Oecologia (2022) 198:407–418
https://doi.org/10.1007/s00442-022-05123-w

POPULATION ECOLOGY – ORIGINAL RESEARCH

# Climate affects the outbreaks of a forest defoliator indirectly through its tree hosts

**Kyle J. Haynes**[1] · **Andrew M. Liebhold**[2,3] · **Jonathan S. Lefcheck**[4] · **Randall S. Morin**[5] · **Guiming Wang**[6]

Received: 17 September 2021 / Accepted: 25 January 2022 / Published online: 9 February 2022

**Abstract**
Although spatial variation in climate can directly affect the survival and reproduction of forest insects and the tree species compositions of forests, little is known about the indirect effects of climate on outbreaks of forest insects through its effects on forest composition. In this study, we use structural equation modeling to examine the direct and indirect effects of climate, water capacity of the soil, host tree density, and non-host density on the spatial extent of *Lymantria dispar* outbreaks in the Eastern USA over a period of 44 years (1975–2018). Host species were subdivided into four taxonomic and ecologically distinct groups: red oaks (Lobatae), white oaks (Lepidobalanus), other preferred hosts, and intermediate (less preferred) hosts. We found that mean annual temperature had stronger effects than mean annual precipitation on the spatial extent of outbreaks, and that indirect effects of temperature (via its effects on oak density) on defoliation were stronger than direct effects. The density of non-host trees increased with increasing precipitation and, consistent with the 'associational resistance hypothesis', defoliation decreased with increasing density of non-host trees. This study offers quantitative evidence that geographic variation in climate can indirectly affect outbreaks of a forest insect through its effects on tree species composition.

**Keywords** Structural equation model · Host composition · *Lymantria dispar* · Climate change · Resource concentration hypothesis · Oaks

## Introduction

Outbreaks of foliage-feeding forest insects often extend over large regions and represent important disturbance factors that can have cascading impacts on forest ecosystem processes (Lovett et al. 2002). Despite extensive research, considerable uncertainty exists about the role of biotic and abiotic agents in triggering these outbreaks (Berryman 1996; Esper et al. 2007). Compared to temporal patterns, more certainty exists about the determinants of outbreaks through space. Specifically, the spatial distribution of host trees is known as a major determinant of the geographical extent of forest defoliator outbreaks (Candau and Fleming 2005; Hartl-Meier et al. 2017), but climate may also constrain the distribution of defoliation as well (Williams and Liebhold 1995; Jepsen et al. 2008; Candau and Fleming 2011), particularly in an era of global change where both temperature and precipitation are expected to vary considerably compared with past climate regimes (Battisti and Larsson 2015). These environmental effects might manifest directly, for example by influencing developmental rates and survival of defoliating insects (Roland et al. 1998; Thompson et al.

---

Communicated by Martin Gossner.

✉ Kyle J. Haynes
haynes@virginia.edu

[1] The Blandy Experimental Farm, University of Virginia, Boyce, VA 22620, USA

[2] USDA Forest Service, Northern Research Station, 180 Canfield Street, Morgantown, WV 26505, USA

[3] Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, 165 00 Praha 6, Suchdol, Czech Republic

[4] Tennenbaum Marine Observatories Network, MarineGEO, Smithsonian Environmental Research Center, Edgewater, MD 21037, USA

[5] USDA Forest Service, Northern Research Station, 3460 Industrial Drive, York, PA 17402, USA

[6] Department of Wildlife, Fisheries and Aquaculture, Mississippi State University, Mail Stop 9690, Mississippi State, MS 39762, USA

2020), or indirectly, by changing tree community composition and abundance. Researchers have simultaneously considered the effects of forest composition and climate on forest insect outbreaks (Cooke and Roland 2000; Gray 2008; Seidl et al. 2011; Robert et al. 2020), but parsing to what extent these factors affect insect outbreaks directly and indirectly would represent a key advance for predicting cascading effects on forest systems worldwide.

In addition to the distribution of hosts, non-hosts may play a role in determining the outbreak range of forest herbivores. The association of biological diversity with stability is a concept that has attracted even the earliest ecologists (e.g., Elton 1927), and has long been applied to insect-plant systems to hypothesize that diverse plant communities are less prone to insect herbivore outbreaks (Pimentel 1961; Jactel et al. 2021). The term "associational resistance" is used to describe increased the resistance of target plants to herbivory when surrounded by non-host plant species (Atsatt and O'dowd 1976; Sholes 2008). Tests of the associational resistance hypothesis have yielded mixed results when applied in forest ecosystems, with some studies confirming negative effects of tree diversity on outbreaks, while others indicating either no effects or even positive impacts (Koricheva et al. 2006). However, Jactel and Brockerhoff's (2007) applied meta-analysis across 119 studies and concluded that tree diversity more often reduces herbivory. Unfortunately, there are statistical issues that constrain quantification of associational resistance from data on insect outbreaks in stands of varying tree diversity. Specifically, host tree density is typically collinear with non-host density, especially at small scales, and this precludes fitting of models using multiple regression to isolate effects of host and non-host plant densities (Guo et al. 2019).

Here we use *Lymantria dispar* as a model system for exploring the direct and indirect effects of climate, host tree density, density of trees that are immune to *L. dispar* feeding, and other abiotic determinants on the spatial extent of *L. dispar* outbreaks. This is an ideal system in part because forest defoliation caused by *L. dispar* has been mapped annually in the eastern USA since the 1970s (Morin and Liebhold 2016). Furthermore, because of the massive extent of defoliation caused by this species, considerable information already exists about site characteristics that are associated with the frequency of outbreaks. *Lymantria dispar* is a highly polyphagous insect known to feed on over two hundred species of trees (Liebhold et al. 1995), and the proportion of forest composition represented by preferred hosts is known to be a primary determinant of forest susceptibility to *L. dispar* (Campbell 1974; Herrick et al. 1986; Davidson et al. 2001). Further, chronically dry sites are known to often support more frequent *L. dispar* outbreaks (Smith 1985). However, most of these studies have evaluated the relationship between environmental factors and *L. dispar* defoliation in individual stands rather than across macroscale regions.

We use the long-running survey of *L. dispar* outbreaks in the eastern United States to test how environment (e.g., temperature, precipitation) and tree community composition simultaneously drive this phenomenon across 253,411 km$^2$ of forest habitat. While we expect both sets of predictors to be important, we hypothesized that forest tree community composition is the more immediate driver, with climate primarily affecting the extent of outbreaks through modulating abundances of host and immune trees and a weaker direct effect on the defoliators themselves.

## Methods

### Study system

Though *L. dispar* is highly polyphagous folivore, it is known to exhibit a range of preferences for host species, with differences in preference closely associated with the tree taxonomic groupings (Liebhold et al. 1995). Of the host species in eastern North America that *L. dispar* prefers, oaks (*Quercus* spp.) are the most abundant (Liebhold et al. 1997). Other common genera of preferred hosts include *Populus* spp. (poplars and aspens) and *Betula* spp. (birches). Less preferred host trees include many maple (*Acer*) species, and most pine (*Pinus*) and hickory (*Carya*) species, and trees that are immune to *L. dispar* feeding include cypresses (*Cupressus* spp.), ashes (*Fraxinus* spp.), and most dogwood species (*Cornus* spp.; Liebhold et al. 1995).

