

# Taylor's law and related allometric power laws in New Zealand mountain beech forests: the roles of space, time and environment

Joel E. Cohen, Jiangshan Lai, David A. Coomes and Robert B. Allen

*J. E. Cohen (cohen@rockefeller.edu), Laboratory of Populations, Rockefeller Univ. and Columbia Univ., 1230 York Ave., Box 20, New York, NY 10065, USA. – J. Lai, State Key Laboratory of Vegetation and Environmental Change, Inst. of Botany, Chinese Academy of Sciences, Beijing, PR China. – D. A. Coomes, Dept of Plant Sciences, Univ. of Cambridge, Downing Street, Cambridge CB2 3EA, UK. – R. B. Allen, Landcare Research, PO Box 40, Lincoln 7640, New Zealand.*

Taylor's law says that the variance of population density of a species is proportional to a power of mean population density. Density–mass allometry says that mean population density is proportional to a power of mean biomass per individual. These power laws predict a third, variance–mass allometry: the variance of population density of a species is proportional to a power of mean biomass per individual. We tested these laws using 10 censuses of New Zealand mountain beech trees in 250 plots over 30 years at spatial scales from 5 m to kilometers. We found that: 1) a single-species forest not disrupted by humans obeyed all three laws; 2) random sampling explained the parameters of Taylor's law at a large spatial scale in 8 of 10 censuses, but not at a fine spatial scale; 3) larger spatial scale increased the exponent of Taylor's law and decreased the exponent of variance–mass allometry (this is the first empirical demonstration that the latter exponent depends on spatial scale), but affected the exponent of density–mass allometry slightly; 4) despite varying natural disturbance, the three laws varied relatively little over the 30 years; 5) self-thinning and recruiting plots had significantly different intercepts and slopes of density–mass allometry and variance–mass allometry, but the parameters of Taylor's law were not usually significantly affected; and 6) higher soil calcium was associated with higher variance of population density in all censuses but not with a difference in the exponent of Taylor's law, while elevation above sea level and soil carbon-to-nitrogen ratios had little effect on the parameters of Taylor's law. In general, the three laws were remarkably robust. When their parameters were influenced by spatial scale and environmental factors, the parameters could not be species-specific indicators. We suggest biological mechanisms that may explain some of these findings.

Understanding the spatial and temporal variability of populations is a fundamental goal of ecology. In response to this challenge, population ecologists have developed and tested power laws that describe general properties of populations. For example, Taylor's law (TL) asserts that the variance of population density of a species is approximately proportional to a power of the mean population density (Taylor 1961, 1984, 1986). This pattern is one of the most widely confirmed power laws in ecology. It has been tested against data of hundreds of species (Taylor 1984, 1986). It has received diverse theoretical interpretations by Taylor, his colleagues, and others (Hanski 1980, 1981, 1982, 1987, Kilpatrick and Ives 2003, Kendal 2004, Eisler et al. 2008, Engen et al. 2008, Kendal and Jørgensen 2011a, b, Cohen et al. 2013, Cohen 2014, Cohen and Xu 2015). How Taylor's power law, and more generally the relationship of variance to mean of population density, varies with the scale of observation, time, disturbance, and other biotic and abiotic environmental characteristics has received insufficient attention.

Taylor's law has found useful applications in studies of genetics, infectious diseases, cancer, number theory, meteorology, and stock and currency trading (Eisler et al. 2008,

Kendal 2004, 2013, Kendal and Jørgensen 2011a, b). These diverse applications suggest that biological mechanisms cannot explain all occurrences of Taylor's law. Even in forests, which are the focus here, the animal behavioral mechanisms that Taylor and colleagues invoked to explain Taylor's law cannot explain why the law holds for trees, which are sedentary.

We focus here on forests because their practical importance has motivated the collection and availability of extensive data which permit testing proposed ecological principles. Tree populations are widely managed for timber, biodiversity, recreation, and carbon sequestration. An ability to provide such services depends upon the demographic and environmental processes that control tree populations. For example, recruitment, self-thinning, disturbance, dispersal of offspring, and competition affect forests at a range of scales and need to be related to Taylor's law if this law is to have a predictive basis in application to forests.

For oak trees in a mixed deciduous forest in Black Rock Forest, New York, USA, Cohen et al. (2012) confirmed Taylor's law and the thinning law or density–mass allometry (DMA): the log of oak density (no. of trees per unit area)

decreased linearly with increasing logarithm of AGB, defined as the mean of above-ground biomass per stem. By combining TL and DMA, Marquet et al. (2005) and, independently, Cohen et al. (2012) predicted theoretically that when TL and DMA hold, then the log variance of population density should decrease linearly with increasing log AGB (variance–mass allometry, or VMA). Cohen et al. (2012) confirmed this prediction empirically for the first time and predicted the slope of VMA from the slopes of TL and DMA. The data of Cohen et al. (2012) were limited to spatial scales ranging from 25 to 75 m on a single hillside in two years, 2007, before girdling some of the trees, and 2010, after girdling. The environmental variation among the contiguous plots in the data of Cohen et al. (2012) was minimal. Lagrue et al. (2015) confirmed TL, DMA and VMA for free-living animals and parasites in New Zealand lakes, showing that these patterns are not limited to forests.

We test Taylor's law, density–mass allometry, and variance–mass allometry in forests on a much larger range of spatial scales, over a much longer time span, over a wider range of environmental conditions, than previously considered in the one prior forest study of all three laws (Cohen et al. 2012). We aim to determine how environmental conditions affect the form and parameters of these laws in forests. In addition to extending empirical ecology, we also aim to test some theoretical explanations of Taylor's law and variance–mass allometry.

We analyzed tree count data collected in forests of mountain beech *Fuscopora cliffortioides*, previously known as *Nothofagus solandri* var. *cliffortioides*, centered on the Craigieburn Range in New Zealand's Southern Alps (43°10'S, 171°35'E). These beech forests are natural monocultures (stands of a single tree species where other woody plants are subcanopy shrubs). These stands have never been subjected to harvesting or other intentional human management. Consequently, in the Craigieburn Range data, the patterns result principally from the dynamics of a single tree species and its environment, and not from interactions of a focal tree species with other tree species or humans. These tree data have been extensively described and analyzed (Wardle and Allen 1983, Harcombe et al. 1998, Coomes and Allen 2007a, 2012, Hurst et al. 2011), but never from the point of view of TL, DMA and VMA.

Our questions are, in summary: do the expected laws hold in these data? If so, could Taylor's law arise from random sampling alone or does Taylor's law indicate the operation of some biological mechanism(s)? If the latter, then what biological mechanisms might be at work? What are the effects of spatial scale; major disturbances; self-thinning and recruitment; elevation and soil chemistry? We now spell out these questions in greater detail.

First, are TL, DMA and VMA valid in a single-species forest without human disruption? A lack of human disruption (e.g. harvesting) might allow background stand-level processes to dominate and structure forest ecosystems, giving stronger evidence for lawful relationships than in previous studies of forests affected by humans.

Second, Cohen and Xu (2015) proved that Taylor's law can arise from random sampling of a skewed probability distribution of population density, with no biological mechanism(s) other than the mechanisms that produced

the probability distribution of population density. Can the observed Taylor's law arise according to the null model of Cohen and Xu (2015)?

Third, Sawyer (1989), in simulated sampling, Yamamura (1990), in empirical observations, and Hanski (1987), in numerical simulations, found that spatial scale affected the parameters but not the form of Taylor's law. The beech forests were observed at much finer ( $5 \times 5$  m) and much larger (90 km<sup>2</sup>) spatial scales than Black Rock Forest. We ask: are there notable trends in the fit and the parameters of TL, DMA and VMA as the area of units of observation changes from  $5 \times 5$  m to  $20 \times 20$  m and the distance between units changes from  $< 20$  m (for subplots in a  $20 \times 20$  m plot) to  $\geq 200$  m to kilometers between plots? A better fit is expected when using the larger units of observation because they more likely robustly represent the mean density of larger trees. One might expect that the uncertainty in estimates of mean density should be lower when using larger units of observation, reducing 'regression dilution' (the reduction in slope of a regression by greater uncertainty in the measurement of the abscissa) and leading to larger estimates of slope in Taylor's law. This expectation is consistent with the predictions of the numerical model of Hanski (1987).

Fourth, the data include 10 censuses over 30 years, 1974–2004, while there were major disturbances from snow, wind, native beetle and related pathogen outbreaks, and earthquake-induced landslides (Wardle and Allen 1983, Allen et al. 1999, Coomes and Allen 2007b, Hurst et al. 2011). On average, 1974 to 1983 was a disturbed phase, 1983 to 1993 a recovery phase, and 1993 to 2004 a disturbed phase (Hurst et al. 2011). The snow damage and pathogen outbreak were more pronounced in the east (1974 to 1985) and the earthquake damage (1994–2004) was much more pronounced in the west (Wardle and Allen 1983, Allen et al. 1999). In simulations, Coomes and Allen (2007b) found that the size–frequency distribution could vary over time because disturbance is stochastic. We ask: does variability over time affect the validity of TL, DMA and VMA? Are the fit and the parameters of TL, DMA and VMA subject to notable variations or trends over time? If so, are the disturbances just described associated with these variations or trends? One might hypothesize that at the end of disturbed phases, the intercept of the DMA allometry would be lower, because disturbances would have destroyed many trees, pushing mean densities in many plots below the self-thinning line until holes in the canopy gradually refilled over time. The residual variance of DMA might also be greater in disturbed periods, but the slope of DMA might not change, depending on how disturbances affected old or young stands.

Fifth, the self-thinning law is intended to describe how population density (stems per ha) declines over time as plants in a stand grow larger, as measured by AGB (Lonsdale 1990, Osawa and Allen 1993, Coomes et al. 2012). By contrast to self-thinning, we say that a stand is 'recruiting' if population density increases over time while AGB decreases, that is, the stand is increasingly occupied by younger, smaller plants. Using the present data, we ask: in plots that are self-thinning, or recruiting, or neither (which we call other), what happens to TL, DMA and VMA? Coomes and Allen (2007a) and Coomes et al. (2012) showed that, in DMA, the intercept of the thinning plots was greater than that of other

plots, presumably because, under classic self-thinning theory, the thinning plots are fully packed with stems whereas other plots sit beneath that thinning line. One might expect the residuals from DMA to have greater variance in the Other plots, because disturbances which knock plots below the thinning line remove variable numbers of stems. One might also expect that recruiting plots would have a lower intercept of DMA, because they are refilling space, but have not yet filled it.

Sixth, the plots in this study ranged widely in elevation above sea level (640 to 1400 m) and soil chemistry. Data for each plot included Ca, the soil concentration of the calcium ion, and C:N, the carbon-to-nitrogen ratio. Clinton et al. (2002) argued that C:N measures nitrogen availability in these soils; when C:N is high, the soil microbes are starved for N and hold N so that it is unavailable to plants. Wisser et al. (1998) showed that soil Ca, magnesium, and C:N are related to understory species composition. Davis et al. (2004) and Smaill et al. (2011) showed that N availability influences seed production. Platt et al. (2004) showed that a mixture of nutrients is important to seedling growth. Coomes and Allen (2007a) showed indirectly that nutrients influence tree growth. Plots with high soil fertility might have high variance in population density as trees are often large and small trees are shaded out, whereas plots with infertile soil might have low variance in population density as most trees might be small. We ask: do these environmental variables affect Taylor's law? Specifically, does adding elevation, Ca, and C:N as independent variables permit more precise prediction of the variance of population density, beyond the influence of mean population density on the variance of population density? Do these environmental variables affect the slope of Taylor's law?

## Methods

The following descriptions of data and theory pertain to all six questions. The section on statistical methods describes the methods used for each question in separate sub-subsections.

### Data

Mountain beech trees were counted in 250 permanent study plots distributed over 9000 ha (90 km<sup>2</sup>) (Supplementary material Appendix 1 Fig. A1). Each 20 × 20 m study plot was subdivided into 5 × 5 m subplots (16 per plot). There were in total 4000 = 16 × 250 subplots of size 5 × 5 m.

