Supporting Information

Slik et al. 10.1073/pnas.1423147112

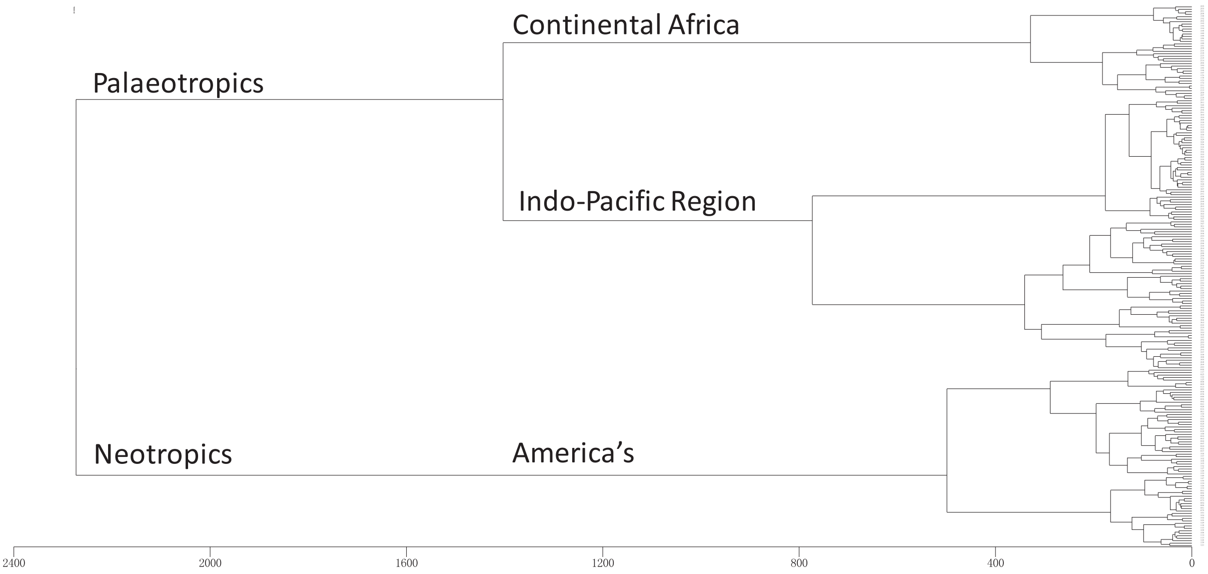

Fig. S1. Results of minimum variance clustering with squared Euclidean distances on square root-transformed relative abundance data at the genus level. Only plots with more than 250 identified individuals were used. Plots were first divided into Paleotropics vs. Neotropics, and then into Continental Africa, Indo-Pacific Region (including Madagascar), and the Americas. The basal internal divisions within these three main regions mainly separate everwet from seasonal and dry forest types. Terminal numbers correspond to location numbers in Dataset S1.

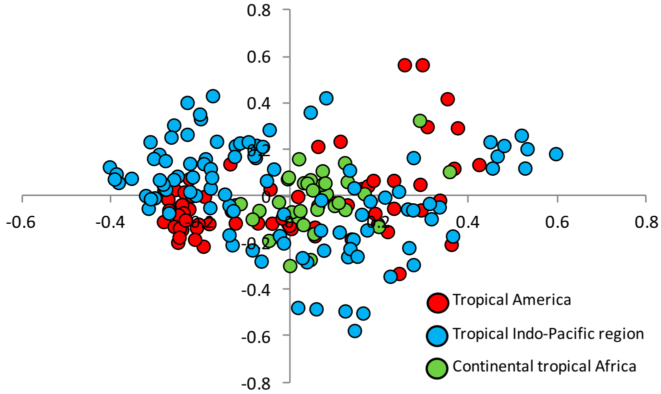

Fig. S2. Environmental variation in climate and soils among sites for the three major tropical regions of the world based on principal components analysis of locations vs. environmental variables [climatic data (BIO1-19) and edaphic data (soil production, easy available water, pH topsoil, organic carbon topsoil, nitrogen topsoil, soil moisture storage, soil drainage, soil depth, C:N ratio topsoil, Cation Exchange Capacity [CEC] clay topsoil, CEC soil topsoil, base saturation topsoil, organic carbon pool, textural class topsoil, textural class subsoil)].

