

LEAF SIZE AND SHAPE AS A PALEOCLIMATE PROXY: INVESTIGATING TAPHONOMIC BIASES WITHIN SPECIES

MARCUS A. BROWN,¹ DANA L. ROYER,¹ VALERIE NAZZARO,² AND LAUREN AZEVEDO-SCHMIDT³

¹Wesleyan University, Department of Earth and Environmental Sciences, Middletown, Connecticut, 06457, USA

²Wesleyan University, Quantitative Analysis Center, Middletown, Connecticut, 06457, USA

³University of California, Davis, Department of Entomology and Nematology, Davis, California, 95616, USA
email: marcusb4@illinois.edu

ABSTRACT: For over a century, leaf size and shape have been used to reconstruct aspects of climate related to temperature and precipitation. These methods are typically calibrated with complete leaves from living forests. An understudied issue is how appropriate these calibrations are to incomplete leaf compression fossils that dominate most fossil leaf assemblages. We sought to test the fidelity of leaf-climate methods with leaves sampled from three taphonomic stages (fresh, surface litter, and buried litter) in two temperate forests. We sampled environments that commonly contain fossil leaves in the geologic record: swamp and river margins. We find that in comparison to fresh leaves, conspecific leaves from the surface litter and buried litter are progressively smaller. As a result, paleoprecipitation estimates using leaf size and shape may be underestimated by as much as 23%. In contrast, traits common in temperature models (e.g., size and number of leaf teeth) show little within-species taphonomic biasing.

INTRODUCTION

The fossil record is a biased storyteller. Because plants do not move and are composed of easily decaying tissues, their preservation as fossils is strongly dictated by taphonomic filters related to transport, decay, and burial (Fig. 1). There is a rich history of work on the topic of leaf taphonomy (e.g., Spicer 1981, 1989; Ferguson 1985, 2005; Burnham 1989; Greenwood 1992). For example, small leaves may be expected to be over-represented in the fossil record because of the greater likelihood of large leaves to break apart (Hagen et al. 2019) and the greater likelihood of sun leaves—which tend to be small (Wylie 1951; Talbert and Holch 1957; Sack et al. 2006)—to reach sites prone to fossilization because sun leaves often travel farther (Spicer 1981; Ferguson 1985) and are more durable (higher leaf mass per area; Hanson 1917; Koch et al. 2004; Sack et al. 2006) than shade leaves.

Evidence for size biasing is mixed (Ferguson 1985): leaves recovered from lake sediments are smaller than their conspecifics in the nearby canopy (Roth and Dilcher 1978) but not smaller relative to nearby surface litter (Spicer 1981; Astorga et al. 2016). Leaves in stream sediment are consistently smaller than leaves from the surrounding canopy (Burnham 1989; Greenwood 1992; Steart et al. 2002; Ellis and Johnson 2013) and forest surface litter (Christophel and Greenwood 1989). Burnham (1989), however, observed no size bias between levee sediment (including back levee and forebank) and the surrounding canopy. These studies all capture the early stages of taphonomy (top of Fig. 1) and associated changes in leaf area within and/or across species (i.e., loss of species in litter). Using scaling based on second-order vein density to investigate later taphonomic stages (bottom of Fig. 1), Hagen et al. (2019) reconstructed ~10% larger leaves in fragmented fossils from the Williston Basin versus complete conspecifics from the same sites.

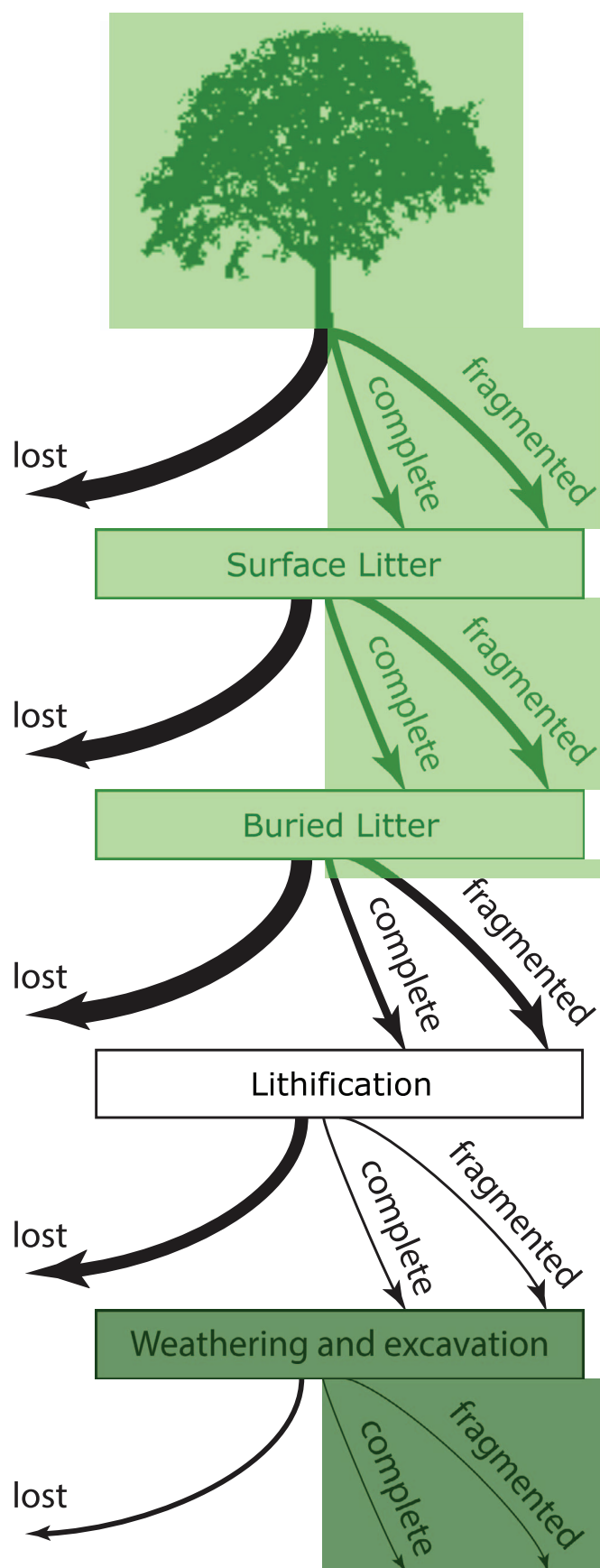
Compared to leaf size, the potential taphonomic biasing of leaf shape is less studied. A taphonomic understanding of size and shape (physiognomy) are both important because physiognomy varies systematically with

climate, especially among woody non-monocot angiosperms (“woody dicots”), making it a valuable tool for reconstructing past environments (Bailey and Sinnott 1915, 1916; Wolfe 1979, 1993; Wilf 1997; Wilf et al. 1998; Peppe et al. 2011). Larger leaves are generally found in wetter climates, while smaller leaves dominate dryer regions (Raunkiaer 1934; Givnish 1984; Wilf et al. 1998; Wright et al. 2017). Leaf margin traits, particularly the presence versus absence of teeth, are closely tied to temperature: plants in colder climates tend to have more toothed margins, while those in warmer regions predominantly have smooth-edged leaves (Wolfe 1979; Peppe et al. 2011).

Paleobotanists have developed multiple proxies for inferring climate from site-mean leaf physiognomy among woody dicots. These include bivariate correlations between mean annual precipitation (MAP) and leaf area (leaf-area analysis; Wilf et al. 1998) and between mean annual temperature (MAT) and the percentage of species with untoothed leaf margins (Wolfe 1979; Wilf 1997). Multivariate methods have been developed too, including the Climate Leaf.

Analysis Multivariate Program (CLAMP; Wolfe 1993; Spicer et al. 2021) and Digital Leaf Physiognomy (DiLP; Peppe et al. 2011). These multivariate techniques focus on both leaf size and aspects of shape, especially related to leaf teeth (their size, abundance, and influence on outline complexity). Critically, all these methods are calibrated to measurements of leaves sampled directly from the canopy; they do not account for taphonomy.