Plots were located along 98 transect lines. The origin of each transect line was located randomly along a stream. The side of the stream sampled by the transect line was also selected randomly. The direction of the transect line (the compass bearing) was then selected as towards the nearest tree line according to a topographic map or, if there was no nearest tree line, to the nearest ridge. Plots were established at 200-m intervals along the transect line until the tree line or a ridge was reached. Each transect line had one to eight plots. Within each plot, the diameter at breast height of each tree was measured and recorded by species and subplot (Hurst and Allen 2007). Very few trees were of species other

than mountain beech, and they are excluded from further analysis here. In 1974, all trees were uniquely tagged at measurement height. Subsequent re-measurements identified recruited and dead trees based on tags during the austral summers starting in 1976, 1978, 1980, 1983, 1985, 1987, 1993, 1999 and 2004 (Hurst et al. 2011). A 'measurement' consisted of the census date, plot number, subplot number, tree identifier and diameter at breast height. In the version of the data we used, 31 127 trees were measured a total of 178 485 times. Our computations excluded all measurements with diameter < 30 mm.

Coomes et al. (2012, p. 48) estimated the above-ground biomass (kg C) of each tree in every plot and census from stem diameter at breast height (cm) and stem height (m), using an empirical relationship developed for mountain beech (Harcombe et al. 1998). Height was estimated for each tree from stem diameter at breast height (cm) and ALT = (elevation [m above sea level] – 640)/100 using an equation developed for mountain beech (Coomes et al. 2012, p. 49, their Eq. 12). Elevation above sea level was measured at the centre of each plot using a barometric altimeter. The lowest plot had elevation 640 m above sea level.

The data used in this paper and the R code used to analyze it and to produce the figures and tables are available through New Zealand's National Vegetation Survey databank (<<https://nvs.landcareresearch.co.nz/>>), date of deposition 23/09/2015, URL <<http://datastore.landcareresearch.co.nz/dataset/allometric-power-laws-in-nz-mountain-beech-forests>>. In addition, the Supplementary material Appendix 2 gives many of the data and summary statistics.

## Theory

Taylor's law (TL) asserts that:

$\log(\text{variance of population density}) = \log(a) + b \times \log(\text{mean of population density})$ ,  $a > 0$ .

All logarithms are to base 10. For each census separately, we examined the mean and variance of population density at two spatial scales: among subplots in a plot; and among plots in a block.

Density–mass allometry (DMA) asserts that:

$\log(\text{mean of population density}) = \log(u) + v \times \log(\text{mean of body mass per individual})$ ,  $u > 0$ .

DMA differs from the other two power laws, TL and VMA, which require spatial subunits for estimates of variability (subplots in plots, or plots in blocks). DMA can be tested using subplots, plots, or blocks as the spatial unit. Because a considerable fraction of subplots had zero trees, we tested DMA using only 20 × 20 m plots.

Variance–mass allometry (VMA) asserts that:

$\log(\text{variance of population density}) = \log(h) + k \times \log(\text{mean of body mass per individual})$ ,  $h > 0$ .

If TL and DMA hold true, then VMA is predicted to hold true with parameter values

$$h = au^b \text{ and } k = bv$$

We tested these predictions using subplots in each plot and using plots in each block.

## Statistical methods

We used ‘significantly’ always to mean ‘nominally statistically significantly’ ( $p < \alpha = 0.05$ ) because  $\alpha = 0.05$  is conventional. The  $p$ -values were not corrected for simultaneous (or multiple) inference, but this made little or no difference to the substantive conclusions drawn. All  $R^2$  in this paper are adjusted for degrees of freedom, so we omit ‘adjusted’ henceforth.

### Question 1. Are the three power laws valid?

Both Taylor’s law and variance–mass allometry require a way of estimating means and variances, that is, a grouping of smaller units of observation into larger units of observation. Each  $20 \times 20$  m plot constitutes a natural group of its 16 subplots of size  $5 \times 5$  m, so we calculated the mean and variance (of population density or AGB) over the 16 subplots for each plot.

At a larger spatial scale, we grouped  $20 \times 20$  m plots into a larger spatial unit which we called ‘blocks’ in conformity with statistical usage. We calculated a mean and variance for the plots in each block. We used Ward’s minimum variance clustering to allocate each of the 250 plots into one of 13 disjoint blocks containing 5–33 plots per block based on geographic distance among plots (Supplementary material Appendix 1 Table A2). Supplementary material Appendix 1 Fig. A1 maps the plots and blocks. This clustering method defines blocks to minimize the within-block sum of squares of distances between plots (Gordon 1999). A silhouette-optimal rule was used to select an appropriate number of blocks (Supplementary material Appendix 1 Fig. A2, Borcard et al. 2011). The silhouette width measures the degree of membership of a plot to its block, based on the average distance between this plot and all plots of the block to which it belongs, compared to the same measure computed for the next closest block (Rousseeuw 1987). Supplementary material Appendix 1 Fig. A3 gives a dendrogram of Ward’s minimum variance clustering for the partitioning of plots.

Taylor et al. (1988) recommended  $\geq 15$  plots per block and  $\geq 5$  blocks. Ten of our 13 blocks were consistent with this recommendation. However, blocks 3, 7 and 13 had 5, 6 and 5 plots each. We plotted these three blocks with a filled circle to distinguish them from the open circles that represented the other ten blocks. The filled circles generally fell along the same linear trends as the open circles (Fig. 1, 3, 4). We concluded that the three blocks with  $< 15$  plots each were not misleading.

Each count of the number of trees in a spatial unit was converted to population density by dividing the number of trees by the area of the spatial unit in hectares (ha).

The mean and variance of the population density of a set of spatial units included the spatial units with 0 population density. Excluding subplots with 0 trees while including subplots with trees would have biased density estimates upwards. Enough spatial units were always included so that neither the mean nor the variance was 0, with the following rare exceptions: three plots in 1999 and four plots in 2004 had 0 trees. The 1994 Arthurs Pass earthquake caused these zeros. These plots were excluded from the tests of TL, DMA and VMA for those censuses only.

We estimated the parameters of each power law in its log-transformed version by ordinary least squares (OLS) linear regression. For example, in TL, we used OLS to estimate the intercept  $\log(a)$  and the slope  $b$ . This application of linear regression assumes, contrary to fact, that the sample mean is known exactly. Because the sample mean is subject to sampling variation, this procedure has been criticized (McArdle 1988, McArdle et al. 1990) and defended (Smith 2009, Lai et al. 2013). Lai et al. (2013) discussed the issues in using OLS regression and cited many relevant references. Here OLS is defensible because the variance of the sample mean is much smaller than the variance of the sample variance. Hence the log mean has much smaller sampling variation than the log variance. The assumption of little variability in the abscissa compared to the variability in the ordinate is more accurate than the alternative assumption in reduced major axis regression that the two variables are symmetrical (Smith 2009). Cohen et al. (2012) found that the substantive conclusions were unaffected by the fitting procedure.

To test for linearity, we extended the linear relationships  $y = a + b \times x$  (where  $x$ ,  $y$  and the coefficients differ for TL, DMA and VMA) to quadratic relationships  $y = a + b \times x + c \times x^2$  and tested whether  $c$ , the coefficient of the quadratic term, differed significantly from zero (following Taylor et al. 1978, Hanski 1987). If it did not, we concluded that the data did not reject a linear relationship of  $y$  to  $x$ . This procedure assumed the appropriateness of the statistical model used to estimate the confidence interval of  $c$ .

For least-squares regression, we used *lm* in R. We computed Ward’s minimum variance clustering and its silhouette information using *hclust* in the *stats* package and *silhouette* in the *cluster* package in R. We computed the moments of frequency distributions using *empMoments* in the *PearsonDS* package in R.

To test at each spatial scale whether the slope of VMA was predicted by the product of the slopes of TL and DMA, or  $k = bv$ , as derived above in Theory, we made bootstrap samples (random samples with replacement) of the 250 plots and, independently, bootstrap samples of the 13 blocks. For each sample, we calculated  $k - bv$ . We took the 2.5 percentile and the 97.5 percentile of these 1000 estimates of  $k - bv$  as the 95% confidence interval of  $k - bv$ . If that CI included zero, then we concluded that we could not reject  $bv$  as a predictor of  $k$ . We did this entire calculation 1000 times for plots, first with 250 plots randomly selected with replacement, and another 1000 times with only 13 plots randomly selected with replacement, so that we could compare the results for plots with the results of 1000 bootstrap samples of size 13 from the 13 blocks. The results are in Supplementary material Appendix 2 Table A6.

### Question 2. Does random sampling in blocks explain Taylor’s law in these data?

Cohen and Xu (2015) proved that if a large random sample from a distribution with finite mean  $M > 0$ , finite variance  $V > 0$ , and finite third and fourth moments is divided randomly into  $N$  blocks, and if the mean  $m_j$  and variance  $v_j$  of the observations in each block are computed, and if the  $\log v_j$  of each block is plotted as a function of the  $\log m_j$  of each block,  $j = 1, 2, \dots, N$ , then TL will hold approximately, and

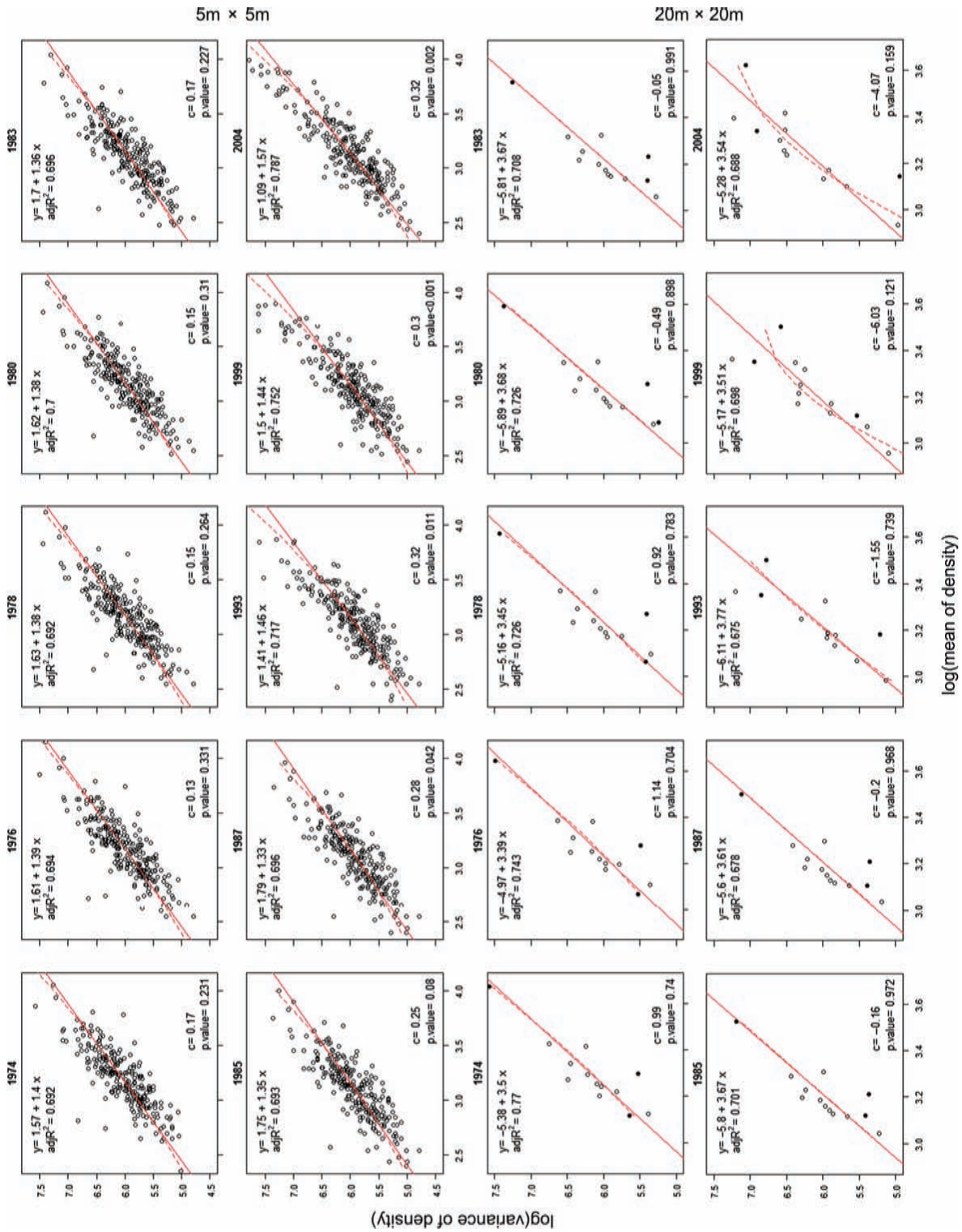


Figure 1. Tests of spatial Taylor's law in 10 censuses using ordinary least-squares regression (solid red line) and quadratic least-squares regression (dashed red line). All logarithms are to base 10. Slope of linear regression is significantly non-zero ( $p < 0.001$ ) in all censuses and both spatial scales; Row 1 and 2: 16 subplots ( $5 \times 5\text{ m}$ ) in each plot. Number of plots is 250 in all censuses except 247 in 1999 and 246 in 2004. Row 3 and 4: 5–33 plots ( $20 \times 20\text{ m}$ ) in each of 13 blocks in all censuses. Solid points mark the three blocks with fewer than 15 plots.