There are two sections of oaks present in North America, red oaks (Lobatae) and white oaks (Lepidobalanus); while oak species in both sections are primary hosts, these two groups have divergent effects on the population dynamics of *L. dispar*. One of the dominant factors affecting the dynamics of *L. dispar* populations when *L. dispar* densities are low, is predation of pupae by small mammals, especially the white-footed mouse (*Peromyscus leucopus*; Elkinton and Liebhold 1990; Elkinton et al. 1996). Red oak acorns are a critical winter food source for these small mammal predators (Wolff 1996b; Jones et al. 1998; Elias et al. 2004; Clotfelter et al. 2007). Red oaks produce acorns in mast seeding events, i.e., in spatially synchronous events with high inter-annual variability, which are a major influence on the population cycles of white-footed mice (Wolff 1996b; Elias et al. 2004). Red oak acorns drop to the forest floor in autumn but do not germinate until spring, making them ideal for building a cache of winter food (Hadj-Chikh et al. 1996). In contrast, because white oak acorns germinate in the fall (Fox 1982), they can only rarely be successfully cached as a winter food source (Steele et al. 2001) and consequently they play a minor role in the dynamics of mammalian predators of *L. dispar*. Another potential source of diverging effects of red oaks and white oaks on *L. dispar* population growth are differences in foliage suitability. By comparing studies on the suitability of different tree species, Liebhold et al. (1995) categorized *L. dispar's* potential host species in North America into three relative ranks, referred to here as "preferred hosts", "intermediate hosts", and "immune". However, published field and laboratory experiments comparing the suitability of foliage based their conclusions on many different measures of *L. dispar* performance, making it difficult to quantify differences in the relative suitabilities of each host species (Liebhold et al. 1995), and there is no clear difference in suitability between the red-oak and white-oak sections. Thus, we had no a priori expectation that differences in the suitability of red and white oak foliage would influence geographic variation in *L. dispar* outbreaks.

## Analysis of direct and indirect effects on *Lymantria dispar* outbreaks

We evaluated geographic variation in the mean annual spatial extent of *L. dispar* outbreaks in the northeastern USA from 1975–2018. Outbreak populations defoliate host trees and annual presence of defoliation provides a measure of habitat suitability for *L. dispar* population growth. The annual presence of defoliation across the study region was determined based on maps sketched during low-altitude reconnaissance flights carried out by State and Federal agencies. The aerial surveys were conducted in July, the month of peak *Lymantria dispar* defoliation, before trees could flush new leaves. Ground checks were completed where the cause of defoliation was uncertain, and defoliation was attributed to *L. dispar* if confirmed by the presence of its life stages (Liebhold et al. 1997). The resulting defoliation maps were compiled into sequential annual raster maps in a geographic information system (Liebhold et al. 1997). The raster maps were initially created with a resolution of 2 × 2 km, and within each cell, defoliation was scored as present or absent (presence if > 50% of raster cell area was defoliated). However, based on the sparseness of tree species composition at fine spatial scales (see next paragraph), we aggregated (summed) the defoliation data (km$^2$ defoliated) across 32 × 32 km quadrats. We then averaged the annual defoliation data across the 44-year study period. Defoliation was thus expressed as the total area in each 32 × 32 km cell that was defoliated, averaged across all years of the survey (1975–2018). Given that these estimates of defoliation are formed from aggregated data on the presence/absence of defoliation, the estimates are best considered relative, not absolute, measures of defoliated area. *Lymantria dispar* has continually expanded its range since the time of its introduction in 1869 so our analysis was limited to the region that had already been invaded by the beginning of the time series (1975). The invaded region was defined by the area of the 1975 US Department of Agriculture *L. dispar* quarantine which is designated at the county level and published annually (US Code of Federal Regulations, Title 7, Chapter III, Sect. 301.45).

Tree species composition and stand treatment estimates were determined across the study area using data from 2013 for each state from the Forest Inventory and Analysis (FIA) program of the U.S. Department of Agriculture (USDA), which is designed to assess forest conditions across the USA (Bechtold and Patterson 2005). Prior to 1999, FIA collected data regionally using a periodic measurement system with sample designs that varied slightly by region. Generally, inventories were conducted in each state every 6–18 years, depending on the state and region. Beginning in 1999, FIA moved to an annual inventory approach where a complete, systematic sample of each state is completed annually (Bechtold and Patterson 2005). Historic data from 11 states in the eastern United States were extracted from the FIA database (DataMart link—https://apps.fs.usda.gov/fia/datamart/datamart.html). The relative abundance of *L. dispar* host trees and trees immune to *L. dispar* feeding showed little change between 1990 and 2013 (Appendix S1, Table S1).

The FIA sampling design is based on a tessellation of the US into 2428 ha hexagons with at least one permanent plot established in each hexagon. For every plot falling in forest land, measurements (species, diameter, height, condition, etc.) are taken for every tree in four 7.32-m fixed-radius subplots. Under the FIA sampling design, each 32 × 32 km quadrat contains approximately 42 FIA plots; however, no tree sampling is performed at plots that are not forested. We determined that ≥ 10 forested survey plots were needed within each quadrat to adequately characterize tree species composition (Bechtold and Patterson 2005) and then selected the finest resolution that would allow most quadrats to meet the criterion. This led to our selection of the 32 × 32 km resolution. Sixty-four of 443 quadrats (16.4%) had too few (< 10) forested FIA survey plots and were excluded from further analysis. In the remaining 379 quadrats, we expressed the abundance of red oaks, white oaks, other preferred hosts, intermediate (less preferred) hosts, and immune tree species (Liebhold et al. 1995) based on their basal areas. Because acorns from red oaks (section Lobatae) are especially critical for the winter survival of granivorous rodents (Shimada and Saitoh 2006), including major predators of *L. dispar*, we subdivided the basal areas of oaks into red oaks and white oaks (section Lepidobalanus) to consider their separate effects on *L. dispar* dynamics.

Geographic variation in climate among the 32 × 32 km quadrats was assessed based on mean annual temperature and precipitation over the 44-year study period (1975–2018). Temperature and precipitation values in 1 km raster format were obtained from the PRISM climate mapping project (http://www.prism.oregonstate.edu/index.phtml), which provides data through interpolation of historical weather data from stations located throughout the conterminous USA. The 1-km$^2$ resolution climate data were spatially aggregated by calculating averages for each 32 × 32 km quadrat.

Given that soil moisture availability appears to play an important role in the distributions of temperate tree species (Iverson and Prasad 1998; Morin et al. 2007), we acquired spatial data on the available water capacity of soil from the USGS STATSGO data base. Available water capacity (AWC) represents the volume of water available to plants if the soil is at its capacity (USDA, NRCS 2019). We acquired the AWC data as a 1:250,000-scale geographic information system (GIS) layer and formed average values of AWC for each 32 × 32 km quadrat. Given high collinearity of AWC taken at different depths, we only used water capacity measured at 100 cm in our analyses.

We examined the direct and indirect effects of environmental factors (temperature, precipitation, and AWC) and host tree composition on *L. dispar* outbreaks using piecewise structural equation modeling (pSEM). This approach is often more appropriate for ecological data than traditional variance–covariance structural equation modeling because it is more flexible in accommodating violations of model assumptions such as multivariate normality and independence of errors (Lefcheck 2016). Here, the ability of pSEM to accommodate spatial autocorrelation was the primary reason this approach was chosen.

Our first step was to formulate an a priori model of expected causal relationships between variables (full model provided in Appendix S2: Figure S1). We postulated direct effects of temperature, precipitation, and AWC on the basal areas of each category of host tree. We also specified directed paths between the basal areas of each tree species group, temperature, and precipitation on *L. dispar* defoliation. Thus, the model is a general test of mediation of environmental effects on defoliation through changing tree abundances. We also specified bivariate correlations between the basal areas of each tree group to account for any potential covariance between the basal areas of different tree groups unaccounted for by the climate variables or soil water capacity.