Prior work has explored how taphonomy may affect leaf-climate methods, principally through either the digital degradation of individual traits in canopy leaves (Royer et al. 2005; Spicer et al. 2005, 2011) or climate reconstruction from univariate techniques based on leaves sampled at different taphonomic stages (Burnham 1989; Greenwood 1992, 2005; Ellis and Johnson 2013). These latter studies did not separate the taphonomic effects on leaf physiognomy due to changes within species versus loss of species in litter and sediment. Thus, in order to better understand the extent of taphonomic biasing on leaf physiognomy solely within species,



we sought to measure the full set of physiognomic traits in DiLP on leaves from three taphonomic stages (fresh leaves, surface litter, buried litter) in three depositional settings at two forests in eastern North America. This study design builds on prior work by quantitatively tracking the taphonomic effects on climate reconstruction due to changes in both leaf size and shape within individual species at multiple depositional settings, sites, and taphonomic stages (light shaded area in Fig. 1). This allows for a detailed view into how taphonomy affects the leaf physiognomy within individual species and complements the fossil-only study of Hagen et al. (2019) (dark shaded area in Fig. 1).

METHODS

Sites

We collected leaves at two sites: Harvard Forest, located near Peterham, Massachusetts, and the Smithsonian Environmental Research Center (SERC), located on the western coast of Maryland near the Chesapeake Bay. Harvard Forest is one of North America's oldest managed forests. It is a humid, cool-temperate forest characterized by relatively low plant diversity. The dominant species include eastern hemlock (*Tsuga canadensis*) and red oak (*Quercus rubra*), with a dense understory (Au et al. 2020). SERC represents a humid, warm-temperate ecosystem influenced by a mixture of freshwater and brackish environments. SERC is comparatively diverse, consisting of 78 species with two distinct forest associations (Driscoll 2016): the upland 'tulip poplar' forest and the wet lowland 'river birch-sycamore' forest (Driscoll et al. 2016).

At both sites, we targeted depositional environments that—in the sedimentary record—commonly contain fossil leaves: swamps, small river tributaries, and larger dynamic rivers (Wing 1988). During the summer of 2019, Azevedo-Schmidt et al. (2022) collected shallowly buried leaf litter in each of these depositional environments ("quarries"). At each quarry, they sampled litter at three or four discrete subquarries. The collected leaves were pressed flat using a leaf press, dried for three days at 50°C, and photographed with labels and a ruler for scale. They analyzed the following five species at Harvard Forest: American beech (*Fagus grandifolia*), black birch (*Betula lenta*), yellow birch (*Betula alleghaniensis*), red maple (*Acer rubrum*), and red oak (*Quercus rubra*); and seven species at SERC: American beech, American hornbeam (*Carpinus caroliniana*), American sycamore (*Platanus occidentalis*), chestnut oak (*Quercus prinus*), southern red oak (*Quercus falcata*), white oak (*Quercus alba*), and sweetgum (*Liquidambar styraciflua*). We found differentiating black and yellow birches at Harvard Forest difficult based on leaf morphology alone, so we combined them for our analyses.

In June 2023, we revisited Azevedo-Schmidt's sites and sampled fresh leaves from the canopy near one subquarry per quarry. We targeted many of the same species that were abundant in Azevedo-Schmidt et al. (2022) but added willow oak (*Quercus phellos*) at SERC. For each species, we aimed to collect leaves from both sunlit and shaded areas of two individuals using a pole pruner with a maximum extension of ~8 m. When this was not possible, we sometimes collected fresh, green leaves from the forest floor. Because we could not access leaves from the top of the canopy, we targeted trees with light gaps to collect "sun" leaves; we acknowledge that these leaves may not be true sun morphotypes.

FIG. 1.—Some taphonomic processes related to the fragmentation of leaves. Most leaves are never captured as fossils, and many fossils are not captured in their complete forms. Our study captures some of the initial taphonomic processes (light shaded area at top); Hagen et al. (2019) captured some of the final stages (dark shaded area at bottom). Arrow thickness broadly reflects leaf fraction. Figure adapted from Hagen et al. (2019).

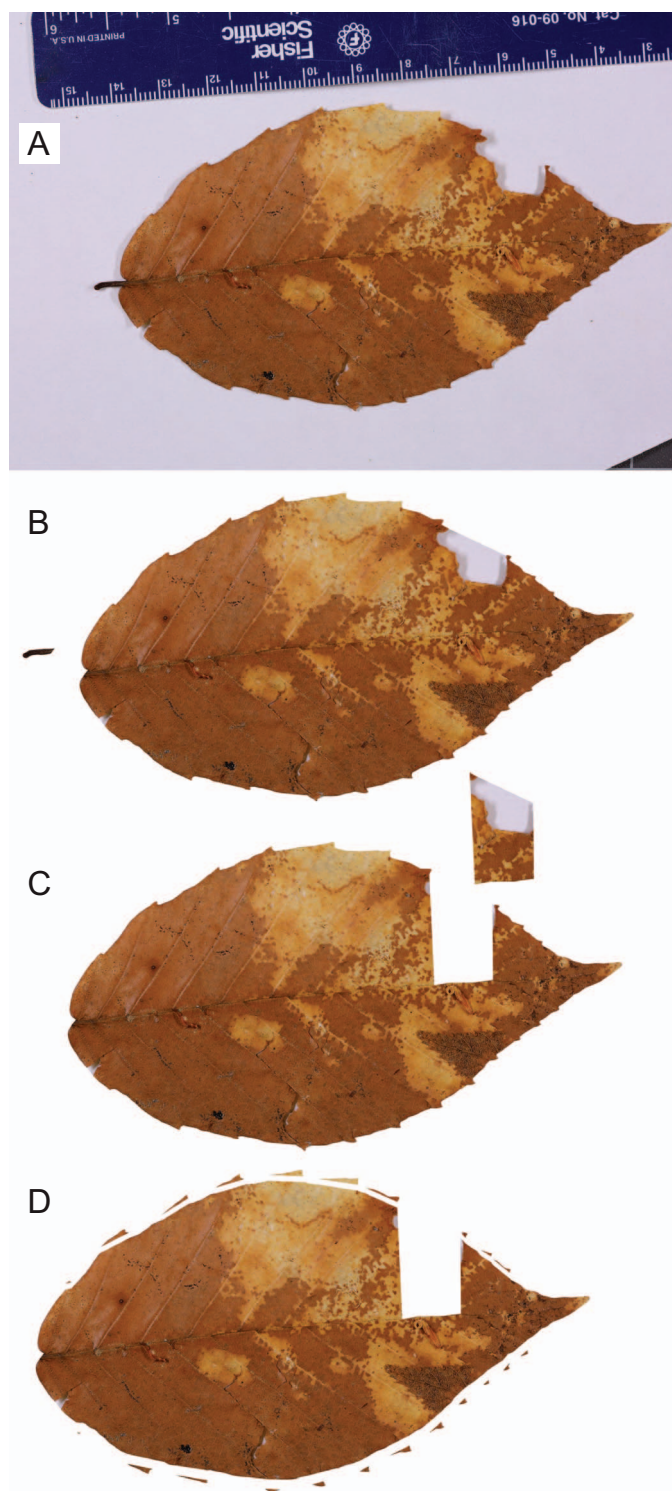


FIG. 2.—Steps in preparing a leaf image for analysis using Digital Leaf Physiognomy (DiLP). **A)** Photograph the leaf (beech leaf #350 from Harvard Forest). **B)** Separate the petiole and reconstruct damaged portions of the leaf margin. Petiole and inferred blade area are measured at this step. **C)** Remove sections that contain damage along the margin. Raw blade area and perimeter are measured at this step; perimeter associated with the leaf interior is excluded. **D)** Separate teeth from the leaf margin. Internal raw blade area and perimeter are measured at this step; perimeter associated with the leaf interior is excluded.

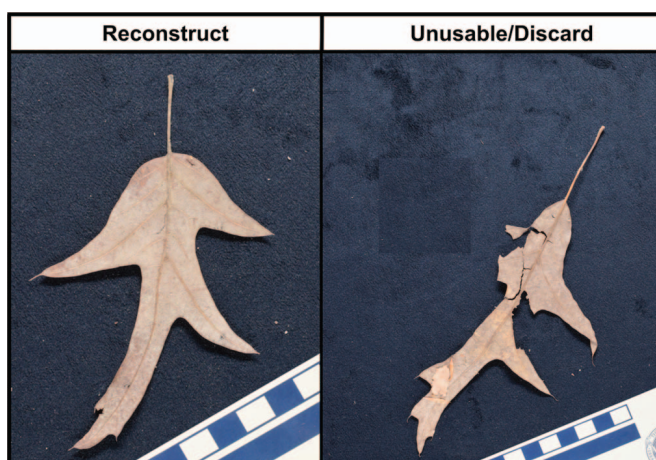


FIG. 3.—Comparison of a reconstructable untoothed leaf (Southern red oak leaf #55 from SERC) with one that should be excluded (Southern red oak leaf #51 from SERC). Using Digital Leaf Physiognomy for untoothed leaves requires that the complete margin be intact or reconstructed.

