

### Supplementary Information I

Table SII: Mean value of each foliar property per plot. In all cases analysis of variance identified significant differences among plots for all log<sub>10</sub> transformed foliar properties (absolute values for δ<sup>13</sup>C), at a Bonferroni (0.05/62= 0.0008) adjusted level.

Plot	lat	long	M <sub>A</sub> (g m <sup>-2</sup> )	C (mg g <sup>-1</sup> )	N (mg g <sup>-1</sup> )	P (mg g <sup>-1</sup> )	Ca (mg g <sup>-1</sup> )	K (mg g <sup>-1</sup> )	Mg (mg g <sup>-1</sup> )	Al (mg g <sup>-1</sup> )	δ <sup>13</sup> C (‰)
AGP-01	-3.74	-70.31	103±31	455±41	20.87±4.24	1.06±0.38	9.76±6.00	9.13±6.00	2.83±1.70	0.19±0.21	-33.11±1.10
AGP-02	-3.74	-70.30	96±19	457±29	19.17±3.43	0.96±0.24	9.92±4.28	6.77±2.63	2.79±1.82	0.21±0.30	-33.06±1.04
ALP-11	-3.95	-73.43	140±51	525±31	17.60±4.75	0.84±0.32	7.26±2.29	6.14±3.67	2.13±0.54	0.05±0.02	-30.02±1.52
ALP-12	-3.95	-73.44	124±19	513±16	20.32±4.73	0.78±0.19	6.90±2.49	4.89±1.56	1.63±0.67	1.30±3.33	-29.99±1.41
ALP-21	-3.95	-73.44	134±27	524±19	20.41±6.21	0.98±0.31	4.62±2.90	4.06±0.86	1.73±0.64	0.04±0.01	-30.36±1.73
ALP-22	-3.95	-73.44	95±23	497±24	21.22±5.39	1.02±0.19	7.60±4.09	6.65±2.50	1.84±0.81	0.06±0.03	-30.73±0.93
ALP-30	-3.95	-73.43	115±26	527±18	20.28±4.64	1.02±0.45	6.60±3.34	3.51±0.98	1.93±0.77	0.06±0.02	-31.01±1.75
BNT-04	-2.63	-60.15	103±25	491±16	19.89±4.27	0.54±0.15	2.19±1.17	3.15±0.94	1.33±0.37	0.09±0.13	-32.13±1.15
BOG-01	-0.70	-76.48	91±26	455±43	25.26±5.81	1.61±0.50	13.15±8.02	9.16±4.04	2.63±1.17	0.07±0.14	-30.94±1.30
BOG-02	-0.70	-76.47	95±24	439±50	23.02±5.25	1.44±0.46	16.16±11.38	13.03±6.92	3.55±3.04	0.05±0.02	-30.76±1.27
CAX-01	-1.74	-51.46	82±19	464±32	23.82±6.33	0.60±0.17	3.06±1.90	2.87±2.23	3.15±1.34	0.09±0.24	-33.49±1.01
CAX-02	-1.74	-51.46	85±19	465±24	22.11±3.78	0.74±0.13	5.78±4.27	2.28±1.03	3.05±1.73	0.07±0.19	-32.35±1.68
CAX-03	-1.73	-51.46	90±27	471±46	20.63±6.11	0.55±0.16	3.69±2.10	1.60±0.94	2.64±1.47	0.15±0.24	-32.61±1.32
CAX-04	-1.73	-51.46	93±17	479±22	19.56±4.62	0.54±0.16	3.76±2.34	1.33±0.67	2.10±0.79	0.39±1.34	-32.10±1.20
CAX-05	-1.72	-51.46	87±34	468±29	19.80±4.89	0.57±0.16	3.60±2.53	1.96±1.09	2.38±1.28	0.19±0.07	-32.59±1.67
CHO-01	-14.39	-61.15	93±32	446±42	22.19±4.43	0.95±0.28	7.67±4.33	11.57±5.65	3.60±1.76	1.11±4.07	-31.85±1.49

CPP-01	-2.19	-47.33	95±20	487±26	18.11±2.86	0.62±0.13	5.23±5.00	4.36±2.68	2.04±1.53	0.00±0.00	-31.97±1.19
CUZ-03	-12.50	-68.96	88±26	438±31	21.79±4.78	1.68±0.68	14.12±4.93	11.73±4.48	2.52±0.97	0.06±0.06	-31.09±1.13
ELD-12	6.10	-61.40	80±16	491±18	20.51±4.34	0.66±0.18	4.48±3.04	5.83±3.40	2.42±1.13	0.13±0.10	-33.36±1.42
ELD-34	6.08	-61.41	78±26	456±46	20.45±5.26	0.80±0.23	11.03±9.61	8.20±2.63	2.69±1.32	0.13±0.06	-32.63±1.84
HCC-21	-14.56	-60.75	76±22	484±24	30.69±3.80	1.03±0.16	12.33±3.87	8.87±2.57	2.70±0.99	0.06±0.02	-30.30±0.73
HCC-22	-14.57	-60.75	90±22	439±45	28.79±6.07	1.32±0.45	15.73±8.62	12.97±9.28	4.78±2.89	1.22±4.00	-30.33±1.29
JAC-04	-2.61	-60.22	127±27	480±38	12.61±3.95	0.48±0.10	3.13±1.91	5.22±3.50	2.54±1.47	0.21±0.06	-32.09±1.04
JAC-12	-2.61	-60.21	103±25	491±28	22.33±5.82	0.60±0.18	2.62±1.55	3.66±1.75	1.97±0.96	0.25±0.55	-31.63±1.33
JAS-02	-1.07	-77.62	113±35	482±35	23.31±5.13	1.03±0.28	7.91±3.88	9.52±4.89	2.00±0.75	0.78±2.77	-30.12±1.00
JAS-03	-1.08	-77.61	103±36	470±39	24.48±6.68	1.24±0.45	11.82±5.90	9.94±4.76	2.34±1.16	1.56±4.86	-30.26±1.06
JAS-04	-1.07	-77.61	120±34	500±33	18.89±3.25	0.83±0.27	7.18±2.80	6.20±4.58	1.61±0.75	3.03±7.17	-30.04±1.19
JAS-05	-1.06	-77.62	94±21	482±33	26.97±4.91	2.10±0.71	12.57±6.82	11.48±4.98	2.74±1.37	0.06±0.03	-29.95±1.08
JRI-01	-0.89	-52.19	120±29	483±19	17.94±4.57	0.57±0.14	5.02±2.94	3.44±2.16	2.21±0.72	0.00±0.00	-30.10±1.60
JUR-01	-8.88	-72.79	97±32	396±28	25.28±5.65	1.56±0.27	11.36±2.88	8.94±3.82	3.34±0.06	0.21±0.07	-31.44±1.85
LFB-01	-14.56	-60.93	79±24	449±28	22.46±4.43	0.90±0.37	5.49±1.64	10.56±4.64	2.67±0.98	2.38±5.11	-32.73±1.60
LFB-02	-14.58	-60.83	96±19	468±34	20.74±4.63	0.78±0.23	4.18±1.74	6.32±2.26	2.04±0.82	6.01±7.57	-31.97±0.98
LOR-01	-3.06	-69.99	101±21	488±18	19.29±3.35	1.07±0.41	4.77±2.58	6.87±2.98	2.22±1.02	0.06±0.05	-32.59±1.75
LOR-02	-3.06	-69.99	93±20	472±29	21.25±6.02	1.14±0.47	5.49±2.55	8.75±6.12	2.26±0.98	0.81±2.19	-33.18±1.17
LSL-01	-14.41	-61.14	77±20	483±15	23.86±4.58	1.18±0.41	6.45±4.09	7.77±2.86	2.15±1.04	1.98±4.36	-31.40±1.85
LSL-02	-14.41	-61.14	73±20	455±31	20.13±2.80	1.06±0.22	6.63±2.56	8.10±4.76	2.28±0.99	10.07±12.48	-31.56±1.01
MBO-01	-1.45	-48.45	90±20	474±28	21.35±5.18	0.74±0.16	2.91±1.69	3.91±1.06	2.32±1.09	2.11±5.81	-31.38±0.79
RES-04	-10.80	-68.77	93±30		19.47±4.08						
RES-05	-10.57	-68.31	85±20	424±45	22.64±4.66	0.95±0.21	13.20±10.24	7.63±4.96	3.36±1.48	0.02±0.03	-31.43±1.68
RES-06	-10.56	-68.30	85±23	436±46	23.20±5.32	1.20±0.38	14.37±10.40	9.01±4.15	4.49±3.25	0.45±1.25	-31.61±1.53