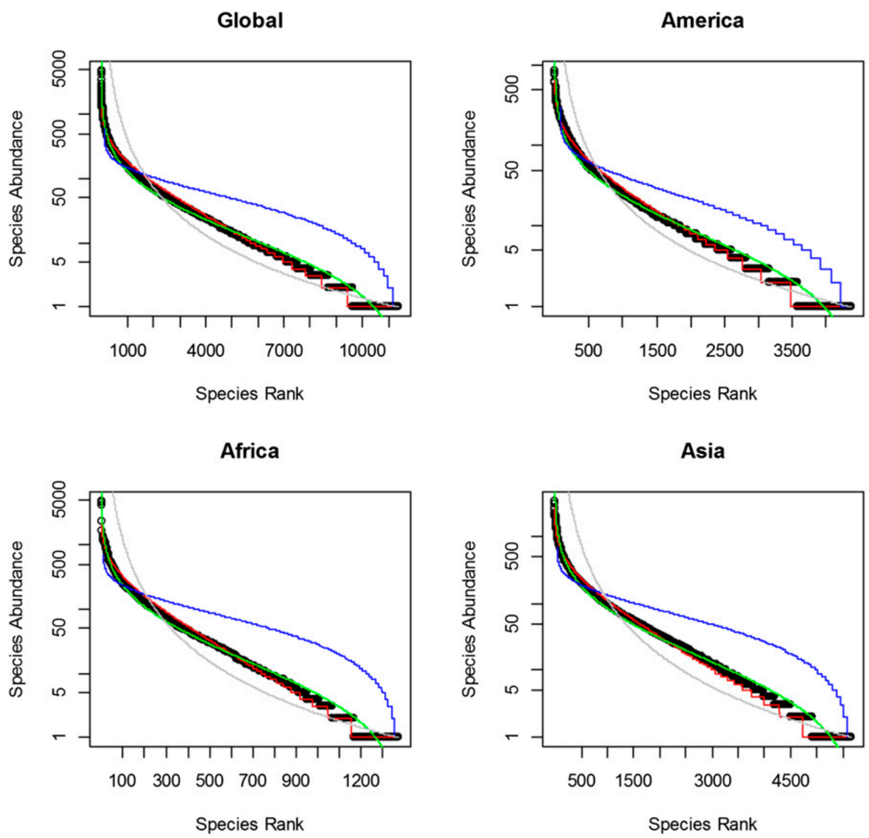

Fig. S3. Species abundance diagrams for our data sets showing the fit of different SAD models: observed (black), log series (red), log normal (green), broken stick (blue), and Pareto (gray).