In August 2023, we returned to the same subquarries at Harvard Forest and sampled leaves from the surface of the leaf litter and similarly in June 2024 at SERC. We targeted species that had been previously collected in buried litter or the canopy. At each location, we filled three gallon-sized plastic bags. The first bag contained a general scoop of existing surface litter within an ~ 0.5 m² area; the remaining bags were filled with species that were underrepresented in the first bag. We dried, pressed, and photographed all leaves following Azevedo-Schmidt et al. (2022) (Figs. 2A, 3). Not all leaf categories (fresh, surface litter, buried litter) were collected for each species-site combination. We assume that any shrinkage within species due to drying (e.g., Blonder et al. 2012) affected the leaf categories (fresh leaves/surface litter/buried litter) proportionately.

Digital Leaf Physiognomy

Digital Leaf Physiognomy (DiLP) is a multivariate, leaf physiognomy-based approach that quantitatively analyzes leaf size and shape to infer past climatic conditions (Peppe et al. 2011). This model relies on quantitative physiognomic measurements, allowing for direct comparisons between living and fossil leaves to assess environmental parameters. DiLP focuses on climatically sensitive variables related to overall size, degree of lobateness or dissection, tooth abundance, and tooth size. Because almost all fossil leaves are fragmented to some degree, the variables in DiLP are designed to be measurable on leaves whose outlines are not completely present, for example tooth count and tooth area normalized to area and perimeter (Table 1; Royer et al. 2005; Peppe et al. 2011).

We processed the leaf images for DiLP following the workflow of Lowe et al. (2024). In short, we used Krita (Krita Foundation, Deventer, Netherlands) to trace each leaf and remove its petiole; for leaves with only minor damage along their margins (Fig. 3), we reconstructed their areas so that we could quantify the original leaf area (“inferred blade area”; Fig. 2B).

For leaves with teeth, damaged sections need to be excluded before calculating variables that include tooth count or perimeter because teeth are likely missing and perimeter cannot be reconstructed as reliably as inferred leaf area; to do this, for each damaged section we excluded all leaf area down to the nearest primary vein (“raw blade area”; Fig. 2C). Following Lowe et al. (2024), we only retained sections of margin containing a minimum of two teeth. Next, teeth were separated from the margins, typically from sinus to sinus (“raw internal blade area”; Fig. 2D), and manually counted. Once prepped, all images were imported into ImageJ (NIH, Bethesda, Maryland, USA) for measurement of the core physiognomic variables: petiole area,

TABLE 1.—Description of the most relevant variables in Digital Leaf Physiognomy (DiLP) (after Royer et al. 2005; Lowe et al. 2024).

Physiognomic variable (abbreviation, units)	Description
*Leaf area (cm ²)	The total area of the leaf, incorporating both the petiole and inferred blade area. Requires leaf to be mostly intact (Fig. 2B).
**Feret diameter ratio (FDR, unitless)	Fraction of the area of a circle with the same diameter as the longest distance across the leaf to the inferred blade area (Fig. 2B).
*Perimeter ratio (unitless) [†]	The ratio of the raw internal blade perimeter to the raw blade perimeter (Fig. 2C, 2D).
Tooth count:blade area (tc:ba, unitless) [†]	The ratio of the total number of intact teeth to the raw blade area (Fig. 2C).
Tooth area:blade area (ta:ba, unitless) [†]	The ratio of the total area of all intact teeth to the raw blade area (Fig. 2C, 2D).
Tooth area:internal perimeter (ta:ip, cm) [†]	The ratio of the total area of all intact teeth to the raw internal blade perimeter (Fig. 2C, 2D).
Average tooth area (avg ta, cm ²) [†]	Average area of a single tooth (Fig. 2C, 2D).
***Tooth count: internal perimeter (tc:ip, cm ⁻¹) [†]	The ratio of the total number of intact teeth to the raw internal blade perimeter (Fig. 2D).

[†] Relevant only for toothed species.
* Used in MAP equation.
** Used in MAT equation.
*** Used in both MAP and MAT equations.

inferred leaf area, and feret diameter (maximum diameter across the leaf) for leaves with only minor damage along their margins (Fig. 3); raw blade area and perimeter; raw blade internal area and perimeter; and the length of cut perimeter (e.g., straight-line cuts in Fig. 2C). See Online Supplemental File Table S1 for all raw data. We then used the dilp package (Butrim et al. 2024) in R (R Core Team 2024) to calculate eight derived variables (Table 1; the subset of variables used in the MAP and MAT models are marked with asterisks). We followed the univariate and multivariate equations in Peppe et al. (2011) to calculate climate from physiognomy.

Data Analysis

To determine the number of leaves per species to analyze in our fresh leaves and surface litter samples, we carried out a sensitivity analysis of the buried litter samples. For every species, we drew 1000 random samples at each level of sampling effort (from one leaf to the full data set) and calculated the medians of each random sample × sampling effort combination; at each level of sampling effort, we then calculated the 2.5 and 97.5 percentile values of the 1000 medians. To standardize scales across variables, we divided the percentile values by their median (50th percentile). We generally found only mild improvements in precision above a sampling effort of 15 leaves (Online Supplemental File Figure S1). We therefore analyzed a minimum of 15 leaves per species in our fresh leaves and surface litter data sets.

We log-transformed two physiognomic variables that showed skewed Q-Q plots (tooth count:internal perimeter and tooth count:blade area). We tested for differences in physiognomy between putative “sun” and “shade” leaves in R using a multivariate linear model: $\text{lm}([\text{physiognomic variable}] \sim [\text{sun or shade}] + [\text{species}] + [\text{quarry}])$. We found no strong patterns; consequently, all analyses of the fresh leaves do not treat sun/shade as a separate factor. A principal component analysis (PCA) identified two strong outlying red oak leaves from Harvard Forest (circled data in Online Supplemental File Figure S2), which we excluded from all analyses. Because the buried litter come from three subquarries per site versus one for the fresh leaves and surface litter, we tested at each quarry whether the physiognomy in the buried litter at the additional subquarries differed from the buried litter at the subquarry sampled for all three leaf types ($\text{lm}([\text{physiognomic variable}] \sim [\text{leaf category}] + [\text{species}] + [\text{quarry} + \text{subquarry combination}])$). These pairwise tests found little support for differences in physiognomy across subquarries within a single quarry: at Harvard Forest, one subquarry comparison differed (1903.4 vs. 1903.3), but in only one of the eight tested variables (perimeter ratio); at SERC, one subquarry comparison (1902.3 vs. 1902.1) differed in two physiognomic variables

(perimeter ratio and tooth area:blade area), and for two other comparisons (1901.3 vs. 1901.2; 1903.3 vs. 1903.2) in one variable (tooth count:blade area and tooth area:internal perimeter, respectively). We therefore included all buried litter in our analyses.

In total, we analyzed 816 leaves from Harvard Forest and 874 from SERC; see Online Supplemental Data File 1 for R code used to run analyses. At each forest, we tested for statistical significance between taphonomic category (fresh leaves/surface litter/buried litter) and leaf physiognomy in R using a multivariate linear model: $\text{lm}([\text{physiognomic variable}] \sim [\text{leaf category}] + [\text{species}] + [\text{quarry}])$. We tested post-hoc the pairwise differences between leaf categories using the fit.contrast function in the gmodels package (Warnes et al. 2024). This allowed us to test for monotonic patterns across categories, for example leaf size declining from fresh leaves to surface litter to buried litter.

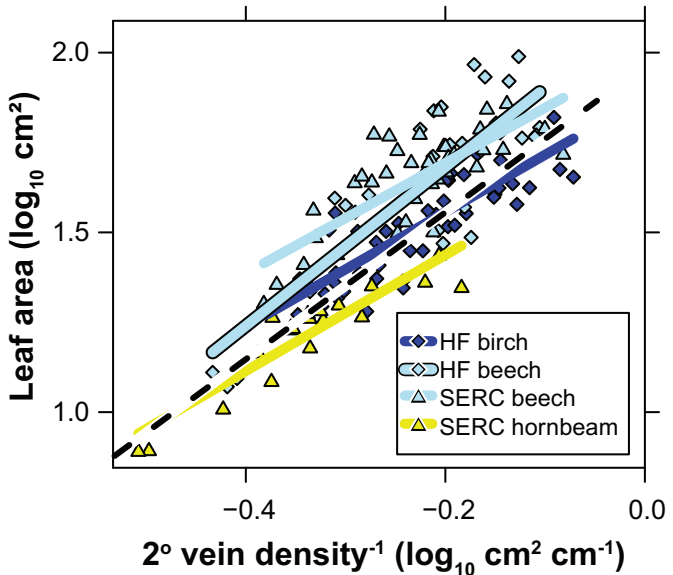


FIG. 4.—Scaling in canopy leaves of four species-site pairs between inverse second-order vein density and leaf area. We used the ordinary least-squares regressions (colored lines) to reconstruct leaf area in the buried litter (Harvard Forest birch: $y = 1.59x + 1.87$, $r^2 = 0.72$, $n = 49$; Harvard Forest beech: $y = 2.21x + 2.12$, $r^2 = 0.70$, $n = 30$; SERC beech: $y = 1.53x + 2.00$, $r^2 = 0.58$, $n = 45$; SERC hornbeam: $y = 1.61x + 1.76$, $r^2 = 0.84$, $n = 20$). For comparison, the black dashed line is the regression based on 485 dicot species from Sack et al. (2012).