To account for spatial autocorrelation, we fit the specified relationships using spatial autoregressive error models for each of the component regressions within the pSEM using the 'errorsarlm' function in the *spatialreg* package (Bivand and Piras 2015) in R version 4.0.3 (R Core Team 2020). Spatial weights were determined as the Euclidean distance between quadrats using the `dnearneigh` function in the *spdep* package (Bivand and Wong 2018). We assessed model assumptions visually, and subsequently all tree abundances, defoliation extent, and AWC were subsequently square-root transformed. We tested for excessive multicollinearity using variance inflation factors (VIFs) computed using the Nagelkerke $R^2$, which we also used as an indicator of the variance explained for each response. An alternative method of accounting for forest land area would be to model the proportion of defoliated forest area, but this exercise did not lead to qualitatively different conclusions and further violated several assumptions of the models. Therefore, we have treated land area using an ANCOVA-like approach. Differences in area of forested land between each grid cells were accounted for by including forest land area (km$^2$) as a covariate in the models with tree basal areas and defoliated area. We specified correlations between the basal areas of each tree group to account for any potential relationships between the basal areas of different tree groups unaccounted for by the climate variables or soil water capacity. The strength of direct vs. indirect effects on defoliation was examined based on standardized regression coefficients obtained by multiplying each coefficient by the ratio of the standard deviation of the predictor over the standard deviation of the response.

The goodness of fit of the structural equation model was evaluated using two measures. First, we applied Shipley's test of directed separation, where the significance (i.e., $P$ value) of all unspecified paths is summarized in a single Fisher's $C$ statistic as in (Shipley 2009). We can compare this statistic to a $\chi^2$-distribution with $2k$ degrees of freedom, where $k$ = the number of tests of directed separation. A non-significant $P$ value would suggest that there are no significant missing paths, i.e., that we have correctly specified the structure of the model. Second, we used the newly proposed metric based on log-likelihood (Shipley and Douma 2020). Here, we consider two models: the proposed model and a saturated model, or one in which all possible paths are included. Then, the individual log-likelihoods of the submodels in each of the saturated and unsaturated models are summed, and their difference is obtained. Finally, this value is multiplied by −2 to yield a $\chi^2$ statistic directly, which can be compared to a $\chi^2$-distribution with the degrees of freedom determined by the difference in the likelihood degrees of freedom of the saturated and unsaturated models. This new test asks whether the saturated model is more likely than the proposed structure, implying that the inclusion of one or more of the missing paths would improve the likelihood of the saturated model given the data. Statistical code for these analyses is provided in Appendix S3.

## Results

Across the study region, 19.8% of the basal area of all forest trees was comprised of preferred *L. dispar* host species, 59.6% was comprised of intermediate hosts, and 20.6% consisted of species that are immune to *L. dispar* damage

(Table 1). Of the preferred *L. dispar* hosts, red oaks, white oaks, and other preferred hosts were similar in their densities, representing 7.7%, 6%, and 6% of the basal area of all forest trees, respectively. The intermediate-host group contained the three most dominant species in the study region, *Acer rubrum*, *Acer saccharum*, and *Tsuga canadensis* (Table 1).

Mean annual temperatures tended to be warmest along the southeastern portion of the study region, along the coast with the Atlantic Ocean (Fig. 1a). The spatial pattern in mean annual precipitation was less coherent than that of temperature, though precipitation tended to be higher in the eastern and mid-latitudinal portion of study region, and in the high-elevation areas of the Appalachian Mountain Range (Fig. 1b; elevation not shown). AWC was perhaps the most spatially coherent environmental variable that we considered, with AWC tending to increase with increasing latitude, though AWC was also high in a small portion of the southeastern corner of the study region (Fig. 1c).

The two dominant groups of *L. dispar* host trees, red oaks and white oaks, differed considerably in their distributions. Although both were relatively rare in the extreme northern and western portions of the study region, white oaks tended to be most abundant in the southern portion of the study region, whereas red oaks were most abundant at mid-latitudes (Fig. 2a, b). The other preferred hosts of *L. dispar* (mainly birches and aspens) were generally most abundant where oak densities were low, in the extreme north and western portions of the study region (Fig. 2c). Intermediate hosts were fairly evenly distributed, except that their

**Table 1** Percentages of the basal area of all forest trees composed of preferred *Lymantria dispar* hosts (subdivided into red oaks, white oaks, and other preferred hosts), intermediate hosts, and trees immune to *L. dispar* feeding (boldface), and the three most dominant species within each group

| Host and immune tree groups and their dominant species | % of basal area of all forest trees |
|---|---|
| Red oaks (section Lobatae) | **7.7** |
| *Quercus rubra* | 5.6 |
| *Quercus velutina* | 1.3 |
| *Quercus coccinea* | 0.7 |
| White oaks (section Lepidobalanus) | **6.0** |
| *Quercus alba* | 3.6 |
| *Quercus prinus* | 2.2 |
| *Quercus bicolor* | 0.1 |
| Other preferred host trees | **6.0** |
| *Betula papyrifera* | 2.2 |
| *Populus tremuloides* | 1.1 |
| *Populus grandidentata* | 0.8 |
| Intermediate host trees | **59.6** |
| *Acer rubrum* | 15.7 |
| *Acer saccharum* | 8.7 |
| *Tsuga canadensis* | 7.6 |
| Immune trees | **20.6** |
| *Abies balsamea* | 5.8 |
| *Betula alleghiensis* | 3.9 |
| *Fraxinus americana* | 3.7 |

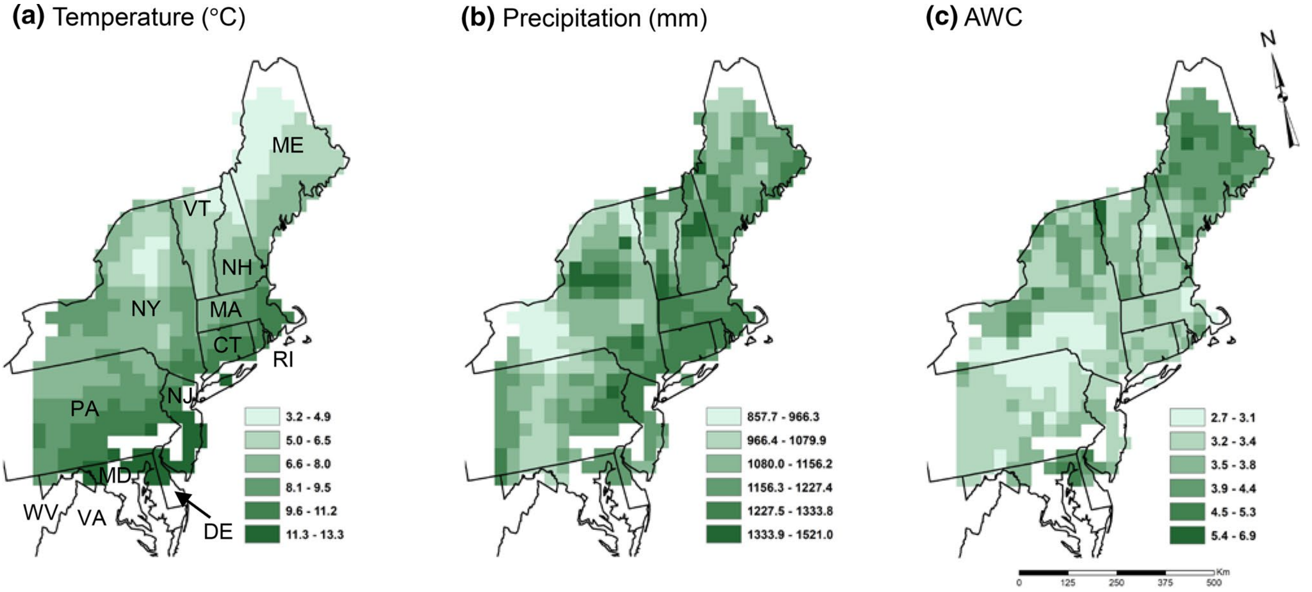

**Fig. 1** Maps of **a** mean annual temperature (°C), **b** mean annual precipitation (mm), and **c** mean available water capacity (AWC) of the soil. State abbreviations: CT, Connecticut; DE, Delaware; ME, Maine; MD, Maryland; MA, Massachusetts; NH, New Hampshire; NJ, New Jersey; PA, NY, New York; Pennsylvania; RI, Rhode Island; VT, Vermont; VA, Virginia; WV, West Virginia