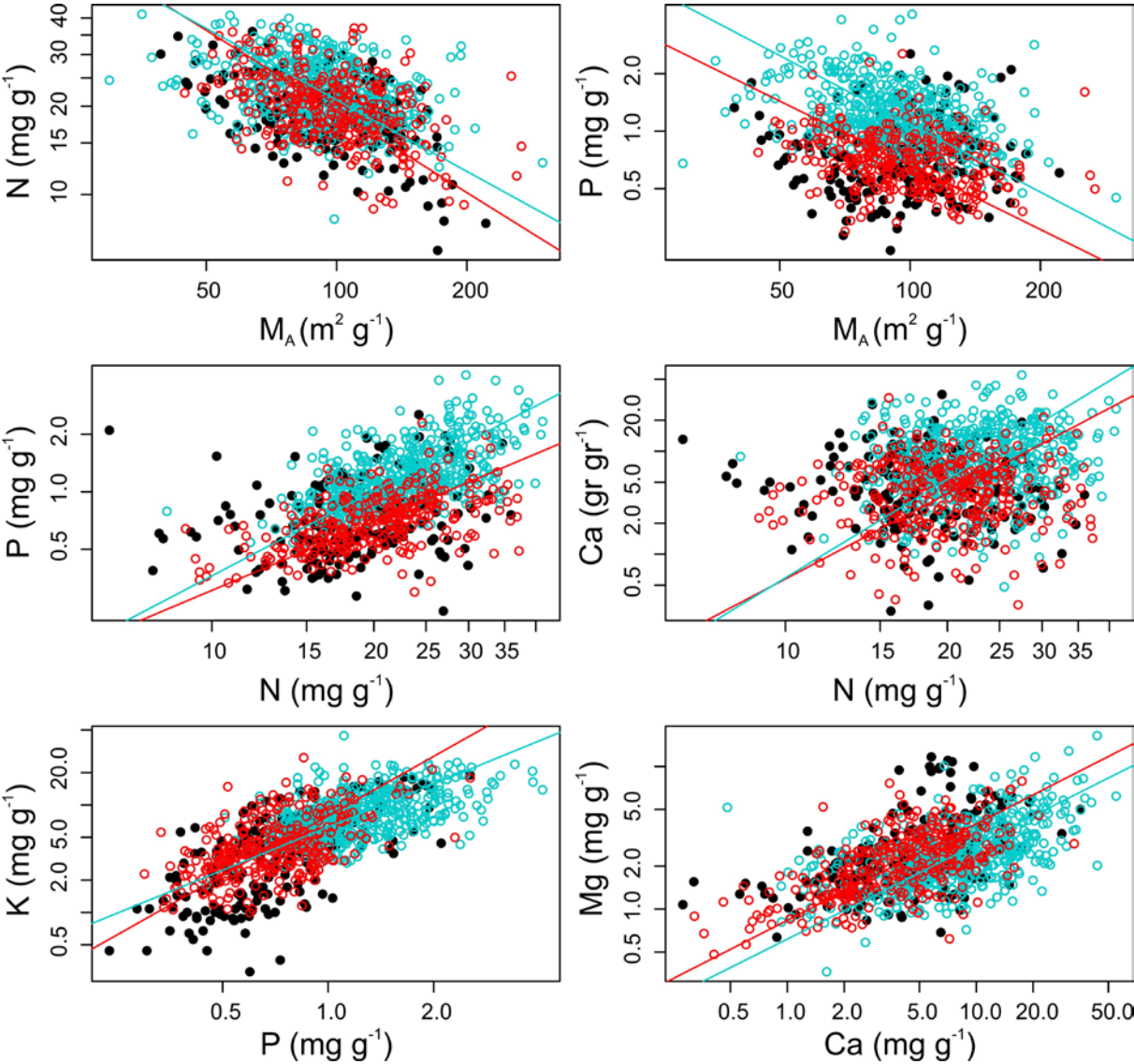
RIO-12	8.11	-61.69	96±27	455±42	16.62±5.43	0.59±0.16	7.43±7.71	4.39±3.08	3.37±1.68	0.15±0.06	-32.59±0.98
SCR-01	1.93	-67.02	124±25	504±11	15.25±2.88	1.40±1.04	1.12±1.00	9.53±7.75	1.04±0.28	0.04±0.01	-31.68±1.52
SCR-04	1.93	-67.04	141±39	512±15	12.48±5.23	1.03±0.53	5.18±3.24	5.69±1.92	1.89±0.75	0.06±0.09	-31.55±1.34
SCR-05	1.93	-67.04	155±55	540±18	16.10±4.13	0.64±0.16	1.41±0.79	5.05±1.54	1.01±0.42	0.14±0.18	-31.65±1.20
SUC-01	-3.25	-72.91	127±56	510±38	19.87±4.30	0.90±0.28	5.77±2.52	6.25±3.14	2.51±1.58	0.17±0.36	-32.00±1.58
SUC-02	-3.25	-72.90	113±29	493±30	18.16±5.52	0.90±0.21	5.37±3.31	6.49±2.52	1.95±0.80	0.08±0.05	-31.01±1.64
SUC-03	-3.25	-72.92	109±28	495±24	18.53±2.82	0.91±0.20	8.72±4.38	5.24±1.94	2.82±1.58	0.07±0.03	-31.65±1.63
SUM-01	-1.75	-77.63	104±33	476±21	25.34±7.41	1.61±0.50	7.56±2.63	8.16±2.68	2.00±0.85	0.48±1.31	-28.93±1.02
TAM-01	-12.84	-69.29	100±21	471±32	22.07±5.13	1.20±0.37	6.55±2.58	8.37±2.87	1.89±0.80	0.06±0.09	-30.48±1.10
TAM-02	-12.83	-69.29	104±27	475±37	22.93±5.49	1.16±0.42	4.98±3.11	8.97±4.57	3.10±2.15	0.07±0.03	-30.36±1.31
TAM-03	-12.84	-69.28	110±15	485±35	18.84±4.25	1.31±0.33	4.17±1.39	8.28±4.76	1.92±1.41	0.06±0.02	-31.34±0.87
TAM-04	-12.84	-69.28	115±25	496±28	21.68±4.67	1.31±0.34	2.23±0.91	7.53±3.01	2.10±0.93	0.34±1.09	-30.60±0.89
TAM-05	-12.83	-69.27	101±24	507±34	23.99±5.50	1.05±0.20	2.79±2.50	6.14±3.17	2.17±1.35	0.05±0.04	-30.42±1.52
TAM-06	-12.84	-69.30	96±21	485±35	24.80±6.68	1.88±0.84	8.44±4.33	8.16±3.54	2.33±1.00	0.06±0.01	-30.19±0.89
TAM-07	-12.83	-69.26	114±22	511±33	21.56±4.99	0.98±0.23	1.91±0.78	6.43±2.28	2.22±0.87	0.22±0.63	-31.06±0.95
TAP-04	-2.85	-54.95	99±31	463±43	22.58±7.70	0.75±0.24	7.52±4.24	3.69±2.21	2.69±1.41	2.70±5.97	-31.79±1.44
TAP-123	-3.31	-54.94	90±19	467±35	21.68±5.59	0.70±0.18	5.23±3.03	2.97±1.24	2.48±1.12	0.22±0.57	-31.04±1.32
TIP-03	-0.64	-76.15	92±48	472±26	28.38±3.28	2.04±0.65	7.15±3.61	7.01±3.05	2.19±1.10	0.10±0.04	-31.40±1.34
TIP-05	-0.64	-76.14	112±40	470±32	21.83±3.78	1.23±0.39	11.53±4.87	8.24±4.27	3.21±0.79	0.09±0.15	-31.14±1.69
YAN-01	-3.44	-72.85	86±29	474±36	19.71±5.29	1.24±0.33	19.09±10.69	9.11±4.33	3.20±1.50	0.11±0.13	-29.92±1.17
YAN-02	-3.43	-72.84	105±14	463±43	20.03±3.74	1.17±0.44	17.50±9.48	8.93±2.84	4.23±0.95	0.28±0.28	-31.15±1.17
YAN-03	-3.44	-72.85	79±25	470±32	22.94±6.07	1.29±0.42	8.54±5.65	11.26±6.07	2.42±1.54	0.05±0.02	-31.39±1.63