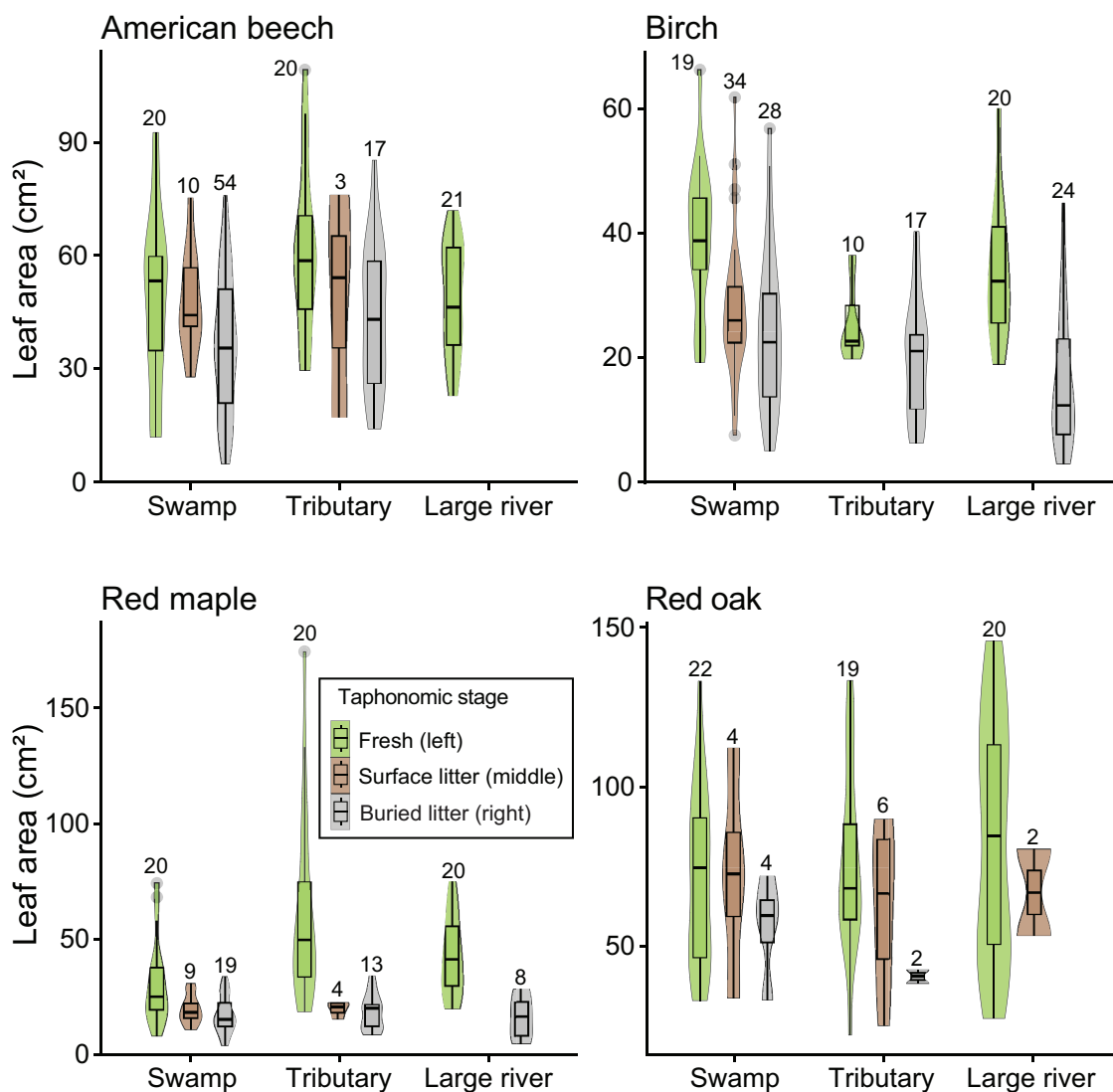


FIG. 5.—Violin boxplots showing a consistent decrease in leaf area across the taphonomic stages (fresh = left; surface litter = middle; buried litter = right) at Harvard Forest. The boxplots span the interquartile range (IQR; 25–75 percentiles); whiskers span the full data range up to $1.5 \times \text{IQR}$, beyond which individual symbols are plotted. Sample size is noted above each boxplot.

Estimating Leaf Area from Second-Order Vein Density

We also sought to test what factors may be driving any disproportionate loss of large leaves in litter. Two proposed factors are that smaller but thicker leaves tend to travel farther in the atmosphere and be more resistant to decay (see Introduction). If the first factor is important, then larger leaves would not be present at any time in the sampled litter layers; if the second factor is important, larger leaves may enter the litter but then subsequently break apart (or, alternatively, enter the litter already fragmented). The known inverse scaling between second-order vein density and leaf area in dicots (Sack et al. 2012) allows us to assess these two factors because we can measure second-order vein density in leaf fragments from the buried litter to estimate their original leaf areas; if these areas are larger than their complete companions, for example, then this would suggest that the second factor is important. This scaling has been tested extensively in living plants (Sack et al. 2012) and applied to fossils (Merkhofer et al. 2015; Hagen et al. 2019; Toumoulin et al. 2020; Baumgartner and Peppe 2021; Munyaka et al. 2025).

We measured second-order vein density in four species-site pairs (birch at Harvard Forest, American hornbeam at SERC, and American beech at both sites) following established methods (Sack et al. 2012; Merkhofer

et al. 2015). In brief, for each leaf we created up to four rectangles in ImageJ, each capturing at least two second-order veins (see Online Supplemental File Figure S3 for example). We computed the vein density of each rectangle, then the leaf average. We created calibrations for each species-site pair based on 20 complete leaves from the canopy, which are all strongly significant (Fig. 4). To test for size differences in the buried litter, at each quarry where the species-site pair is present we measured second-order vein density on 15–30 complete-to-nearly-complete leaves whose areas could be directly measured (species-site total of 15–69) and 11–18 fragmented leaves whose areas could not be directly measured (species-site total of 11–45; see Online Supplemental Data File 2 for R code used to run analyses). For paleobotany, the latter category represents information that is typically lost. We used vein scaling to estimate area even for the complete leaves to account for any built-in biases in the calibrations (emphasizing accuracy over precision; see also Hagen et al. 2019).

RESULTS

We found a clear reduction in leaf area along the taphonomic stages from fresh leaves to surface litter to buried litter. At Harvard Forest, the

TABLE 2.—Outcomes of post-hoc tests comparing differences in leaf physiognomy between taphonomic stages. For each physiognomic variable (see Table 1 for definitions), the first row represents changes between fresh leaves and surface litter, the second row between surface and buried litter, and the third row the overall change from fresh leaves to buried litter. Each cell contains information about direction of change (arrows) and *p*-value (in boldface when < 0.05). At each site, a modified model was run to assess the primacy of leaf area in driving the observed patterns. For leaf area, the modified model additionally includes tooth area:internal perimeter as a factor. For all other variables where the *p*-value of the standard model is < 0.05 , the modified model additionally includes leaf area as a factor.