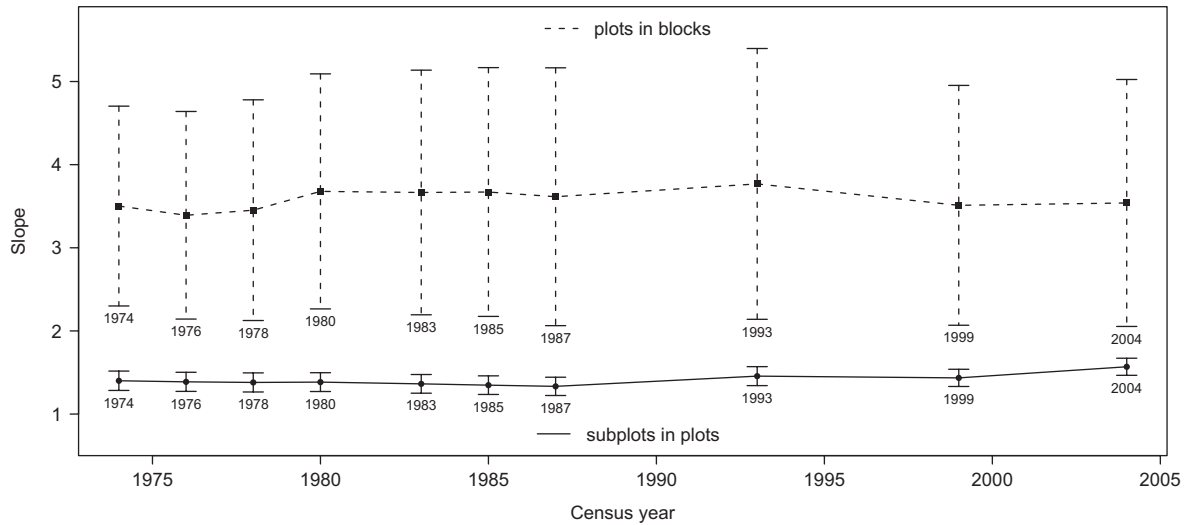


Figure 2. Slopes of Taylor's law in each census. Error bars show the 95% confidence interval of the slope coefficient. Spatial units are subplots in plots (solid line, solid error bars) and plots in blocks (dashed line, dashed error bars).

the slope  $b$  of TL can be predicted from the moments of the underlying probability distribution. We used their formulas to estimate  $b$  and the unbiased estimator of its sample variance when all blocks are weighted equally.

#### Questions 3 and 4. Spatial scale and temporal change

To test whether spatial scale and census year influenced the slope of TL, we did an analysis of covariance (Supplementary Appendix 2 Table A4) in which the dependent variable was log variance of population density and the independent variables were: log mean of population density, unit (2 categories, for subplots in plots versus plots in blocks), census (10 categories), log mean  $\times$  unit, log mean  $\times$  census, unit  $\times$  census, and log mean  $\times$  unit  $\times$  census. We also did a linear regression (separately for plots and for blocks) in which the dependent variable was the slope  $b$  in a census year and the independent variable was the calendar year.

#### Question 5. Self-thinning and recruitment

To test whether plots that were self-thinning, recruiting, or other over the duration of the study had different parameters of TL, DMA and VMA, we defined a  $20 \times 20$  m plot to be self-thinning if it had progressively fewer trees that were progressively bigger, i.e. if and only if its density and AGB satisfied

$$\text{density}(1974) \geq \text{density}(1987) \geq \text{density}(2004)$$

$$\text{AGB}(1974) \leq \text{AGB}(1987) \leq \text{AGB}(2004)$$

The four plots that did not have a positive number of trees in each of these three years were excluded from the following analysis. Of the 246 plots with nonzero trees in all three of these years, 149 satisfied this definition of self-thinning.

We also said that a plot was recruiting if and only if it had progressively more trees that were progressively smaller, i.e. if and only if

$$\text{density}(1974) \leq \text{density}(1987) \leq \text{density}(2004)$$

$$\text{AGB}(1974) \geq \text{AGB}(1987) \geq \text{AGB}(2004)$$

Of the 246 plots with nonzero trees in all three of these years, only seven satisfied this definition of recruiting. The relative rarity of recruiting plots is partly a consequence of our definition because it would be remarkable for a plot to recruit continuously for 30 years. Coomes and Allen (2007a) observed that many plots disturbed in 1974–1984 recruited trees in 1984–1993.

We defined a plot that was neither self-thinning nor recruiting to be other. Of the 246 plots with nonzero trees in all three of these years, 90 satisfied this definition of other.

Supplementary material Appendix 2 Table A3 gives the indicator variables for the three dichotomous variables Other, Recruiting and Self-thinning. (We use initial capitals to refer to these indicator variables, and lower case to describe the plots thus classified.) For example, plot 2 was self-thinning, so Other = 0, Recruiting = 0, and Self-thinning = 1 for this plot. Because, by definition for every plot, Other + Recruiting + Self-thinning = 1, only two of these three indicator variables were linearly independent, so we picked Other as the reference category and investigated the influence of Recruiting and Self-thinning relative to Other. We did not define self-thinning, recruiting, or other subplots or blocks.

We then analyzed TL, DMA and VMA with a linear model that had a dependent variable equal to log variance of population density (for  $5 \times 5$  m subplots of each  $20 \times 20$  m plot), and had, as independent variables: log mean of population density, Self-thinning, Recruiting, Self-thinning  $\times$  log mean, and Recruiting  $\times$  log mean. We asked whether the coefficients of Self-thinning  $\times$  log mean and Recruiting  $\times$  log mean differed from zero, representing any influence of the indicator variables on the slope of TL. Analogous models were used to assess the influence of the indicator variables on the slope of DMA and VMA.

#### Question 6. Environmental factors

Soil calcium (Ca) and C:N were measured in 1992. The values for Ca measured the calcium available to plants in the upper 10 cm of mineral soil (micrograms of available calcium per gram of dried soil) (Wiser et al. 1998). The soil

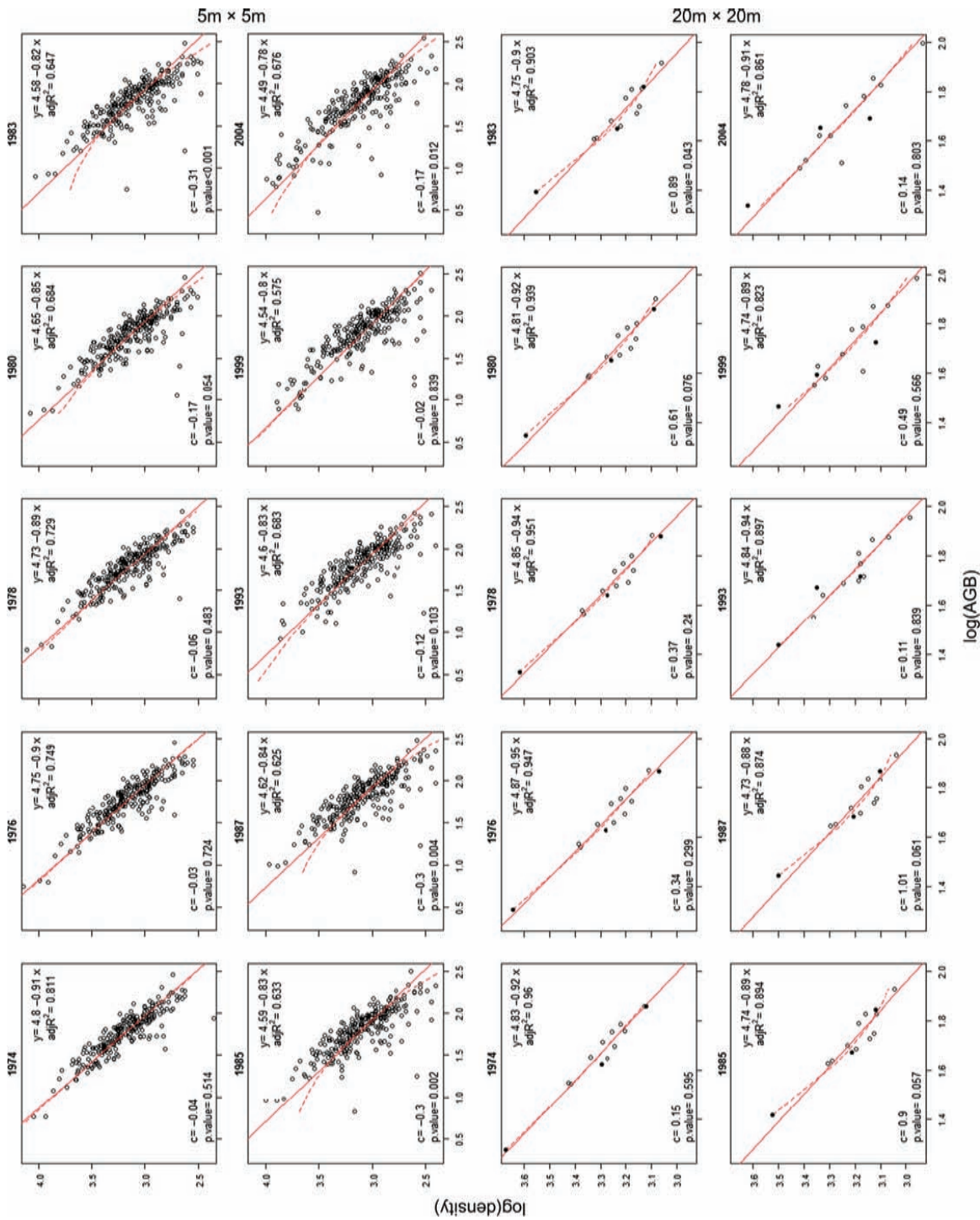


Figure 3. Tests of density–mass allometry in 10 censuses using ordinary least-squares regression (solid red line) and quadratic least-squares regression (dashed red line). AGB = average above-ground biomass (kg C) per tree. Density = number of trees per ha. All logarithms are to base 10. Slope of linear regression is significantly non-zero ( $p < 0.001$ ) in all censuses and both spatial scales. Row 1 and 2: 16 subplots ( $5 \times 5\text{ m}$ ) in each plot. Number of plots is 250 in all censuses except 247 in 1999 and 246 in 2004. Row 3 and 4: 5–33 plots ( $20 \times 20\text{ m}$ ) in each of 13 blocks in all censuses. Solid points mark the 3 blocks with fewer than 15 plots.

