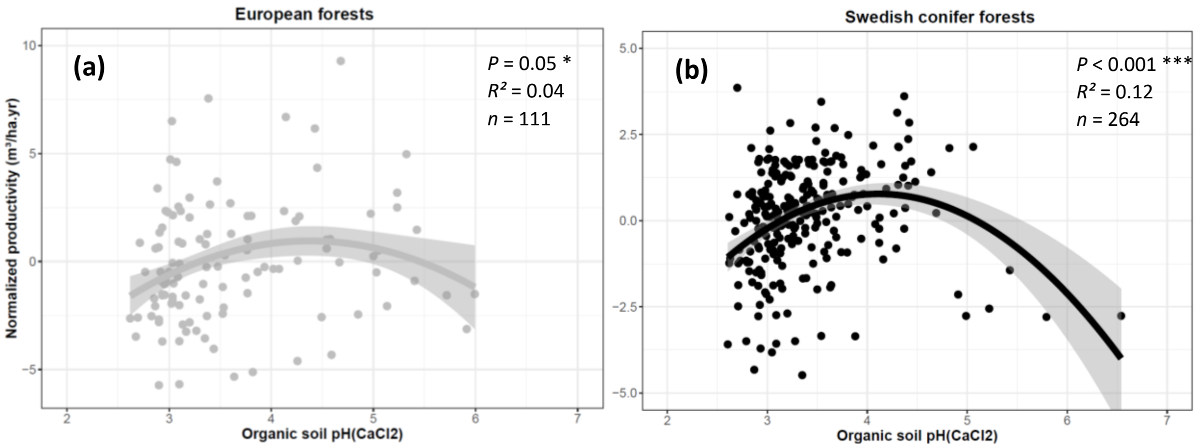

Figure S3 Normalized productivity versus organic soil pH_{CaCl2} for (a) forests in the European ICP Forests database with the necessary data, and (b) conifer forests in the Swedish database. Shaded area around the regression curve represents 95% confidence intervals. Parameter estimates for the full dataset are shown in Table S21.

The adjusted metric versus multiple regressions

Table S23 Comparison of nutrient metric abilities to explain variation in productivity across different natural gradients in soil characteristics and productivity in Sweden. The adjusted metric refers to the metric presented in the current paper (Eqs. 5-7 in 4), whereas the regression equation represents a multiple regression model using the same soil variables as the adjusted metric (Eq. 8). All three metrics were calibrated using data of southern Sweden. The gradients represent regional sets of data points from the Swedish dataset used in the current study that did not show large variation in climate, such that productivity could be used as a response variable to test the metrics against without the necessity to normalize for climate. More details regarding the gradients including selection of the data points and their position within Sweden are given in Van Sundert et al. (2018). Note that for Norway spruce, no gradient in total exchangeable bases (TEB) without substantial variation in climate was found, so that only for Scots pine, there was a gradient in TEB. Errors represent the s.e.m. Errors represent the s.e.m.

Gradient in Sweden	Explanatory power of metric presented in Van Sundert et al. (2018)	Explanatory power of adjusted metric	Explanatory power of regression equation
Norway spruce: soil moisture gradient	slope = 1.6 ± 0.4 *P* < 0.001 *** R^2 = 0.13 *n* = 133	slope = 2.3 ± 0.4 *P* < 0.001 *** R^2 = 0.26 *n* = 74	slope = 1.1 ± 0.2 *P* < 0.001 *** R^2 = 0.27 *n* = 74
Norway spruce: productivity gradient	slope = 1.6 ± 0.4 *P* < 0.001 *** R^2 = 0.15 *n* = 79	slope = 3.1 ± 0.6 *P* < 0.001 *** R^2 = 0.33 *n* = 51	slope = 1.6 ± 0.3 *P* < 0.001 *** R^2 = 0.30 *n* = 51
Scots pine: soil moisture gradient	slope = 1.4 ± 0.2 *P* < 0.001 *** R^2 = 0.21 *n* = 142	slope = 2.1 ± 0.4 *P* < 0.001 *** R^2 = 0.27 *n* = 83	slope = 1.0 ± 0.2 *P* < 0.001 *** R^2 = 0.27 *n* = 83
Scots pine: TEB gradient	slope = 1.1 ± 0.3 *P* < 0.001 *** R^2 = 0.20 *n* = 60	slope = 1.7 ± 0.6 *P* = 0.01 * R^2 = 0.24 *n* = 21	slope = 0.9 ± 0.3 *P* = 0.001 * R^2 = 0.26 *n* = 21
Scots pine: productivity gradient	slope = 1.9 ± 0.3 *P* < 0.001 *** R^2 = 0.35 *n* = 68	slope = 2.7 ± 0.6 *P* < 0.001 *** R^2 = 0.31 *n* = 44	slope = 1.4 ± 0.3 *P* < 0.001 *** R^2 = 0.34 *n* = 39