Physiognomic variable (units)	Harvard Forest		SERC	
	(Standard model)	(Modified model)	(Standard model)	(Modified model)
*Leaf area (cm ²)	↓ <0.001 ↓ 0.005 ↓ <0.001	↓ <0.001 ↓ 0.005 ↓ <0.001	↓ <0.001 ↓ <0.001 ↓ <0.001	↓ <0.001 ↓ <0.001 ↓ <0.001
**Feret diameter ratio (unitless)	↑ <0.001 ↓ <0.001 ↑ 0.05	↓ <0.001 ↓ <0.001 ↑ <0.001	↓ <0.001 ↓ 0.23 ↓ <0.001	↓ <0.001 ↓ <0.001 ↓ 0.33
*Perimeter ratio (unitless)	↓ <0.001 ↓ 0.47 ↓ <0.001	↓ <0.001 ↓ <0.001 ↓ <0.001	↑ 0.24 ↑ 0.83 ↑ 0.04	↑ 0.85
Tooth count:blade area (cm ²)	↑ 0.02 ↑ 0.01 ↑ <0.001	↓ 0.01 ↑ 0.06 ↓ 0.41	↑ 0.64 ↑ <0.001 ↑ <0.001	↑ 0.04 ↑ 0.54
Tooth area:blade area (unitless)	↑ 0.001 ↓ <0.001 ↑ 0.10	↑ 0.15 ↓ <0.001 ↓ <0.001	↑ 0.97 ↓ 0.27 ↓ 0.10	↓ 0.86 ↓ 0.03 ↓ 0.002
Tooth area:internal perimeter (cm)	↑ 0.58 ↓ <0.001 ↓ <0.001	↓ <0.001 ↓ <0.001 ↓ <0.001	↓ 0.02 ↓ <0.001 ↓ <0.001	↓ 0.86 ↓ 0.03 ↓ 0.002
Average tooth area (cm ²)	↑ 0.002 ↓ <0.001 ↓ 0.82	↑ <0.001 ↓ <0.001 ↓ <0.001	↓ 0.35 ↓ 0.009 ↓ <0.001	↑ 0.95 ↑ 0.51 ↓ 0.001
***Tooth count:internal perimeter (cm ⁻¹)	↓ 0.06 ↑ 0.002 ↑ 0.22	↑ 0.09	↓ <0.001 ↑ <0.001 ↑ 0.13	↓ 0.06

* Used in MAP equation.

** Used in MAT equation.

*** Used in both MAP and MAT equations.

average leaf size within species in the surface litter was 7.5% smaller than in the fresh leaves ($p < 0.001$), declining further by 40% in the buried litter ($p = 0.005$). Together, this constitutes an overall reduction of 45% ($p < 0.001$) from fresh to buried leaves (Fig. 5, Table 2). We found a stronger pattern at SERC: surface litter is 46% smaller than fresh leaves ($p < 0.001$) and buried litter 28% smaller than surface litter ($p < 0.001$). Combined, this represents an overall decline of 61% ($p < 0.001$; Fig. 6). Across both sites, the average overall decline in swamps (40%) was less than in the tributaries (50%) and dynamic rivers (54%).

Using second-order vein scaling (Fig. 4), we find that in three of our four tested species-site pairs, complete-to-nearly-complete leaves from the buried litter have smaller areas than the reconstructed original leaf areas of fragments (Fig. 7). The magnitude of this effect scales with leaf size: the species with the largest leaves show the largest effect (American beech at both sites) while the smallest leaf species shows a mild reversed effect (American hornbeam at SERC).

The other physiognomic variables did not exhibit as strong or consistent a trend across both sites as leaf area (Table 2, Online Supplemental Files Figures S4–S17). Only two variables showed consistent and significant changes between fresh leaves and buried litter (tooth count:blade area and tooth area:internal perimeter), but not in all the intervening taphonomic stages (Table 2). We questioned whether the significant patterns in these variables were due to their covariance with leaf area. In other words, were these more scattered significant patterns driven by a stronger pattern in leaf area? A principal component analysis (PCA) of our data set supports this viewpoint (Fig. 8). For example, tooth count to blade area appears anti-correlated with leaf area at

both sites. Furthermore, if we include leaf area as a factor in our model for explaining variables other than leaf area, the significance of the taphonomic factor (fresh leaves/surface litter/buried litter) often weakens, including for both tooth count:blade area and tooth area:internal perimeter (Table 2). The reverse is not true: if these other variables are added to the leaf area model, the taphonomic factor mostly remains as significant (see Table 2 for model that includes tooth area:internal perimeter). This clearly implicates leaf area as the dominant physiognomic variable linked to taphonomic stage.

The univariate precipitation model (leaf area analysis; Peppe et al. 2011) indicates that the 45% and 61% reductions in leaf area observed at Harvard Forest and SERC, if representative of all woody dicots at these sites, leads to an underestimation of 16–23% in mean annual precipitation (MAP). The same trend but smaller is observed in the DiLP model, ranging from 6–19%. This smaller change relative to the univariate model is due to the addition of two traits—perimeter ratio and tooth count:internal perimeter—with only the former showing a significant taphonomic effect between the fresh leaves and buried litter, but in opposite directions in the two forests (Table 2).

The impact of taphonomy on leaf physiognomy did not lead to large changes in MAT in the multivariate DiLP (0–1.5°C), again assuming if our observed changes are representative of all woody dicots. This is because only one variable at one site in the DiLP MAT model changed significantly (feret diameter ratio at SERC; Table 2). Because we tracked single species along a taphonomic sequence, the scoring for toothed vs. untoothed did not change, meaning no change in estimated MAT with the univariate leaf-margin analysis.

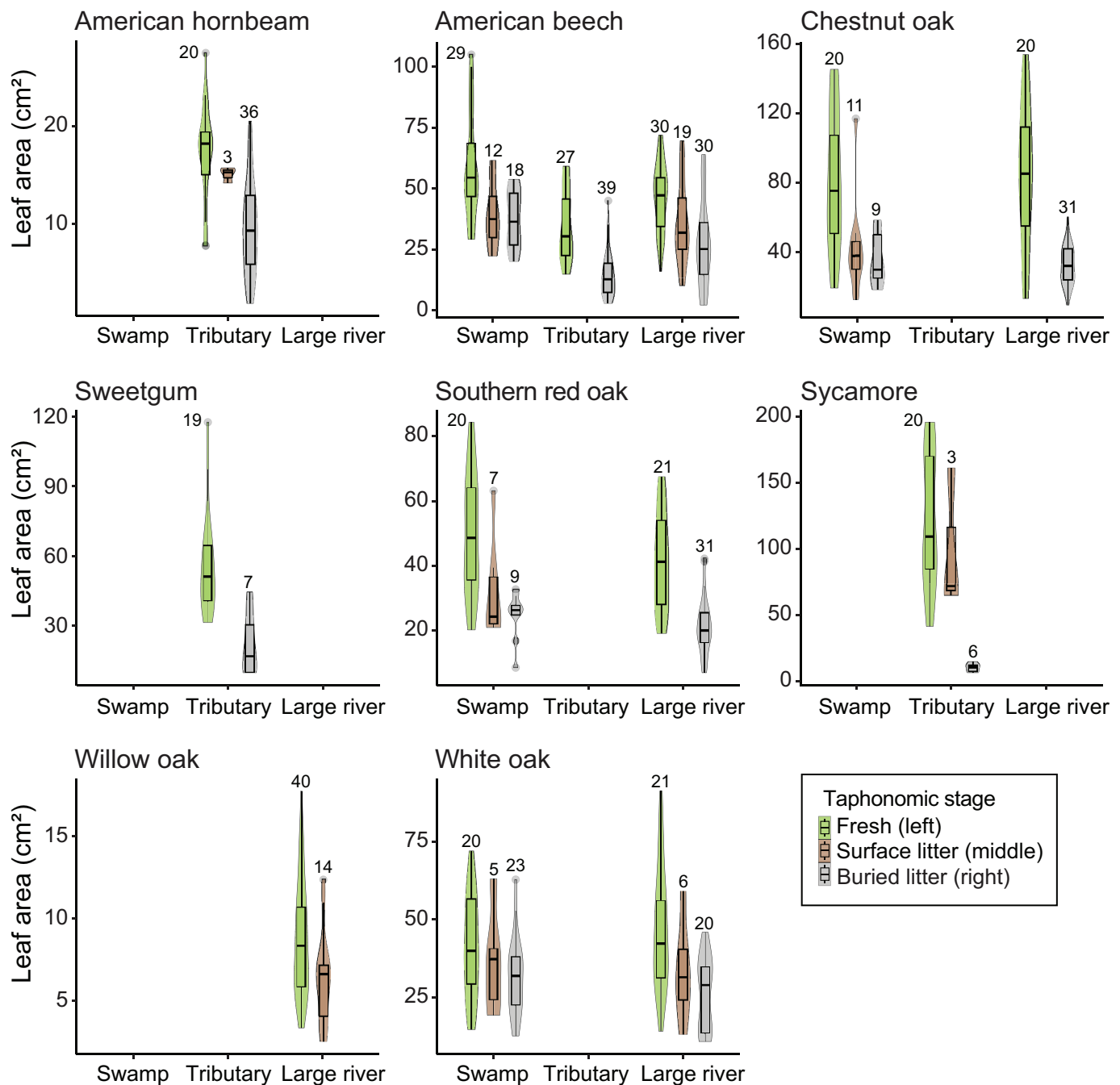


FIG. 6.—Violin boxplots showing a consistent decrease in leaf area across the taphonomic stages (fresh = left; surface litter = middle; buried litter = right) at the Smithsonian Environmental Research Center. The boxplots span the interquartile range (IQR; 25–75 percentiles); whiskers span the full data range up to $1.5 \times \text{IQR}$, beyond which individual symbols are plotted. Sample size is noted above each boxplot.