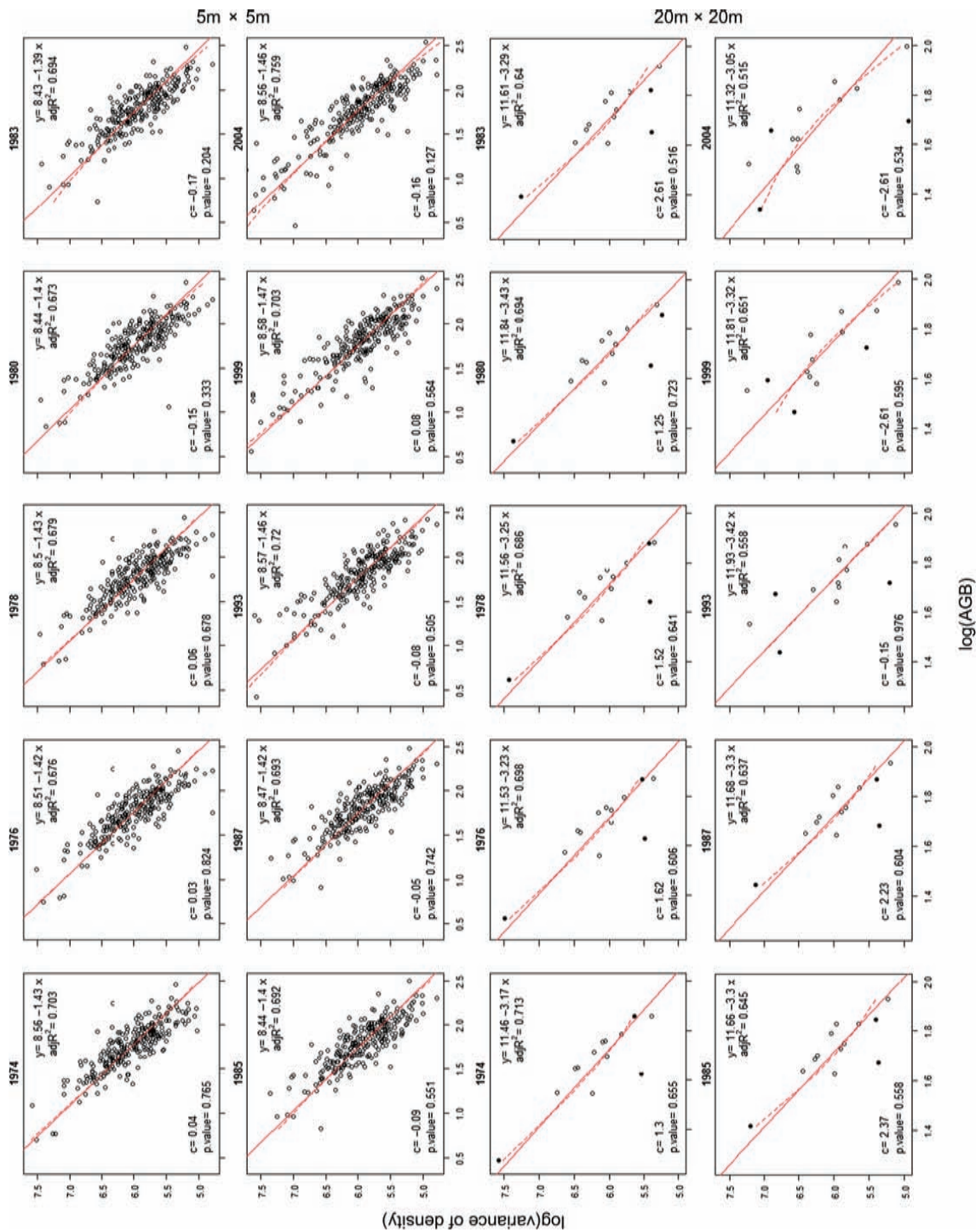


Figure 4. Tests of variance-mass allometry in 10 censuses using ordinary least-squares regression (solid red line) and quadratic least-squares regression (dashed red line). AGB = average above-ground biomass (kg C) per tree. All logarithms are to base 10. Slope of linear regression is significantly non-zero ( $p < 0.001$ ) in all censuses and both spatial scales. Row 1 and 2: the variance is among 16 subplots ( $5 \times 5$  m) in each plot while the AGB is for plots. Number of plots is 250 in all censuses except 247 in 1999 and 246 in 2004. Row 3 and 4: the variance is among 5–33 plots ( $20 \times 20$  m) in each block while the AGB is for blocks. Number of blocks is 13 in all censuses. Solid points mark the three blocks with fewer than 15 plots.



was collected in eight representative samples from each plot and pooled in a single composite soil sample per plot. The composite sample was dried, sieved and the nutrients were extracted. Calcium was recovered from the soil using a Bray 2 extract. A subsample of the sieved soil for each plot was finely ground, and its N and C concentrations (by percentage) were determined using a CNS elemental analyzer. The values for C:N are the dimensionless ratios of carbon to nitrogen values as percentages of the upper 10 cm of mineral soil. These environmental variables are listed by plot in Supplementary material Appendix 2 Table A1.

For any quantitative attribute  $X$  of a plot, such as its concentration of a chemical in the soil, we defined the standardized version of  $X$  as  $Y = (X - \text{sample mean [over all plots] of } X) / (\text{unbiased estimate of sample standard deviation [over all plots] of } X)$ .

To test whether environmental factors influenced the variance of population density, beyond the effect of mean population density, we fitted this linear model:

$$\text{Model 1: log variance of population density} = \log(a) + b \times \log \text{ mean of population density} + c \times [\text{standardized ALT}] + d \times [\text{standardized Ca}] + e \times [\text{standardized C:N}].$$

To anticipate our results, we found that the coefficients  $c$  of standardized ALT and  $e$  of standardized C:N were not significantly different from zero in most cases, but the coefficient  $d$  of standardized Ca was significantly greater than zero in a majority of cases. This showed that an increased calcium concentration was associated with a greater variance of population density. To test whether Ca influenced the slope of TL, that is, whether it influenced the rate of change of the variance of population density with respect to the mean population density, we fitted:

$$\text{Model 2: log variance of population density} = \log(a) + b \times \log \text{ mean of population density} + d \times [\text{standardized Ca}] + f \times [\text{standardized Ca}] \times \log \text{ mean of population density}.$$

Model 2 included the term  $f \times [\text{standardized Ca}] \times \log \text{ mean of population density}$  for the interaction of standardized Ca and log mean. The null hypothesis in model 2 was that there was no interaction between standardized Ca and the log mean of population density, that is, that standardized Ca did not affect the slope of TL.

## Results

### Question 1. Do Taylor's law (TL), density–mass allometry (DMA), and variance–mass allometry (VMA) hold?

#### Taylor's law

In general, TL was not rejected by these data. In 20 tests of the linearity of the relationship between log variance and log mean (null hypothesis: the quadratic coefficient is zero) for 10 censuses  $\times$  2 spatial units (subplots in plots plus plots in blocks), four of the  $p$ -values were  $< 0.05$  and occurred in the last four censuses, only for subplots in plots (in the

second row and last four columns in Fig. 1). Where there was evidence for nonlinearity ( $p < 0.05$ ), the quadratic coefficient ranged from 0.28 to 0.32, indicating that log variance was a convex function of log mean (Supplementary material Appendix 2 Table A6).

#### Density–mass allometry

For plots (rows 1 and 2 in Fig. 3), linearity of  $\log(\text{density})$  as a function of  $\log(\text{AGB})$  was not rejected in six of 10 censuses (Supplementary material Appendix 2 Table A6). In the remaining four censuses (1983, 1985, 1987, 2004), the nonlinearity was due to plots that fell far below the linear relationship of the majority of plots. In these apparent outliers, the density was exceptionally low, given the mean biomass. These low densities may reflect external events like landslides or epidemics that eliminated many trees. The slopes of DMA for plots fell between  $-0.91$  (in 1974) and  $-0.78$  (in 2004).

For blocks (rows 3 and 4 in Fig. 3), linearity of  $\log(\text{density})$  as a function of  $\log(\text{AGB})$  was not rejected in nine of 10 censuses. The exceptional census, in 1983, was only marginally significantly nonlinear ( $p = 0.043$ ), so DMA was confirmed in blocks. The slopes of DMA for blocks fell between  $-0.96$  (in 1976) and  $-0.88$  (in 1987). The CIs of the DMA slopes for the ten censuses all included  $-1$  (Supplementary material Appendix 2 Table A6), like the results of Cohen et al. (2012), but the CIs of the last six censuses (1983–2004) also included  $-3/4$ .

#### Variance–mass allometry

VMA described well the linear dependence with negative slope of the log variance of density on the log AGB (Fig. 4). The 20 tests of the null hypothesis of linearity never rejected it with  $p < 0.05$  (Supplementary material Appendix 2 Table A6).

The slope of VMA was acceptably predicted by the product of the slopes of TL and DMA, that is, the 95% CI of  $k - bv$  from 1000 bootstrap samples included zero for subplots in plots and for plots in blocks when each bootstrap sample was of size 13, for all ten censuses, that is, approximately  $k \approx bv$ . However, for subplots in plots, the 95% CI of  $k - bv$  did not include, but fell strictly below, zero when each of 1000 bootstrap samples was of size 250, for all ten censuses, that is,  $k < bv$  in these estimates (Supplementary material Appendix 2 Table A6).

Because we had 13 blocks and 250 plots, we do not know whether the difference in results is due to sample size or spatial scale. When we used consistent sample sizes (13) for both spatial scales, we obtained consistent results: approximately  $k \approx bv$ .

### Question 2. Could random sampling in blocks explain the slope of TL?

To start with an illustrative example (Supplementary material Appendix 1 Fig. A4), in the 2004 census, the frequency histograms of density for the  $5 \times 5$  m subplots ( $N = 3936 = (250 - 4) \times 16$ ) had mean density = 1860 trees per hectare, standard deviation = 2670 trees per hectare, skewness = 4.43, and kurtosis = 31.99 (Supplementary material Appendix 2 Table A5). The predicted value

of  $b$ , asymptotically for large numbers of blocks, was  $1860 \times 4.43/2670 \approx 3.09$ . For the  $N=246$  plots in the 2004 census, the estimated slope of TL was 1.57 (Fig. 1) and the 95% confidence interval (CI) was 1.47 to 1.67 (Supplementary material Appendix 2 Table A6). The predicted slope 3.09 lay far outside the 95% CI of the estimated  $b$  in this example. The hypothesis that random sampling in blocks could explain the slope of TL was rejected in this example.

In every case, for  $5 \times 5$  m subplots, the asymptotic slope predicted from random sampling of a single distribution exceeded the estimated slope of TL, and also exceeded the upper limit of the CI of the slope of TL in every census (Supplementary material Appendix 2 Table A5).

However, for  $20 \times 20$  m plots in blocks, the CI of the slope of TL included the asymptotic slope from random sampling of a single distribution for eight of 10 census years (except for 1993 and 1999). In eight of 10 census years, the observed TL had a slope indistinguishable from the slope predicted by random samples of plots.

In summary, the analysis rejected the null hypothesis that the slope of TL could be explained by random sampling of a single underlying distribution when the spatial units were subplots, but not when the spatial units were  $20 \times 20$  m plots.

Mean population density varied among subplots in a plot by about 1.5 orders of magnitude, from approximately  $\log_{10}$  mean number of stems per ha = 2.5 to approximately 4.0. Mean population density varied much less among plots in a block, by about half of one order of magnitude, from approximately  $\log_{10}$  mean number of stems per ha = 3.0 to 3.5 or 3.6.

### Question 3. Does spatial scale affect TL, DMA and VMA?

The parameters of TL, DMA and VMA were strongly affected by spatial scale (Fig. 5). The differences in slope and intercept between spatial scales were notably greater for TL and VMA than for DMA. The spatial scale that had the higher intercept had a lower slope than the other spatial scale (except for a few overlaps for DMA). At all 10 censuses, TL had higher intercepts and lower slopes at the smaller spatial

scale, while VMA and (usually) DMA were reversed. We now describe the details.