Table S24 Tests of variable implementation in regression equation 8, presented as an alternative to the nutrient metric in this paper. Associations between residuals of normalized productivities and soil variables used for the equation are shown. Abbreviations: SOC = soil organic carbon concentration; soil C:N ratio = soil carbon to nitrogen ratio. For the Swedish data, a validation subset of southern Swedish forests was used instead of the dataset of entire Sweden to avoid heteroscedasticity-induced artifacts (see Van Sundert et al. (2018)). For the grassland dataset, mineral soil data were used to calculate the metric because no organic layer data were available. Errors represent the s.e.m.

Dataset	ln SOC (%)	ln soil C:N ratio	pH
European spruce forests (*n* = 23)	$P = 0.14$	slope = -6 ± 2 $P = 0.03$ * $R^2 = 0.17$	$P = 0.31$
European pine forests (*n* = 22)	$P = 0.22$	$P = 0.67$	slope = 1.3 ± 0.7 $P = 0.06$ (*) $R^2 = 0.21$
European beech forests (*n* = 24)	slope = -2.5 ± 0.8 $P = 0.006$ ** $R^2 = 0.27$	$P = 0.22$	slope = 1.2 ± 0.6 $P = 0.05$ (*) $R^2 = 0.12$
Grasslands worldwide (*n* = 68)	$P = 0.80$	$P = 0.47$	$P = 0.70$

Applications of the soil-based metric and future prospects

Table S25 Ability of mineral vs organic soil C:P ratio and soil total P to explain spatial variation in normalized productivity of European forests in the ICP Forests database. Aqua regia extractable P was taken here as the best available proxy for soil total P, such that actual total P as derived from the acid digestion method may have been underestimated (ISO 11466, 1995; Ivanov, 2012). Errors represent the s.e.m.

Dataset	ln(mineral soil C:P ratio)	Organic soil C:P ratio	ln(mineral soil total P) (mg kg^{-1})	Organic soil total P (mg kg^{-1})
European forests	P = 0.54 n = 106	slope = -0.005 ± 0.002 ic = 2.2 ± 0.9 P = 0.004 ** R^2 = 0.06 n = 109	P = 0.54 n = 106	P = 0.18 n = 113
European spruce forests	P = 0.96 n = 42	slope = -0.008 ± 0.004 ic = 4 ± 2 P = 0.03 * R^2 = 0.09 n = 42	P = 0.15 n = 42	P = 0.31 n = 43
European pine forests	slope = 1.3 ± 0.5 ic = -6 ± 2 P = 0.01 * R^2 = 0.18 n = 29	P = 0.34 n = 30	P = 0.18 n = 29	P = 0.66 n = 30
European beech forests	slope = -3.9 ± 0.8 ic = 19 ± 4 P < 0.001 *** R^2 = 0.45 n = 27	slope = -0.010 ± 0.004 ic = 4 ± 2 P = 0.03 * R^2 = 0.14 n = 28	P = 0.81 n = 29	P = 0.42 n = 29

References in supporting information

Anderegg, W. R. L., Schwalm, C., Biondi, F., Camarero, J. J., Koch, G., Litvak, M., ...Pascala, S. (2015).

Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. *Science*, 349(6247), 528–532.