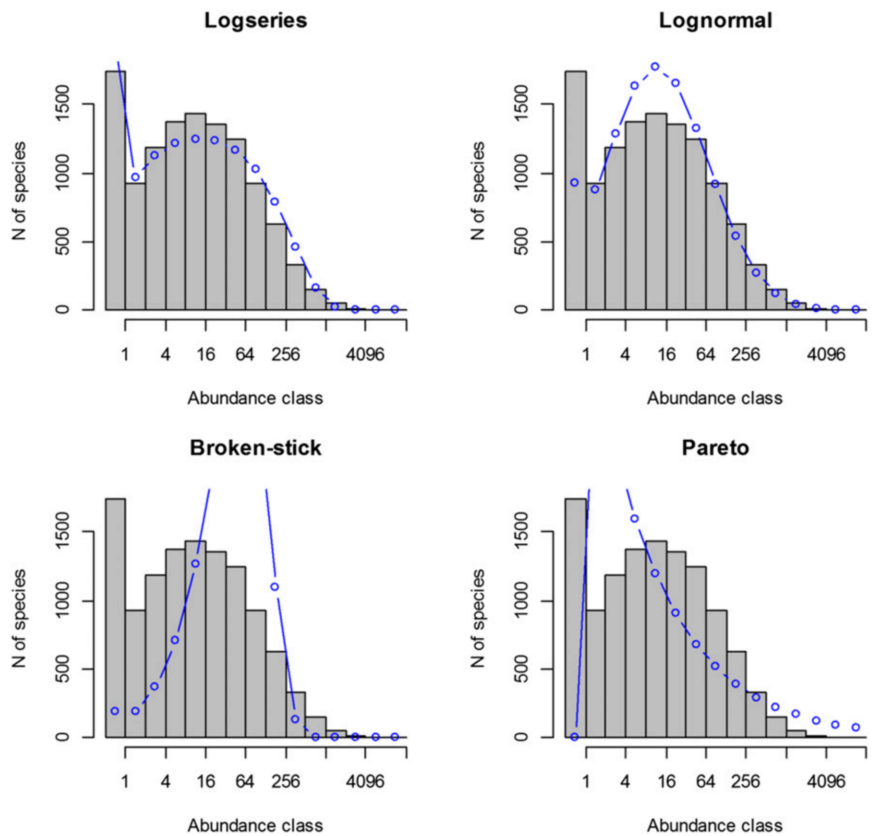

Fig. S4. Preston plots showing the observed frequency of species in octaves of abundance (bars) and the predicted frequency by four different SAD models (blue line) for the global data set.

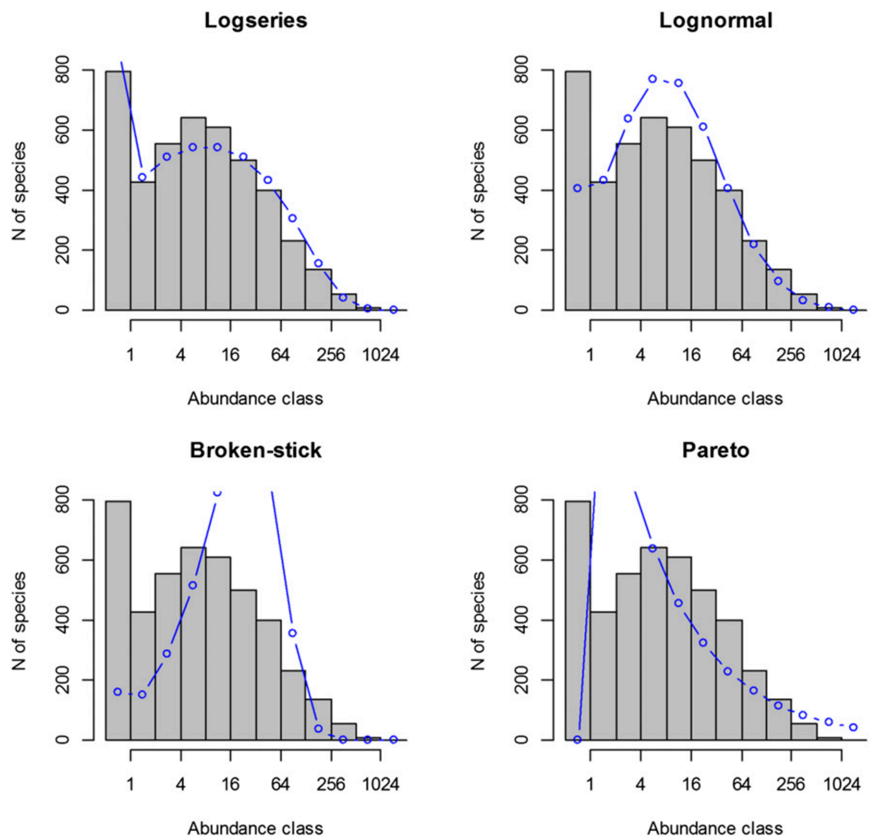

Fig. S5. Preston plots showing the observed (bars) frequency of species in octaves of abundance, and the predicted frequency by four different SAD models (blue line) for the American data set.