Table SI2: Pairwise relationships of key foliar properties using the raw data, classified in low and high fertility sites. *Slope* of the SMA, Pearson's *r* correlation coefficient, *sig* the significance of the correlation, and *n* the number of cases used. Boldface indicates significant difference ( $p < 0.05$ ) in slope or elevation and/or shift across the SMA axis. The (--) sign indicates that the respective estimates can not be made. Significance codes: \*\*\* < 0.001, \*\* < 0.01, \* < 0.05.

$\log_{10}(y)$	$\log_{10}(x)$	low fertility					high fertility					sig. of difference in		
		slope	slope 95%ci	r	sig	n	slope	slope 95%ci	r	sig	n	slope	elevation	shift
C	M <sub>A</sub>	0.278	(0.250 0.308)	0.376	***	302	0.300	(0.275 0.328)	0.194	***	477	0.262	0.138	0.001
N	M <sub>A</sub>	-0.919	(-0.830 -1.018)	-0.431	***	305	-0.810	(-0.745 -0.880)	-0.353	***	489	0.059	<0.001	<0.001
P	M <sub>A</sub>	-1.116	(-1.001 -1.245)	-0.259	***	305	-1.187	(-1.093 -1.289)	-0.386	***	481	0.378	<0.001	<0.001
Ca	M <sub>A</sub>	-2.584	(-2.311 -2.890)	-0.139	*	305	-2.553	(-2.338 -2.789)	-0.176	***	482	0.869	<0.001	<0.001
K	M <sub>A</sub>	-2.057	(-1.840 -2.300)	-0.155	**	305	-1.617	(-1.485 -1.760)	-0.319	***	482	0.001	(--)	(--)
Mg	M <sub>A</sub>	-1.739	(-1.561 -1.937)	-0.295	***	305	-1.703	(-1.559 -1.859)	-0.185	***	482	0.765	0.119	0.001
Al	M <sub>A</sub>	(--)	(--)	0.048	0.398	308	(--)	(--)	0.020	0.665	481	(--)	(--)	(--)
$\delta^{13}\text{C}$	M <sub>A</sub>	0.163	(0.146 0.182)	-0.316	***	291	0.161	(0.147 0.176)	-0.128	**	482	0.837	<0.001	0.002
N	C	(--)	(--)	-0.062	0.274	309	(--)	(--)	-0.038	0.401	483	(--)	(--)	(--)
P	C	(--)	(--)	0.038	0.509	309	-3.933	(-3.601 -4.296)	-0.178	***	482	(--)	(--)	(--)
Ca	C	-9.122	(-8.220 -10.122)	-0.372	***	309	-8.476	(-7.879 -9.117)	-0.579	***	483	0.257	<0.001	<0.001
K	C	-7.366	(-6.598 -8.223)	-0.187	**	309	-5.354	(-4.942 -5.800)	-0.446	***	483	<0.001	(--)	(--)