Binkley, D., & Högberg, P. (2016). Tamm review: revisiting the influence of nitrogen deposition on Swedish forests. *Forest Ecology and Management*, 368, 222–239.

Cools, N., De Vos, B. (2016). Part X: Sampling and analysis of soil. In: UNECE ICP Forests Programme Coordinating Centre (ed.), Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Eberswalde: Thünen Institute of Forest Ecosystems. available at http://www.icp-forests.net/page/icp-forests-manual, last access: 24-11-2018.

De Vries, W., Reinds G. J., Posch M., Sanz, M. J., Krause, G. H. M., Calatyud, V., … Vel, E. M. (2003). Intensive Monitoring of Forest Ecosystems in Europe. Technical Report. Brussels, Geneva: EC, UN/ECE.

Dobbertin, M. & Neumann, M. in Manual on methods and for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Vol. [Published online 05/2010] 1–29 (UNECE ICP-Forests Programme Co-ordinating Centre., 2010).

Fleck, S., Cools, N., De Vos, B., Meesenburg, H., & Fischer, R. (2016). The level II aggregated forest soil condition database links soil physicochemical and hydraulic properties with long-term observations of forest condition in Europe. *Annals of Forest Science*, 73(4), 945–957.

ICP Forests. (2010). Manual on methods and for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Hamburg: UNECE ICP Forests Programme Co-ordinating Centre.

IIASA, & FAO. (2012). Global Agro-ecological Zones (GAEZ v3.0). Laxenburg, Rome: International Instutute for Applied Systems Analysis, Food and Agricultural Organization of the United Nations.

ISO 11466. (1995). Soil Quality – Extraction of trace elements soluble in aqua regia. Geneva: International Organization for Standardization. available at www.iso.ch, last access: 24-11-2018.

Ivanov, K., Zaprjanova, P., Petkova, M., Stefanova, V., Kmetov, V., Georgieva, D., & Angelova, V. (2012). Comparison of inductively coupled plasma mass spectrometry and colorimetric determination of total and extractable phosphorus in soils. *Spectrochimica Acta, Part B: Atomic Spectroscopy*, 71-72, 117–122.

Jandl, R., Smidt, S., Mutsch, F., Fürst, A., Zechmeister, H., Bauer, H. & Dirnböck, T. (2012). Acidification and nitrogen eutrophication of Austrian forest soils. *Applied and Environmental Soil Science*, 2012, 1–9.

Kirk, G. J. D., Bellamy, P. H., & Lark, R. M. (2009). Changes in soil pH across England and Wales in response to decreased acid deposition. *Global Change Biology*, 3111–3119.

Novotný, R., Buriánek, V., Šrámek, V., Hunová, I., Skorepová, I., Zapletal, M., & Lomský, B. (2015). Nitrogen deposition and its impact on forest ecosystems in the Czech Republic – change in soil chemistry and ground vegetation. *iForest - Biogeosciences and Forestry*, 10 (1), 48–54.

Olsson M. T., Erlandsson M., Lundin L., Nilsson T., Nillson A., & Stendahl J. (2009). Organic carbon stocks in Swedish podzol soils in relation to soil hydrology and other site characteristics. *Silva Fennica*, 43, 209–222.

Rosseel, Y. (2012). Lavaan: an R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36.

Solberg, S., Dobbertin, M., Reinds, G. J., Lange, H., Andreassen, K., Garcia Fernandez, P., Hildingsson, A., & de Vries, W. (2009). Analyses of the impact of changes in atmospheric deposition and climate on forest growth in European monitoring plots: a stand growth approach. *Forest Ecology and Management*, 258(8), 1735–1750.

Stendahl, J., Johansson, M. B., Eriksson, E., Nilsson, A., & Langvall, O. (2010). Soil organic carbon in Swedish spruce and pine forests – differences in stock levels and regional patterns. *Silva Fennica*, 44, 5–21.

Van Sundert, K., Horemans, J. A., Stendahl, J. & Vicca, S. (2018). The influence of soil properties and nutrients on conifer forest growth in Sweden, and the first steps in developing a nutrient availability metric. *Biogeosciences*, 15, 3475–3496.