DISCUSSION

We find a strong and consistent reduction in leaf area within species across three taphonomic stages at both sites, with an overall reduction of 45% at Harvard Forest and 61% at SERC (Figs. 6, 7, Table 2). The consistency of these patterns suggests a chronic size bias within species in forested systems. This is in keeping with Roth and Dilcher (1978), who found that leaves in lake sediments were, on average, 38% smaller than nearby fresh leaves of the same species. In contrast, Astorga et al. (2016) found minimal size sorting within species between leaves in lake sediment versus litter from the surrounding forest.

The magnitude of the leaf-size bias is largest within river settings (~52%). We suspect that this is related to smaller leaves more easily being transported down rivers than larger leaves (Greenwood 1992) and a greater transport distance for leaves in river versus swamp sediment (Ferguson 1985), increasing the odds for larger leaves to break apart (e.g., Hagen et al. 2019). Our observed size bias in swamps (~40%) is like within-species patterns reported for a single lake (38%; Roth and Dilcher 1978), suggesting similar taphonomic processes.

Vein scaling suggests that complete leaves from buried litter are typically smaller than the reconstructed areas of leaf fragments from the same

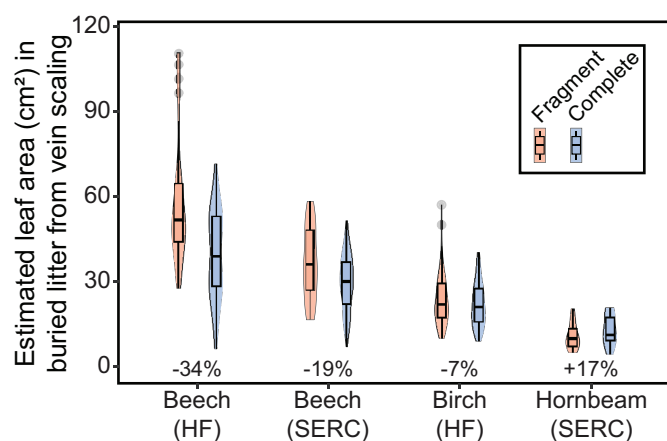


FIG. 7.—Estimated leaf area of fragments and complete-to-nearly-complete leaves of four species-site pairs using second-order vein density (see Fig. 4 for species-site calibrations). The percentage difference between the two categories are included as labels. The boxplots span the interquartile range (IQR; 25–75 percentiles); whiskers span the full data range up to $1.5 \times \text{IQR}$, beyond which individual symbols are plotted.

litter (Fig. 7). This difference is strongest in the species with the largest leaves (American beech: 19–34%), but the signal is noisy with an overall effect across the four species-site pairs of 11%. This supports several ideas: (1) some large leaves were deposited in our quarries, but then subsequently broke apart; (2) large leaves were more likely to break apart than small leaves; (3) because the size difference is smaller than our observed differences between fresh leaves and buried litter (45–61%), either larger leaves were disproportionately excluded from our depocenters or disproportionately fragmented severely enough that vein scaling was not possible; and (4) because the overall effect is similar to the fossil study of Hagen et al. (2019) (11% vs. 10%), size biasing after burial may be minimal (compare with Fig. 1).

Other physiognomic variables exhibited some change across taphonomic stages, but unlike leaf area none displayed a strong monotonic pattern in both sites across the three taphonomic stages (Table 2). Because many of these other physiognomic variables are collinear with leaf area, their (weak) connection with taphonomic stage is likely driven by the stronger connection in leaf area.

Our measured declines in leaf area—if representative of all woody dicot species at our sites—lead to a 16% and 23% underestimation of mean annual precipitation (MAP) at Harvard Forest and SERC, respectively, when using leaf-area analysis. This potentially points to a chronic biasing in the method (falsely dry). One way to partly address this taphonomic effect is to calibrate the relationship between leaf area and MAP with buried litter at sites that are good candidates for fossilization (see also Roth and Dilcher 1978). The taphonomic effect on MAP with DiLP is smaller: a 6% and 19% underestimation at Harvard Forest and SERC, respectively. This is because the additional physiognomic variables in the MAP model are largely unaffected by taphonomic stage, diluting out the large effect in leaf area.

In contrast to MAP, MAT reconstructions are largely unaffected by taphonomic biasing as captured by our study design. This is because the single trait used in leaf-margin analysis (percentage of species without teeth) is unaffected by taphonomic stage within species; and the additional characters in the DiLP MAT model (feret diameter ratio, number of teeth:internal perimeter) do not show significant monotonic changes at both sites (Table 2). This ultimately is good news for the prospects of reconstructing MAT from leaf physiognomy, so long as the species captured in a fossil deposit are representative of the standing vegetation.

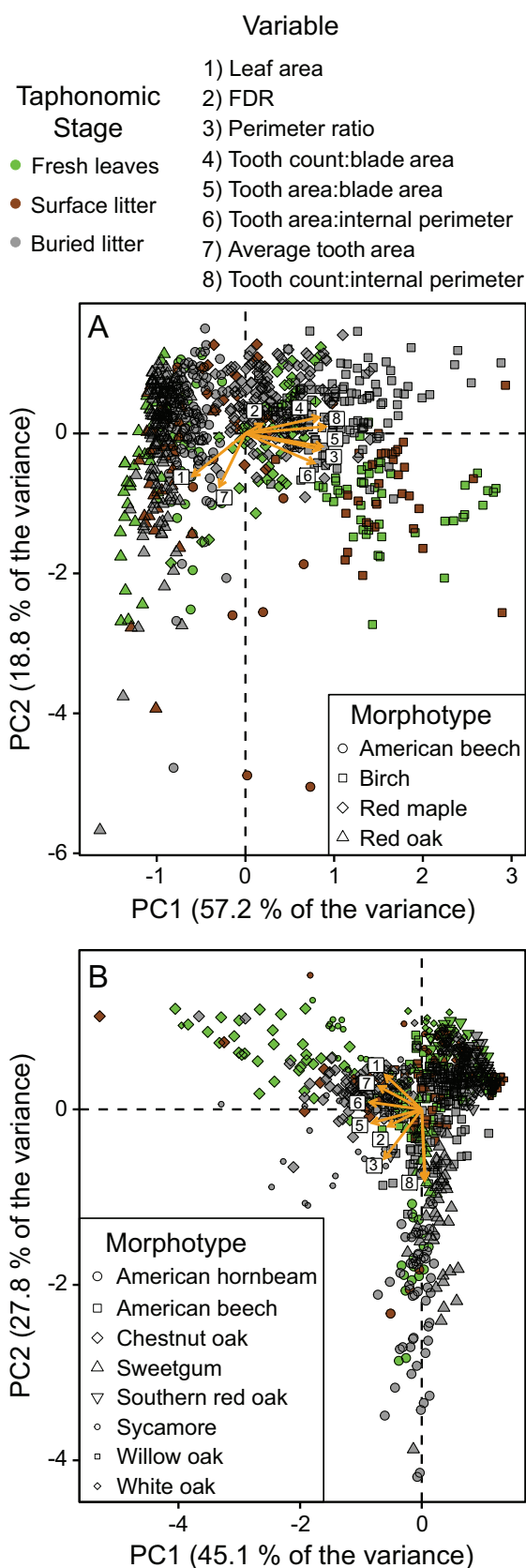


FIG. 8.—Principal Component Analysis (PCA) of physiognomic traits across taphonomic stages. A) Harvard Forest. B) Smithsonian Environmental Research Center. Each symbol is an individual leaf. Abbreviations for physiognomic variables are defined in Table 1.

CONCLUSION

Our findings highlight the importance of considering taphonomic processes when using leaf physiognomy to infer past climates. The consistent reductions in leaf area within species across both sites suggest a general size biasing in forested environments, particularly in river settings. This bias has implications for climate reconstructions, especially for MAP, which may be systematically underestimated. MAP models based on multiple traits (e.g., DiLP) appear less affected. MAT estimates are probably more robust to within-species taphonomic effects. Recalibrating MAP models with buried litter from depositional settings could improve their accuracy.