#### Taylor's law

Visually (Fig. 2) and statistically (Supplementary material Appendix 2 Table A4), the TL slope was larger for plots in blocks than for subplots in plots. For subplots in plots, the slope ranged from a low of 1.33 (in 1987) to a high of 1.57 (in 2004). By contrast, for plots in blocks, the slope ranged from 3.39 (in 1976) to 3.77 (in 1993) (Fig. 1, 2). According to the analysis of covariance, only three factors significantly affected log variance of population density with  $p < 0.05$ : log mean population density, unit (that is, plot or block), and log mean  $\times$  unit, the first and last of these each with  $p < 2 \times 10^{-16}$ .

#### Density–mass allometry

In each census, the CI of the slope of DMA for blocks included or strongly overlapped with the slope of DMA for plots (Supplementary material Appendix 1 Fig. A5). Although the slope of DMA for plots was less negative than the slope of DMA for blocks in every census, the difference in the slope of DMA between plots and blocks was not significant in any individual census. Because we had far fewer blocks (13) than plots (250 usually), the CIs for blocks were much wider than the CIs for plots. Hence the systematic difference in slopes between blocks and plots did not attain statistical significance ( $p = 0.265$ ;  $R^2 = 0.046$ ).

#### Variance–mass allometry

When variance was calculated over plots in blocks, the CI of the slope  $k$  of VMA contained the slope  $b \times v$  predicted by composing TL and DMA (see Theory above) for all 10 censuses (Supplementary material Appendix 2 Table A6). For example, in 1974,  $b = 3.50$ ,  $v = -0.92$ ,  $b \times v = -3.22$ ,  $k = -3.17$ , and the CI of  $k$  was  $(-4.42, -1.91)$ , which included  $-3.22$ . The predicted slopes were very close to the estimated values with an average 0.07 absolute difference (Supplementary material Appendix 1 Fig. A6).

By contrast, when variance was calculated over subplots in plots, in every census the predicted slope from  $b \times v$  was less negative than the estimated slope  $k$  of VMA, and exceeded the upper limit of the CI of the slope of VMA (Supplementary material Appendix 2 Table A6, Fig. A6).

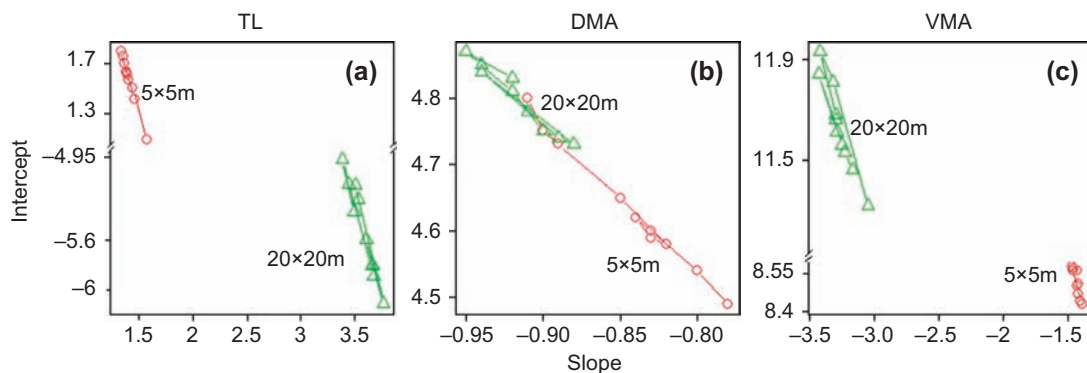


Figure 5. Intercept and slope of Taylor's law (a, TL), density–mass allometry (b, DMA), and variance–mass allometry (c, VMA) in 10 censuses (one marker per census) at a small spatial scale (red circle for  $5 \times 5$  m subplots in  $20 \times 20$  m plots) and a large spatial scale (green triangle for  $20 \times 20$  m plots in blocks). The lines connect points that represent successive censuses.

Thus the slope of VMA could be predicted as the product of the slopes of TL and DMA accurately for plots in blocks and not at all for subplots in plots. Whether this difference is due to spatial scale or difference in sample sizes between plots and blocks is not clear.

**Question 4. Are the fit and the parameters of TL, DMA and VMA subject to notable variations or trends over time? If so, are disturbances associated with these variations or trends?**

Despite wide variations in the degree of natural disturbance over the 30 years of observation, the three power laws varied relatively little in time.

**Taylor's law**

When the census was treated as a categorical variable, there was no significant evidence (Supplementary material Appendix 2 Table A4) that the census year affected the slope  $b$  of TL (Fig. 2). However, when the slope  $b$  of TL was modeled as a linear function of the numerical value of the calendar year of the census, the slope  $b$  of TL was estimated to increase, on average, by about 0.005 per year ( $p = 0.02$ ,  $R^2 = 0.44$ ) when using subplots in plots. This rate of increase amounts to an increase in the slope  $b$  of TL by nearly 0.15 over 30 years, a non-trivial increment in the slope of TL. The slight increase in the slope  $b$  of TL by about 0.003 per year when using plots in blocks was not significant ( $p = 0.49$ ,  $R^2 = -0.057$ ).

**Density-mass allometry**

The slope of DMA as a function of calendar year for plots was significantly positive ( $p < 0.001$ ;  $R^2 = 0.7952$ ), though small (Supplementary material Appendix 2 Table A7). The CIs of the DMA slopes for all ten censuses excluded  $-1$  (Supplementary material Appendix 2 Table A6), unlike the results of Cohen et al. (2012), but the CIs of five of ten censuses (1983, 1985, 1993, 1999, 2004) included  $-3/4$ , as predicted by metabolic scaling theory (West et al. 2009).

The slope of DMA as a function of calendar year for blocks was not significantly different from zero, though also slightly positive. In summary, the slope of DMA increased over time in both plots and blocks, but the magnitude of these effects was small and sometimes not significant.

**Variance-mass allometry**

For blocks, the slope of VMA declined very slightly from 1974 to 1993 and then rose slightly, but wide CIs masked any difference in slope among censuses. The regression of the slope of VMA as a function of calendar year did not have a slope significantly different from zero.

The CIs were much smaller for subplots in plots than for plots in blocks, presumably because there were many more plots than blocks. For plots, the slope of VMA declined slightly with increasing calendar year, with marginal statistical significance ( $p = 0.032$ ;  $R^2 = 0.39$ ; Supplementary material Appendix 2 Table A10). In every census, the slope for subplots in plots was less negative than the slope for plots in blocks (Supplementary material Appendix 2 Table A6).

**Question 5. Do self-thinning and recruitment affect TL, DMA and VMA?**

We defined self-thinning, recruiting, and other as properties of plots but not subplots or blocks. The slope of TL was not significantly influenced by whether plots were self-thinning, recruiting, or other, though the intercept of TL may have been influenced by these attributes. The slopes and intercepts of DMA and VMA were influenced by whether plots were self-thinning, recruiting, or other. In general, for both DMA and VMA, self-thinning plots had higher intercepts and lower slopes, and recruiting plots had lower intercepts and higher slopes, than other plots. The details follow for each power law separately.

**Taylor's law**

Self-thinning plots had a lower intercept for TL (a lower variance of population density in  $5 \times 5$  m subplots for a given mean density) than other plots in eight of 10 censuses (all but the last two censuses), but the difference was significant only in 1985 and 1987. Self-thinning plots had a higher slope of TL than other plots in eight of 10 censuses (all but the last two), but the difference was significant (marginally:  $p = 0.027$ ) only when it was most negative, in 2004 (Supplementary material Appendix 2 Table A3, which gives both linear models and ANOVA). Where the variable Self-thinning raised (or lowered) the intercept of TL, its interaction with log mean density had an effect of opposite sign on the slope of TL. The variable Recruiting had no significant effect on the intercept of TL, raising the intercept in six censuses and lowering it in four, compared to other plots, and no significant interaction with log mean density to change the slope of TL.

**Density-mass allometry**

By contrast with the results for TL, the indicators Self-thinning and Recruiting and their interactions with the independent variable generally had a significant influence on the intercepts and slopes of DMA and VMA (Supplementary material Appendix 2 Table A3).

In the linear models for DMA, the coefficient of Self-thinning was positive in all 10 censuses ( $p < 0.05$ ), meaning that self-thinning plots had, on the average, higher log density than other plots. The coefficient of Recruiting was significantly different from zero in eight of the ten censuses, and was negative in seven of these eight. Where the coefficient was negative, recruiting plots had, on the average, lower log density than other plots. The positive coefficients of Recruiting occurred in the last three censuses, 1993–2004.

Of the 20 coefficients of Self-thinning and Recruiting in all 10 censuses, 18 differed significantly from zero. Self-thinning always increased, and Recruiting generally decreased, the log mean density in DMA, relative to Other plots. Wherever Self-thinning increased the intercept of DMA, its interaction with log AGB decreased the slope of DMA. Wherever Recruiting decreased the intercept of DMA, its interaction with log AGB increased the slope of DMA. These countervailing effects pivoted the DMA regression line around a center of observations of (log AGB, log mean of density).

### Variance–mass allometry

For VMA, Self-thinning increased the intercept in nine of 10 censuses, but significantly in only four of ten censuses (1983–1993). Recruiting decreased the intercept in nine of 10 censuses, but significantly in only four of 10 censuses (1974–1980). Thus in most cases, for a given log AGB, self-thinning plots had higher variance, and recruiting plots had lower variance, than did other plots. With only two exceptions in 20 cases, the interactions of these indicator variables with log AGB had the opposite effect on the slope of VMA, lowering the slope where the indicator raised the intercept and raising the slope where the indicator lowered the intercept. These countervailing effects pivoted the VMA regression line around a center of observations of (log AGB, log variance of density).

### Question 6. Do elevation and soil chemistry affect TL?

In model 1, the coefficients  $c$  of standardized ALT were always positive but never significantly different from zero, except for a case of marginal significance in 1980 when  $p = 0.0497$  (Supplementary material Appendix 2 Table A8). The coefficients  $e$  of standardized C:N were always negative but never significantly different from zero, except in 2004 when  $p = 0.0185$ . By contrast, the coefficient  $d$  of standardized Ca was greater than zero in all 10 censuses, and was significantly positive in the most recent seven of the 10 censuses. Thus an increased Ca concentration was associated with a greater variance of population density, after controlling for other factors. The coefficient  $d$  of Ca generally increased in time, from 0.02 in 1974 to 0.05 in 2004.

In model 2, there was no significant interaction between standardized Ca and the log mean of population density, that is, the standardized Ca did not affect the slope of TL in any census (Supplementary material Appendix 2 Table A9). The standardized Ca was highly collinear (linear correlation  $> 0.99$ ) with the interaction term of the standardized Ca with log mean density. When the interaction term was removed, the standardized Ca strongly affected the variance of population density.

## Discussion

### Question 1. Validity of TL, DMA and VMA and their mutual dependence

A spatial Taylor's law (TL) (Fig. 1), density–mass allometry (DMA) (Fig. 3), and variance–mass allometry (VMA) (Fig. 4) usefully described bivariate relationships of each pair of log mean density, log variance of density, and mean above-ground biomass as properties of tree populations in the absence of other tree species and human disruption. The internal dynamics of mountain beech forests, environmental disturbances unrelated to humans (e.g. wind throw, earthquakes and landslides) and enduring or slowly changing spatial variability in the environment (here soil chemistry and elevation) generated TL, DMA and VMA. Human disruption was not necessary for the validity of TL, DMA and VMA.

TL and DMA have a long history in ecology, but VMA was unknown until Marquet et al. (2005) and independently Cohen et al. (2012) derived VMA from TL and DMA mathematically. Prior to Cohen et al. (2012), VMA was never tested empirically. It is now clear that any two of the three mathematical equations for TL, DMA and VMA algebraically determine the third, if one ignores the error structure, i.e. the deviations or residuals from the log-log linear relations.