Mg	C	-6.200	(-5.629 -6.830)	-0.506	***	309	-5.635	(-5.213 -6.092)	-0.492	***	483	0.131	0.989	<0.001
$\delta^{13}\text{C}$	C	0.617	(0.552 0.690)	-0.264	***	293	(--)	(--)	-0.001	0.986	475	(--)	(--)	(--)
P	N	1.187	(1.080 1.305)	0.540	***	309	1.476	(1.374 1.586)	0.600	***	482	<0.001	(--)	(--)
Ca	N	(--)	(--)	0.067	0.237	309	3.176	(2.907 3.471)	0.143	**	483	(--)	(--)	(--)
K	N	2.223	(1.990 2.484)	0.136	*	309	2.006	(1.841 2.186)	0.281	***	483	0.152	<0.001	<0.001
Mg	N	1.871	(1.674 2.092)	0.120	*	309	2.112	(1.936 2.304)	0.233	***	483	0.093	0.067	<0.001
$\delta^{13}\text{C}$	N	(--)	(--)	-0.029	0.618	296	0.200	(0.183 0.219)	-0.162	***	483	(--)	(--)	(--)
Ca	P	2.295	(2.059 2.559)	0.231	***	312	2.141	(1.963 2.336)	0.211	***	487	0.327	<0.001	<0.001
K	P	1.839	(1.671 2.024)	0.512	***	312	1.332	(1.235 1.437)	0.526	***	487	<0.001	(--)	(--)
Mg	P	1.552	(1.391 1.733)	0.170	**	312	1.410	(1.292 1.538)	0.212	***	487	0.178	<0.001	<0.001
$\delta^{13}\text{C}$	P	0.148	(0.133 0.166)	-0.241	***	295	0.135	(0.123 0.147)	-0.257	***	478	0.186	<0.001	<0.001
Mg	Ca	0.676	(0.619 0.739)	0.612	***	312	0.665	(0.617 0.717)	0.538	***	488	0.771	<0.001	<0.001
$\delta^{13}\text{C}$	Ca	(--)	(--)	-0.076	0.194	295	(--)	(--)	-0.085	0.063	479	(--)	(--)	(--)
Ca	K	1.248	(1.121 1.389)	0.277	***	312	1.577	(1.451 1.714)	0.355	***	488	0.001	(--)	(--)
Mg	K	0.844	(0.757 0.940)	0.242	***	312	1.049	(0.968 1.136)	0.437	***	488	0.002	(--)	(--)
$\delta^{13}\text{C}$	K	(--)	(--)	0.008	0.894	295	0.099	(0.091 0.108)	-0.112	*	479	(--)	(--)	(--)
$\delta^{13}\text{C}$	Mg	-0.096	(-0.086 -0.108)	0.121	*	295	(--)	(--)	-0.015	0.737	479	(--)	(--)	(--)

Figure SI1: SMA regressions lines using the raw data. Red circles indicate records on low fertility and blue circles on high fertility sites. Black dots represent measurements on unclassified fertility sites. Differences in slope, elevation or shifts along the SMA axis are summarized in table SI2. Note some overlap from species found on both low and high fertility sites is not explicitly taken into account.



## Supplementary Information II

*Assessing the validity of the REML approach to the partitioning of variance in the data set to genetic, plot and residual components and the associated derivation of species and plot effects.*

Although the REML approach used here is well established in the social sciences literature, it may at first seem unlikely that “species” and “plot” effects can be reliably determined in such a unbalanced situation as exists as here, *viz.* with little replication across plots and especially with approximately only half the species being observed on only one occasion. To help allay such concerns we have therefore undertaken numerical simulations to allow an empirical assessment of the validity of the approach. In brief, the verification procedure consisted of:

1. Generation of a artificial population of approximately the same number of species as in the data set with the sampled species attributes ( $M_A$ , N, P etc) co-varying in a way similar to that implied by our analysis (600 species in all).
2. Generation of random (residual) variations in the sampled properties (i.e. within species variation and measurement error), with the properties co-varying in a way similar to that implied by the data.
3. Generation of 60 “plot effects” of the same magnitude as that implied by the analysis
4. Combination of (1), (2) and (3) to generate a simulated dataset of “raw observations” with species randomly distributed across plots in a manner similar to that occurring in our analysis, and with the total sample size consisting of 1200 trees.
5. Application of the identical REML approach as is used in the analysis here to determine if the REML approach could indeed:
  - a) partition the variance within the dataset correctly,
  - b) allow the accurate retrieval of species and plot effects,
  - c) provide derived species and plot effects of sufficient accuracy to allow correct estimates of the nature of bivariate relationships to be obtained.

## **1. Computer generation of 600 artificial species with different but realistically autocorrelated traits**

We first took 285 species for which a combined suite of 12 traits had all been measured (including some not reported here, but to be reported elsewhere<sup>1</sup>, for example leaf area, xylem density, Huber values (etc)) and generated a new (artificial) population of species for which traits varied both within the one species and between different species in a way implied by our data set. This was achieved using the procedure of Taylor and Thompson (1986)<sup>2</sup> using the Routine “RNDAT” available in the IMSL Fortran Library<sup>3</sup>. This empirical procedure generates a pseudorandom sample with approximately the same moments as the given sample with the sample obtained being essentially the same as if from a Gaussian kernel estimate of the sample density<sup>4</sup>.

## **2. Computer generation of residual variance component.**

Of the 285 species that sampled for all traits of interest, we then retrieved a second subset which was characterised by having been recorded within any of the one plots sampled at least twice. This subset consisting only of observations for the species within the plots for which they occurred more than once, yielded a dataset consisting of 52 species and a total of 143 trees. We then computed a pooled variance-covariance matrix from the observations using the IMSL Fortran Library subroutine “COVPL”. This routine uses modified Givens rotations to compute the Cholesky decomposition of the pooled sums of squares and cross-products matrix (Golub and van Loan 1983)<sup>5</sup> and is particularly applicable in cases such as was the case here, where it was suspected that each of the plot/species combinations represented a different population mean, but with identical variance-covariance matrices.