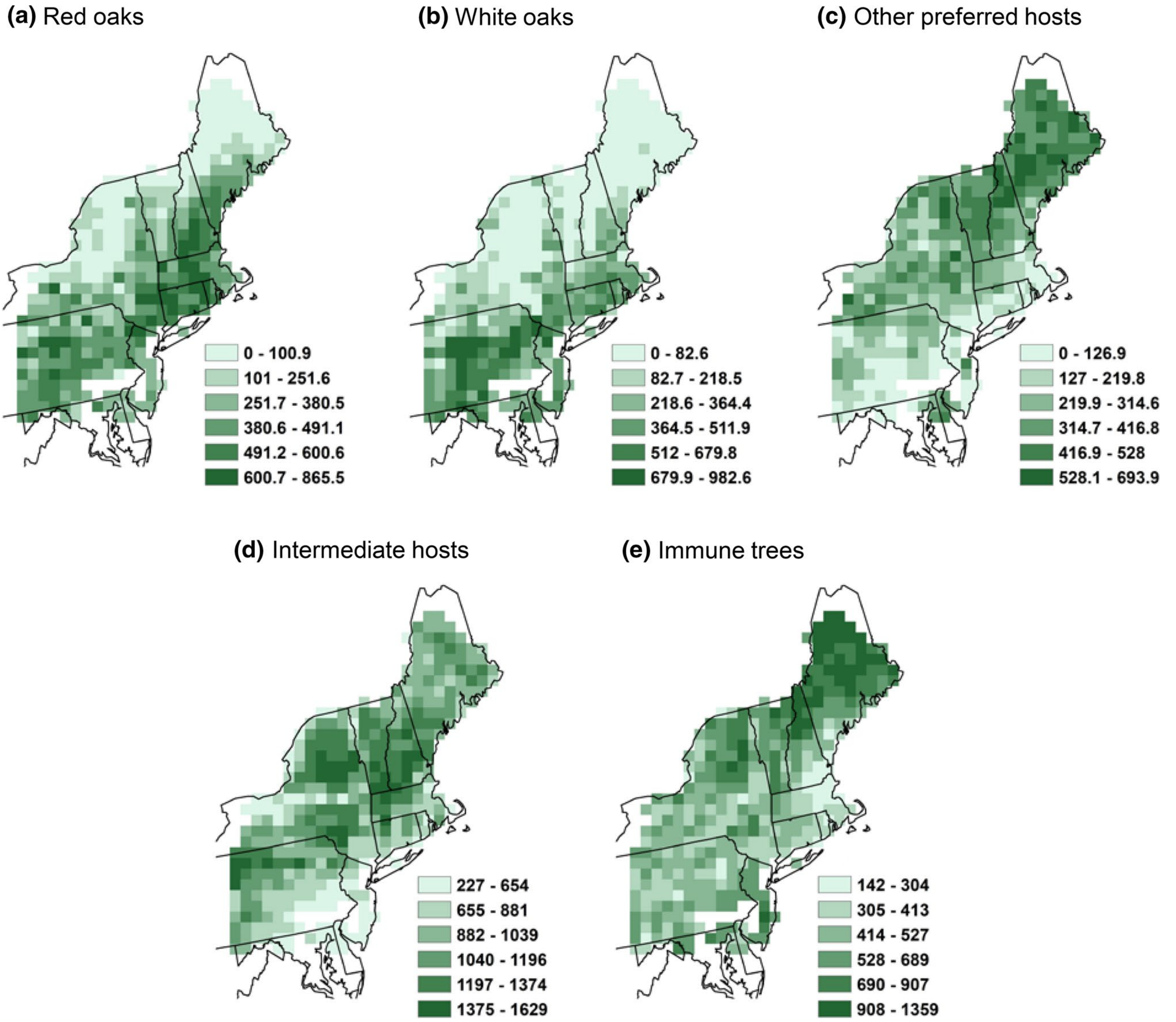

**Fig. 2** Maps of the mean basal areas (m$^2$) of **a** red oaks (Lobatae), **b** white oaks (Lepidobalanus), **c** other preferred *Lymantria dispar* hosts, **d** intermediate host trees, and e) trees that are immune to *L. dispar* feeding. All displayed values are square-root transformed

densities were low in the southeastern portion of the study region (Fig. 2d). Tree species that are immune to *L. dispar* feeding occurred throughout the study region but reached their highest densities in the extreme north (Fig. 2e).

Our original structural equation model contained excessive multicollinearity (VIF > 10) due to correlations between forest area and the basal area of intermediate hosts and temperature. Understanding effects of temperature was a central goal of our study, and accounting for differences in amount of forest area among quadrats was necessary to eliminate the bias for cells with more forested area to have more area defoliated. In contrast, the basal area of intermediate host trees was not expected to strongly affect *L. dispar* densities. Thus, we reduced multicollinearity to an acceptable level (maximum VIF = 8.48) by removing intermediate host trees from the structural equation model. Its removal resulted in no qualitative differences (See Appendix S4: Fig. S4).

The simplified structural equation model (intermediate hosts removed) fit the data well (based on our two metrics of goodness-of-fit: $C_2 = 0.74$, $P = 0.69$; and $\chi^2_1 = 0.16$, $P = 0.69$), and revealed both direct and indirect relationships between climate and area of forest defoliated by *L. dispar* (Fig. 3). The model explained an extremely high proportion of variance ($R^2 = 0.85$) in defoliation. We found a significant and positive direct relationship between temperature and defoliation, i.e., areas with higher temperatures (Fig. 1a) tended to have increased defoliation (Fig. 4, Table 1). Similarly, increasing temperature increased the abundance of