Is there some sense in which two of these three power laws are fundamental and the third is derived from them? In classical allometric power laws, the independent variable is individual body mass and the dependent variable is some individual measurement like basal metabolic rate. Given the successful history of allometric power laws, it is tempting to view DMA and VMA as fundamental because they are both related to body mass, and TL as derived.

However, DMA and VMA differ from allometric power laws in important respects: the independent variable in DMA and VMA is mean body mass (AGB), a characteristic of a population, not an individual, and the dependent variables are also population characteristics (mean density in DMA, variance of density in VMA). Therefore it is necessary to consider whether the residuals from any two laws proposed to be fundamental could reproduce the residuals observed in the third. For example, if mean body mass determined both mean density and variance of density, but the mean and variance were independent, conditional on the mean body mass, then one would expect the relation (TL) between mean density and variance of density to be looser (have lower coefficient of determination  $R^2$ ) than the relations of mean body mass to mean density (DMA) and of mean body mass to the variance of density (VMA).

The only evidence we know that bears on which power laws are fundamental comes from simultaneous tests of TL, DMA and VMA (Cohen et al. 2012, Lagrue et al. 2015, this paper).

Cohen et al. (2012) confirmed TL, DMA and VMA in studies of oak population density in Black Rock Forest, NY. In 2007, before some trees were girdled, TL had  $R^2$  of 0.811, DMA 0.976 (using plots, not subplots), and VMA 0.754. In 2010, after girdling, TL had  $R^2$  of 0.807, DMA 0.765 (using plots), and VMA 0.686. Girdling reduced the  $R^2$  of all three relationships. Before and after girdling, VMA had the lowest  $R^2$  of the three relationships, suggesting that VMA was derived from TL and DMA.

Lagrue et al. (2015) confirmed TL, DMA and VMA in studies of free-living species (distinguishing those with and those without parasites) and parasitic species in New Zealand lakes. TL was a much tighter relationship (with  $R^2$  of 0.96 to 0.98, depending on the group of species considered) than either DMA ( $R^2$  of 0.08 to 0.71) or VMA ( $R^2$  of 0.05 to 0.76). It was impossible empirically to recover the tight linear relationship of TL from the loose linear relationships of DMA and VMA. TL must have been fundamental, and either DMA or VMA could have been derived from the combination of TL with the other.

In the present study, the  $R^2$ -values of TL (Fig. 1), DMA (Fig. 3), and VMA (Fig. 4) were similar (Supplementary material Appendix 2 Table A6). For subplots in plots, the ranges of  $R^2$  for TL, DMA and VMA all included 0.69 to 0.76 and there was no consistent ordering of  $R^2$  among TL,

DMA and VMA. For plots in blocks, by contrast, in all ten censuses, DMA was a tighter relationship (had a higher  $R^2$ ) than TL, and TL had a higher  $R^2$  than VMA. These results suggest that, for New Zealand mountain beech forests at the spatial scale of plots in blocks, TL and DMA determined VMA.

In summary of these three studies, except possibly at a small spatial scale, VMA appears to be derived from TL and DMA.

### Question 2. Random sampling as a model for TL

Random sampling from a single underlying probability distribution failed to explain the parameters of TL at a fine spatial scale ( $5 \times 5$  m subplots of plots) and largely succeeded at a larger scale ( $20 \times 20$  m plots in blocks) (Supplementary material Appendix 2 Table A5). Unfortunately, our study confounded the spatial scale, the number of observations, and the range of values of log mean density. We have 250 plots (for most censuses) but only 13 blocks, while mean density varied over a much wider range among subplots in plots (typically 1.5 orders of magnitude) than among plots in blocks (half an order of magnitude or less).

Taylor et al. (1988) recommended that the range of values of log mean should be as wide as possible. A future experimental design that controlled spatial scale, number of observations, and the range of values of log mean density would help to understand why the random-sampling model failed at a fine spatial scale and succeeded for most censuses at a large spatial scale.

### Question 3. Spatial scale

Spatial scale strongly influenced the parameters, but not the form, of Taylor's law (Fig. 1, 2). The slope of TL was notably smaller at a small spatial scale than at a large spatial scale. A broad empirical consequence is that empirical tests of Taylor's law and of models of TL need to specify the spatial scale(s) of observation and should, if possible, include a wide range of spatial scales.

That TL slope depends on spatial scale within a single species contradicts Taylor's (1961, 1986, p. 2) suggestion that the slope is a species-specific measure of aggregation or clustering, independent of spatial scale, with spatial TL slope  $b = 1$  for randomly distributed individuals (as in a Poisson distribution) and  $b > 1$  for clustered individuals. While a Poisson spatial distribution implies  $b = 1$ , the converse is false. Taylor's mistaken suggestion rests on a misunderstanding of the TL slope:  $b$  is simply the elasticity of the variance of population density with respect to the mean population density. The elasticity is expressed mathematically as  $d \log(\text{variance}) / d \log(\text{mean})$  (Engen et al. 2008). Intuitively, if the mean population density increases by 1%, the variance of population density increases by approximately  $b$  percent, according to TL. The size of  $b$  may be determined by many processes in addition to, or other than, clustering.

Hanski (1987) simulated a numerical model of spatially (not temporally) correlated local populations using a multivariate lognormal distribution. Numerically, the lower the spatial cross-correlation between local populations, the higher the spatial TL slope: high correlations among local

populations produced spatial TL slopes less than 2, and low correlations produced spatial TL slopes greater than 2. This prediction of Hanski's (1987) model would be consonant with our results if there were higher spatial cross-correlations in density among  $5 \times 5$  m subplots in  $20 \times 20$  m plots than among  $20 \times 20$  m plots in blocks. This supposition is testable but not yet tested. The plots in blocks were separated by  $\geq 200$  m and up to several km, and plots in blocks had higher TL slopes than subplots in plots.

Models of Engen et al. (2008) and Cohen and Xu (2015) were less successful than Hanski's (1987), though they suggested partial interpretations of the differences in spatial TL slope at different spatial scales. Engen et al. (2008) investigated theoretically a model that predicted that, for the slope to exceed 2, "we must have highly regular spatial patterns that produce autocorrelation functions that increase substantially in certain intervals", such as "a chessboard pattern with different constant densities in black and white squares". Even this extreme and unrealistic example gave a slope that exceeded 2 only transiently and, in their numerical example, only slightly (to  $\sim 2.03$ ). This model did not account for our observation of a spatial TL slope that was always greater than 3.39 for  $20 \times 20$  m plots in blocks. By contrast, this model predicted a spatial TL slope between 1 and 2, as we observed at a fine spatial scale. The detailed assumptions of Engen et al. (2008) remain untested with our data. A random sampling model of Taylor's law (Cohen and Xu 2015) failed to predict the slope at a fine spatial scale and succeeded at a large spatial scale, but this difference could be due to different numbers of observations.

Different models of Taylor's law may have value at different spatial scales. A broad theoretical consequence of this suggestion is that it is desirable to test multiple models of Taylor's law.

The combination of demographic stochasticity and environmental spatial variation may help account for the difference in slopes of TL at different spatial scales (cf. Hanski 1982, 1987). Demographic stochasticity probably played a more important role in  $5 \times 5$  m subplots than in  $20 \times 20$  m plots. For example, in these same study plots, Allen et al. (1999, pp. 711–712) found that the distribution of the intensity of damage from a severe earthquake and consequent landslides and rockslides depended on the spatial scale of the units of observation. After the earthquake, 100% of trees were killed on 21% of  $5 \times 5$  m subplots, on 15% of  $10 \times 10$  m subplots (unions of four subplots each of size  $5 \times 5$  m) and only 7% of  $20 \times 20$  m plots. The greater frequency of 100% tree mortality at smaller spatial scales was accompanied by a greater frequency of subplots with no mortality at all, and a smaller proportion of subplots with partial mortality (more than zero mortality but less than total mortality). Thus the variance of mortality following an earthquake was greater in smaller spatial units. These observations suggest that the partial stem mortality in a plot usually arose from complete destruction of trees in some of its subplots and little mortality in other of its subplots, rather than from a uniform partial mortality across each of its subplots. That mortality in a plot is the average of the mortality across its subplots (and likewise for recruitment) might account for our finding that mean density varied much less widely among plots in blocks than among subplots in plots. By contrast, spatial

variation in durable or slowly changing attributes like elevation and soil chemistry could dominate spatial demographic stochasticity for  $20 \times 20$  m plots in blocks, perhaps driving an increase in the variance of density. This qualitative suggestion remains to be made quantitative.

Spatial scale also affected DMA and VMA. In forests, the self-thinning power law (Lonsdale 1990) holds when density-dependent mortality of trees reduces mean population density while mean plant mass (AGB) increases. The self-thinning power law is mathematically equivalent to DMA. In self-thinning forests, if younger stands with higher stem densities were scattered among many patches of older stands with lower density, one might expect apparent clustering of trees and  $b > 1$ . Heterogeneity in density among patches in recruiting forests would equally justify  $b > 1$  (as we observed). These qualitative suggestions remain to be made quantitative. The slope of DMA for subplots in plots was less negative than the slope of DMA for plots in blocks in every census. However, the difference in the DMA slope between plots and blocks was not significant in any census, largely because the CIs for blocks were much wider than the CIs for plots.

Our observations appear to be the first empirical demonstration that the slope of VMA depends on the spatial scale of observation. The slope of VMA for subplots in plots was significantly less negative than the slope for plots in blocks (Supplementary Appendix 2 Table A6) in every census. This difference was expected from the theoretical demonstration that the slope of VMA is the product of the slopes of TL and DMA.

#### Question 4. Temporal trends and effects of natural disturbance

Censuses 1–5 (1974–1983) spanned a disturbed phase, censuses 5–8 (1983–1993) a recovery phase, and censuses 8–10 (1993–2004) another disturbed phase (Wardle and Allen 1983, Hurst et al. 2011). Yet the form and the parameters of the three power laws (Supplementary material Appendix 2 Table A6) were at most weakly associated with these levels of disturbance. The wide confidence intervals of the slopes of the spatial TL (Fig. 2) almost masked temporal trends in the slope, but at both spatial scales, the slope  $b$  of TL increased with the calendar year of the census (significantly when using subplots in plots, not significantly when using plots in blocks) (Supplementary material Appendix 2 Table A4). The slope of DMA increased over time in both plots and blocks, but the magnitude of these effects was small and sometimes not significant (Supplementary material Appendix 2 Table A7). There was no significant trend in the slope of VMA as a function of calendar year for plots in blocks, and a marginally significant decrease in the slope of VMA for subplots in plots (Supplementary material Appendix 2 Table A10). Since the slope  $k$  of VMA is, in theory, the product of the slopes  $b$  of TL and  $v$  of DMA ( $k = bv$ ), if  $b$  increased and  $v$  changed little or not at all, one would expect  $k$  to decrease (since  $v < 0$ ).

#### Question 5. Self-thinning and recruitment

The slope of TL was not significantly influenced by whether plots were self-thinning, recruiting, or other, though the

intercept of TL may have been influenced by these attributes. The slopes and intercepts of DMA and VMA were influenced by whether plots were self-thinning, recruiting, or other. In general, for both DMA and VMA, self-thinning plots had higher intercepts and lower slopes, and recruiting plots had lower intercepts and higher slopes, than other plots.

#### Question 6. Elevation and soil chemistry

Neither the elevation above sea level nor the soil carbon-to-nitrogen ratio significantly affected the slope of TL in any census, with two marginal exceptions. Soil calcium concentration was associated with a greater variance of population density, increasing the intercept of TL in all 10 censuses, significantly so in seven. Forests on lower slopes have higher available soil calcium (Wiser et al. 1998) and higher biomass mortality rates (Harcombe et al. 1998). If such mortality is more pronounced at small spatial scales (Allen et al. 1999), then one might expect higher variance of population density, as observed.