The variance-covariance matrix so derived was then used to generate pseudorandom numbers from the multivariate normal distribution, characterised with the generated population having a mean vectors value of all zero and deviance of one via a Cholesky factorisation using the IMSL library routines “CHFAC” and “RNMVN”.

---

<sup>1</sup> Patiño, S. *et al.*: Relationships among ecologically important dimensions of plant trait variation in 53 Neotropical forests as affected by species, soils and climate (in preparation).

<sup>2</sup> Taylor, M. S., and J. R. Thompson: Data based random number generation for a multivariate distribution via stochastic simulation, *Computational Statistics & Data Analysis*, **4**, 93-101, 1986.

<sup>3</sup> IMSL Fortran Library Users Guide STAT/LIBRARY Volume 2, Visual Numerics, Houston, USA.

<sup>4</sup> Thompson, J.R: *Empirical Model Building*, John Wiley & Sons, New York, 1989.

<sup>5</sup> Golub, G. H. and C. F. Van Loan: *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Maryland, 1983.



### 3. Computer generation of plot variance component.

As trait correlation is not such an issue for “plot effects”, we simply first generated 60 pseudorandom numbers (mean of zero and variance of one) from a standard normal distribution for each trait of interest using the IMSL library routine “RMMOA” employing the acceptance/rejection technique of Kinderman and Ramage<sup>6</sup>.

### 4. Generation of individual tree values

For each trait of interest we then took what we believed to be the inferred variances for “species”, “plot” and “residual” effects from our own REML analysis, and created a dataset with the same apportionment of variability for that trait according to:

$$\sigma_{p,t}^2 = \sigma_{s,t}^2 f_{p,t} / f_{s,t} \quad \text{and} \quad \sigma_{r,t}^2 = \sigma_{s,t}^2 f_{r,t} / f_{s,t}$$

where  $\sigma_{p,t}^2$  is the plot variance for trait  $t$ ,  $\sigma_{s,t}^2$  is the species variance for trait  $t$  (estimated from the population of 600 generated in Section 1)  $\sigma_{r,t}^2$  is the residual variance for trait  $t$ , and  $f_{p,t}$ ,  $f_{s,t}$  and  $f_{r,t}$  are the fractions of the variance attributable to each of the three components for each trait as originally inferred from our analysis.

For each tree, we then generated a value for each trait, this being that which would occur where there to be no plot effect as

$$\Xi_{\bullet,t} = S_t + z_{r,t} / \sigma_{r,t}$$

where  $\Xi_{\bullet,t}$  is the value for the trait observed in the absence of any plot affect,  $S_t$  is the characteristic value of the trait for that species (Section 1) and  $z_{r,t}$  is the pseudorandom standard normal random deviate describing the residual (within species) variation for the trait in question (from Section 2).

In order to generate 1200 trees (similar to that in our dataset) from 600 species (also similar to that in our dataset) we allowed 300 species to occur only once (again similar to as in our dataset) with 150 species occurring twice, 75 trees occurring four times, 50 species occurring five times and 25 species occurring six times.

---

<sup>6</sup> Kinderman, A.J., and J.G. Ramage: Computer generation of normal random variables, *Journal of the American Statistical Association*, **71**, 893-896, 1976.

#### 4. Generation of dataset across plots.

Each of the 1200 trees was then randomly assigned to each of the 60 plots with its overall value for each trait then being defined as:

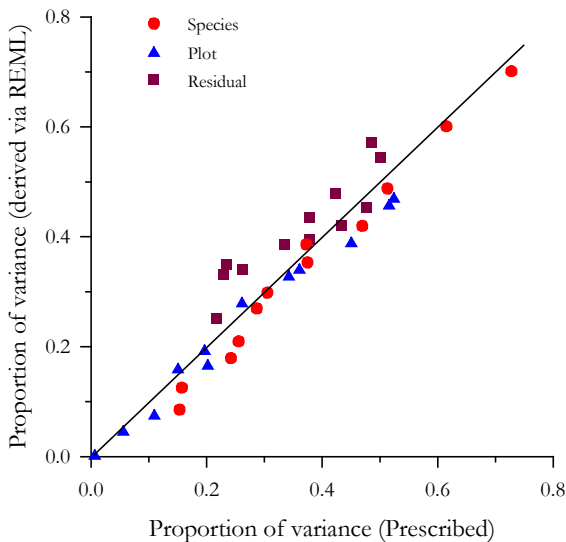
$$\Xi_{p,t} = \Xi_{\bullet,t} = S_p + z_{p,t}/\sigma_{p,t}$$

where  $z_{p,t}$  is the pseudorandom standard normal random deviate describing the plot effect for the trait in question (from Section 3).

This gave on average 20 trees per plot (similar to as in our dataset), though, depending on the simulation (which gave different results for every run), the number of trees in any one plot could vary from as little as 14, to as much as 26. This again replicating the unbalanced and somewhat random nature of our sampling strategy. Just by chance, some of the “more abundant” species always occurred in some plots more than once, though it was also the case that this “within plot overlap” was less than in our dataset.

#### 5. Results.

##### a) Variance partitioning



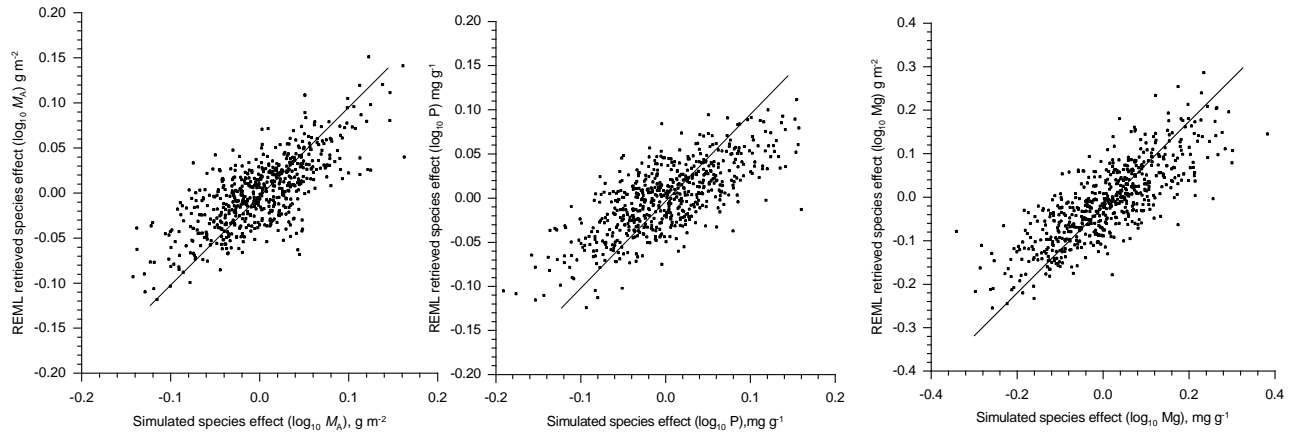
**Figure 1.** Relationship between the proportion of trait variance prescribed to the species, plot and residual terms in the initial data simulation, and that actually retrieved by the REML model