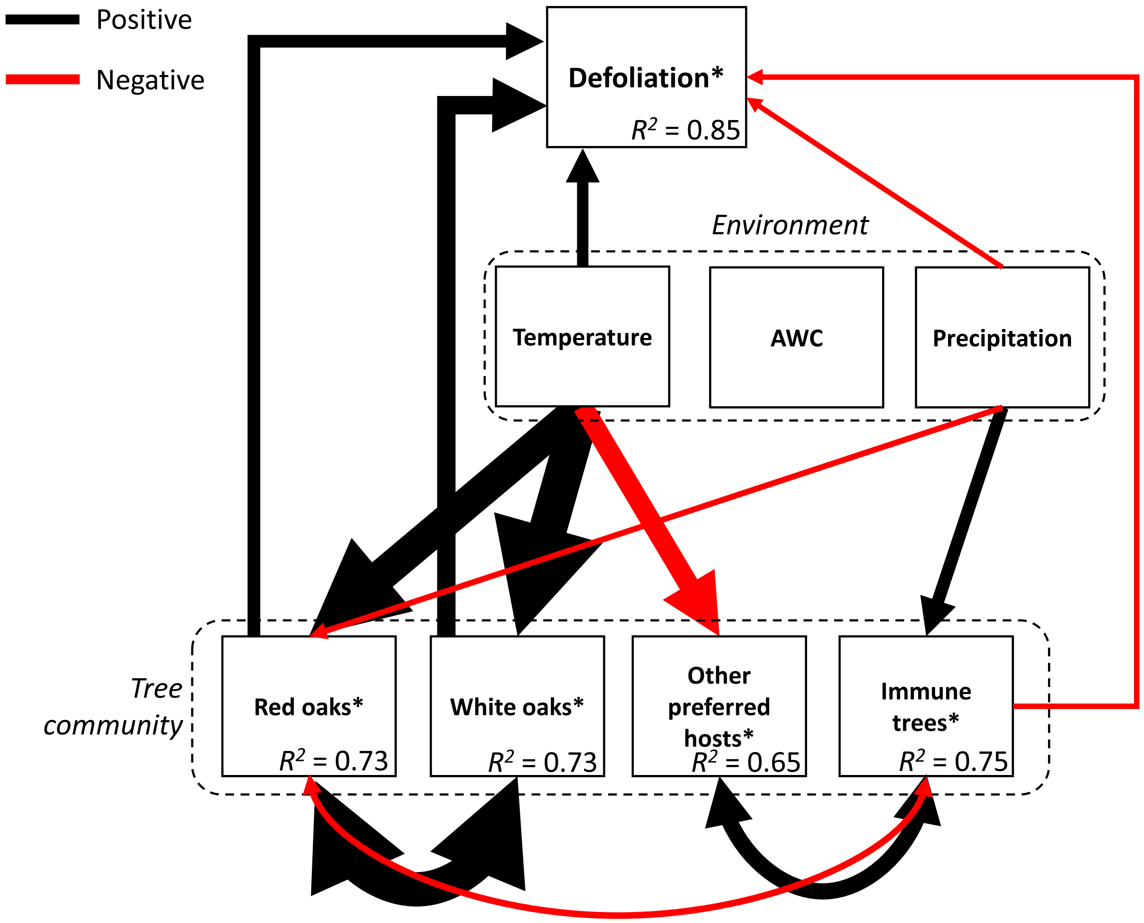

**Fig. 3** Structural equation model examining relationships between (square-root transformed) mean annual spatial extent of *Lymantria dispar* defoliation (km$^2$), climate (mean annual temperature and precipitation), soil available water capacity (AWC), and (square-root transformed) basal areas (m$^2$) of red oaks, white oaks, other preferred *L. dispar* hosts, intermediate hosts, and trees that are immune to *L. dispar* feeding. Straight arrows represent hypothesized causal paths that were statistically significant ($P < 0.05$), whereas curved arrows represent correlations between categories of trees that were significant. Non-significant paths and correlations are not shown. To indicate the strength of each relationship, the thicknesses of the arrows were scaled based on their standardized regression coefficients. The $R^2$ values for each component model are displayed in the boxes of each response variable. The covariate forest area (km$^2$) was used as a predictor variable in all variables representing areas: tree group basal areas and defoliated area. For each of these predictor variables, asterisks indicate if forest area was a significant predictor

both red and white oak hosts approximately equally (based on comparison of their standardized regression coefficients) and reduced other hosts to a lesser degree (Table 2).

The indirect effect of temperature on defoliation through tree groups—computed by multiplying the standardized regression coefficients of the adjacent paths—showed that the effects of temperature on defoliation were stronger through white oaks (0.572 × 0.279 = 0.160) than red oaks (0.489 × 0.193 = 0.094). Critically, the combined indirect effect of temperature through both sections of oaks (0.160 + 0.094 = 0.254) was stronger than the direct effect of temperature on defoliation (0.172).

Further, we found a negative effect of precipitation on *L. dispar* defoliation, but precipitation was a weaker predictor of both defoliation and tree abundances than was temperature (Table 2, Fig. 3). However, increasing precipitation was associated with increasing basal area of immune trees, and, in turn, the basal area of these trees had a significant negative effect on *L. dispar* defoliation (Table 2, Fig. 3). Based on standardized regression coefficients, the direct effect of precipitation on defoliation was weak (−0.081), and the indirect effect of precipitation on defoliation through its effect on immune trees was even weaker (0.190 × −0.084 = −0.017). Soil water holding capacity was not significantly related to the abundances of any tree group. The covariate, forest area (km$^2$) in each grid cell, as expected, was a significant predictor of the abundances of all tree groups and *L. dispar* defoliation (Table 2).

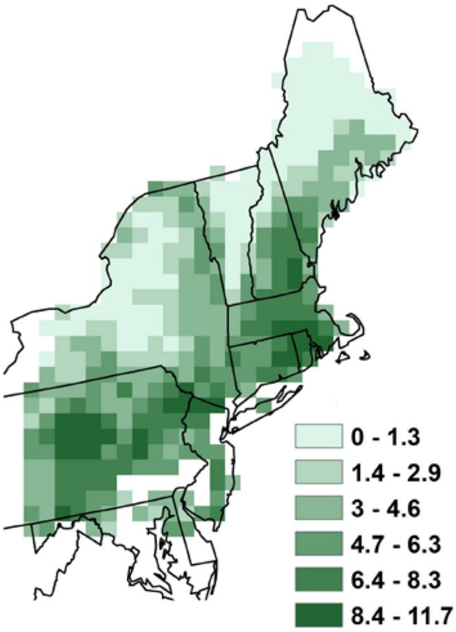

**Fig. 4** Map of the (square-root transformed) mean annual spatial extent (km$^2$) of *Lymantria dispar* defoliation from 1975 to 2018

## Discussion

Climatic conditions, mainly through the effects of temperatures on phenology, are major drivers of the distributions of temperate tree species distributions in North America (Morin et al. 2007). Climate is also thought to constrain the distribution and outbreak ranges of defoliating insect species through its direct effects on insect survival, phenology, or reproduction (Martinat 1987; Candau and Fleming 2005; Johnson et al. 2010; Battisti and Larsson 2015). Using piecewise structural equation modeling (pSEM), we revealed novel quantitative evidence that geographic variation in climate (primarily mean annual temperature) indirectly affects the spatial extent of forest defoliation by a forest insect via its effects on tree species composition more than it directly affects insect populations. Our finding suggests that, while climate has an overarching role in determining the mean spatial extent of *L. dispar* outbreaks, it does this primarily by altering the degree to which forests are composed of preferred host species. Although climate has long been known to be a major determinant of tree species composition, the extent to which climate affects the extent of forest insect

**Table 2** Results of spatial error models examining putative causal relationships between climate variables (30-year annual mean temperatures and precipitation), basal areas (m$^2$) of trees categorized by *Lymantria dispar* feeding preferences or their effects on populations of *L. dispar* predators (red oaks vs. white oaks), area of forest (km$^2$), and mean area (km$^2$) of forest defoliated by *L. dispar* per year