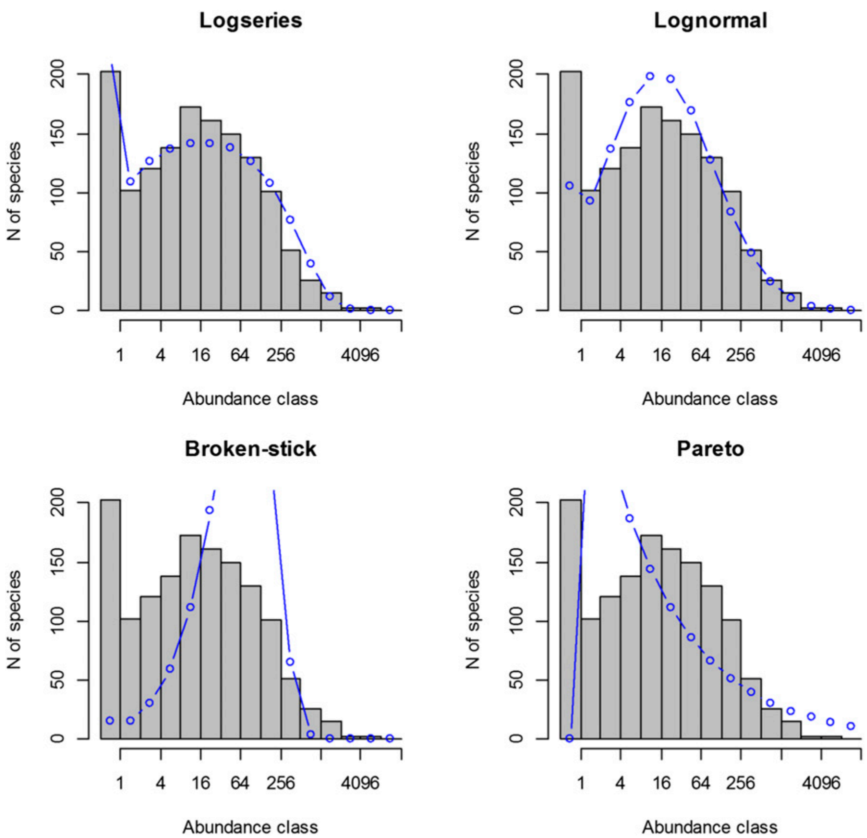

Fig. S6. Preston plots showing the observed (bars) frequency of species in octaves of abundance, and the predicted frequency by four different SAD models (blue line) for the African data set.