For the twelve traits simulated, there were considerable differences in the degree of inferred partitioning of their variances between plots, species and that allocated to the residual (within species and measurement error) terms input into the simulations. Nevertheless, as is shown in Figure 1, irrespective of the way in which the variation was assigned, the REML technique retrieved a partitioning of variance close to the true value. There was, however, a slight trend for the variance associated with the species and plot effects to be underestimated and the proportion of variance associated with the residual terms to be overestimated. Nevertheless, it is clear that the relative partitioning between the environmental and

genetic effects as inferred by the REML technique in the main paper must reflected their true values to a high degree of accuracy.

## b) Species and plot effects

Generally speaking all traits gave similar results, and so for illustrative purposes, we concentrate here on  $M_A$ , P and Mg; these reflecting contrasts in apparent variance partition (Fig. 1 of the main text); phosphorus having one of the strongest environmental components, magnesium the strongest genetic component and  $M_A$  being somewhat intermediate between the two. Results are shown in Figure 2.



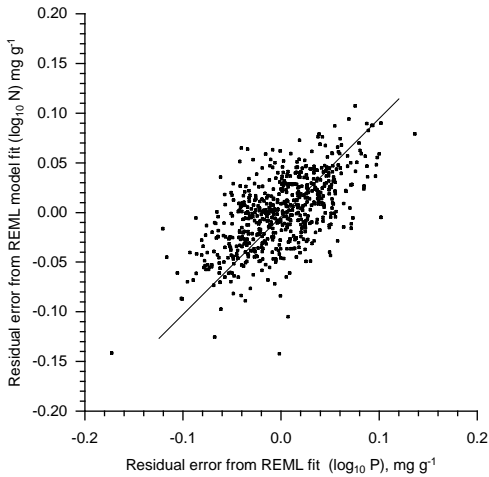
**Figure 2.** Relationship between simulated (known) species effects and those retrieved by the REML procedure for leaf mass per unit leaf area ( $M_A$ ), phosphorus and magnesium.

This shows that the REML procedure does a reasonable good job of retrieving the species effects, though in all cases there does seem to be some overestimation at the lowest values and underestimation at the highest values, the severity of this effect increasing as the relative magnitude of the genetic component declines (i.e being greatest for P). This is, in fact, what is expected from the procedure used to retrieve the species means when such terms are treated as random effects. This output represents empirical Bayes "posterior" estimates with a shrinkage of ordinary least squares (OLS) estimates towards the general mean<sup>7</sup>. Nevertheless, especially when it is realised that we have added noise to even our single species occurrences through addition of a random variance component,  $z_{r,t}$ , (this then making it virtually impossible to retrieve the true species mean even with substantial replication), it is clear that the REML procedure does a remarkably good job of retrieving the prescribed species values despite the markedly unbalanced "experimental" set up employed.

It is also worth noting that strong correlations were found to exist between the various traits in terms of the errors associated with the REML retrievals. This is shown in Figure 3, where the

<sup>7</sup> Snijders, T. A. B. and Bosker, R. J.: *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, SAGE, 266 pp., 1999.

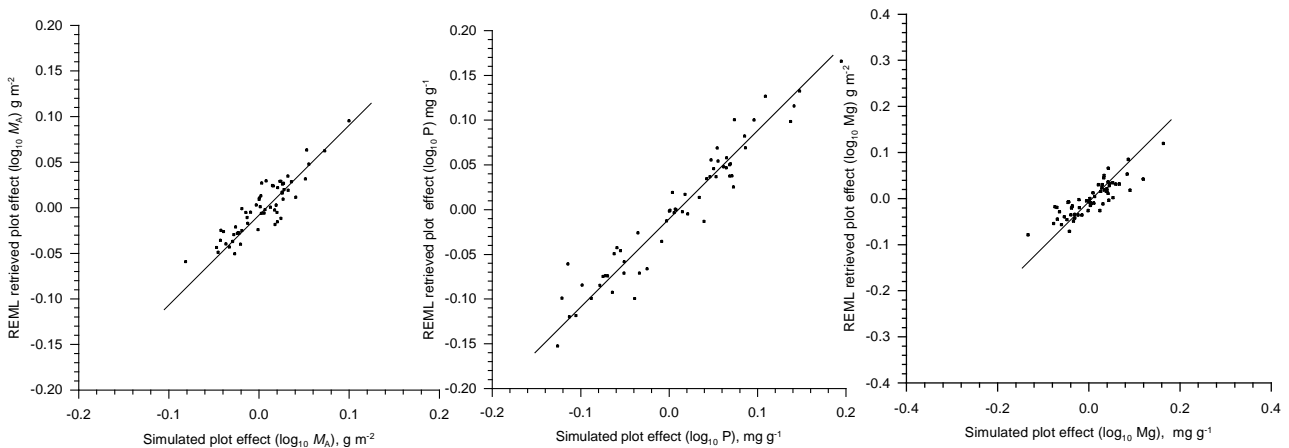
residuals (REML value less actual simulated value) are shown for N and P. Such correlations exist



**Figure 3.** Relationships between residuals from the REML model fit (i.e. fitted versus actual (simulated) values for nitrogen and phosphorus species effects

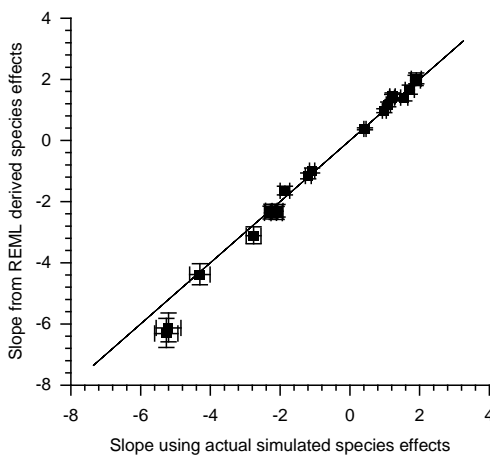
because of the within-species trait covariance structure been has taken into account when simulating the residual term  $z_{r,t}$  (Section 2) That is to say, traits vary in a co-ordinated way within a species\_ in much the same way that they vary between species.