| Response variable | Predictor | Coef | SE | *z* value | Std. Coef | *P* Lobatae), white oaks (Lepidobalanus)alue |
|---|---|---|---|---|---|---|
| Red oaks | Temperature | 47.720 | 6.697 | 7.126 | 0.488 | **< 0.001** |
| Red oaks | Precipitation | − 0.180 | 0.086 | − 2.090 | − 0.091 | **0.037** |
| Red oaks | AWC | 1.167 | 13.957 | 0.084 | 0.003 | 0.933 |
| Red oaks | Forest area | 0.298 | 0.040 | 7.509 | 0.317 | **< 0.001** |
| White oaks | Temperature | 62.099 | 7.269 | 8.543 | 0.572 | **< 0.001** |
| White oaks | Precipitation | − 0.184 | 0.096 | − 1.928 | − 0.084 | 0.054 |
| White oaks | AWC | 14.842 | 15.708 | 0.945 | 0.034 | 0.345 |
| White oaks | Forest area | 0.320 | 0.045 | 7.139 | 0.307 | **< 0.001** |
| Other preferred hosts | Temperature | − 20.908 | 5.209 | − 4.014 | − 0.305 | **< 0.001** |
| Other preferred hosts | Precipitation | 0.015 | 0.068 | 0.220 | 0.011 | 0.826 |
| Other preferred hosts | AWC | − 18.935 | 11.182 | − 1.693 | − 0.069 | 0.090 |
| Other preferred hosts | Forest area | 0.117 | 0.032 | 3.658 | 0.177 | **< 0.001** |
| Immune | Temperature | 1.568 | 7.082 | 0.221 | 0.015 | 0.825 |
| Immune | Precipitation | 0.399 | 0.0903 | 4.416 | 0.190 | **< 0.001** |
| Immune | AWC | 11.791 | 14.565 | 0.810 | 0.029 | 0.418 |
| Immune | Forest area | 0.413 | 0.041 | 9.979 | 0.416 | **< 0.001** |
| Defoliation | Red oaks | 0.003 | 0.001 | 4.376 | 0.193 | **< 0.001** |
| Defoliation | White oaks | 0.003 | 0.001 | 6.501 | 0.279 | **< 0.001** |
| Defoliation | Other preferred hosts | < 0.001 | 0.001 | 0.449 | 0.015 | 0.653 |
| Defoliation | Immune | − 0.001 | 0.001 | − 2.042 | − 0.084 | **0.041** |
| Defoliation | Temperature | 0.219 | 0.074 | 2.949 | 0.172 | **0.003** |
| Defoliation | Precipitation | − 0.002 | 0.001 | − 2.42 | − 0.081 | **0.016** |
| Defoliation | Forest area | 0.001 | 0.001 | 2.551 | 0.104 | **0.011** |

The basal areas of each tree group, area of forest defoliated, and available water capacity of soils (AWC) were square-root transformed outbreaks directly versus indirectly through its effects on forest composition is rarely if ever known. While effects of forest management (e.g., selective harvesting) on tree-species composition cannot be completely discounted, it is unlikely that it explains the geographic variation in tree distributions observed in this study. Forest-stand treatments were limited to an estimated 9% of the forest land across our study area within 5 years of the forest inventory plot visits (USDA Forest Service 2021).

Defoliation increased with the abundances of both red and white oaks, consistent with previous research indicating that high oak abundance predisposes forest stands to *L. dispar* defoliation (Herrick et al. 1986; Liebhold et al. 1994). Despite the greater importance of red oaks, compared to white oaks, in affecting the population cycles of the white-footed mouse (Wolff 1996a; Elkinton et al. 1996; Jones et al. 1998; Elias et al. 2004; Clotfelter et al. 2007), the chief predator of *L. dispar* (Elkinton and Liebhold 1990; Elkinton et al. 1996), the effects of red-oak abundance and white-oak abundance on *L. dispar* defoliation were approximately equal (Table 1). While temporal variation in small mammal populations may profoundly affect *L. dispar* dynamics (Elkinton et al. 1996), spatial variation in mammal abundance may play a lesser role in determining spatial variation in *L. dispar* densities. In contrast, densities of hosts, either white or red oaks, may strongly drive spatial variation in *L. dispar* abundance.

The positive direct relationship detected here between mean annual temperature and defoliation (Fig. 3) suggests a similarly positive relationship between temperature and *L. dispar* population growth. Considerable information exists about direct effects of temperature on *L. dispar* population dynamics. Low winter temperatures are known to decrease survival of over-wintering egg masses (Andresen et al. 2001). Conversely, high summer temperatures adversely affect the development and survival of *L. dispar* larvae (Thompson et al. 2017), but the portion of *L. dispar*'s range where temperatures regularly become hot enough to negatively affect its larvae, *L. dispar*'s southeastern range edge (the coastal plain of Virginia), is outside of the current study region. Thus, it is not surprising that we only detected a positive effect of increasing temperature on defoliation. Though our analysis indicates a direct positive influence of temperature on defoliation (Fig. 3), it is possible that some of this relationship arose through indirect impacts on natural enemy populations. For example, Reilly et al. (2014) found a negative effect of increasing temperature on larval mortality caused by the entomopathogenic fungus *Entomophaga maimaiga* and this could conceivably contribute to a positive association of temperature with defoliation.

One possible explanation for the fact that defoliation did not increase with the abundance of other, non-oak host trees is that these other species represent a relatively small fraction of host tree basal area (~30%). In addition, the highest abundances of these other preferred host trees (mainly aspens and birches) tended to occur in colder portions of the study region (Figs. 1, 2). In our study region, much of the state of Maine, and smaller northern or high-elevation portions of New Hampshire, Vermont, and New York were predicted to have low climate suitability for *L. dispar* due to a lack of degree days sufficient for *L. dispar* to complete prediapause development prior to the onset of winter (Gray 2004). In addition, cold winter temperatures along the U.S.-Canadian border (in the Great Lakes region) have been found cause high rates of overwintering mortality in *L. dispar* egg masses (Streifel et al. 2019). That aspens and birches tend to occur where the climate is of low suitability for *L. dispar* may explain why these species have not been defoliated as frequently and extensively as oaks.

Consistent with Jactel and Brockerhoff (2007) finding that increasing tree diversity reduces herbivory on focal species in most study systems, higher abundances of trees that are immune to *L. dispar* feeding leads to a reduction in defoliation by *L. dispar*. In our study region, there was a significant positive relationship between the abundances of immune trees and annual precipitation. Associational resistance from immune trees, combined with the effect of spatial variation in climate on the abundance of these trees, indicates that regional climate variation may influence the strength of associational resistance (Fig. 3).

One possible explanation for why associational resistance could occur with an herbivore species as highly polyphagous as *L. dispar* is this species' poor dispersal ability. Adult females are flightless, and passive wind-borne dispersal of neonates ("ballooning") is the primary means of larval dispersal (Mason and McManus 1981; Liebhold and Tobin 2006). Ballooning larvae may be less likely to be transported to a preferred host-tree species in forest stands with high densities of immune trees. Given the limited dispersal ability of *L. dispar*, however, further research carried out at finer spatial resolution than the 32 km raster cells used in this study is probably necessary to fully evaluate whether host-tree species that are preferred by *L. dispar* can benefit from associational resistance (i.e., occurring near immune trees).

Our finding that higher long-term mean temperatures had positive direct and indirect effects on the mean annual spatial extent of *L. dispar* outbreaks suggests that future climate warming may cause increased *L. dispar* defoliation within our study region. In using predictions from spatial regression models to make predictions about effects of environmental changes through time, we are carrying out a 'space-for-time substitution', a long-used practice in ecology that nonetheless can lead to erroneous conclusions because it assumes spatially replicated data represent measures of local equilibrium conditions and that site-level temporal dynamics can be ignored (Damgaard 2019). Our predictions about the effects of climate change on forest defoliation should be interpreted with caution, but nonetheless they represent fruitful avenues for future research.

The strongest effects of temperature on defoliation were indirect effects mediated by positive effects of temperature on oak abundance. But given that the region is undergoing a decline in oak dominance, due to a host of factors (e.g., fire suppression, deer browsing, invasive species; (Horsley et al. 2003; Nowacki and Abrams 2008; Dey 2014) including reduced growth and increased mortality of oaks due to *L. dispar* defoliation (Morin and Liebhold 2016), the most potent pathway through which rising temperatures could increase *L. dispar* defoliation—by increasing oak abundance—may be blocked. However, it is possible that increasing temperatures along the northern edges of the study region could make these areas more climatically suitable for *L. dispar* evidenced by phenological modeling (Gray 2004) and by the direct effect of increasing temperature on defoliation found in this study. As the northern portions of the study region are likely to become more suitable for *L. dispar* with continued warming, aspen and birch species that remain abundant in these areas in future decades may become more susceptible to frequent and widespread defoliation by *L. dispar*.