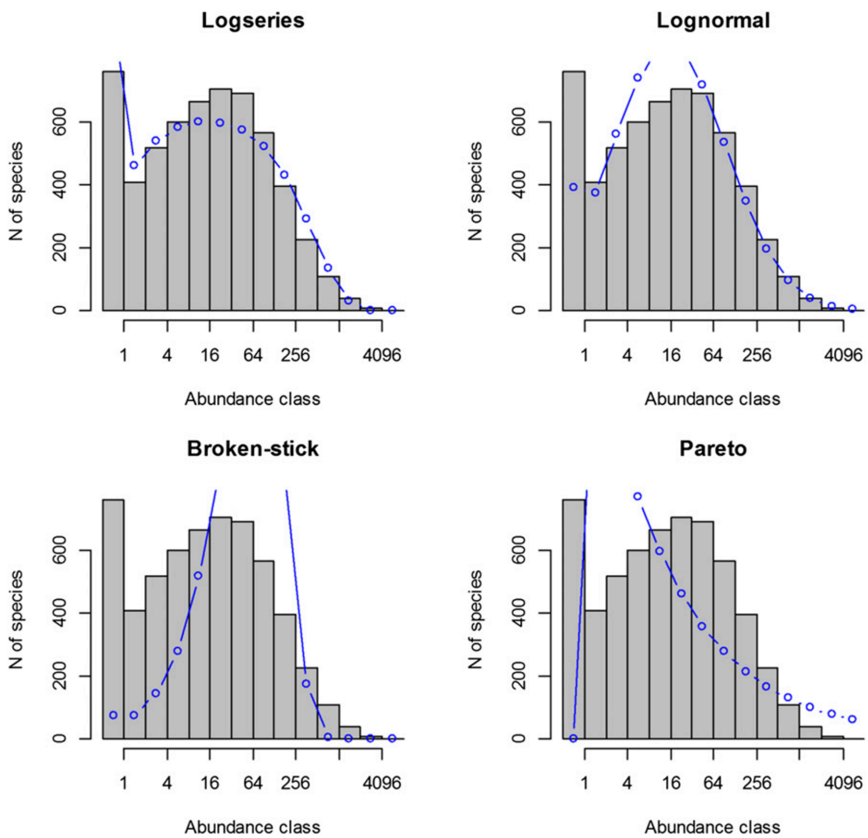

Fig. S7. Preston plots showing the observed (bars) frequency of species in octaves of abundance, and the predicted frequency by four different SAD models (blue line) for the Asian data set.

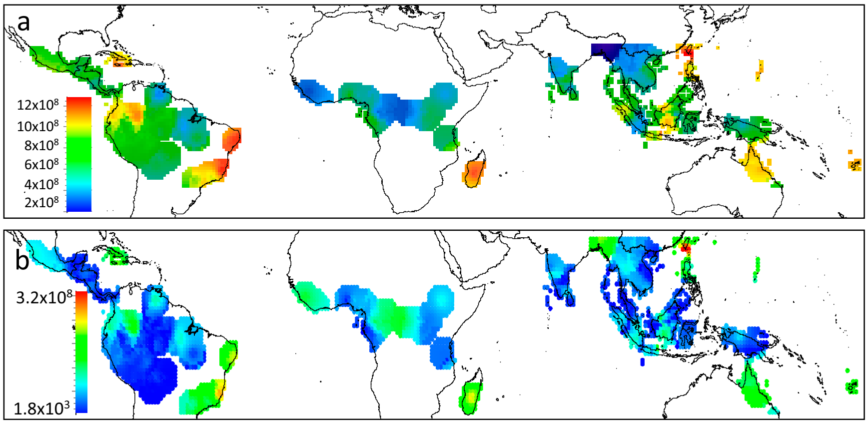

Fig. S8. Stem number estimates (*A*) and their uncertainty (*B*) for individual 1° grid cells. Uncertainty was assessed with a jackknifing procedure, whereby the extrapolation was recalculated 100 times, each time by randomly removing an observation (location) from the analysis. By comparing the observed values with the average prediction obtained with the jackknifing procedure, the uncertainty is given by the difference between the two values. Through regression analysis we subsequently developed an error function that was used to calculate the absolute error in all grid cells for which we made stem density predictions.

Table S1. Number of stems (*N*) and tree species (*S*) recorded in the global data set and in each tropical region

Data sets	N	S	f_1	f_2	$\hat{C}_n$, %	Mean $\hat{C}_n$, %*
Pantropical	657,630	11,371	1,743	930	91.6	95.6 ± 0.3
Tropical America	116,754	4,375	797	429	91.5	94.6 ± 0.7
Continental Africa	117,902	1,376	203	102	90.8	96.8 ± 0.5
Indo-Pacific	422,974	5,672	759	406	91.7	95.8 ± 0.5

The number of singletons (f_1) and doubletons (f_2) are shown, as well as the sample coverage ($\hat{C}_n$ = percentage of the total number of individuals in an assemblage that belong to the species represented in the sample). We also indicate mean (±SE) sample coverage per location for each data set.
*Mean sample coverage per location did not differ among regions ($F_{2,204}$ = 2.33, P = 0.10).