ACKNOWLEDGMENTS

We thank the editors and reviewers for their constructive feedback. We thank Matthew Gilbert, Aoife Hogan, Anya Le, Xander Lord, Sophia Samant, Michaela Somers, Hannah Steinberg, Zoe Todd, and Katie Yin for their aid in collecting data from the leaf specimens. We are grateful to Patrick Megonigal at the Smithsonian Environmental Research Center (SERC) and Audrey Barker-Plotkin at Harvard Forest (HF) for site access and logistical support. Funding for this research was provided by the McNair Scholars Program.

SUPPLEMENTAL MATERIAL

Data are available from the PALAIOS Data Archive: <https://www.sepm.org/supplemental-materials>.

REFERENCES

- ASTORGA, G.A., JORDAN, G.J., AND BRODRIBB, T., 2016, Towards understanding the fossil record better: insights from recently deposited plant macrofossils in a sclerophyll-dominated subalpine environment: *Review of Palaeobotany and Palynology*, v. 233, p. 1–11, doi: [10.1016/j.revpalbo.2016.06.004](https://doi.org/10.1016/j.revpalbo.2016.06.004).
- AU, T.F., MAXWELL, J.T., NOVICK, K.A., ROBESON, S.M., WARNER, S.M., LOCKWOOD, B.R., PHILLIPS, R.P., HARLEY, G.L., TELEWSKI, F.W., THERRELL, M.D., AND PEDERSON, N., 2020, Demographic shifts in eastern US forests increase the impact of late-season drought on forest growth: *Ecography*, v. 43, p. 1475–1486, doi: [10.1111/ecog.05055](https://doi.org/10.1111/ecog.05055).
- AZEVEDO-SCHMIDT, L., MEINEKE, E.K., AND CURRANO, E.D., 2022, Insect herbivory within modern forests is greater than fossil localities: *Proceedings of the National Academy of Sciences of the United States of America*, v. 119, e2202852119, doi: [10.1073/pnas.2202852119](https://doi.org/10.1073/pnas.2202852119).
- BAILEY, I.W. AND SINNOTT, E.W., 1915, A botanical index of Cretaceous and Tertiary climates: *Science*, v. 41, p. 831–834, doi: [10.1126/science.41.1066.831](https://doi.org/10.1126/science.41.1066.831).
- BAILEY, I.W. AND SINNOTT, E.W., 1916, The climatic distribution of certain types of angiosperm leaves: *American Journal of Botany*, v. 3, p. 24–39, doi: [10.1002/j.1537-2197.1916.tb05397.x](https://doi.org/10.1002/j.1537-2197.1916.tb05397.x).
- BAUMGARTNER, A. AND PEPPE, D. J., 2021, Paleoenvironmental changes in the Hiwegi Formation (lower Miocene) of Rusinga Island, Lake Victoria, Kenya: *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 574, e110458, doi: [10.1016/j.palaeo.2021.110458](https://doi.org/10.1016/j.palaeo.2021.110458).
- BLONDER, B., BUZZARD, V., SIMOVA, I., SLOAT, L., BOYLE, B., LIPSON, R., AGUILAR-BEAUCAGE, B., ANDRADE, A., BARBER, B., BARNES, C., BUSHEY, D., CARTAGENA, P., CHANEY, M., CONTRERAS, K., COX, M., CUETO, M., CURTIS, C., FISHER, M., FURST, L., GALLEGOS, J., HALL, R., HAUSCHILD, A., JEREZ, A., JONES, N., KLUKAS, A., KONO, A., LAMB, M., MATTHAI, J.D.R., MCINTYRE, C., MCKENNA, J., MOSIER, N., NAVABI, M., OCHOA, A., PACE, L., PLASSMANN, R., RICHTER, R., RUSSAKOFF, B., AUBYN, H.S., STAGG, R., STERNER, M., STEWART, E., THOMPSON, T.T., THORNTON, J., TRUJILLO, P.J., VOLPE, T.J., AND ENQUIST, B.J., 2012, The leaf-area shrinkage effect can bias paleoclimate and ecology research: *American Journal of Botany*, v. 99, p. 1756–1763, doi: [10.3732/ajb.1200062](https://doi.org/10.3732/ajb.1200062).
- BURNHAM, R.J., 1989, Relationships between standing vegetation and leaf litter in a paratropical forest: implications for paleobotany: *Review of Palaeobotany and Palynology*, v. 58, p. 5–32, doi: [10.1016/0034-6667\(89\)90054-7](https://doi.org/10.1016/0034-6667(89)90054-7).
- BUTRIM, M., LOWE, A., FLYNN, A., BAUMGARTNER, A., PEPPE, D., AND ROYER, D., 2024, dilp: reconstruct paleoclimate and paleoecology with leaf physiognomy: R package version 1.1.0, <https://cran.r-project.org/web/packages/dilp/index.html>.
- CHRISTOPHEL, D.C. AND GREENWOOD, D.R., 1989, Changes in climate and vegetation in Australia during the Tertiary: *Review of Palaeobotany and Palynology*, v. 58, p. 95–109, doi: [10.1016/0034-6667\(89\)90079-1](https://doi.org/10.1016/0034-6667(89)90079-1).
- DRISCOLL, A.G., 2016, The effect of treefall gaps on the spatial distribution of three invasive plants in a mature upland forest in Maryland: *The Journal of the Torrey Botanical Society*, v. 143, p. 349–358, doi: [10.3159/torrey-d-15-00022.1](https://doi.org/10.3159/torrey-d-15-00022.1).
- ELLIS, B. AND JOHNSON, K.R., 2013, Comparison of leaf samples from mapped tropical and temperate forests: implications for interpretations of the diversity of fossil assemblages: *PALAIOS*, v. 28, p. 163–177, doi: [10.2110/palo.2012.p12-073r](https://doi.org/10.2110/palo.2012.p12-073r).
- FERGUSON, D.K., 1985, The origin of leaf-assemblages—new light on an old problem: *Review of Palaeobotany and Palynology*, v. 46, p. 117–188, doi: [10.1016/0034-6667\(85\)90041-7](https://doi.org/10.1016/0034-6667(85)90041-7).
- FERGUSON, D.K., 2005, Plant taphonomy: ruminations on the past, the present, and the future: *PALAIOS*, v. 20, p. 418–428, doi: [10.2110/palo.2005.P05-25p](https://doi.org/10.2110/palo.2005.P05-25p).
- GIVNISH, T.J., 1984, Leaf and canopy adaptations in tropical forests, in E. Medina, H.A. Mooney, and C. Vázquez-Yanes (eds.), *Physiological Ecology of Plants of the Wet Tropics: The Hague, Junk, Tasks for Vegetation Science*, v. 12, p. 51–84.
- GREENWOOD, D.R., 1992, Taphonomic constraints on foliar physiognomic interpretations of Late Cretaceous and Tertiary palaeoclimates: *Review of Palaeobotany and Palynology*, v. 71, p. 149–190, doi: [10.1016/0034-6667\(92\)90161-9](https://doi.org/10.1016/0034-6667(92)90161-9).
- GREENWOOD, D.R., 2005, Leaf margin analysis: taphonomic constraints: *PALAIOS*, v. 20, p. 498–505, doi: [10.2110/palo.2004.P04-58](https://doi.org/10.2110/palo.2004.P04-58).
- HAGEN, E.R., ROYER, D.L., MOYE, R.A., AND JOHNSON, K.R., 2019, No large bias within species between the reconstructed areas of complete and fragmented fossil leaves: *PALAIOS*, v. 34, p. 43–48, doi: [10.2110/palo.2018.091](https://doi.org/10.2110/palo.2018.091).
- HANSON, H.C., 1917, Leaf-structure as related to environment: *American Journal of Botany*, v. 4, p. 533–560, doi: [10.2307/2435253](https://doi.org/10.2307/2435253).
- KOCH, G.W., SILLETT, S.C., JENNINGS, G.M., AND DAVIS, S.D., 2004, The limits to tree height: *Nature*, v. 428, p. 851–854, doi: [10.1038/nature02417](https://doi.org/10.1038/nature02417).
- LOWE, A.J., FLYNN, A.G., BUTRIM, M.J., BAUMGARTNER, A., ROYER, D.L., AND PEPPE, D.J., 2024, Reconstructing terrestrial paleoclimate and paleoecology with fossil leaves using digital leaf physiognomy and leaf mass per area: *Journal of Visualized Experiments*, v. 212, e66838, doi: [10.3791/66838](https://doi.org/10.3791/66838).
- MERKHOFFER, L., WILF, P., HAAS, M.T., KOOYMAN, R.M., SACK, L., SCOFFONI, C., AND N. RUBÉN CÚNEO, 2015, Resolving Australian analogs for an Eocene Patagonian paleo-rainforest using leaf size and floristics: *American Journal of Botany*, v. 102, p. 1160–1173, doi: [10.3732/ajb.1500159](https://doi.org/10.3732/ajb.1500159).
- MUNYAKA, V., OGINGA, K., COTE, S., HEAD, J., LUKENS, W., LUTZ, J., KINYANJUI, R., MANTHI, F., McNULTY, K., HALL, A., AND TEGART, A., 2025, Insights on the paleoclimate and paleoecology of an early Miocene hominoid site: a multiproxy study from Koru, Western Kenya: *Paleoceanography and Paleoclimatology*, v. 40, e2025PA005152, doi: [10.1029/2025PA005152](https://doi.org/10.1029/2025PA005152).
- PEPPE, D.J., ROYER, D.L., CARIGLINO, B., OLIVER, S.Y., NEWMAN, S., LEIGHT, E., ENIKOLOPOV, G., FERNANDEZ-BURGOS, M., HERRERA, F., ADAMS, J.M., CORREA, E., CURRANO, E.D., ERICKSON, J.M., HINOJOSA, L.F., et al., 2011, Sensitivity of leaf size and shape to climate: global patterns and paleoclimatic applications: *New Phytologist*, v. 190, p. 724–739, doi: [10.1111/j.1469-8137.2010.03615.x](https://doi.org/10.1111/j.1469-8137.2010.03615.x).
- R CORE TEAM, 2024, R: a language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria, www.R-project.org.
- RAUNKIAER, C., 1934, *The Life Forms of Plants and Statistical Plant Geography*: Clarendon, Oxford, 632 p.
- ROTH, J. L. AND DILCHER, D. L., 1978, Some considerations in leaf size and leaf margin analysis of fossil leaves: *Courier Forschungsinstitut Senckenberg*, v. 30, p. 165–171.
- ROYER, D.L., WILF, P., JANESKO, D.A., KOWALSKI, E.J., AND DILCHER, D.L., 2005, Correlations of climate and plant ecology to leaf size and shape: potential proxies for the fossil record: *American Journal of Botany*, v. 92, p. 1141–1151, doi: [10.3732/ajb.92.7.1141](https://doi.org/10.3732/ajb.92.7.1141).
- SACK, L., MELCHER, P.J., LIU, W.H., MIDDLETON, E., AND PARDEE, T., 2006, How strong is intracanalopy leaf plasticity in temperate deciduous trees?: *American Journal of Botany*, v. 93, p. 829–839, doi: [10.3732/ajb.93.6.829](https://doi.org/10.3732/ajb.93.6.829).
- SACK, L., SCOFFONI, C., MCKOWN, A.D., FROLE, K., RAWLS, M., HAVRAN, J.C., TRAN, H., AND TRAN, T., 2012, Developmentally based scaling of leaf venation architecture explains global ecological patterns: *Nature Communications*, v. 3, e837, doi: [10.1038/ncomms1835](https://doi.org/10.1038/ncomms1835).
- SPICER, R.A., 1989, The formation and interpretation of plant fossil assemblages: *Advances in Botanical Research*, v. 16, p. 95–191, doi: [10.1016/S0065-2296\(89\)60240-2](https://doi.org/10.1016/S0065-2296(89)60240-2).
- SPICER, R.A., 1981, The sorting and deposition of allochthonous plant material in a modern environment at Silwood Lake, Silwood Park, Berkshire, England: *U.S. Geological Survey Professional Paper* 1143, p. 1–77, doi: [10.3133/pp1143](https://doi.org/10.3133/pp1143).
- SPICER, R.A., BERA, S., DE BERA, S., SPICER, T.E.V., SRIVASTAVA, G., MEHROTRA, R., MEHROTRA, N., AND YANG, J., 2011, Why do foliar physiognomic climate estimates sometimes differ from those observed? Insights from taphonomic information loss and a CLAMP case study from the Ganges Delta: *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 302, p. 381–395, doi: [10.1016/j.palaeo.2011.01.024](https://doi.org/10.1016/j.palaeo.2011.01.024).
- SPICER, R.A., HERMAN, A.B., AND KENNEDY, E.M., 2005, The sensitivity of CLAMP to taphonomic loss of foliar physiognomic characters: *PALAIOS*, v. 20, p. 429–438, doi: [10.2110/palo.2004.p04-63](https://doi.org/10.2110/palo.2004.p04-63).
- SPICER, R.A., YANG, J., SPICER, T.E.V., AND FARNSWORTH, A., 2021, Woody dicot leaf traits as a paleoclimate proxy: 100 years of development and application: *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 562, e110138, doi: [10.1016/j.palaeo.2020.110138](https://doi.org/10.1016/j.palaeo.2020.110138).
- STEART, D.C., BOON, P.I., GREENWOOD, D.R., AND DIAMOND, N.T., 2002, Transport of leaf litter in upland streams of Eucalyptus and Nothofagus forests in south-eastern Australia: *Archiv für Hydrobiologie*, v. 156, p. 43–61, doi: [10.1127/0003-9136/2002/0156-0043](https://doi.org/10.1127/0003-9136/2002/0156-0043).
- TALBERT, C.M. AND HOLCH, A.E., 1957, A study of the lobing of sun and shade leaves: *Ecology*, v. 38, p. 655–658, doi: [10.2307/1943135](https://doi.org/10.2307/1943135).