The ability of the model to retrieve the plot effects is shown in Figure 4. This shows the model to have good ability to retrieve the correct values, though for magnesium (for which the overall variance associated with species location was small) some compression of the mean effect is apparent due to the Bayes retrieval technique.



**Figure 4.** Relationship between simulated (known) plot effects and those retrieved by the REML procedure for leaf mass per unit leaf area ( $M_A$ ), phosphorus and magnesium

### c) SMA slopes



**Figure 5.** Relationships between slopes of all significant bivariate relationships for species effects from the REML model fit compared with those of the input data

Although Figures 2 and 4 suggest that there is some error associated with the REML retrievals of the species and plot effects, Figure 3 provides a practical example illustrating that these errors are almost inevitably correlated. As mentioned above, this is because the variance-covariance matrices are quite similar in structure for the (between) species effects as for the “residual” within-species variation. This correlation of “errors” means that standard methods of dealing with observational errors in fitting

slopes by SMA regression<sup>8</sup> cannot be used for the regression fits of the REML derived values, even though standard errors for the species effect estimates can be obtained from the REML model output<sup>9</sup>. Rather, we have simply tested the validity of the slopes obtained for our species effects (assumed as in the main manuscript to be without error) by comparing them with those using the actual species dependent values as input into the simulation (Figure 5). This shows an excellent agreement between the two approaches with the exception of three bivariate relationships with steeply negative slopes. The three points in Figure 5 all involve the relationships between the concentrations of the three cations examined, [Ca], [K], and [Mg] and [C], and would seem to arise as a consequence of within-species co-variation of [C] with [Ca], [K] or [Mg] being less than is observed between different species. Nevertheless, for the majority of traits of interest, the simulations here indicate that the slope of the bivariate relationships presented for the species effects as derived through the REML procedure are very close to their true value, and certainly within the quoted confidence intervals; this also being the case for the plot effects (data not shown).

## 6. Conclusions

Even when the considerable observed within-species variability is taken into account, the above simulations show that the REML technique employed provides an excellent estimate of the true partitioning of the variation of the various traits examined. Although the individual species and plot effects are (inevitably) retrieved with some error and a small bias at extreme values, because traits varying in a similar way both between and within species, the SMA slopes obtained for both the species and plot effects (assuming zero error) are excellent estimates of their true values.

---

<sup>8</sup> Warton, D.I. et al. Bivariate line-fitting methods for allometry. *Biological Reviews*, 81, 259-291, 2006.

<sup>9</sup> Gelman, A. and Hill, J.: *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 648pp, 2006.

### Supplementary Information III

Figure SIII 1: Maps illustrating the spatial patterning of the plot effect contribution to each foliar property. Moran's I spatial autocorrelation statistic along with a Monte Carlo significance test (999 permutations) are given at the bottom right part of each map. For all nine plot-environmental effect contributions a significant (0.005, global Bonferroni adjusted) spatial structure was identified.