The influence of calcium concentration on the intercept increased with time. The distinctly limited, but nonzero, effect of environmental variables on TL in these New Zealand forests of mountain beech is consistent with the limited, but nonzero, effects of environmental variables on laboratory populations of the bacterium *Pseudomonas fluorescens* (Kaltz et al. 2012) and on Japanese coastal populations of intertidal barnacles, *Chthamalus* spp. (Fukaya et al. 2013, 2014).

#### Questions for future research: spatial autocorrelation and temporal Taylor's law

To investigate the impact of synchrony (spatial correlation in population density) on Taylor's law and the other power laws, and to test the details of Hanski's (1987) numerical model, it would be desirable to examine the spatial correlations of population density among subplots and plots as a function of the distance between them, within and between transects and blocks. The distance between a pair of subplots or plots could be measured spatially (as the distance on the ground) or conceptually in terms of differences in quantitative attributes. For example, the distance between two plots could be measured by differences in elevation and calcium concentration.

It would also be desirable to test the temporal Taylor's law at distinct spatial scales (subplot, plot and block) and dynamic stochastic population models that lead to TL (Cohen et al. 2013, Cohen 2014).

#### Conclusions

From the apparent jumble of untended New Zealand mountain beech forests subjected to wind throw, earthquakes and landslides has emerged surprising order. Taylor's law, density–mass allometry, and variance–mass allometry are valid in these forests on a much larger range of spatial scales, over a much longer time span, over a wider range of environmental conditions, than previously demonstrated. Random sampling from a skewed distribution can account for the form and slope of Taylor's law at a large spatial scale (among plots in blocks), but not at a fine spatial scale

(among subplots in plots). The parameters of Taylor's law and variance–mass allometry, but not density–mass allometry, are strongly affected by the spatial scale of observation. Spatial scale was known to affect the slope of Taylor's law. Here we show for the first time that spatial scale also affects the slope of variance–mass allometry. The demographic processes of self-thinning and recruitment and the calcium concentration in the soil were important influences on the parameters, but not the form, of some of these power laws. Where parameters depend on environmental conditions or scale of observation, the parameters cannot be species-specific characteristics.

*Acknowledgements* – J. E. Cohen and J. Lai contributed equally to this article. JEC acknowledges U.S. National Science Foundation grant DMS-1225529 and the assistance of Priscilla K. Rogerson. RBA was supported by Landcare Research. This project benefited from many years of input by staff of the former New Zealand Forest Service, Forest and Range Experiment Station, Forest Research Institute and currently Landcare Research. We thank Aaron McGlinchy of Landcare Research for his help in making our data publicly accessible.

## References

- Allen, R. B. et al. 1999. Immediate damage by an earthquake to a temperate montane forest. – *Ecology* 80: 708–714.
- Borcard, D. et al. 2011. Numerical ecology with R. – Springer.
- Clinton, P. W. et al. 2002. Nitrogen storage and availability during stand development in a New Zealand *Nothofagus* forest. – *Can. J. For. Res.* 32: 344–352.
- Cohen, J. E. 2014. Stochastic population dynamics in a Markovian environment implies Taylor's power law of fluctuation scaling. – *Theor. Popul. Biol.* 93: 30–37.
- Cohen, J. E. and Xu, M. 2015. Random sampling of skewed distributions implies Taylor's power law of fluctuation scaling. – *Proc. Natl Acad. Sci. USA* 112: 7749–7754.
- Cohen, J. E. et al. 2012. Allometric scaling of population variance with mean body size is predicted from Taylor's law and density–mass allometry. – *Proc. Natl Acad. Sci. USA* 109: 15829–15834.
- Cohen, J. E. et al. 2013. Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. – *Proc. R. Soc. B* 280: 20122955.
- Coomes, D. A. and Allen, R. B. 2007a. Effects of size, competition and altitude on tree growth. – *J. Ecol.* 95: 1084–1097.
- Coomes, D. A. and Allen, R. B. 2007b. Mortality and tree-size distributions in natural mixed-age forests. – *J. Ecol.* 95: 27–40.
- Coomes, D. A. et al. 2012. A general integrative framework for modelling woody biomass production and carbon sequestration rates in forests. – *J. Ecol.* 100: 42–64.
- Davis, M. R. et al. 2004. The influence of N addition on nutrient content, leaf carbon isotope ratio, and productivity in a *Nothofagus* forest during stand development. – *Can. J. For. Res.* 34: 2037–2048.
- Eisler, Z. et al. 2008. Fluctuation scaling in complex systems: Taylor's law and beyond. – *Adv. Phys.* 57: 89–142.
- Engen, S. et al. 2008. A general model for analyzing Taylor's spatial scaling laws. – *Ecology* 89: 2612–2622.
- Fukaya, K. et al. 2013. Variable processes that determine population growth and an invariant mean–variance relationship of intertidal barnacles. – *Ecosphere* (4, Article 48): 1–20. <<http://dx.doi.org/10.1890/ES12-00272.1>>.
- Fukaya, K. et al. 2014. Effects of spatial structure of population size on the population dynamics of barnacles across their elevational range. – *J. Anim. Ecol.* 83: 1334–1343.
- Gordon, A. D. 1999. Classification. – Chapman and Hall/CRC.
- Hanski, I. 1980. Spatial patterns and movements in coprophagous beetles. – *Oikos* 34: 293–310.
- Hanski, I. 1981. Coexistence of competitors in patchy environment with and without predation. – *Oikos* 37: 306–312.
- Hanski, I. 1982. On patterns of temporal and spatial variation in animal populations. – *Ann. Zool. Fenn.* 19: 21–37.
- Hanski, I. 1987. Cross-correlation in population dynamics and the slope of spatial variance–mean regressions. – *Oikos* 50: 148–151.
- Harcombe, P. A. et al. 1998. Spatial and temporal patterns in stand structure, biomass, growth and mortality in a monospecific *Nothofagus solandri* var. *cliffortioides* (Hook, f.) Poole forest in New Zealand. – *J. Sust. For.* 6: 313–345.
- Hurst, J. M. and Allen, R. B. 2007. A permanent plot method for monitoring indigenous forests – field protocols. – Manaaki Whenua, Landcare Research.
- Hurst, J. M. et al. 2011. Size-specific tree mortality varies with neighbourhood crowding and disturbance in a montane *Nothofagus* forest. – *PloS ONE* 6: e26670.
- Kaltz, O. et al. 2012. Bacterial microcosms obey Taylor's law: effects of abiotic and biotic stress and genetics on mean and variance of population density. – *Ecol. Process.* 1: 1–6.
- Kendal, W. S. 2004. Taylor's ecological power law as a consequence of scale invariant exponential dispersion models. – *Ecol. Complex.* 1: 193–209.
- Kendal, W. S. 2013. Fluctuation scaling and 1/f noise: shared origins from the Tweedie family of statistical distributions. – *J. Basic Appl. Phys.* 2: 40–49.
- Kendal, W. S. and Jørgensen, B. 2011a. Taylor's power law and fluctuation scaling explained by a central-limit-like convergence. – *Phys. Rev. E* 83: 066115.
- Kendal, W. S. and Jørgensen, B. 2011b. Tweedie convergence: a mathematical basis for Taylor's power law, 1/f noise and multifractality. – *Phys. Rev. E* 84: 066120.
- Kilpatrick, A. M. and Ives, A. R. 2003. Species interactions can explain Taylor's power law for ecological time series. – *Nature* 422: 65–68.
- Lagroe, C. et al. 2015. Parasitism alters three power laws of scaling in a metazoan community: Taylor's law, density–mass allometry and variance–mass allometry. – *Proc. Natl Acad. Sci. USA* 112: 1791–1796.
- Lai, J. S. et al. 2013. The allometry of coarse root biomass: log-transformed linear regression or nonlinear regression? – *PloS ONE* 8: e77007.
- Lonsdale, W. M. 1990. The self-thinning rule: dead or alive? – *Ecology* 71: 1373–1388.
- Marquet, P. A. et al. 2005. Scaling and power-laws in ecological systems. – *J. Exp. Biol.* 208: 1749–1769.
- McArdle, B. H. 1988. The structural relationship - regression in biology. – *Can. J. Zool.* 66: 2329–2339.
- McArdle, B. H. et al. 1990. Variation in the size of animal populations – patterns, problems and artifacts. – *J. Anim. Ecol.* 59: 439–454.
- Osawa, A. and Allen, R. B. 1993. Allometric theory explains self-thinning relationships of mountain beech and red pine. – *Ecology* 74: 1020–1032.
- Platt, K. H. et al. 2004. Mountain beech seedling responses to removal of below-ground competition and fertiliser addition. – *N. Z. J. Ecol.* 28: 289–293.
- Rousseeuw, P. J. 1987. Silhouettes – a graphical aid to the interpretation and validation of cluster-analysis. – *J. Comput. Appl. Math.* 20: 53–65.
- Sawyer, A. J. 1989. Inconstancy of Taylor-*b* – simulated sampling with different quadrat sizes and spatial distributions. – *Res. Popul. Ecol.* 31: 11–24.

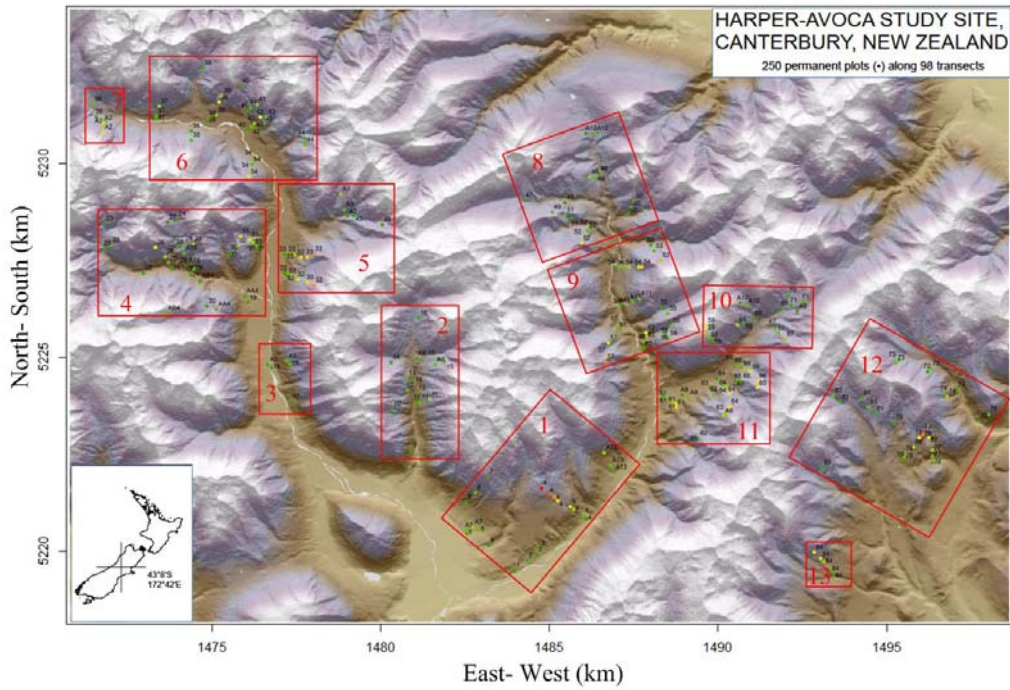
- Smaill, S. J. et al. 2011. Climate cues and resources interact to determine seed production by a masting species. – *J. Ecol.* 99: 870–877.
- Smith, R. J. 2009. Use and misuse of the reduced major axis for line-fitting. – *Am. J. Phys. Anthropol.* 140: 476–486.
- Taylor, L. R. 1961. Aggregation, variance and the mean. – *Nature* 189: 732–735.
- Taylor, L. R. 1984. Assessing and interpreting the spatial distributions of insect populations. – *Annu. Rev. Entomol.* 29: 321–357.
- Taylor, L. R. 1986. Synoptic dynamics, migration and the Rothamsted insect survey: Presidential Address to the British Ecological Society, December 1984. – *J. Anim. Ecol.* 55: 1–38.
- Taylor, L. R. et al. 1978. The density-dependence of spatial behaviour and the rarity of randomness. – *J. Anim. Ecol.* 47: 383–406.
- Taylor, L. R. et al. 1988. Specificity of the spatial power-law exponent in ecology and agriculture. – *Nature* 332: 721–722.
- Wardle, J. A. and Allen, R. B. 1983. Dieback in New Zealand *Nothofagus* forests. – *Pac. Sci.* 37: 397–404.
- West, G. B. et al. 2009. A general quantitative theory of forest structure and dynamics. – *Proc. Natl Acad. Sci. USA* 106: 7040–7045.
- Wiser, S. K. et al. 1998. Community structure and forest invasion by an exotic herb over 23 years. – *Ecology* 79: 2071–2081.
- Yamamura, K. 1990. Sampling scale dependence of Taylor power law. – *Oikos* 59: 121–125.