- TOUMOULIN, A., KUNZMANN, L., MORAWECK, K., AND SACK, L., 2020, Reconstructing leaf area from fragments: testing three methods using a fossil Paleogene species: *American Journal of Botany*, v. 107, p. 1786–1797, doi: [10.1002/ajb2.1574](https://doi.org/10.1002/ajb2.1574).
- WARNES, G.R., BOLKER, B., LUMLEY, T., AND JOHNSON, R.C., 2024, gmodels: various R programming tools for model fitting, version 2.19.1, doi: [10.32614/CRAN.package.gmodels](https://doi.org/10.32614/CRAN.package.gmodels).
- WILE, P., 1997, When are leaves good thermometers? A new case for Leaf Margin Analysis: *Paleobiology*, v. 23, p. 373–390, doi: [10.1017/s0094837300019746](https://doi.org/10.1017/s0094837300019746).
- WILE, P., WING, S.L., GREENWOOD, D.R., AND GREENWOOD, C.L., 1998, Using fossil leaves as paleoprecipitation indicators: an Eocene example: *Geology*, v. 26, p. 203–206, doi: [10.1130/0091-7613\(1998\)026%3C0203:uflapi%3E2.3.co;2](https://doi.org/10.1130/0091-7613(1998)026%3C0203:uflapi%3E2.3.co;2).
- WING, S.L., 1988, Depositional environments of plant bearing sediments: *The Paleontological Society Special Publications*, v. 3, p. 1–13, doi: [10.1017/s2475262200004858](https://doi.org/10.1017/s2475262200004858).
- WOLFE, J.A., 1993, A method of obtaining climatic parameters from leaf assemblages: *U.S. Geological Survey Bulletin*, v. 2040, p. 1–71, doi: [10.3133/b2040](https://doi.org/10.3133/b2040).
- WOLFE, J.A., 1979, Temperature parameters of humid to mesic forests of eastern Asia and relation to forests of other regions of the Northern Hemisphere and Australasia: *U.S. Geological Survey Professional Paper*, v. 1106, p. 1–27, doi: [10.3133/pp1106](https://doi.org/10.3133/pp1106).
- WRIGHT, I.J., DONG, N., MAIRE, V., PRENTICE, I.C., WESTOBY, M., DÍAZ, S., GALLAGHER, R.V., JACOBS, B.F., KOOYMAN, R., LAW, E.A., LEISHMAN, M.R., NIINEMETS, Ü., REICH, P.B., SACK, L., et al., 2017, Global climatic drivers of leaf size: *Science*, v. 357, p. 917–921, doi: [10.1126/science.aal4760](https://doi.org/10.1126/science.aal4760).
- WYLIE, R.B., 1951, Principles of foliar organization shown by sun-shade leaves from ten species of deciduous dicotyledonous tree: *American Journal of Botany*, v. 38, p. 355–361, doi: [10.2307/2437823](https://doi.org/10.2307/2437823).

Received 23 September 2025; accepted 8 January 2026.