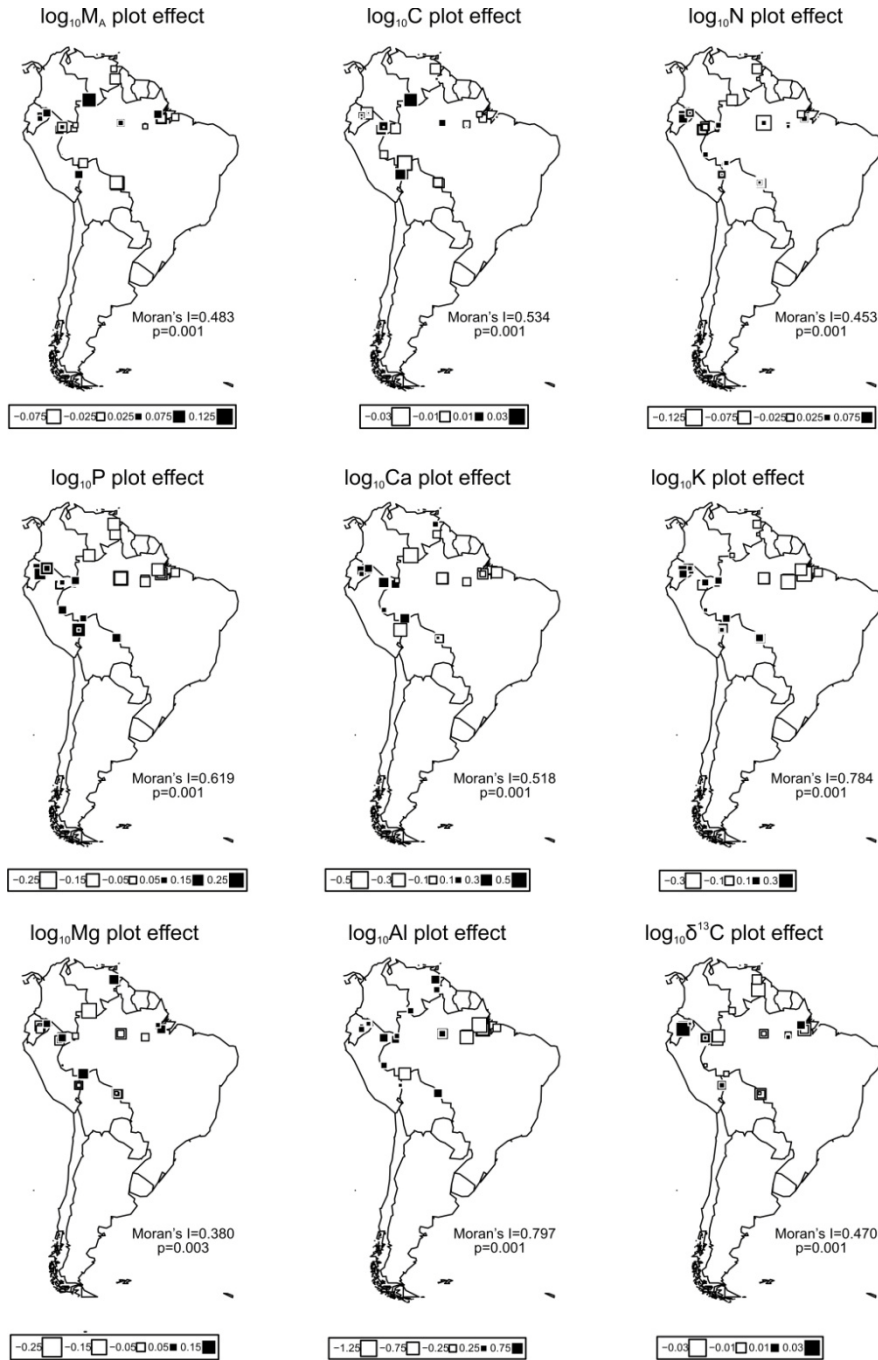
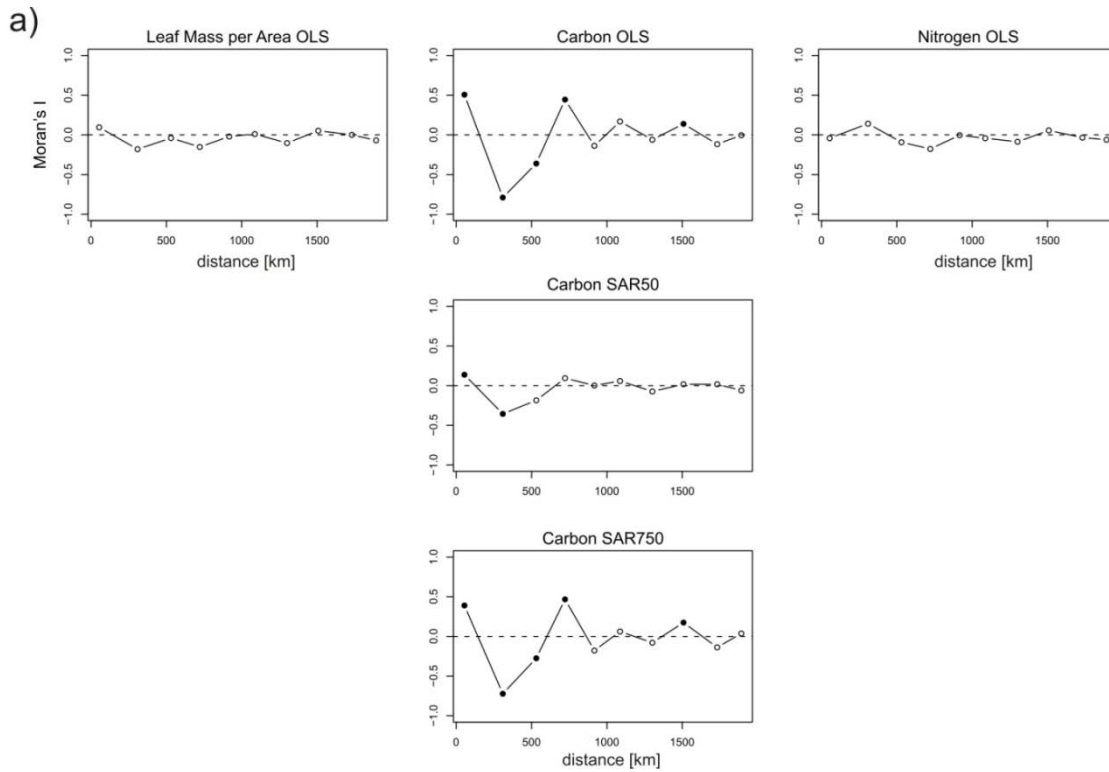
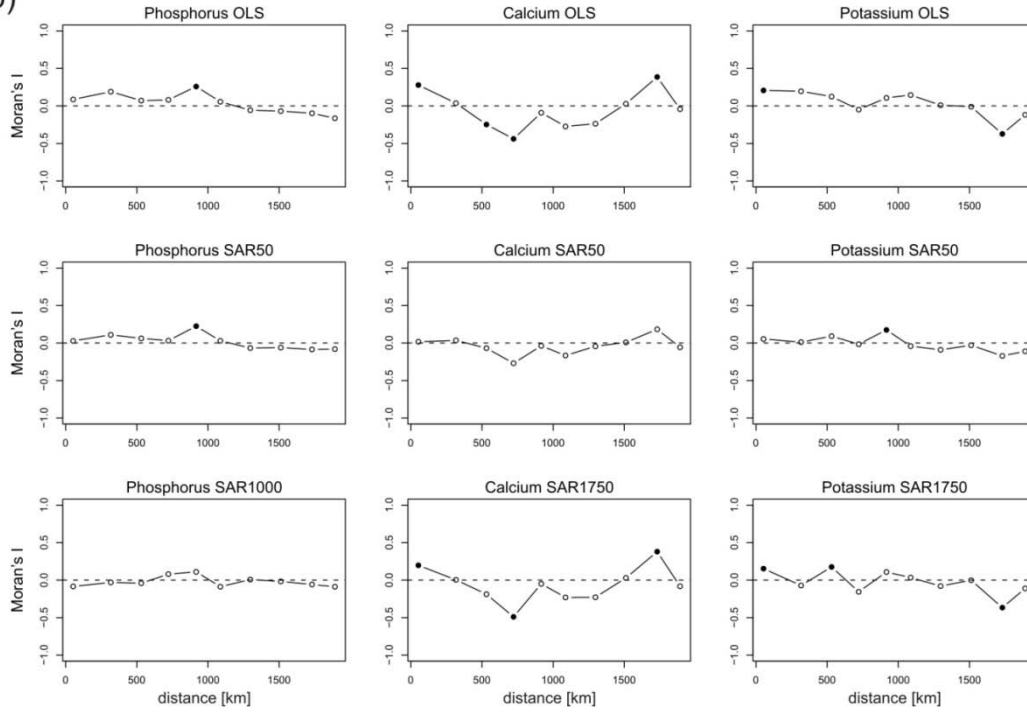


Figure SIII 2. Correlograms of the residuals from non-spatial least square regressions and spatial autoregressive models at a common fine (SAR 50) and at a variant medium scale for each foliar property of interest. For non-spatially autocorrelated residuals (see also Table 5 for global Moran's test) only the correlogram from the OLS residuals are illustrated (for example  $M_A$ , [N] and [Mg]). In cases where a spatial pattern was identified the residuals of the fine scale SAR50 km and medium scale SAR models are also given. Solid circles indicate a significant spatial signal at the specific distance class ( $j$ ), adjusted through a progressive Bonferroni test  $0.05/j$ . See text for details.



b)



c)