Supplementary material (available online as Appendix oik-02622 at <[www.oikosjournal.org/appendix/oik-02622](http://www.oikosjournal.org/appendix/oik-02622)>). Appendix 1–2.



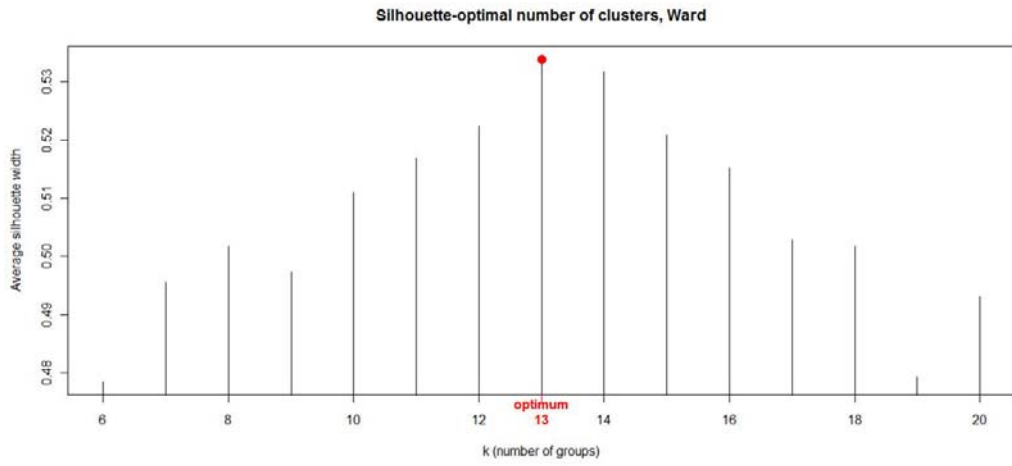
1 **SUPPLEMENTARY FIGURES**

2 Figure S1. 250 plots were assigned to 13 blocks so that each block contained 5 to 33  
3 plots (see Supplementary Appendix A2 Table S2), as recommended by a  
4 silhouette-optimal rule (Figure S2) for Ward's minimum variance clustering based on  
5 geographic distance. Each point represents one plot. Each red frame represents one  
6 block and red number (1-13) is the order of blocks. The area of forest differs among  
7 blocks. The plot co-ordinates are in kilometers according to the New Zealand  
8 Transverse Mercator Projection.



9  
10

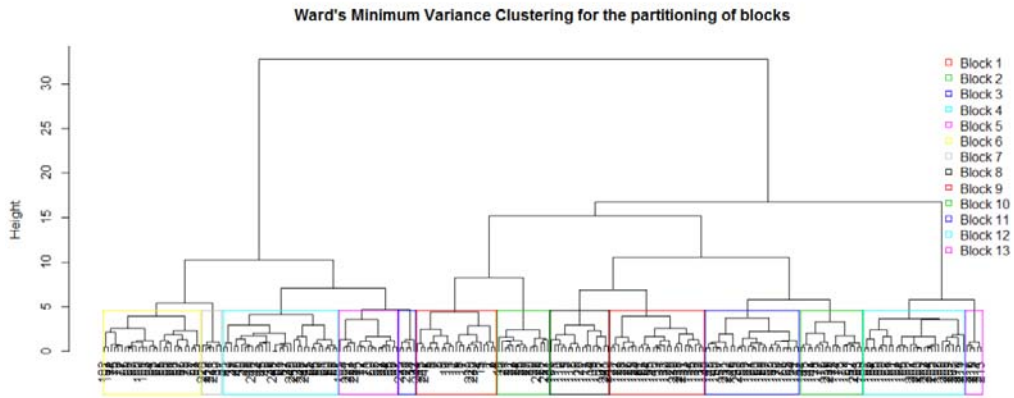
11 Figure S2. Silhouette-optimal rule for Ward's minimum variance clustering based on  
12 geographic distance was used to create blocks. The greater the value of average  
13 silhouette width, the better the plot is clustered.



14

15

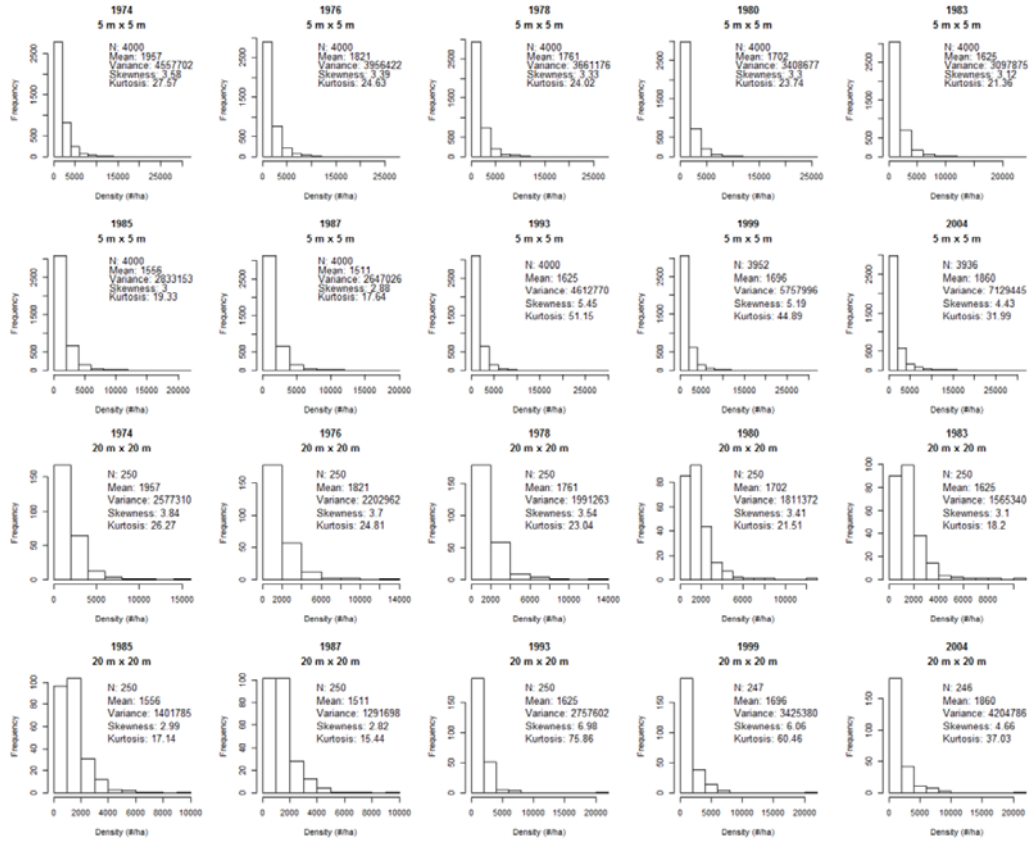
16 Figure S3. Dendrogram with boxes around the 13 blocks selected for the partitioning  
17 of the 250 plots. In the dendrogram, the y-axis is the distance metric between clusters.  
18 For example, if two clusters merged at a height  $y$ , it means that the distance between  
19 those clusters was  $y$ .



20

21

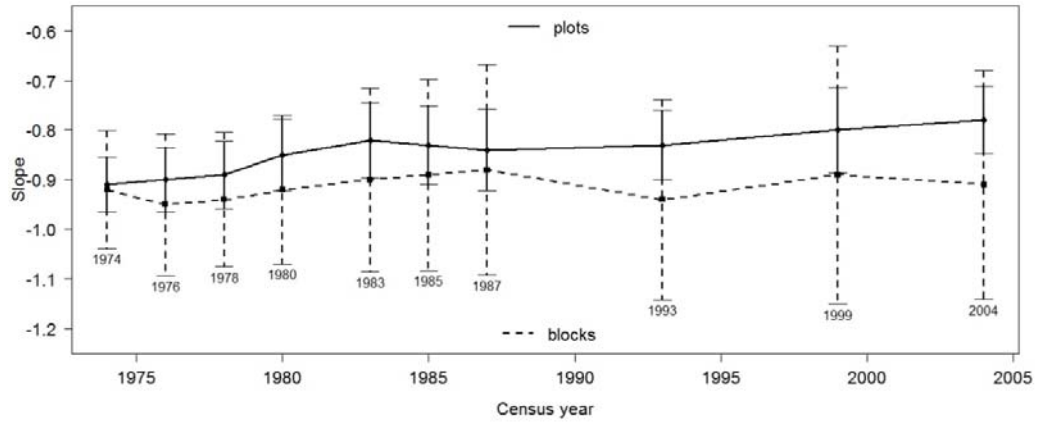
22 Figure S4. Frequency histograms and sample moments of tree density in 10 censuses  
 23 for subplots and plots. The values for the kurtosis are the standardized fourth moment,  
 24 not the excess kurtosis.



25

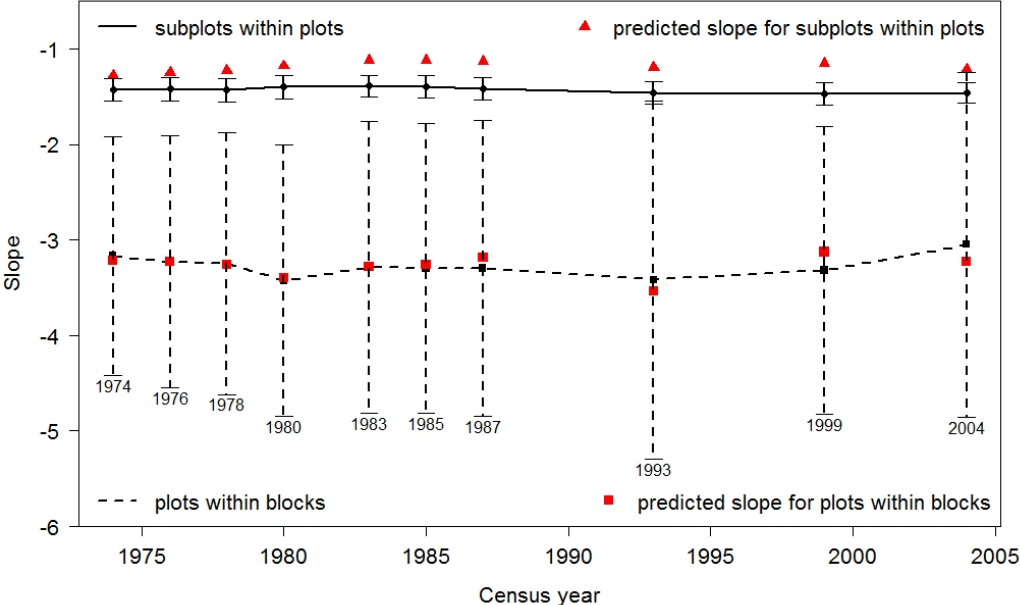
26

27 Figure S5. Slopes of density-mass allometry in each census, for 250 plots (for most  
28 censuses) (solid line), and 13 blocks (dashed line). Error bars show the 95%  
29 confidence interval of the slope.



30  
31

32 Figure S6. Slopes of variance-mass allometry and their predicted values from Taylor's  
33 law and density-mass allometry in each census, for 5 m × 5 m subplots in plots (solid  
34 line), and 20 m × 20 m plots in blocks (dashed line). Error bars show the 95%  
35 confidence interval of the slope.



36