The results of our study have broader implications for our understanding of how climate, and changes in climate, might influence the intensity of insect outbreaks. Our results suggest that changes in climate may have greater indirect influences (via effects on the tree-species composition of forests) on forest defoliator outbreaks than direct influences (e.g., on insect physiology or survival). Because changes in forest composition can be expected to be lagged in response to climate change (Bertrand et al. 2011; Schneider et al. 2015), changes in defoliator outbreak intensity likely will not immediately follow changes in climate, but instead the changes in outbreaks may be quite delayed. Developing accurate long-term projections of changes in forest composition, therefore, is key to predicting the intensity of defoliator outbreaks in the coming decades.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s00442-022-05123-w.

**Acknowledgements** We thank Eugene Luzader for assistance with data. AML received support from grant EVA4.0, No. CZ.02.1.01/0.0/0.0/16_019/0000803 financed by OP RDE.

**Author contributions statement** KH and AL conceived the study, KH, AL, JL and RM developed the methodology, JL and KJ performed the statistical analyses, and all authors wrote the manuscript.

**Funding** AML received support from grant EVA4.0, No. CZ.02.1.01/0.0/0.0/16_019/0000803 financed by OP RDE.

**Availability of data and materials** Upon acceptance, the datasets will be deposited in Dryad.

## Declarations

**Conflict of interest** Not applicable.

**Ethics approval** Not applicable.

**Consent to participate** Not applicable.

**Consent for publication** Not applicable.

## References

Andresen JA, McCullough DG, Potter BE et al (2001) Effects of winter temperatures on gypsy moth egg masses in the Great Lakes region of the United States. Agric for Meteorol 110:85–100. https://doi.org/10.1016/S0168-1923(01)00282-9

Atsatt PR, O'dowd DJ (1976) Plant defense guilds. Science 193:24–29. https://doi.org/10.1126/science.193.4247.24

Battisti A, Larsson S (2015) Climate change and insect pest distribution range. In: Björkman C, Niemelä P (eds) Climate change and insect pests. CABI, Boston, MA, pp 1–15

Bechtold WA, Patterson PL (eds) (2005) The enhanced forest inventory and analysis program—national sampling design and estimation procedures. https://doi.org/10.2737/SRS-GTR-80

Berryman AA (1996) What causes population cycles of forest Lepidoptera? Trends Ecol Evol (amst) 11:28–32. https://doi.org/10.1016/0169-5347(96)81066-4

Bertrand R, Lenoir J, Piedallu C et al (2011) Changes in plant community composition lag behind climate warming in lowland forests. Nature 479:517–520. https://doi.org/10.1038/nature10548

Bivand R, Piras G (2015) Comparing implementations of estimation methods for spatial econometrics. J Stat Softw 63:1–36. https://doi.org/10.18637/jss.v063.i18

Bivand RS, Wong DWS (2018) Comparing implementations of global and local indicators of spatial association. TEST 27:716–748. https://doi.org/10.1007/s11749-018-0599-x

Campbell RW (1974) The gypsy moth and its natural enemies. USDA Forest Service, Washington, DC, USA

Candau J-N, Fleming RA (2005) Landscape-scale spatial distribution of spruce budworm defoliation in relation to bioclimatic conditions. Can J for Res 35:2218–2232. https://doi.org/10.1139/x05-078

Candau J-N, Fleming RA (2011) Forecasting the response of spruce budworm defoliation to climate change in Ontario. Can J for Res 41:1948–1960. https://doi.org/10.1139/x11-134

Clotfelter ED, Pedersen AB, Cranford JA et al (2007) Acorn mast drives long-term dynamics of rodent and songbird populations. Oecologia 154:493–503. https://doi.org/10.1007/s00442-007-0859-z

Damgaard C (2019) A Critique of the space-for-time substitution practice in community ecology. Trends Ecol Evol (amst) 34:416–421. https://doi.org/10.1016/j.tree.2019.01.013

Davidson CB, Johnson JE, Gottschalk KW, Amateis RL (2001) Prediction of stand susceptibility and gypsy moth defoliation in Coastal Plain mixed pine–hardwoods. Can J for Res 31:1914–1921. https://doi.org/10.1139/x01-124

Dey DC (2014) Sustaining oak forests in eastern North America: regeneration and recruitment, the pillars of sustainability. For Sci 60:926–942. https://doi.org/10.5849/forsci.13-114

Elias SP, Witham JW, Hunter ML (2004) *Peromyscus leucopus* abundance and acorn mast: population fluctuation patterns over 20 years. J Mammal 85:743–747. https://doi.org/10.1644/BER-025

Elkinton J, Liebhold AM (1990) Population dynamics of gypsy moth in North America. Annu Rev Entomol 35:571–596. https://doi.org/10.1146/annurev.ento.35.1.571

Elkinton JS, Healy WM, Buonaccorsi JP et al (1996) Interactions among gypsy moths, white-footed mice, and acorns. Ecology 77:2332–2342. https://doi.org/10.2307/2265735

Elton CA (1927) Animal ecology. Sidgwick and Jackson, London

Esper J, Büntgen U, Frank DC et al (2007) 1200 years of regular outbreaks in alpine insects. Proc Biol Sci 274:671–679. https://doi.org/10.1098/rspb.2006.0191

Fox JF (1982) Adaptation of gray squirrel behavior to autumn germination by white oak acorns. Evolution 36:800. https://doi.org/10.2307/2407893

Gray DR (2004) The gypsy moth life stage model: landscape-wide estimates of gypsy moth establishment using a multi-generational phenology model. Ecol Model 176:155–171. https://doi.org/10.1016/j.ecolmodel.2003.11.010

Gray DR (2008) The relationship between climate and outbreak characteristics of the spruce budworm in eastern Canada. Clim Change 87:361–383

Guo Q, Fei S, Potter KM et al (2019) Tree diversity regulates forest pest invasion. Proc Natl Acad Sci USA 116:7382–7386. https://doi.org/10.1073/pnas.1821039116

Hadj-Chikh LZ, Steele MA, Smallwood PD (1996) Caching decisions by grey squirrels: a test of the handling time and perishability hypotheses. Anim Behav 52:941–948. https://doi.org/10.1006/anbe.1996.0242

Hartl-Meier C, Esper J, Liebhold A et al (2017) Effects of host abundance on larch budmoth outbreaks in the European Alps. Agric for Entomol 19:376–387. https://doi.org/10.1111/afe.12216

Herrick OW, Gansner DA, Gansner DA (1986) Rating forest stands for gypsy moth defoliation. Research Paper NE-583. US Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, Broomall, PA, p 4

Horsley SB, Stout SL, deCalesta DS (2003) White-tailed deer impact on the vegetation dynamics of a northern hardwood forest. Ecol Appl 13:98–118. https://doi.org/10.1890/1051-0761(2003)013[0098:WTDIOT]2.0.CO;2

Iverson LR, Prasad AM (1998) Predicting abundance of 80 tree species following climate change in the eastern united states. Ecol Monogr 68:465–485. https://doi.org/10.1890/0012-9615(1998)068[0465:PAOTSF]2.0.CO;2

Jactel H, Brockerhoff EG (2007) Tree diversity reduces herbivory by forest insects. Ecol Lett 10:835–848. https://doi.org/10.1111/j.1461-0248.2007.01073.x