Table S2. Tree species richness estimated from our sample by different nonparametric methods with the SPADE program and their respective variations (SE and lower and upper 95% confidence intervals)

Estimator/model	Estimate	SE	95% CI		%
Pantropical					
Chao1	13,004.4	103.1	12,814.5	13,219.2	87.4
Chao1-bc	13,001.7	79.1	12,853.9	13,164.2	87.5
ACE	12,661.1	45.3	12,575.3	12,752.9	89.8
ACE-1	12,989.9	72.8	12,853.3	13,139.1	87.5
First-order jackknife	13,114.0	59.0	13,002.1	13,233.6	86.7
Second-order jackknife	13,927.0	102.3	13,734.3	14,135.4	81.6
Gamma-Poisson model	13,236.8	128.2	13,001.9	13,505.5	85.9
Gamma-Poisson-UMLE	13,228.7	127.0	12,996.0	13,494.7	86.0
Gamma-Poisson-CMLE	13,232.0	127.4	12,998.6	13,498.8	85.9
Africa					
Chao1	1,578.0	37.5	1,516.8	1,665.8	87.2
Chao1-bc	1,575.1	36.9	1,514.8	1,661.5	87.4
ACE	1,538.8	15.1	1,511.7	1,571.2	89.4
ACE-1	1,586.7	45.1	1,515.1	1,695.1	86.7
First-order jackknife	1,579.0	20.1	1,543.2	1,622.5	87.1
Second-order jackknife	1,680.0	34.9	1,618.9	1,756.4	81.9
Gamma-Poisson model	1,630.1	52.8	1,545.8	1,756.2	84.4
Gamma-Poisson-UMLE	1,637.0	57.3	1,546.6	1,775.2	84.1
Gamma-Poisson-CMLE	1,641.4	59.0	1,548.6	1,784.3	83.8
America					
Chao1	5,115.3	69.1	4,991.9	5,263.5	85.5
Chao1-bc	5,112.7	68.8	4,989.7	5,260.3	85.6
ACE	4,959.3	30.4	4,902.6	5,022.1	88.2
ACE-1	5,104.8	50.4	5,012.5	5,210.4	85.7
First-order jackknife	5,172.0	39.9	5,097.5	5,254.2	84.6
Second-order jackknife	5,540.0	69.2	5,412.1	5,683.6	79.0
Gamma-Poisson model	5,211.8	83.2	5,064.0	5,391.3	83.9
Gamma-Poisson-UMLE	5,201.6	82.9	5,054.4	5,380.7	84.1
Gamma-Poisson-CMLE	5,204.8	83.4	5,056.7	5,385.0	84.1
Asia					
Chao1	6,381.5	67.8	6,260.5	6,527.3	88.9
Chao1-bc	6,378.8	67.6	6,258.2	6,524.1	88.9
ACE	6,232.3	29.7	6,177.0	6,293.5	91.0
ACE-1	6,374.8	47.1	6,288.4	6,473.3	89.0
First-order jackknife	6,431.0	39.0	6,358.4	6,511.3	88.2
Second-order jackknife	6,784.0	67.5	6,659.4	6,924.3	83.6
Gamma-Poisson model	6,480.3	85.2	6,329.8	6,665.2	87.5
Gamma-Poisson-UMLE	6,474.6	82.6	6,328.3	6,653.6	87.6
Gamma-Poisson-CMLE	6,477.8	83.2	6,330.6	6,658.0	87.6

The estimated percentage of species recorded (species sampled/species estimated × 100) is also indicated. CI, confidence interval.

Table S3. Difference in Akaike's information criterion values between the best SAD model (i.e., with the lowest AIC value, ΔAIC = 0) and the rest of models

	ΔAIC	df
Global		
Log series	0	1
Log normal	1,225.5	2
Pareto	2,570.7	1
Broken stick	15,104.8	0
America		
Log series	0	1
Log normal	524.5	2
Pareto	853	1
Broken-stick	4,033.4	0
Africa		
Log series	0	1
Log normal	151.8	2
Pareto	367.6	1
Broken stick	2,091.2	0
Asia		
Log series	0	1
Log normal	576.2	2
Pareto	1,722.7	1
Broken stick	6,978.9	0

Note that in all cases the log series was the best model for the data. The number of free parameters within each model (df) are also indicated.

Table S4. The impact of error in stem numbers for the different regions on the number of tree species estimated to occur in each region using Fisher's log series

Error ratio in stem numbers	Species estimate Africa	Species estimate America	Species estimate Asia	Species estimate global
1.9	4,765.9 (3.0%)	19,164.7 (3.1%)	19,607.8 (3.1%)	41,770.5 (3.1%)
1.8	4,754.1 (2.8%)	19,116.1 (2.8%)	19,557.7 (2.9%)	41,664.9 (2.8%)
1.6	4,728.3 (2.2%)	19,010.5 (2.3%)	19,448.7 (2.3%)	41,434.9 (2.3%)
1.5	4,714.2 (1.9%)	18,952.6 (2.0%)	19,388.9 (2.0%)	41,308.8 (2.0%)
1.4	4,699.1 (1.6%)	18,890.7 (1.6%)	19,325.0 (1.6%)	41,174.1 (1.6%)
1.3	4,682.9 (1.2%)	18,824.2 (1.3%)	19,256.4 (1.3%)	41,029.3 (1.3%)
1.2	4,665.4 (0.9%)	18,752.4 (0.9%)	19,182.3 (0.9%)	40,873.0 (0.9%)
1.1	4,646.4 (0.5%)	18,674.3 (0.5%)	19,101.8 (0.5%)	40,703.1 (0.5%)
1.0	4,625.6	18,588.8	19,013.5	40,516.9
0.9	4,602.5 (−0.5%)	18,494.2 (−0.5%)	18,916.0 (−0.5%)	40,311.2 (−0.5%)
0.8	4,576.8 (−1.0%)	18,388.6 (−1.1%)	18,807.0 (−1.1%)	40,081.1 (−1.1%)
0.7	4,547.6 (−1.7%)	18,268.8 (−1.7%)	18,683.3 (−1.7%)	39,820.3 (−1.7%)
0.6	4,513.8 (−2.4%)	18,130.5 (−2.5%)	18,540.6 (−2.5%)	39,519.3 (−2.5%)
0.5	4,474.0 (−3.3%)	17,966.9 (−3.3%)	18,371.8 (−3.4%)	39,163.2 (−3.3%)
0.4	4,425.2 (−4.3%)	17,766.7 (−4.4%)	18,165.2 (−4.5%)	38,727.4 (−4.4%)
0.3	4,362.3 (−5.7%)	17,508.6 (−5.8%)	17,898.9 (−5.9%)	38,165.5 (−5.8%)
0.2	4,273.6 (−7.6%)	17,144.8 (−7.8%)	17,523.5 (−7.8%)	37,373.7 (−7.8%)
0.1	4,122.0 (−10.9%)	16,522.9 (−11.1%)	16,881.8 (−11.2%)	36,019.9 (−11.1%)

The impact of the error is minimal, with a doubling of stem numbers resulting in an increase of only ~3.0% in estimated species numbers, and a decline in tree numbers of 90% resulting in a decline in estimated species numbers of only ~11%.

Other Supporting Information Files

Dataset S1 (XLSX)