Jactel H, Moreira X, Castagneyrol B (2021) Tree diversity and forest resistance to insect pests: patterns, mechanisms, and prospects. Annu Rev Entomol 66:277–296. https://doi.org/10.1146/annurev-ento-041720-075234

Jepsen JU, Hagen SB, Ims RA, Yoccoz NG (2008) Climate change and outbreaks of the geometrids *Operophtera brumata* and *Epirrita autumnata* in subarctic birch forest: evidence of a recent outbreak range expansion. J Anim Ecol 77:257–264. https://doi.org/10.1111/j.1365-2656.2007.01339.x

Johnson DM, Büntgen U, Frank DC et al (2010) Climatic warming disrupts recurrent Alpine insect outbreaks. Proc Natl Acad Sci USA 107:20576–20581. https://doi.org/10.1073/pnas.1010270107

Jones CG, Ostfeld RS, Richard MP et al (1998) Chain reactions linking acorns to gypsy moth outbreaks and Lyme disease risk. Science 279:1023–1026. https://doi.org/10.1126/science.279.5353.1023

Koricheva J, Vehviläinen H, Riihimäki J et al (2006) Diversification of tree stands as a means to manage pests and diseases in boreal forests: myth or reality? Can J for Res 36:324–336. https://doi.org/10.1139/x05-172

Lefcheck JS (2016) Piecewisesem: piecewise structural equation modelling in R for ecology, evolution, and systematics. Methods Ecol Evol 7:573–579. https://doi.org/10.1111/2041-210X.12512

Liebhold AM, Tobin PC (2006) Growth of newly established alien populations: comparison of North American gypsy moth colonies with invasion theory. Popul Ecol 48:253–262. https://doi.org/10.1007/s10144-006-0014-4

Liebhold AM, Elmes GA, Halverson JA, Quimby J (1994) Landscape characterization of forest susceptibility to gypsy moth defoliation. For Sci 40:18–29

Liebhold AM, Gottschalk KW, Muzika RM et al (1995) Suitability of North American tree species: a summary of field and laboratory tests. USDA, PA

Liebhold AM, Gottschalk KW, Luzader ER et al (1997) Gypsy moth in the United States, an atlas. USDA Forest Service Radnor, PA

Lovett GM, Christenson LM, Groffman PM et al (2002) Insect defoliation and nitrogen cycling in forests. Bioscience 52:335. https://doi.org/10.1641/0006-3568(2002)052[0335:IDANCI]2.0.CO;2

Martinat PJ (1987) The role of climatic variation and weather in forest insect outbreaks. In: Barbosa P, Schultz J (eds) Insect outbreaks. Academic Press, New York, pp 241–268

Mason CJ, McManus ML (1981) Larval dispersal of the gypsy moth. The gypsy moth: research toward integrated pest management. U.S. Department of Agriculture, Washington, DC, pp 161–202

Morin RS, Liebhold AM (2016) Invasive forest defoliator contributes to the impending downward trend of oak dominance in eastern North America. Forestry 89:284–289. https://doi.org/10.1093/forestry/cpv053

Morin X, Augspurger C, Chuine I (2007) Process-based modeling of species' distributions: what limits temperate tree species' range boundaries? Ecology 88:2280–2291. https://doi.org/10.1890/06-1591.1

Nowacki GJ, Abrams MD (2008) The demise of fire and "mesophication" of forests in the eastern united states. Bioscience 58:123–138. https://doi.org/10.1641/B580207

Pimentel D (1961) Species diversity and insect population outbreaks. Ann Entomol Soc Am 54:76–86. https://doi.org/10.1093/aesa/54.1.76

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Reilly JR, Hajek AE, Liebhold AM, Plymale R (2014) Impact of *Entomophaga maimaiga* (Entomophthorales: Entomophthoraceae) on outbreak gypsy moth populations (Lepidoptera: Erebidae): the role of weather. Environ Entomol 43:632–641. https://doi.org/10.1603/EN13194

Robert L-E, Sturtevant BR, Kneeshaw D (2020) Forest landscape structure influences the cyclic-eruptive spatial dynamics of forest test caterpillar outbreaks. Ecosphere 11:e03096. https://doi.org/10.1002/ecs2.3096

Roland J, Mackey BG, Cooke B (1998) Effects of climate and forest structure on duration of forest tent caterpillar outbreaks across Central Ontario, Canada. Can Entomol 130:703–714. https://doi.org/10.4039/Ent130703-5

Schneider RR, Devito K, Kettridge N, Bayne E (2015) Moving beyond bioclimatic envelope models: integrating upland forest and peatland processes to predict ecosystem transitions under climate change in the western Canadian boreal plain. Ecohydrology 9:899–908. https://doi.org/10.1002/eco.1707

Seidl R, Schelhaas M-J, Lexer MJ (2011) Unraveling the drivers of intensifying forest disturbance regimes in Europe. Glob Change Biol 17:2842–2852. https://doi.org/10.1111/j.1365-2486.2011.02452.x

Shimada T, Saitoh T (2006) Re-evaluation of the relationship between rodent populations and acorn masting: a review from the aspect of nutrients and defensive chemicals in acorns. Popul Ecol 48:341–352. https://doi.org/10.1007/s10144-006-0012-6

Shipley B (2009) Confirmatory path analysis in a generalized multilevel context. Ecology 90:363–368. https://doi.org/10.1890/08-1034.1

Shipley B, Douma JC (2020) Generalized AIC and chi-squared statistics for path models consistent with directed acyclic graphs. Ecology 101:e02960. https://doi.org/10.1002/ecy.2960

Sholes ODV (2008) Effects of associational resistance and host density on woodland insect herbivores. J Anim Ecol 77:16–23. https://doi.org/10.1111/j.1365-2656.2007.01317.x

Smith HR (1985) Wildlife and the gypsy moth. Wildl Soc Bull 13:166–174

Steele MA, Turner G, Smallwood PD et al (2001) Cache management by small mammals: experimental evidence for the significance of acorn-embryo excision. J Mammal 82:35–42. https://doi.org/10.1644/1545-1542(2001)082%3c0035:CMBSME%3e2.0.CO;2

Streifel MA, Tobin PC, Kees AM, Aukema BH (2019) Range expansion of *Lymantria dispar dispar* (L.) (Lepidoptera: Erebidae) along its north-western margin in North America despite low predicted climatic suitability. J Biogeogr 46:58–69. https://doi.org/10.1111/jbi.13474

Thompson LM, Faske TM, Banahene N et al (2017) Variation in growth and developmental responses to supraoptimal temperatures near latitudinal range limits of gypsy moth *Lymantria dispar* (L.), an expanding invasive species. Physiol Entomol 42:181–190. https://doi.org/10.1111/phen.12190

Thompson LM, Powers SD, Appolon A et al (2020) Climate-related geographical variation in performance traits across the invasion front of a widespread non-native insect. J Biogeogr. https://doi.org/10.1111/jbi.14005

USDA, NRCS (2019) National soil survey handbook. https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrcs142p2_054242. Accessed 5 Nov 2020

Williams DW, Liebhold AM (1995) Herbivorous insects and global change: potential changes in the spatial distribution of forest defoliator outbreaks. J Biogeogr 22:665. https://doi.org/10.2307/2845968

Wolff JO (1996a) Coexistence of white-footed mice and deer mice may be mediated by fluctuating environmental conditions. Oecologia 108:529–533. https://doi.org/10.1007/BF00333730

Wolff JO (1996b) Population fluctuations of mast-eating rodents are correlated with production of acorns. J Mammal 77:850. https://doi.org/10.2307/1382690